DETERMINATION OF THE CONSENSUS TARGET SEQUENCES RECOGNIZED BY ZNFO, A NOVEL OOCYTE-SPECIFIC ZINC FINGER TRANSCRIPTION FACTOR IN CATTLE

MINGXIANG Zhang

Follow this and additional works at: https://researchrepository.wvu.edu/etd

Recommended Citation
Zhang, MINGXIANG, "DETERMINATION OF THE CONSENSUS TARGET SEQUENCES RECOGNIZED BY ZNFO, A NOVEL OOCYTE-SPECIFIC ZINC FINGER TRANSCRIPTION FACTOR IN CATTLE" (2018). Graduate Theses, Dissertations, and Problem Reports. 7286.
https://researchrepository.wvu.edu/etd/7286

This Thesis is brought to you for free and open access by The Research Repository @ WVU. It has been accepted for inclusion in Graduate Theses, Dissertations, and Problem Reports by an authorized administrator of The Research Repository @ WVU. For more information, please contact ian.harmon@mail.wvu.edu.
DETERMINATION OF THE CONSENSUS TARGET SEQUENCES RECOGNIZED BY ZNFO, A NOVEL OO CYTE-SPECIFIC ZINC FINGER TRANSCRIPTION FACTOR IN CATTLE

MINGXIANG ZHANG

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

Master of Science
in Reproductive Physiology

Jianbo Yao, Ph.D., Chair
Ida Holásková, Ph.D.
Vagner Benedito, Ph.D.

Division of Animal and Nutritional Sciences
Morgantown, West Virginia

2018

KEYWORDS: ZNFO, zinc finger motif, ZNFO binding element, oocyte, cattle, embryogenesis, maternal-effect gene

Copyright 2018 Mingxiang Zhang
ABSTRACT

Determination of the Consensus Target Sequences Recognized by ZNFO, A Novel Oocyte-specific Zinc Finger Transcription Factor in Cattle

Mingxiang Zhang

Most of the early embryonic loss in livestock species occurs at or close to maternal to embryonic transition (MET), a process in which embryos switch from maternal factors to the factors encoded by their own genome. Maternal factors stored at oocytes during the process of folliculogenesis as well as oocyte maturation have essential role in regulation of early development including MET. Zinc finger proteins are one of the largest protein families in eukaryotes. The classic two cysteine and two histidine residue (C2H2) zinc finger proteins compose the largest transcription factor repertoire in mammals. Around one-third of them have a N-terminal KRAB domain. The zinc finger domain can make contact with DNA to locate zinc finger proteins in certain positions of the genome while the KRAB domain can interact with KAP1, which serves as a scaffold for various factors. ZNFO is a novel oocyte-specific transcription factor identified from a bovine oocyte library. It is a classic C2H2 zinc finger protein, having an N-terminal KRAB domain and a C-terminal zinc finger domain which is composed of nine zinc finger motifs. Our previous work revealed an essential function of ZNFO during early embryonic development in cattle. Depletion of ZNFO blocks the development of embryos to blastocyst stage. As a maternal transcription factor, ZNFO may repress certain target gene(s) to orchestrate post-fertilization events in cattle. Defect of ZNFO in vivo may cause early embryonic loss in cattle. To investigate the molecular function of ZNFO, we characterized the binding property of ZNFO to DNA. Using a cyclic amplification and selection of targets (CASTing) assay, we identified a potential ZNFO binding element (ZBE), ATATCCTGN$_{5}$ANCCC. To confirm the binding specificity of ZNFO to the identified element, an electrophoretic mobility shift assay (EMSA) was performed using IDye 700 labeled probes containing the target sequence and purified Halo-tagged ZNFO fusion protein. A competitive binding assay was also performed using 10-fold and 100-fold molar excess of unlabeled cold competitors containing the target element and two mutated elements. The results confirmed the interaction between ZNFO and ZBE, and showed that both ATATCCTG and CCC are critical for the binding of ZNFO to ZBE. Further analysis of promoter regions of candidate bovine genes that contain ZBE may lead to the discovery of specific genes regulated by ZNFO.
Introduction

Fusion of gametes is actually the start of all sexually reproducing organisms. Genetic as well as epigenetic information is passed to next generation through germ cells. Primordial germ cells (PGCs), which are originally from proximal epiblast cells, are the common precursor of both female and male gametes. In mammals, the germ cell fate is not pre-determined but induced through a model called epigenesis\textsuperscript{1}, through which PGCs specification is induced by extraembryonic signal. During germ cell specification, the integration of three key events determines the germ cell fate: 1) repression of mesoderm genes, 2) reacquisition of pluripotency as well as 3) genome-wide epigenetic reprogramming. In mice, a tripartite transcription factor complex including BLIMP1, also known as PRDM1\textsuperscript{2-3}, PRDM14\textsuperscript{4}, and AP2γ\textsuperscript{5}, has been identified as a key regulatory machinery. This tripartite transcription factor complex can be induced by either BMP-SMAD pathway or WNT3-β CATENIN pathway. These three genes together can repress somatic genes, while PRDM1 and PRDM14 are responsible for upregulation of germ cell and pluripotency genes as well as the initiation of genome-wide reprogramming and germ cell migration\textsuperscript{6}.

PGCs will migrate from the extraembryonic tissues to sexually indifferent gonadal ridges developing inside embryo, and eventually differentiate into sperm or oocyte depending on the chromosome. Even though differentiated from the common precursor, sperm differs significantly from oocyte in structure. For its mobility, sperm lose its cytoplasm. During the process of fertilization, sperm delivers paternal nucleus to oocyte, which contributes both maternal nucleus and cytoplasm, harboring various maternal factors. Maternal factors define an essential aspect of oocyte quality. Absence of certain maternal factor(s) in oocyte may reduce quality of oocyte causing early embryonic loss. Post-fertilization events highly depend on maternal factors.

ZNFO is a maternal transcription factor expressed exclusively in bovine oocyte. As other KRAB-containing zinc finger proteins, ZNFO has a conserved N-terminal KRAB domain. In addition, it also has a nine-motif zinc finger domain, which can make contact with DNA. GAL4 luciferase assay showed that ZNFO has an intrinsic repressive function\textsuperscript{7}. ZNFO also has an essential role in regulating early embryonic development in cattle, which was demonstrated by RNA inference mediated knock-down assay\textsuperscript{7}. ZNFO may repress its target gene(s) to orchestrate post-fertilization events during early embryonic development in cattle. Any defect of ZNFO in vivo may result in embryonic mortality in cattle. In order to elucidate the molecular function of ZNFO, we induced the expression of zinc finger domain of ZNFO (ZNFO-ZNFs), which was fused with a His\textsubscript{6} Halo tag. Then we performed a CASTing assay followed by gel shift assay to identify the consensus sequence recognized by ZNFO as ZNFO binding element (ZBE).
DEDICATION

To the one I have always admired: my father.
ACKNOWLEDGEMENTS

Foremost, I would like to show my deepest gratitude to my advisor Dr. Jianbo Yao. In the past several years, he has continually provided me his wisdom, encouragement, humor, teachings, and support. He is not only my mentor of science but also my mentor of life. With his great support as well as patience, I gradually catch what my science is and how to deal with frustration.

Second, I want to express my gratitude to my committee members: Dr. Vagner Benedito and Dr. Ida Holaskova for serving on my graduate committee and for each providing me their own teachings and support throughout my master study.

My thanks also go to my lovely lab mates, Dr. Jian Wang, Dr. Liyuan Fu, and Dr. Prasanthi Koganti for their great support and discussions as well as dear friendship during my graduate life. This great lab family make me grow and teaches me to be stronger. I really miss the time where everyone was in the office.

Lastly, and ever sincerely, I want to thanks my family members far away in China. I received a great support from my parents when I decided to purse my degree far away from home. Science always gave me a hard time and I always received encouragement from my parents. I also want to thanks my girlfriend, Yuexin Wang, who is also far away in China, but her voice was useful to cure my frustration.

Further, the financial supports from the USDA and the support of the Division of Animal and Nutritional Sciences at West Virginia University are greatly appreciated.
Identification of the DNA Binding Element for ZNFO .................................................. 22
Confirmation of Target Consensus of ZNFO ............................................................... 22
DISCUSSION .................................................................................................................. 22
FIGURES ....................................................................................................................... 25
TABLES ......................................................................................................................... 32
References ................................................................................................................... 33
LIST OF FIGURES

REVIEW OF LITERATURE-Figure 1. Zinc Finger Structure .................................................. 16
REVIEW OF LITERATURE-Figure 2. KAP1 Structure ....................................................... 17
REVIEW OF LITERATURE-Figure 3. The KRAB-ZFP/KAP1 repressor complex ............. 18

FIGURE 1. Clone of DNA binding domain into pH6HTNHis6Halo® T7 vector .................. 25
FIGURE 2. Expression of His6Halo ZNFO DNA binding domain fusion protein. .............. 26
FIGURE 3. Consensus sequences recognized by ZNFO .................................................... 28
FIGURE 4. Confirmation of Consensus Target Bound to ZNFO ....................................... 29
FIGURE 5. Mutation of Consensus Target Bound to ZNFO ............................................ 30
FIGURE 6. CASTing assay of ZNFO .................................................................................. 31

LIST OF TABLES

Table 1. Primers/oligos Used in This Study .................................................................... 32
REVIEW OF LITERATURE

INTRODUCTION

The journey from a fertilized oocyte to a zygote, and further to an embryo undergoes a successive regulation. The intrinsic quality of oocytes has fundamental impacts on the development of embryos. There are four key levels of oocyte competence after fertilization, among which the ability to develop to the blastocyst stage is considered as key marker of oocyte competence by most laboratories. Most early embryonic developmental blocks occur at or close to the maternal to zygotic transition stage, which happens at the eight-cell stage in cattle and the two-cell stage in mice. It is the maternal contribution from the oocyte that has the most important impact on early embryonic development. Maternal effect genes may affect multiple processes including pronuclear formation and fusion, the first cell division, and embryonic gene transcription. Even though it is suggested that each individual maternal effect gene may have fundamental effect on early embryonic development, the function and mechanism of those genes are largely unknown.

In mammals like eukaryotes, transcription occurs on a chromatin template. Chromatin has a compact organization in which most of the DNA sequence are structurally inaccessible and functionally inactive. DNA cannot be directly packaged into chromatin but through three levels including nucleosome, 10-nm fiber and 30-nm fiber. Among these three levels, nucleosome is the fundamental subunit of chromatin. A mono-nucleosome is a beadlike structure in which around 200bp DNA is organized by a histone-formed octamer. The first step of transcription in mammals is to open the chromatin, which involves disassembling the octamer. Various proteins are required in the process of transcription. In general, those proteins can be classified as two categories: RNA polymerase and transcription factor. A eukaryotic RNA polymerase doesn’t have a sigma factor which can read DNA to find and bind to promoter region. In order to initiate transcription at promoter region in eukaryotes, a RNA polymerase needs various protein to load on DNA. These proteins are defined as basal transcription factors. Thus, in eukaryotes, RNA polymerase together with those basal transcription factors form the basal transcription apparatus. Transcription factors can also bind to other sequence called enhancer, which will determine the expression of promoter or determine whether the expression is all cell types or certain type of cells. The interaction between transcription factor and enhancer can either silence or activate the promoter expression.

Zinc finger is one of the most abundant DNA-binding motifs in eukaryotic transcription factors (TFs). Zinc finger transcription factors (ZFP) play important roles in various cellular functions including cell growth, proliferation, development, apoptosis as well as intracellular signal transduction. A canonical zinc finger motif, the C2H2 motif, has four conserved residues which are two cysteines and two histidines. The KRAB (Kruppel-Associated Box) domain is a potential repression module located at the amino-terminal of most C2H2 zinc finger proteins. And through the interaction with KAP-1 (KRAB-Associated Protein 1), the KAP1-KRAB-ZFP system works as a well-characterized mammalian euchromatin silencer. The vast occurrence of KRAB-ZNF protein in mammalian proteomes, as well as the variation of DNA targets that they can recognize, makes this gene-specific silencer family a potential master repressive regulator during many cellular processes. However, the downstream regulation mechanism of zinc finger protein is still obscure, despite of wide existence of KRAB-ZNF transcription factors.

This review will include the following parts: 1) germ cell formation, basically the generation
of primordial germ cells (PGCs), and differentiation of PGCs as well as oogenesis, 2) maternal contributions on early embryonic development, basically the process from fertilization to implantation and maternal regulation of early embryonic development, especially maternal regulation factors and the events involved in maternal-to-zygotic transition (MZT), and 3) eukaryotic transcription including DNA organization of chromatin and transcription apparatus in eukaryotes as well as zinc finger transcription factors with an emphasis on the classic C2H2 KRAB-containing zinc finger transcription factors and the mechanism of their downstream regulation and their function performed during the early embryonic development.

GERM CELL FORMATION AND OOGENESIS

1. Formation of Germ Cells

Fusion of gametes, sperms and oocytes, is actually the start of all sexually reproducing organisms. Genetic as well as epigenetic information is passed to next generation through germ cells. Primordial germ cells are the common precursor of all germ cells. In most animal species, not including mammals, PGCs are determined by germ plasm. Germ plasm refers to a specific cytoplasm which contains specific proteins and mRNAs as well as organelles and localizes in certain cells of the early embryo. In mammals, the germ cell fate is not pre-determined but using a model termed as epigenesis. During the pre-gastrulation period, germline specification occurs in the epiblast due to signaling induction. These cells are termed as PGCs precursors. Those precursors then move into extraembryonic tissues and are determined their PGC fate. In normal embryos, PGCs will finally migrate from the extraembryonic tissues to the sexually indifferent gonadal ridges developing inside embryo. At this time, each PGC still has potential to develop into either sperm or oocyte in spite of the sex chromosomes (XX or XY) it has until sex differentiation.

Origin and Migration of PGCs

Mammalian PGCs are originally from proximal epiblast (PEpi) cells. The epiblast cells are not irreversibly allocated to germ cell lineage but a site-specific manner actually determines the germ cell fate in the potential epiblasts. Therefore, it is the extrinsic signals from surrounding somatic cells that induce epiblast to give rise to germ cell lineage. In mammalian germline studies, mice is actually the main animal model. Mouse PGCs were first localized as a distinct cluster of about 40 alkaline phosphatase-positive cells around 7.25 days post coitum (dpc) at the base of the incipient allantois. Actually, these cells are shown to originated from a few, around 6, the most proximal epiblast cells, which are immediately adjacent to the extraembryonic ectoderm (ExE). The specification of PGCs are accompanied by their motility. In mice, the motility of PGCs may be from their inception (7.25 dpc) to the time when they colonize the developing genital ridge (11.5 dpc). At 7.5 dpc, PGCs move through the primitive streak into the definitive endoderm. There are also some PGCs end up in the allantois or parietal endoderm. PGCs then incorporate into the endoderm of the developing hindgut. In the next 2 days, the PGCs are still associated with hindgut. An active, directed migration of PGCs from the dorsal axis of the hindgut to the developing genital ridges begins at 9.0-9.5 dpc. By 12.5 pdc, PGCs are no longer motile. PGCs stimulate local connective tissue to proliferate and will aggregate with each other as well as somatic cells of the urogenital ridges forming the primitive sex cord. Once the PGCs enter the developing genital ridges, PGCs continue proliferating through mitosis for around 2-3 days. Although by 12.5 dpc, differentiation of somatic components is quite obvious, the germ cells are still sex-undistinguishable. Around 14.5 dpc, however, primordial germ cells in fetal ovaries have developed to oogonia.
Molecular Regulation of PGCs Induction and Migration

During germ cell specification, the integration of three key events actually determines the germ cell fate: repression of mesoderm genes, reacquisition of pluripotency as well as genome-wide epigenetic reprogramming. In order to elucidate some basic mechanisms governing germ cell induction, various studies were performed to mutate genes encoding certain growth factors, adhesion molecules, and transcription factors.

A tripartite transcription factor complex was identified to be a key regulatory machinery. The transcription complex is composed of three factors: B-lymphocyte induced maturation protein 1 (BLIMP1, also known as PRDM1) and PRDM14, two PR (PRD1-BF1 and RIZ) domain-containing transcriptional factors, as well as AP2γ. Knockout studies of Prdm1, Prdm14 or Tflape in mice indicate the impairment of PGCs specification. At approximately 6.25 dpc, BLIMP1 is expressed in a few pluripotent epiblast cells then followed by upregulation of PRDM14 and AP2γ. These three genes together can repress somatic genes, while PRDM1 and PRDM14 are responsible for upregulation of germ cell and pluripotency genes as well as the initiation of genome-wide reprogramming and germ cell migration.

One required signaling pathway for this tripartite transcription network is BMP-SMAD signaling pathway which begins around 5.5 dpc. Bone morphogenetic protein (Bmp)4 and Bmp8b are from ExE, while Bmp2 is emitted from the visceral endoderm (VE). BMP4 signal actually acts through a receptor complex which is composed of ALK3 (or ALK6) as well as one of type II BMP receptor to phosphorylate SMAD1/5. The phosphorylated SMAD1/5 then form a complex with SMAD 4 to regulate the expression of target genes. BMP2, which is emitted from VE has a highly similar structure as BMP4 but is less efficient and therefore provides extra BMP signal to induce a PGC fate. Bmp8b, a signal from ExE, can suppress the development of anterior visceral endoderm (AVE), which secret inhibitory signal to PGCs development. Thereby, Bmp8b signal restricts the inhibitory signal from AVE and ensure the PGCs induction properly among proximal epiblast. The exact target genes of BMP-SMAD signal is not clear yet. However, mutations in Bmp4, Bmp8b, SMAD1 as well as SMAD5 show impaired PGCs induction. It demonstrates the importance of BMP-SMAD pathway. Both BLIMP1 and PRDM 14 are induced by BMP-SMAD signaling pathway. A second signaling pathway required is WNT3-β-CATENIN pathway. Actually, BMP4 signaling pathway has an activation role for the expression of WNT3 directly or indirectly. WNT3 is expressed in both posterior visceral endoderm around 5.5 dpc and posterior epiblast around 5.75 dpc. The expression of WNT3 allows the translocation of β-CATENIN from cytoplasm to nucleus, which will activate T gene. The T gene encodes the T-box transcription factor T- BRACHYURY, a classical and highly conserved mesoderm factor, will bind to the putative enhancer of BLIMP1 and PRDM14 and activate them directly.

The migration of PGCs is also regulated by various factors. Many studies were performed on certain genes to clarify their functions during PGCs migration. IFITMS and IFITM3 proteins are from the interferon-induced transmembrane protein gene family. It has been reported that BMP4 signal from the extra-embryonic ectoderm can induce the expression of IFITMS. IFITMS and IFITMS3 initiate mPGCs migration through repulsion. Germ cells expressing IFITIM3 are repelled from surrounding cells which express IFITMS. KITL belongs to short-chain helical cytokine family. An allelic series of mutation illustrate the requirement of KITL for PGCs proliferation and migration. In addition, integrin β1-/- murine PGCs were shown disability to colonize the gonad, and blocking antibody against E-cadherin, which has no expression when PGCs are in hindgut but upregulated when migrating form hindgut toward gonad, will cause the
defective PGC-PGC coalescence\textsuperscript{34}. Above all, integrin $\beta$1 as well as E-cadherin perform their function through the germ cell-extracellular matrix and germ-germ cell interaction respectively.

2. **Oogenesis and Oocyte Maturation**

In all mammalian female fetus, due to absence of a Y chromosome which contains a gene called Sex Determining Region (SRY), female pathway will actively set up female reproduction tract development. The process of oogenesis is actually from the differentiation of PGCs into oogonia after sexual differentiation\textsuperscript{52}. Each oogonia proceed mitosis two diploid primary oocytes. Then each primary oocyte undergoes a meiotic division prophase I (leptotene, zygotene, pachytene, and diplotene) until dictyate arrest which stops shortly before LH surge. Each immature oocyte is contained within a follicle, a real function unit of ovary. In most species, follicles are observed in cortex. There are four different types of ovarian follicles within cortex: primordial follicles, primary follicles, secondary follicles as well as antral follicles. Defects of either oogenesis or folliculogenesis can significantly impair the female fertility.

**Germ Cell Cyst Formation, Breakdown and Primordial Follicle Formation**

During early embryonic development, germ cell cyst is a very important developmental stage. In fetal ovaries, germ cells undergo division synchronously. However, these mitotic divisions are accompanied by incomplete cytokinesis which makes the daughter cells remain connected by ring canals, a kind of intercellular bridges\textsuperscript{53}. In mice, germline cysts form between 10.5 to 13.5 dpc, following the arrival of PGCs to genital gonad. The germ cell cyst is considered as an evolutionary conserved stage in both invertebrates and vertebrates\textsuperscript{54-57}. When the formation of germ cell cyst is complete, plenty of germ cell syncytia containing most oocyte nuclei which are arrested at diplotene stage of meiotic prophase I can be identified\textsuperscript{58}. In addition, those germ cell syncytia are connected through ring canals and they are further organized into long ovigerous cords which are surrounded by pre-granulosa cells and stromal mesenchymal cells\textsuperscript{53,59}. The breakdown of germ cell cyst, a developmentally programmed step, occurs primarily between 20.5 to 22.5 dpc, just prior to onset of primordial follicle formation in mice\textsuperscript{53}. The breakdown of cyst is attributed to the combined effect of oocyte nuclei degradation and pre-granulosa cell invasion\textsuperscript{53}. During this process, a series of programed events take place: retracting or cleaving the cytoplasmic bridges between remaining nuclei, granulosa cell completely surrounding the remaining nuclei as well as a basement membrane completely encompassing and delineating each newly formed primordial follicle\textsuperscript{59}. This process is a transition from the cysts into primordial follicles.

**Folliculogenesis**

Primordial follicles have a function as reserve (also refer to an initiation pool of primordial follicles) in ovaries which are established shortly after birth in rodent\textsuperscript{53}, and during fetal life in most domestic species\textsuperscript{50,61}. Establishment of this reserve is defined as the first stage of the folliculogenesis, while the end of female reproduction life span is marked by exhaustion of this pool due to apoptosis or development followed by ovulation or atresia\textsuperscript{62}. Once the primordial follicles are activated to grow, the primordial follicle enlarges caused by an increase in size of the oocyte and conversion of the squamous granulosa cells into cuboidal granulosa cells. In this stage, the follicle is known as primary follicle which is characteristic by the presence of zona pellucida\textsuperscript{62}. Zona pellucida is composed of ZP1, ZP2, and ZP3, all of which are glycoprotein. All of these three protein are essential to folliculogenesis and defect in any of three proteins will cause sterile or partially sterile in mice\textsuperscript{63-65}. As the granulosa cells proliferate to form multiple layers as well the
acquirement of an outer layer of thecal cells, primary follicles transit to pre-antral follicles. Antral follicle is next stage of folliculogenesis, which is characteristic by the presence of a single antral cavity. The antral cavity is filled with follicular fluid, which is composed of water, steroid hormones in high concentrations secreted by granulosa cells, electrolytes as well as serum proteins. The fate of most antral follicles are undergoing atresia, while follicle-stimulating hormone (FSH) can regulate the survived antral follicle(s) to grow to pre-ovulatory stage. Upon the stimulation of LH surge, the pre-ovulatory follicle(s) can ovulate an oocyte for fertilization. And the remaining theca and granulosa cells differentiated into a temporary endocrine structure call corpus luteum (CL) which is the resource of progesterone.

**Oogenic Maturation**

Oocyte maturation is one of the final steps during oocyte development. It actually refers to the resumption of meiosis just prior to ovulation and subsequent fertilization. During the follicles development, the oocyte also develops. The formation of follicular antrum approximately corresponds to the final developmental stage of oocyte. In order to support fertilization and further embryonic development, it accumulates the required molecular and cytoplasmic machinery during this stage of follicular development. In addition, the granulosa cells differentiate two distinct lineages in the aspects of anatomy and function during this time: the mural granulosa cells which not only form the wall of follicles but also have a steroidogenic role and cumulus cells which are closely association with oocytes. Through gap junction, cumulus cells can form a structure call cumulus-oocyte complex (COC) with oocyte. Apart from having function on maintaining oocyte meiotic arrest through a direct substance transfer, cumulus cells also involved in the induction of oocyte resumption upon LH surge. The next stage of oocyte development is to undergo metaphase I and extrusion of the first polar body before the other arrest which is actually after the formation of metaphase II. The cell cycle will be arrested in metaphase II until the subsequent fertilization.

Oocyte maturation can be clarified in three aspects: meiotic maturation, cytoplasmic maturation and molecular maturation. In terms of meiotic maturation, a series of programmed nuclear events which are associated with activation of cell cycle machinery upon the escape from those inhibitory factors. In respect of cytoplasmic maturation, the oocyte modifies the transcription and translation machinery to stop the synthesis RNAs and proteins. In addition, a re-distribution of organelles is another modification. With regard to molecular maturation, specific mRNA of proteins are assumed accumulated in the last few days before ovulation. The accumulation of these specific molecules may be essential for those post-fertilization events, such as embryonic genome activation.

**Molecular Signaling Factors, and Communication Networks**

Development of ovaries is regulated by various factors as well as communication networks. The inference of any of those regulators may interrupt the final fertilization or embryonic development. The real unit of ovary is individual follicles. As indicated above the formation of primordial follicles actually serves as a reserve. To some degree the formation of primordial follicles is the beginning of ovary development. It actually comes from the breakdown of germ cell cyst. On 16 dpc, upon the binding of surface protein jagged-1 to the receptor, the neurogenic locus notch homolog protein 2 (NOTCH2) signaling pathway is activated in granulosa cell. A NOTCH2-mediated signaling pathway is essential for the breakdown of germ cell cyst. In terms of primordial follicle formation, factor in germline alpha (FIGLA) is shown to affect
folliculogenesis. The FIGLA null ovary shows the inability to interact with granulosa cells. In addition, FIGLA also are involved in the upregulation of ZP1, ZP2, ZP3 genes which are involved in the formation of zona pellucida.

The careful orchestration of oocyte arrest as well as the following oocyte maturation also relies on various factors. The factors involved in keeping oocyte meiotic arrest come from ovary. If oocytes are removed from ovary, they undergo spontaneous maturation. Cyclic adenosine monophosphate (cAMP) is the most important intracellular signaling molecules that responsible for meiotic arrest of oocytes. Actually, cAMP signaling serves as an inhibitor of oocyte maturation. Various pieces of evidence have been shown. For example, an incubation with cAMP analogs, such as dibutyryl cAMP or phosphodiesterases (PDEs) antagonists which blocks the metabolism of cAMP, such as 3-isobutyl-1-methylxanthine or milrinone can block the isolated oocyte maturation. Purine hypoxanthine which is produced in the follicles is another inhibitor of meiosis in mammalian oocytes. It plays a role as a PDE inhibitor which might block metabolism of cAMP to maintain it in meiotic-arresting levels and can inhibit meiosis of either denuded oocytes or encased oocytes which are in follicles. In addition, gap junctions between oocytes and surrounding cumulus cells also play a very important role in maintaining meiotic arrest. In ovary, connexin 43(Cx43)-containing gap junctions are the primary member. An in mammal, gap junctions transport nutrients and metabolic molecules to oocytes. It has been observed that in vitro culture oocyte surrounded by cumulus cells will be maintained in meiotic arrest while interference of contact between oocytes and cumulus cells during cumulus expansion coincides with oocytes maturation. This observation indicates that gap junction may transport inhibitory factors mentioned above to maintain oocyte arrest. Exposure of follicles to LH signal can also lead to maturation of oocyte. Some paracrine factors which belong to epidermal growth factor (EGF) family form outer thecal cells and murine granulosa cells can promote cumulus cells expansion and oocyte maturation. Actually the communication between oocytes and surrounding somatic cells are bidirectional. Oocytes can also secret growth factors that direct the differentiation and function of cumulus cells. Those factors are called oocyte-secreted factors. Growth-differentiation factor 9 (GDF9) and bone morphogenetic protein 15 (BMP15) are two oocyte-secret factors. They activate signaling pathways in cumulus cells which are required for cumulus cells differentiation and distinctive phenotype maintenance. It in turn provides a suitable microenvironment for oocyte development. This bidirectional communication is actually a mechanism that oocytes increase their own intrinsic quality.

MATERNAL CONTRIBUTION TO EARLY EMBRYONIC DEVELOPMENT

The process of oocyte molecular maturation indicates that the intrinsic quality of oocytes has essential impacts on the post-fertilization events. Whereas sperm cells are differentiated for the motility to deliver the haploid nucleus, the oocyte develops a complex cytoplasm containing cytoplasmic enzymes, mRNAs, organelles, and metabolic substrates during the process of oocyte development. Those molecules stored in oocyte cytoplasm are involved in the regulation of various post-fertilization events. The first important transition post-fertilization is maternal-to-embryonic transition (MET) in which the embryo switch from maternal molecules accumulated during the oocyte development to the factors synthesized by embryonic genome after the activation of embryonic genome. Key developmental events during the MET includes deletion of maternal transcripts, epigenetic reprogramming or chromatin remodeling, and finally the activation of embryonic genome. In the past few decades, maternal encoded proteins have been identified in mounting numbers to regulate various key events during early embryonic development.
1. Oocyte Intrinsic Quality Contributes to Early Embryonic Development

During fertilization, oocyte not only delivers nucleus but also serves as a warehouse which stores various maternal factors required for early embryonic development. Ability to develop to the blastocyst stage is the key marker of oocyte competence commonly used by most laboratories. It is speculated that the failure to activate the embryonic genome may be the main cause of early embryonic loss since among those early embryos not able to reach blastocyst stage most are blocked at or close to MET stage that occurs at eight-cell stage in cow. The activation of embryonic genome relies on the storage of RNAs and proteins in competent oocytes rather than embryonic genome. The treatment of transcription inhibitors during the first few days following fertilization can cause the normal cleavage at the four- to eight-cell stage. The accumulation of specific molecules refers to oocyte capacitation, one of the most important indication to oocyte intrinsic quality.

The accumulated specific molecules are actually the product of maternal genes. Maternal-effect genes are genes affect the development of offspring regardless of their own genotype. Conventional knock-out assay were performed to identify first a few maternal-effect genes in mice. In terms of domestic animals, siRNAs or antibodies were injected to inhibit the expression of certain genes in mRNA or protein level, respectively to determine the function of certain maternal-effect genes. Here are a few examples of maternal-effect genes introduced in either mice or cattle.

Zygote Arrest 1 (Zar1) is a maternal-effect gene which was first identified in mice through subtractive hybridization and cDNA library screening and cDNA library screening. The expression level of Zar1 is really high in growing oocytes of ovaries as well as one-cell embryos but dramatically low in two-cell embryos and was absent from four-cell embryos. Zar1-/− mice have normal ovarian development as well folliculogenesis, and Zar1-/− oocytes can be fertilized but the embryonic development blocks at the one-cell stage. And additional analysis indicates that the block is caused by an incomplete fertilization in which maternal and paternal genomes do not unite but are still separate in discrete pronuclei. In Zar1 null mice, the synthesis of a group of protein called the transcription-requiring complex (TRC), which is the maker of embryonic genomic activation, was analysis to determine the function of Zar1. The results indicate embryonic genome activation occurs in Zar1 null mice. Zar1 contains an atypical PHD (plant homeodomain) motif which can be identified in either transcriptional regulators or chromatin-modulating complex subunit. However, cytoplasmic localization of Zar1 made the detailed mechanism through which Zar1 achieve its function remain to be identified.

*JY-I*, a bovine maternal-effect gene identified through expressed sequence tags (ESTs) from cDNA libraries of bovine oocytes. *JY-I* encodes for a bovine-specific secreted protein and *JY-I* mRNA and protein are expressed exclusively in oocyte during folliculogenesis. Maximal level of *JY-I* mRNA can be detected in the germinal vesicle stage oocytes while during the early embryonic development the level starts decrease and undetectable at 16-cell stage embryo. Embryo culture experiments with transcription inhibitor α-amanitin which block the transcription during the first and second embryonic cell cycles. The detection of *JY-I* during this period indicates that *JY-I* mRNA is derived from oocyte. Microinjection of siRNA targeting *JY-I* into bovine zygote stage embryo results in a dramatic decrease in proportion of embryos which can develop into 8-16 cell stage embryo which corresponds to maternal to embryonic transition in cattle and further blastocyst stage compared to the uninjected, sham-injected, and negative-control siRNA-injected
embryos. However, the development of JY-1-deficient embryos can be rescued once exogenous JY-1 protein was given. In addition, microinjection of siRNA targeting JY-1 into cumulus-enclosed oocytes was performed and then meiotic arrest was maintained for 48 hours before in vitro maturation (IVM). A dramatic reduction in cumulus expansion, decreased proportion of oocytes developing to metaphase II (MII) as well as reduced proportions of embryos that developed to the 8-16 cell and blastocyst stages following IVM. Addition of recombinant JY-1 protein during oocyte culture can rescue the reduction of cumulus cell expansion as well as reduced proportion of oocytes reach MII stage but no effect on embryonic development.

**Kpna7 (importin α8)** is another bovine maternal-effect gene which was identified through EST analysis from the same library used for identification of JY-1 and encodes a protein which belongs to the importin α family. In cattle, Kpna7 mRNA is expressed exclusively in ovaries, immature or mature oocytes as well as embryos at stages before embryonic genome activation. Similarly, Kpna7 protein is in abundant in GV stage, MII stage oocyte as well as early cleavage stage embryos followed by a dramatic decrease in morula and blastocyst stage. siRNA-mediated Kpna7 knock-down assay performed in early embryo reduced developmental potential to blastocyst stage. In addition, compared to other members in importin α family, Kpna7 possess a stronger binding affinity to nucleoplasm 2 (NPM2), a maternal-effect nuclear protein, indicating an essential role of Kpna7 in transporting nuclear proteins.

### 2. Maternal Control of Early Embryonic Development via Epigenetic Reprogramming

In order to promote the accurate early embryonic development, various events need to be reset. Epigenetic modifications is used to regulate those impotent events. Epigenetics refers to heritable changes in gene expression or cellular phenotypes that occur without the alterations in DNA sequence. Fusion of sperm and oocyte actually brings two haploid genomes that has dramatically different organization together. The package of sperm chromatin is predominantly an extremely dense protamines in addition to a modest amount of sperm-specific histone variants. One necessary of epigenetic reprogramming is believed to resolve this dramatic differences between maternal and paternal chromatin and ensure that the highly differentiated germ cells are transition to a totipotent one cell which are ready for the following cleavage. Those epigenetic reprogramming events rely on the stored reprogramming factor in oocytes since the silence of embryonic genome. Various maternal proteins has been reported to involve in the epigenetic reprogramming of early embryonic development.

The first developmental event is global demethylation followed by a wave of de novo methylation which is mediated by de novo methyltransferase Dnmt3A and Dnmt3b. DPPA3 is one maternal protein that regulates developmental events via epigenetic reprogramming. In mice, the expression of DPPA3 can be detectable from the process of germ-cell specification at 7.25 dpc specifically in the founder population of PGCs to 15.5 dpc in both male and female gonad. In male, the expression of DPPA3 is not detectable while DPPA3 is expressed in either immature or mature oocytes and remained detectable in early embryos after fertilization. It was demonstrated that DPPA3 interacted with Ran binding protein 5 (RanBP5), a nuclear transport shuttle protein, which facilitates the nuclear transport of DPPA3. In unfertilized oocytes, DPPA3 was widely distributed in the cytoplasm but excluded from the metaphase chromosomes. After fertilization, a dramatic decrease of cytoplasmic localization of DPPA3 can be observed and DPPA3 is predominantly localized in two pronuclei. In addition, maternal genome of zygotes derived from DPPA3-defective oocytes was demethylated while that from the zygotes derived
from normal oocytes was not\textsuperscript{104}, indicating that DPPA3 play a fundamental role in inhibiting active demethylation of maternal genome which will block the further development of zygotes.

Histone proteins directly contact with DNA to form the unit of chromatin, the nucleosome. The core histone tend to form two types of subcomplexes: a very stable H3\textsubscript{2}-H4\textsubscript{2} tetramer and two H2A-H2B dimer. Each histone also has conserved N-terminus tail (H2A and H2B have C-terminus as well) containing covalent site exposed for post-translational modification. For example, serine or threonine can be phosphorylated; lysine residues can be acetylated, methylated, or ubiquitinated; arginine residues methylated\textsuperscript{106,107}. In addition, some residues can be mono-, di-, or tri-methylated, which may add the complexity to post-translational modification. Histone post-translational modifications can regulate the contacts between nucleosome and chromatin upregulating or repressing the expression of certain genes via compacting or relaxing DNA fiber respectively.

Histone post-translational modification is one kind of epigenetic regulation during the early embryonic development. Paternal chromatin is packed with protamines, which make the structure of paternal chromatin more condensed for its motility. However, after fertilization, protamines of paternal chromatin are replaced by maternal stored histones that are devoid of most modifications while maternal chromatin inherited most histone modifications from oocytes\textsuperscript{108,109}. Genomic imprinting is a epigenetic phenomenon that allows only one allele from either paternal or maternal genome expressed. And the silence of the other allele is mediated predominantly via DNA methylation. However, it has been shown that maternal repressive modification of histone is another option of genomic imprinting\textsuperscript{110}. In mice, maternal tri-methylation at lysine 27 of histone H3(H3K27me3) was identified as a DNA methylation-independent imprinting mechanism\textsuperscript{110}. 76 genes with paternal allele-specific DNase I were identified through an integrated analyses of DNA methylome and tri-methylation at lysine 27 of histone H3 (H2K27me3) chromatin immunoprecipitation followed by sequencing in mice zygotes and morula embryos\textsuperscript{110}. In addition, ectopic removal of H2K27me# induces the expression of maternal allele.

Histone variants also play essential roles in epigenetic regulation of early development of embryos. H3.3 is such a maternal histone proteins involved in the process of chromatin remodeling: shortly after fertilization, maternal H3.3 is deposited onto the sperm chromatin. Lin et al have demonstrated that RNAi-mediated knock-down of H3.3 in fertilized mice zygotes blocks early development of embryo at morula stage which can be rescued by exogenous H3.3 mRNA\textsuperscript{111}. In addition, deficiency of H3.3 can lead to over-condensation of chromosomes, high level of aneuploidy and mis-segregation of chromosome as early as two-cell stage\textsuperscript{111}. Moreover, markers of open chromatin, such as H3K36me2 have a significant reduced level in H3.3-deficient embryos\textsuperscript{111}. It can be concluded that H3.3 can mediate a balance between open and condensed structure of chromosome, which are critical for embryonic development\textsuperscript{111}.

3. \textit{Activation of Embryonic Genome}

Embryonic genome activation (EGA) is a process in which an embryo begins to transcribe its newly formed genome and gradually develops independently rather than relies on maternal factors. The EGA occurs at two-cell stage in mice\textsuperscript{112}, 8- to 16- cell stage in cow\textsuperscript{113}. Proper activation of embryonic genome is essential for the development\textsuperscript{114}. Embryonic genome will transcribe and produce factors to regulate the development of embryos instead of maternal factors once it is activated. This process is called maternal-to-embryonic transition (MET). Actually, two ordered molecular activities are encompassed in maternal-to-embryonic transition: maternal clearance,
deletion of maternal mRNAs and proteins accumulated during the oocyte maturation, followed by set-up of embryonic regulation which is actually the consequence of embryonic genome activation\textsuperscript{114}.

**Maternal Clearance**

A combination of three mRNA features can influence mRNA stability as well as protein synthesis: the sequence of mRNA, the 7-methyguanylate (m\textsuperscript{7}G) cap at the 5’ end (5’ cap) and the length of the 3’ poly (A) tail. During the period of MET, maternal clearance can be achieved through either decapping or inferencing poly(A) tail. The studies of maternal clearance are mainly performed in lower low vertebrate animals, such as Drosophila, Xenopus or zebrafish. Here, my review is mainly on low vertebrate animals. In Drosophila, the translation of Smaug is activated by the PAN GU Kinase\textsuperscript{115}. Smaug can recognize its target via the interaction between a sterile alpha motif (SAM) domain and stem loop structures called Smaug recognition elements (SREs)\textsuperscript{116,117}. In general Smaug can recruit the CCR4/POP2/NOT-deadenylase complex which initiates poly(A) tail shortening and the final elimination of mRNA\textsuperscript{118}. MicroRNAs can also regulate translation, deadenylation, and mRNA destabilization through the interaction with target mRNAs\textsuperscript{119}. In zebrafish, miR-430 is transcribed from embryonic genome\textsuperscript{120}. MiR-403 can induce deadenylation and destabilize mRNA of several hundred transcripts\textsuperscript{121}. In terms of decapping, the detailed mechanisms which are used to activate decapping and their regulation of maternal clearance have not been clarified yet.

**Embryonic Genome Activation**

In most mammals, transcription from the embryonic genome occurs in two phases: a minor wave of genome activation which occurs before cleavage and a major activation occurs at 2-cell stage in mice\textsuperscript{122} which is thought to involve as many as \(~800\) and \(~3500\) genes, respectively\textsuperscript{123,124}. In general, two contrasting models of activation are widely accepted and have been the focus of research related to embryonic genome activation. In the first model, the “nucleocytoplasmic (N/C) ratio” model, the rapid and synchronous cell division during the early development increases the quantity of nuclear material causing an increased nucleocytoplasmic (N/C) ratio due to the constant cytoplasm volume\textsuperscript{125,126}. It will finally dilute the relative level maternal factors, a barrier for embryonic genome activation. The other model is “maternal clock”, a mechanism independent of cell division but certain quantity and activity of maternal factors determine the activation of embryonic genomes. In mammals, multiple factors are involved in activation of embryonic genomes including maternal factor genes, chromatin remodeling. A growing number of maternal factors have been shown to regulate the activation of embryonic genome including Mater\textsuperscript{92}, Zar I\textsuperscript{12}, Oct4\textsuperscript{127}, and Sox2\textsuperscript{128}. Chromatin remodeling, as mention in previous part, certain part of chromatin can be exposed to or protect from the transcriptional machinery causing the activation or repression of genes. A proper spatiotemporal expression of certain genes that are time-critical as well as repressing non-required genes is essential in triggering normal embryonic development. The repressive transcriptional state of the chromatin structure may be a selective mechanism to ensure the correct expression profile during the development\textsuperscript{129}. In order to further identify regulation of embryonic genome activation, a combination of in situ DNase I sensitivity assays and in vitro transcription analysis were applied\textsuperscript{130}. During the early 1-cell stage to late 1-cell stage, an opposite changes between in situ DNase I sensitivity in vitro transcription analysis, in which increase transcriptional activity corresponds to decreased DNase I sensitivity, indicates that it is maternal factors that not chromatin structure initiate the minor activation wave at 1-cell stage. In
terms of chromatin remodeling factors, the deficiency of them can also block the activation. Maternal deletion of Brg1, encoding a subunit of SWI/SNF chromatin remodeling complex BRG1, can reduce the level of dimethyl H3K4, a mark for transcriptionally active chromatin\textsuperscript{15}. Maternal mutation of Brg1 can block embryonic development at early cleavage stages and reduces transcription for \textasciitilde30\% of genes\textsuperscript{15}.

**Eukaryotic Transcription Regulation and Zinc Finger Transcription Factor**

In eukaryotes, the transcription of the genome is chromatin-based. The status of chromatin can determine the “switch-on” or “switch-off” of certain genes. So open status of chromatin is one prerequisite of genome transcription initiation. And the open chromatin can only be found in the cells in which they are expressed or potentially expressed. In the process of transcription, various proteins are involved. RNA polymerase is one protein, the others are called transcription factors. The function of transcription factors includes recognizing cis-acting sites on DNA, recognizing RNA polymerase or incorporated into an initiation complex. A eukaryotic RNA polymerase cannot read the DNA but contact DNA through a group of protein called basal transcription factors. Together with RNA polymerase, basal transcription factors form the basal transcription apparatus. In addition, through recognizing short conserved sequences, basal transcription apparatus can interact with the promoter of certain individual gene. Enhancer sequence is another type of sequence which not only determine whether promoter is expressed or not but also determine the expression of promoter in which cell type. The location of enhancers is not restrict to being near the promoter but shows variability. It can be upstream, inside a gene, or beyond the end of a gene. In addition, enhancers can be either orientation relative to the genes. The organization of DNA-binding domain is the major principle used to classify transcription factors\textsuperscript{131}. In human and mice genomes, the C2H2 zinc finger, homeodomain, and helix-loop-helix DNA binding domain dominate more than 80\% of transcription factor repertoire\textsuperscript{22,132}.

1. *Initiation of Protein-coding Gene in Eukaryotes*

In eukaryotes, RNA polymerase II is mainly responsible for synthesizing heterogeneous RNA (hnRNA), which is the precursor for most mRNA. To initiate transcription, RNA polymerase II requires basal transcription factors termed as TF\textsubscript{II}X at all promoters. It has been revealed that five basal factors including TF\textsubscript{II}B, TF\textsubscript{II}D, TF\textsubscript{II}E, TF\textsubscript{II}F and TF\textsubscript{II}H for this initiation process as well as a sixth TF\textsubscript{II}A to potentiate the magnitude of transcription\textsuperscript{133-135}. In terms of promoter, a shortest sequence which are required for the initiation of transcription is defined as core promoter. A typical pol II core promoter contains three elements: an initiator (Inr), a region which is composed of CA flanked by pyrimidines and occupies position form -3 to +5, a TATA box, which locates 25bp upstream of the Inr, the start point of transcription, or downstream promoter elements (DPE), which are located at +28 to +32 and often identified in TATA box-less promoters.

The first step in transcription initiation is recognition of TATA-box through TATA binding protein (TBP). Among the five factors mentioned above, TF\textsubscript{II}D has TBP as one subunit as well as up to 14 other TBP associated proteins (TAFs). So TF\textsubscript{II}D serves as a positioning factor for pol II transcription apparatus and it can bend DNA once binding to DNA to form a platform form the remaining factors. TF\textsubscript{II}B then enters and extend the interaction with sequence flanking the TATA box\textsuperscript{136}. Another function of TF\textsubscript{II}B is to recruits Pol II - TF\textsubscript{II}F complex. TF\textsubscript{II}F has two subunits: the larger one, RAP74, having a ATP-dependent DNA helicase activity which may be involved in
melting activity during transcription initiation, and a smaller one RAP38, which can interact with pol II. TF_E and TF_H are the last factors assembled in the pol II initiation complex. TF_E can extend the boundary region which is protected downstream by another turn of double helix, while TF_H has multiple independent enzymatic activities including ATPase as well as helicases of both polarities. Transcription can only initiated when TF_E and TF_H join the complex indicating that the requirement of these two factors to melt DNA allowing the movement of RNA polymerase II escape from promoter region.

2. **Activators and Repressors**

In addition to basal transcription factors, various other proteins are also involved serving as either activator or repressor during the initiation of eukaryotic transcription. In terms of mechanism, three classes of activators can be identified: true activators, anti-repressors and architectural proteins, while repressors can affect the chromatin structure, sequester an activator in the cytoplasm, mask the activation domain of a activator or compete the interaction with enhancers.

The true activators are actually real transcription factors which have direct or indirect interaction through a coactivator with basal transcription apparatus. An activator need at least a DNA-binding domain which can locate the activators into the vicinity of the promoter. For those transcription factors directly interacting with basal transcription, they have another domain covalently connected to the DNA binding domain which can activate basal transcription apparatus. And coactivator can serve the same function in the case of indirect interaction between activators and basal transcription apparatus. The function of a true activator can be achieved through one of the following mechanism: tissue-specific synthesis, converting to active form by covalent modification, ligand binding, binding of inhibitor which can affect the DNA binding ability or cleavage from inactive precursor. The anti-repressors are actually certain types of activators which can recruit histone modification enzymes or chromatin remodel complexes once binding to the enhancer. The anti-repressors can only act on chromatin template but not DNA template. Architecture proteins serve to bend the DNA template which can either facilitate cooperative formation or prevent the formation of inhibitory complex. Yin-Yang1 (YY1) is such an architecture transcription factor. YY1 binds to active enhancer elements as well as promoter-proximal elements and forms a dimer which facilitates the looping of DNA, mediating the enhancer-promoter interaction.

In order to process the function, transcription factors need to be transport from cytoplasm where they are synthesized to nucleus where they can interact with chromatin or other factors. Binding to the functional domain involved in the transportation process can sequester this activator. Nuclear factor κB (NF-κB) is a transcription factor which involved in inflammatory reaction. The function of NF-κB can be repressed by PDLIM1, a member of LIM protein, through sequestering p65 subunit of NF-κB, blocking its nuclear translocation. Another mechanism to inhibit an activator is to its activation domain. One example using this mechanism is yeast Cyc8-Tup1 corepressor complex. It only contributes modestly to the expression of target genes a through corepressor function but regulates primarily by masking and inhibiting the transcriptional activation domain of other proteins. In terms of chromatin-mediated silencing, the mechanism actually involved in converting the closed status of chromatin, euchromatin, into a closed status, heterochromatin, which can block exposure of promoter to the transcription machinery. The best example is KAP1-KRAB-ZF transcription factor which will be discussed in detail in next section.
3. KRAB Containing Zinc Finger Protein Mediated Transcription Silencing Complex

The C2H2 Class of Zinc Finger Proteins

As one of the major protein motif that interacts with nucleic acids, zinc finger was first identified in the *Xenopus* TFIII A. The zinc finger is a small peptide mini-domain which can form a protrusion finger-like secondary structure through interaction with zinc ion. In addition, adjunct individual zinc finger can combine to form a zinc finger domain which can interact with DNA. The two cysteine and two histidine residue (C2H2) zinc finger proteins are the largest family of TFs in human genome and may be the predominant family of regulatory proteins in all mammalian genomes. C2H2 zinc finger proteins contain from 1 to more than 30 fingers. In general, the sequence of each finger can be organized in following formation: φ-X-Cys-X_{(2,4)}-Cys-X_{3-φ-X_{5-φ-X_{2}-His}}-X_{(3-4)}-His. In this case, φ represents a hydrophobic residue while X represents any amino acid. Each zinc finger motif may have 30 amino-acid residues, 25 of which will be involved in forming the finger-like domain and two cysteine and two histidine residues tetrahedrally coordinate a zinc ion. The other 5 residues set up linkers between consecutive fingers. In terms of tertiary structure, each finger is composed of two antiparallel β-sheets followed by an α-helix forming a left-handed ββα-module. Each individual zinc finger is actually a DNA interaction motif which can be linked tandemly in a linker, polar fashion to recognize DNA sequence. In general, the variation in key amino acids as well the number of zinc finger motif allow the each zinc finger domain to recognize different DNA sequence. In each zinc finger motif, α-helix contributes primary contacts to the interaction with major groove of DNA. Through specific hydrogen-bond, the amino acids at helical positions -1, 3, and 6 can interact with three successive bases (a triplet) on one strand of the DNA, while a interaction does exist between helical position 2 and the noncoding strand.

KRAB-containing Zinc Finger Protein and KAP1 Corepressor

In addition to the C-terminus zinc finger domain which function as DNA binding domain, the KRAB (Krüppel-associated box) domain has been identified in nearly one-third of all C2H2 zinc finger proteins. KRAB domain consists of around 75 amino acid residues and folds into two amphipathic helices. KRAB domain is divided into two subdomain: KRAB A-box and the KRAB B-box. It has been reported that while KRAB A box is necessary and sufficient to repress the transcription, it is considered a weaker repression domain, but when combined with KRAB B box the repressive activity is enhanced. In mammals, KRAB domain can be further divided into three subfamilies: one carrying the classical A box only (KRAB A), another containing classical KRAB A box as well as the classical B box (KRAB A+B), the other carrying a highly divergent KRAB B box, named b with the classical KRAB A box (KRAB A+b).

KAP1 (KRAB-associated protein 1, also known as Tif1β, TRIM28 or KRI1P) is a corepressor protein that interact with KRAB-containing zinc finger transcription factors. As a member of TIF1 (transcription intermediary factor) family, the overall structure of KAP1 includes an N-terminal tripartitite motif which contains an RBCC (RING (really interesting new gene) finger, two B-box zinc fingers, and a coiled-coil) domain, a central TIF1 signature sequence (TSS) domain consisted of a 25-amino acid tryptophan- and phenylalanine-rich sequence, as well as a C-terminal combination plant homeodomain (PHD) and bromodomain. Specifically, KAP1 also contains a central HP1 (heterochromatin protein 1)-binding domain. The N-terminal RBCC domain of KAP1 is a high affinity protein interaction domain which stretches from amino acid 20 to 377.
and is required and necessary for the interaction with KRAB domain of KRAB-ZFPs. In addition, all three RBCC subdomains, Ring finger, two B-box zinc fingers as well as a coiled-coil domain contribute to the interaction with KRAB domain\textsuperscript{154}. Even though RBCC is a common domain in TIF1 family, only KAP1 can interact with KRAB-ZFPs\textsuperscript{154}. TSS domain is adjunct to the RBCC domain, deletion of TSS domain can abrogates transcriptional repression mediated by TIF\textsubscript{y}\textsuperscript{152}. HP1-binding domain is a hydrophobic \textit{P}x\textit{V}x\textit{L} pentapetide which is located between amino acids 486 to 497\textsuperscript{155}. The HP 1-binding domain plays a key role in transcription regression. And the detail will be discussed in next section about the mechanism of KAP1-KRAB-ZFPs. The C-terminal tandem PHD and bromomdomain (called the PB domain) locate between amino acids 618 and 835 and they are required to obtain the maximum levels of repression\textsuperscript{156}. The PB domain of KAP1 can interact with two enzymes involved in chromatin modification:Mi2\textalpha, an isoform of Mi2 protein found in the NuRD (nucleosome remodeling and histone deacetylation) complex and SETDB1 (SET domain, bifurcated 1), a H3K9me3-specific histone methyltransferase\textsuperscript{153}.

\textit{Molecular Mechanism of KAP1-KRAB-ZFPs Mediated Transcriptional Repression}

In general, mechanism of KAP1-KRAB-ZFPs repression is to convert the chromatin to a closed stature. KRAB-ZFPs can be located in the genome through the recognition of certain DNA sequence by zinc-finger DNA binding domain. And through KRAB domain, the KRAB-ZFPs can recruit KAP1, which can coordinate the assembly of a macromolecular complex containing factors involved in chromatin remodeling, such as Mi2\textalpha, SETDB1, and HP1. The macromolecular complex can create an epigenetically stable and inheritable heterochromatic microenvironment\textsuperscript{157,158}. In this regression process, KAP1 acts as a scaffold for a silencing complex including the histone methyltransferase SETDB1 (or ESET)\textsuperscript{159}, the nucleosome remodeling and deacetylation (NuRD) complex\textsuperscript{156}, heterochromatin protein1 (HP1)\textsuperscript{158} and DNA methyltransferase\textsuperscript{160}. SETDB1, a SET domain protein, has histone H3-K9-specific methyltransferase activity. The methylation of lysine 9 of Histone 3 (H3-MeK9) mark establishes a high-affinity binding site for the recruitment of the HP1 family of heterochromatin proteins\textsuperscript{161}. The recruitment of SETDB1 requires a post-translational modification in which The PHD domain of KAP1 serves as an intramolecular E3 ligase of KAP1 to direct specific sumoylation modifications of particular lysine residues in bromodomain\textsuperscript{162}. NuRD complex is shown to have both histone deacetylase and ATPase chromatin remodeling activities. Through the interaction between KAP1 and Mi-2\textalpha, a subunit of NuRD complexes, NuRD is recruited to the promoter region of certain genes and process the repression of gene expression\textsuperscript{156}. It has been revealed that only limited amount of endogenous KPA1 shows interaction with NuRD complexes, indicating that a minor role of NuRD complexes\textsuperscript{156}. Heterochromatin protein1 (HP1) has essential role for the regulatory function of KAP1-KRAB-ZFPs complex. Deposition of HP1 can establish a de novo microenvironment of heterochromatin which has essential role in the repression activity of KAP1-KRAB-ZFPs\textsuperscript{158}. The role of DNA methyltransferase will be discussed in detail in the next section.

\textit{Function of KAP1-KRAB-ZFPs: Early Embryonic Development and Differentiation}

As the most abundant transcriptional repressive system, KAP1-KRAB-ZFPs complex involves various process including neoplastic transformation, stem cell pluripotency, cell cycle, differentiation, metabolism and early embryonic development. This section will focus on the involvement of embryonic development.

Dysfunction of KRAB-ZFPs can result in severe mutant phenotypes during embryogenesis. One example is \textit{Chato} mutation, which in mice causes functional defects of a KRAB-ZFP,
Zfp568<sup>163</sup>. The disruption of Zfp568 can cause defects in convergent extension during development. Disruption of co-repressor KAP1 can also cause severe outcome during embryogenesis. Targeted deletion of KAP1 in mice can be lethal and KAP1-null embryo cannot pass the egg cylinder stage, prior to the onset of gastrulation<sup>164</sup>.

During early embryonic development, de novo DNA methylation which follows a genome-wide DNA demethylation is an essential aspect of the epigenetic reprogramming during the early embryonic development. De novo DNA methyltransferase including DNA methyltransferase (DNMT) 3a and DNMT 3b as well as DNMT3L can recognize nonmethylated DNA and catalyze de novo cytosine methylation. As mentioned in previous section, KAP1 can serve as scaffold which can recruit various factors including DNA methyltransferase. In order to elucidate contribution of KAP1-KRAB-ZNP to the establishment of de novo methylation during early development, studies was performed using murine embryonic stem cells (ESCs)<sup>160</sup>. Quenneville et al. used a lentiviral vector-based system which is suitable for monitoring both DNA methylation and epigenetic silencing. In this system, two sets of vector were used: a Tet repressor (tTR) fused different parts of ZFP57 vector as well as a PGK.GFP expression cassette downstream of TetO repeats with or without a 2 kb KvDMR1 imprinted control region (ICR) as intervening sequence<sup>160</sup>. The operation principle is that the interaction between tTR fusion protein and TetO can be repressed by the presence of doxycycline but activated when it was omitted<sup>165</sup>. Different murine ESC lines were engineered to produce various of tTR derivatives including tTR.ZNF, which is only zinc finger domain is fused with tTR, tTR.KRAB, which can produce can produce a recombinant protein containing only KRAB domain or tTR.ZFP57, a ZFP57 full length fusion protein for the transduction with TetO.ICR.PGK.GFP vector in the presence of doxycycline or absence of doxycycline. After three weeks, GFP expression as well as ICR DNA methylation was examined respectively by fluorescence-activated cell sorting (FACS) and pyrosequencing<sup>160</sup>. The results showed a decreased level of GFP expression as well as increased level of ICR region in tTR-KRAB and tTR-ZFP57 group when doxycycline was omitted<sup>160</sup>, indicating that KAP1-KAb-ZNFs mediated cytosine methylation in ESCs. In addition, transcription repression driven by the deposition of repressive histone marks is prior to DNA methylation and doesn’t reflect DNA methylation<sup>165,166</sup>. Genome-wide analyses reveal that KAP1-KRAB-ZNFs mediated DNA methylation spreads over short distances from KAP1-binding site so as to involve nearby CpG island<sup>160</sup>. Moreover, it has been revealed that in adult mice liver cells, the methylation status of CpG islands correlates with their proximity to KAP1-binding sites in ESCs<sup>160</sup>. All together, the results indicate that KRAB-ZNPs together with co-repressor KAP1 may contribute to the de novo methylation establishment during early embryogenesis and DNA methylation patterns are maintained through development<sup>160</sup>.
Figure 1. Zinc Finger Structure. (A) An individual zinc finger motif. (B) Interaction between zinc finger motif and DNA. (Reproduced from Klug, 2010).
Figure 2. KAP1 Structure. (Reproduced from Cheng et al., 2014)
Figure 3. The KRAB-ZFP/KAP1 repressor complex. (Reproduced from Lupo et al., 2017)
DETERMINATION OF THE CONSENSUS TARGET SEQUENCES RECOGNIZED BY ZNFO, A NOVEL OOCYTE-SPECIFIC ZINC FINGER TRANSCRIPTION FACTOR IN CATTLE

INTRODUCTION

The intrinsic quality of oocytes has a great contribution to reproductive success. During early embryonic development, it is the maternal factors from the oocyte that play significantly important roles in the regulation of post-fertilization events until embryonic genome activation. Several observations have demonstrated the dependence of reproductive success on oocyte intrinsic quality. Successful rate in assisted reproductive technologies (ART) is significantly higher when women of advanced reproductive age use the oocytes from young donor compared with their own oocytes and in vivo matured oocytes has higher developmental competence compared with in vitro counterparts. In addition, the embryo mortality occurs predominantly during the early developmental stages in various species including dairy cow, women and horse. Oocyte quality refers to the ability of oocyte to resume meiosis, cleave and develop to blastocyst after fertilization, which is a key marker of oocyte competence. Oocyte capacitation, which refers to the accumulation of maternal factors is one of the most important indication of oocyte intrinsic quality.

Maternal to embryonic transition (MET) can be defined as a process in which the regulation of development is switched from maternal factors to molecules produced by embryonic genome following the activation of embryonic genome. Key events in MET include depletion of maternal transcripts, epigenetic reprogramming or chromatin remodeling and the final activation of embryonic genome. Infertility can be attributed to the failure of embryonic genome activation since early embryos that are not able to reach blastocyst are blocked at or close to MET stage which occurs at eight-cell stage in cow. As the first important transition following fertilization, MET also relies on maternal factors. A growing list of maternal factors have been identified using conventional knock-out assay in mice or micro-injection in domestic animals. One example of maternal factor in cattle is newborn ovary homeobox encoding gene (NOBOX). RNA interference (RNAi)-mediated defect of NOBOX can cause apparent down-regulation of genes related to signal transduction, cell cycles as well transcription regulation during embryonic genome activation (EGA).

Zinc finger (ZNF) is a small peptide mini-domain which can form a protrusion finger-like secondary structure through the interaction with zinc ion. Adjunct individual zinc finger can be combined to form a zinc finger nucleic acid interaction domain. Zinc finger proteins (ZNFs) are one of the largest protein superfamilies in eukaryotes. In addition, the classic two cysteine and two histidine residue (C2H2) zinc finger proteins compose the largest transcription factor repertoire in mammalian genome. Structurally, each C2H2 zinc finger motif has a general sequence form: φ-X-Cys-X-(2-4)-Cys-X3-φ-X3-φ-X2-His-X3-4-His, in which φ represents a hydrophobic residue while X represents any amino acid, and two cysteine and two histidine can interact with zinc ion forming secondary structure. In addition to the C-terminus zinc finger domain which function as a DNA binding domain, the KRAB (Krüppel-associated box) domain has been identified in nearly one-third of all C2H2 zinc finger proteins. KRAB domain has around 75 amino acid residues and folds into two amphipathic helices. KRAB domain can be further divided into two subdomains: KRAB A domain, which is required and sufficient for the repressive regulation and its activity is enhanced when KRAB B is present. In order to proceed with its repressive function, KAP1 (KRAB-associated protein 1) corepressor is required. KAP1 served as a scaffold to recruit...
various factors including SETDB1 (SET domain, bifurcated 1)\textsuperscript{159}, NuRD (nucleosome remodeling and histone deacetylation) complex\textsuperscript{156}, heterochromatin protein1\textsuperscript{158} (HP1) and DNA methyltransferase\textsuperscript{160}. ZNFO, an oocyte-specific maternal factor identified from a previously constructed bovine oocyte cDNA library, was shown to play an essential role in bovine early embryonic development. Previous work revealed that the ZNFO cDNA contains a 2,145 bp open reading fragment (ORF) encoding a protein of 714 amino acids, composed of a conserved N-terminal KRAB domain and a nine-motif zinc finger domain\textsuperscript{7}. RNAi-mediated knock-down of ZNFO can impair the development of bovine embryos to blastocyst stage\textsuperscript{7} indicating the requirement of ZNFO for early embryonic development. As a KRAB containing ZNFs, ZNFO can interact with KAP1 to perform a repressive regulatory function\textsuperscript{7}. Based on these previous data, it is speculated that ZNFO mediates downstream activity of potential target genes through a cis-acting ZNFO recognition sequence. Herein, we report the identification of a potential ZNFO DNA binding element (ZBE) using a cyclic amplification and selection (CASTing) assay followed by an electrophoretic mobility shift assay (EMSA). The study provides essential information for future identification of candidate genes regulated by ZNFO, which will help better understanding the molecular mechanism through which ZNFO regulates early embryonic development.

MATERIALS AND METHODS

Plasmid Construction

For expression of a His\textsubscript{6}Halo tag fusion recombination protein, the entire zinc finger domain (ZNF) coding sequence from ZNFO, which spans from 1233 bp to 2145 bp was cloned into pH6HTN His\textsubscript{6}Halo Tag\textsuperscript{®} T7 vector (Promega, Madison, WI). ZNFO-XbaI forward primer and ZNFO-NotI reverse primer (Table 1) were designed to amplify the zinc finger domain coding sequence of ZNFO by polymerase chain reaction (PCR). The reaction mixture contains 10 ng of a ZNFO plasmid containing a full-length cDNA, 10 µM of each deoxynucleoside triphosphate (dNTPs) and 5 units of Taq polymerase (Promega, Madison, WI). The PCR was carried out using the following program: an initial denaturation of 3 min at 95 °C followed by 35 cycles of 30 seconds at 95°C, 30 seconds at 60°C and 1 min at 72°C and then a final extension at 72°C for 10 min. Totally, 20 cycles were performed. The PCR products were purified using the QIAquick nucleotide removal kit (Qiagen) according to the manufacturer’s instructions. The purified PCR products were digested using restriction enzymes XbaI and NotI (NEB) followed by another round of purification using the same nucleotide removal kit as described before. Finally, the digested DNA was ligated to the vector digested with the same enzymes by incubation at 4 °C overnight in a reaction containing 3 units of T4 DNA ligase (Promega, Madison, WI). The clone (named His\textsubscript{6}Halo-ZNFO-ZNF) was completed, then sequenced to confirm no sequence errors were introduced.

Expression and Purification of Fusion Protein His\textsubscript{6}Halo-ZNFO

In order to express His\textsubscript{6}Halo-ZNFO-ZNF, the plasmid His\textsubscript{6}Halo-ZNFO-ZNF was transformed into Escherichia coli KRX (Promega, Madison, WI) strain and induced according to the manufacturer’s instructions. First, cells were incubated overnight at 37°C in Luria-Berta (LB) media containing 0.4% of glucose as well as ampicillin (100µg/ml) to set up a start culture. Then an induction culture media, which is composed of Luria-Berta media containing 0.05% of glucose, 0.1% of rhamnose and 100µg/ml ampicillin, was seeded with one tenth of start culture and incubated at 23°C for 18 hours. Bacterial pellets were collected by centrifugation and lysed in Halo tag purification buffer (50mM HEPES, 150mM NaCl, pH 7.5) containing 0.1mg/ml of lysozyme.
(Thermo Fisher Scientific, Waltham, MA) and 0.01U/ml of RQ1 RNase-Free DNase (Promega, Madison, WI) on ice, followed by sonication for 5 sec, 12 times (5 sec on, 5 sec off) at 50% amplitude. The soluble protein and insoluble debris were then separated by centrifugation. The supernatant was incubated with equilibrated HaloLink resin (Promega, Madison, WI) at 4°C overnight. After up to six times wash with Halotag purification buffer, the recombinant protein was either maintained on resin and stored at 4°C or cleaved by TEV protease (Promega, Madison, WI), followed by removal of TEV protease by HisLink protein purification resin (Promega, Madison, WI).

**Cyclic Amplification and Selection of Target (CASTing) Assay**

A library of single-strand oligonucleotides containing a 20 bp random core sequence flanked on each side by 21 bp nucleotides (5’-AAAGCTGAGTATGCCGAGCTG-N$_{20}$-GTTACGCACCTGGTTGAATGT-3’) was generated. Double-strand oligonucleotides were prepared by incubating 400 pmol of library in a PCR buffer (Promega, Madison, WI) containing 10 µM of each dNTPs, and 5 units of Taq polymerase and amplified using the following program: 3 min at 95°C, 10 min at 37°C followed by 20 min at 70°C. Totally, 20 cycles were performed. The double-strand oligonucleotides were purified using QIAquick nucleotide removal kit (Qiagen). The capture reaction was performed by mixing the library with 160 ng of purified His$_6$Halo-ZNFO-ZNF protein immobilized on HaloLink resin in binding buffer containing 25 mM Tris (pH 7.5), 100mM KCl, 1mM MgCl$_2$, 5mM Dithiothreitol (DTT), 5% glycerol, 100µM ZnCl$_2$, 0.1% tween20, 100µg/ml poly (dI-dC) and 1mg/ml bovine serum albumin (BSA). After incubating for 60 min at room temperature with rotation, the resin was washed five times with cold wash buffer which is the same as capture buffer but without poly (dI-dC) and then boiled for 10 min in 40 µl sterile H$_2$O. The eluted nucleotides were then amplified by PCR using the flanking primers (Table 1) and subsequently used for a second round of capture. After six rounds of capture, the PCR products were purified by QIAquick nucleotide removal kit (Qiagen) and cloned into pGEM-T easy vector (Promega, Madison, WI). Random clones were picked and sequenced. The sequences were trimmed and aligned using the multiple sequence alignment software, Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/).

**Electrophoretic Mobility Shift Assay (EMSA)**

Oligonucleotides used as labeled probe or competitor were designed based on the identified consensus sequences and ordered from Integrated DNA Technologies (Skokie, Illinois). Both the sense and antisense oligonucleotides (Table 1) used for the probe were labeled with IRDye700 at 5’end and oligonucleotides used for competitors have the exact sequences as probe but without labeling. In order to generate the double-stranded probe or competitors, a “-1°C/cycle” annealing program was performed in PCR machine, which starts with a 5-min incubation at 95°C, followed by 70 cycles, with temperature decreasing one degree each cycle. EMSA was performed in a 20-µl reaction at a final concentration of 10mM HEPES (pH 7.9), 50µg poly(dl-dC), 5% glycerol, 5µg/µl BSA, 50mM KCl, 5mM MgCl$_2$, and 1mM ZnCl$_2$, EMSA probe containing core consensus sequence. The reactions were conducted by incubating IRDye700-labeled probe with 160 ng of purified His$_6$Halo-ZNFO-ZNF protein at 20°C for 30 min. For competitor assay, purified protein was incubated on ice with cold competitors for 10 min before addition of probe. The IRDye 700 labeling allows the probe to be visualized on an Odyssey system (Li-COR, Lincoln, NE).

**RESULTS**

**Production of Recombinant ZNFO DNA Binding Domain Protein**
In order to express a Halo-tagged fusion protein containing the DNA binding domain of ZNF6, the sequence coding for the C-terminal zinc finger domain was cloned into pH6HTN His₆-Halo Tag® T7 vector, which has a N-terminal His₆-Halo tag followed by a TEV cleavage site as well as a multiple cloning site (Fig. 1). The clone, named His₆-Halo-ZNFO-ZNF, was confirmed by sequencing and was subsequently used for protein induction. The construct was transformed into Escherichia coli KRX strain and was induced to express the recombinant protein. The fusion protein was purified and examined by a 12% SDS PAGE gel (Fig. 2A). The purified protein corresponds to the estimated size of fusion protein. The vector that only expresses the His₆-Halo tag was used as a control.

Identification of the DNA Binding Element for ZNFO

To determine the consensus target sequence of ZNFO, a CASTing assay was performed using the purified fusion protein, which was immobilized on HaloLink Resin. Double-stranded oligonucleotides captured by ZNFO protein went through six rounds of CASTing assay. Each successive round of selection enriched the oligonucleotide core sequence. As shown in Fig 2B, the ZNFO captured oligonucleotides were enriched and then maintained with each round as compared to the His₆-Halo control in which an initial weak signal was detected after the first round and then completely diminished by the second round. The purified products from the final round of CASTing were cloned into pGEM-T easy vector. A total of 43 sequences were obtained and analyzed (Fig. 3A). Through the analysis, a consensus DNA binding sequence (ZBE), ATATCCTGN₅ANCCC, was identified (Fig. 3B).

Confirmation of Target Consensus of ZNFO

To confirm the binding of ZNFO to the identified consensus sequence, we performed electrophoretic mobility shift assay (EMSA) using eluted ZNFO-ZNF protein and the IRDye700 labeled probe containing the consensus DNA binding sequence as well as cold competitors. As shown in Fig. 5, ZNFO-ZNF protein bound to the EMSA probe causing a shift. The bound complex was competed out by 10-fold and 100-fold molar excess of unlabeled cold competitors. Further competitive binding assays using cold competitors containing mutated DNA binding elements showed that they fail to compete for the probe indicating that both ATATCCTG and CCC of ATATCCTGN₅ANCCC are essential for ZNFO binding (Fig. 6).

DISCUSSION

Cyclic amplification and selection of target (CASTing) assay is one of the key methods used in our research. CASTing assay was first used to identify myogenin consensus binding site in 1991. CASTing assay first pulls down the target DNA sequence interacting with the protein of interest from a random library in the first round and then DNA pulled down by the protein is amplified as a secondary library which is much more purified than the random one. Through incubation with the secondary library, protein can continuously pull down target DNA sequences. And through a few rounds of CASTing assay the products are purified from a random library. In order to pull down the target DNA, the protein and DNA complex must be able to be separated from the supernatant. In this case, a tag which can immobilized on certain resins, like GST or an universal bead which is immobilized with a specific antibody is required. In our research, we introduced a His₆-Halo® Tag to DNA binding domain of ZNFO which immobilize zinc finger domain of ZNFO on the HaloLink Resin. Through interaction, it separates the target DNA from the random library (Fig. 6).
Electrophoretic mobility shift assay (EMSA) is another key technology in our study. EMSA is a rapid and sensitive method to detect the interaction between protein and nucleic acid\textsuperscript{174}. It is based on the observation that the electrophoretic mobility of protein-nucleic acid complex is typically less compared with free nucleic acid with exact the same sequence, causing a “shift” phenomenon. In the process of EMSA, either purified artificial protein induced from Escherichia coli or nuclear extract which need to perform a super-shift assay using a specific antibody can be applied. In terms of labeling, various labels have been applied in EMSA, such as radiolabeled probe usually $^{32}$p\textsuperscript{175}, biotinylated probe\textsuperscript{176}, as well as fluorescent labeled probe. In our study of ZNFO zinc finger domain. Zinc finger domain was cleaved from HaloLink Resin when a TEV protease was introduced which allows the purified Zinc finger motif to be applied in EMSA. In this case, a super shift assay is not required. In terms of labeling, IRDye 700 labeling allows the interaction can be detected on gel, avoiding the transfer efficiency issue and making it much more sensitive compared with biotin.

The identified ZNFO binding element, ATATCCTGN\textsubscript{5}ANCCC, appears to have two core sequences, ATATCCTG and CCC. However, our competitive binding assay with competitors containing mutated binding elements is limited. Only a single nucleotide mutation in each of the two core sequences was made. It is possible that some of the nucleotides in these two core sequences are not essential for ZNFO binding. A more comprehensive competitive binding assay with mutations covering every nucleotide of the core sequences will provide a better picture of the essential sequence required for ZNFO binding.

During early embryonic development, there are various characterized or uncharacterized factors regulating post-fertilization. Similar binding property of two different factors which are expressed in same tissue, may indicate overlapping regulatory function. For example, the binding element of LHX8 is TGATTG, which is slightly different from another known homeoprotein NOBOX (TAATTG/A)\textsuperscript{177}. And it was speculated that LHX8 and NOBOX may have similar function. For ZNFO, characterizing its DNA binding property can promote the understanding of molecular regulatory mechanism during early embryonic development in cattle. However, to elucidate the regulatory network during embryonic development in cattle, a ChIP-Seq experiment is required to identify downstream genes of ZNFO.
SUMMARY

As shown in our previous results, ZNFO has an essential physiological function in the regulation of early embryonic development in cattle\(^7\). As a transcription factor, ZNFO might be involved in a complicated regulatory network, which orchestrates post-fertilization events during early stage development in cattle. In this study, our objective was focused on the DNA binding property of the zinc finger domain (ZNF) of ZNFO. We identified a consensus sequence which can be captured by the ZNF of ZNFO (ATATCCTGN\(_5\)ANCCC) using CASTing approach followed by gel shift assay. We further confirmed that both ATATCCTG and CCC are essential for the interaction between ZNFO and the consensus sequence through a competitive binding assay using two mutant competitors, ATATCCTGN\(_5\)ANCTC and ATATTCTGN\(_5\)ANCCC.

This study is the first step to elucidate the target gene(s) regulated by ZNFO. In general, either activators or repressors need to recognize certain element(s) in genome, and then recruit transcription machinery or repressive complex, respectively. For example, RNA polymerase II use TFIIID as positioning factor that recognizes TATA box or downstream promoter element (DPE). The nine-motif ZNF domain of ZNFO is the positioning domain, which can locate repressive complex in genome. With this consensus element information, we will align promoter region or enhancer region in bovine genome to identify candidate target genes of ZNFO. The genes having ZBE in the promoter region or enhancer region could be the candidate genes regulated by ZNFO.

Further experiments using Chromatin immunoprecipitation (ChIP) assay is required to confirm ZNFO target genes. ChIP is to cross-link chromatin, which preserves DNA-protein interaction. Following DNA fragmentation, a specific antibody is used to pull-down the complex containing protein of interest. ChIP assay can be followed by PCR or deep sequencing. Since ZNFO is oocyte-specific, we plan to perform ChIP assay using oocytes. Once the complex is pulled-down, we will perform PCR amplification based on the candidate genes identified through data mining using the consensus element information. In addition, ChIP-Seq will be performed to systematically identify ZNFO target genes.
Figure 1. Clone of DNA binding domain into pH6HTNHis6Halo® T7 vector. The whole C-terminal zinc finger domain was cloned into pH6HTNHis6Halo® T7 vector using XbaI and NotI sites.
Figure 2. Expression of His\textsubscript{6}Halo ZNFO DNA binding domain fusion protein. CASTing assays were performed and binding and amplification were done with His\textsubscript{6}Halo or His\textsubscript{6}Halo-ZNFO DNA binding domain (ZNFO-ZNF) fusion protein. (A) His\textsubscript{6}Halo protein and His\textsubscript{6}Halo tagged ZNFO-ZNF protein were induced in Escherichia coli KRX strain using rhamnose. Coomassie blue staining of His\textsubscript{6}Halo and His\textsubscript{6}HaloZNFO-ZNFs bound to HaloLink Resin. (B) Results from PCR amplification of bound DNA in CASTing assay.
Figure 3. Consensus sequences recognized by ZNFO. (A) Forty-three sequences were selected to analyze the potential consensus element bound by ZNFO. (B) The alignment of sequences revealed one potential consensus target recognized by ZNFO, ZNFO binding element (ZBE), Figure 3B was generated using MEME suite (http://meme-suite.org).
Figure 4. Confirmation of Consensus Target Bound to ZNFO. A IRDye 700 labeled ZNFO probe containing core sequence ATATCCTGN₅ANCCC was incubated with purified ZNFO-ZNF protein. A 10-fold or 100-fold molar excess of unlabeled ATATCCTGN₅ANCCC sequence was used as a competitor. The interaction between ZNFO-ZNF and probe causes the protein-DNA complex moving slowly compared with free probe, which is observed as a shift. When specific competitor was introduced, the interaction between ZNFO-ZNF and probe was competed out, causing the shift weaker or barely detectable.
**Figure 5. Mutation of Consensus Target Bound to ZNFO.** A IRDye 700 labeled ZNFO probe containing core sequence ATATCCTGN_5ANCCC was incubated with purified ZNFO-ZNF protein. A 10-fold or 100-fold molar excess of unlabeled ATATCCTGN_5ANCCC, ATATCCTGN_5ANCTC, ATATCTGN_5ANCTC sequence was used as a target element competitor, mutation 1 element competitor and mutation 2 element competitor, respectively. As shown in figure, mutant competitors cannot compete out the interaction between ZNFO and probe while specific competitors can compete out the interaction between ZNFO and probe.

<table>
<thead>
<tr>
<th>Competitor</th>
<th>Target</th>
<th>Mutation 1</th>
<th>Mutation 2</th>
<th>Protein</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>10X</td>
<td>100X</td>
<td></td>
</tr>
</tbody>
</table>
**Figure 6. CASTing assay of ZNFO.** The His6Halo tag was fused with DNA binding domain of ZNFO to pull down consensus recognized by ZNFO from a random library

Oligo library: random sequences
AAAGCTGAGTATGCCGAGCTGNNNNNNNNNNNNNNNNTACGACCTGTTGAATGT
TTTCGACTCATACGCTCGACNNNNNNNNNNNNNNNCAATCGTGACCAAACCTTACA

Binding

Halo-tagged ZNFO protein:

Wash, boil and PCR
Repeat 5 more times
Clone final round products
Sequence analysis

Identify sequences:
AAAGCTGAGTATGCCGAGCTGNNNNAAACCCCNNNNNNTACGACCTGTTGAATGT
TTTCGACTCATACGCTCGACNNNNNTTGGGCGNNNNNNNNCAATCGTGACCAAACCTTACA
### Table 1. Primers/oligos Used in This Study

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Primer Sequences*5' (to 3')</th>
<th>Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZNF- Zenly-Nba1-F</td>
<td>GGGCCCTCTAGATATGCTAAAGCTTTTAAACCCAGTC</td>
<td>ZNF-ZNF clone</td>
</tr>
<tr>
<td>ZNF-Net1-R</td>
<td>GGCCGCCCGCCGCTCAAGAGGTCTCTGCAATAG</td>
<td>ZNF-ZNF clone</td>
</tr>
<tr>
<td>ZNF-EMSA-IRDy7700-F</td>
<td>IRDy7700ACACGCACTGTAACGGGTTAACACAGATATACCTCAGCACATAC</td>
<td>EMSA probe</td>
</tr>
<tr>
<td>ZNF-EMSA-IRDy7700-R</td>
<td>IRDy7700ACACGCACTGTAACGGGTTAACACAGATATACCTCAGCACATAC</td>
<td>EMSA probe</td>
</tr>
<tr>
<td>ZNF-EMSA-WT-F</td>
<td>GTATGCGCTGCATATCCGTATTTAAAACGGCTTACGGACCTG</td>
<td>EMSA target element competitor</td>
</tr>
<tr>
<td>ZNF-EMSA-WT-R</td>
<td>ACCAGTGTCGCTCAACCGGTTAAAGCGATATCACGCTCGGCAATAC</td>
<td>EMSA target element competitor</td>
</tr>
<tr>
<td>ZNF-EMSA-Mutation1-F</td>
<td>GTATGCGCTGCATATCCGTATTTAAAACGGCTTACGGACCTG</td>
<td>EMSA mutation element1 competitor</td>
</tr>
<tr>
<td>ZNF-EMSA-Mutation1-R</td>
<td>ACCAGTGTCGCTCAACCGGTTAAAGCGATATCACGCTCGGCAATAC</td>
<td>EMSA mutation element1 competitor</td>
</tr>
<tr>
<td>ZNF-EMSA-Mutation2-F</td>
<td>GTATGCGCTGCATATCCGTATTTAAAACGGCTTACGGACCTG</td>
<td>EMSA mutation element2 competitor</td>
</tr>
<tr>
<td>ZNF-EMSA-Mutation2-R</td>
<td>ACCAGTGTCGCTCAACCGGTTAAAGCGATATCACGCTCGGCAATAC</td>
<td>EMSA mutation element2 competitor</td>
</tr>
<tr>
<td>LOC703-1072F</td>
<td>AAAGCTGAGTATGCGGACTG</td>
<td>CASTing</td>
</tr>
<tr>
<td>LOC703-1072R</td>
<td>ACATTCAACAGCTGATAC</td>
<td>CASTing</td>
</tr>
</tbody>
</table>
References


140. Ono, R., Kaisho, T. & Tanaka, T. PDLIM1 inhibits NF-κB-mediated inflammatory signaling by sequestering the p65 subunit of NF-κB in the cytoplasm. Scientific Reports 5, 18327 (2015).
158. Sripathy, S.P., Stevens, J. & Schultz, D.C. The KAP1 Corepressor Functions To Coordinate the Assembly of De Novo HP1-Demarcated Microenvironments of Heterochromatin


