Maintaining Proper Levels of DNA Methylation Marks Through the TET Family is Critical for Normal Embryo Development in Pigs

Kyungjun Uh

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Animal and Poultry Sciences

Kiho Lee
Alan D. Ealy
Willard Eyestone
Michelle L. Rhoads
Hehuang Xie

August 5, 2020
Blacksburg, VA

Keywords: DNA methylation, TET family, preimplantation embryo, pluripotency

Maintaining Proper Levels of DNA Methylation Marks Through the TET Family is Critical for Normal Embryo Development in Pigs

Kyungjun Uh

ABSTRACT

DNA methylation is one of the principal epigenetic modifications that plays an essential role in transcriptional regulation. After fertilization, mammalian embryos undergo dynamic changes in genome-wide DNA methylation patterns and the changes are essential for normal embryo development. Ten-eleven translocation (TET) methylcytosine dioxygenases are implicated in DNA demethylation by catalyzing the conversion of 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC). The three members of TET protein family, TET1, TET2, and TET3, are highly expressed in preimplantation embryos in a stage-specific manner. Previous studies demonstrated that TET proteins are involved in diverse biological processes such as gene regulation, pluripotency maintenance, and cell differentiation by mediating 5mC oxidation. My dissertation research was conducted to elucidate the mechanistic roles of TET proteins in epigenetic reprogramming of mammalian embryos using porcine embryos as a model.

The first set of studies focused on the relationship between TET proteins and pluripotency. To understand the role of TET proteins in establishing pluripotency in preimplantation embryos, CRISPR/Cas9 technology and TET-specific inhibitors were applied. *TET1* depletion unexpectedly resulted in an increased expression of *NANOG* and *ESRRB* genes in blastocysts, although the DNA methylation levels of *NANOG* promoter were not changed. Interestingly, transcript abundance of *TET3* was increased in blastocysts carrying inactivated *TET1*, which might have had an effect on the increase of *NANOG* and *ESRRB*. When the activity of TET enzymes was inhibited to eliminate the compensatory increase of *TET3* under the absence of functional TET1, the expression levels of *NANOG* and *ESRRB* were decreased and methylation level of *NANOG* promoter was increased. In addition, ICM specification was impaired by the inhibition of TET enzymes. These results suggest that the TET family is a critical component of the pluripotency network of porcine embryos by regulating expression of genes involved in pluripotency and early lineage specification. In the next set of studies, the presence of *TET3* isoforms in porcine oocytes and cumulus cells was investigated to dissect the gene structure of *TET3* that could assist in understanding mechanistic actions of *TET3* in the DNA demethylation process. Among the three *TET3* isoforms identified in cumulus cells, only the *pTET3L* isoform, which contains CXXC domain that carry DNA binding property, was verified in mature porcine oocytes. Expression level of the *pTET3L* isoform was much higher in mature oocytes compared to that in somatic cells and tissues. In addition, the transcript level of this isoform was significantly increased during oocyte maturation. These results suggest that *pTET3L* isoform is predominantly present in mature porcine oocytes and that CXXC domain may play an important role in DNA demethylation in zygotes. In a follow-up study, the role of the *TET3* CXXC domain in controlling post-fertilization demethylation in porcine embryos was investigated by injecting *TET3* GFP-CXXC into mature porcine oocytes. The injected CXXC was

exclusively localized in the pronuclei, indicating that the CXXC domain may localize TET3 to the nucleus. The CXXC overexpression reduced the 5mC level in zygotes and enhanced the DNA demethylation of the *NANOG* promoter in 2-cell stage embryos. Furthermore, the transcript abundance of *NANOG* and *ESRRB* was increased in blastocysts derived from GFP-CXXC overexpressing zygotes. These results provide an evidence that the CXXC domain of *TET3* is critical for post-fertilization demethylation of porcine embryos and proper expression of pluripotency related genes in blastocysts. In the last set of studies, the impact of MBD proteins on porcine embryo development was examined under the hypothesis that competitive binding of MBD and TET proteins to 5mC contributes to the proper maintenance of DNA methylation levels in embryos. Cloning of porcine *MBD1*, *MBD3*, and *MBD4* from mature oocytes indicates that the genes are highly conserved among different species, implying the involvement of porcine MBD proteins in the maintenance of DNA methylation. *MBD1* overexpression in oocytes impaired preimplantation development of porcine embryos, suggesting that the *MBD1* overexpression may have negatively affected porcine embryo development because proper DNA methylation levels were not preserved under the high level of MBD1.

Collectively, the studies in my dissertation demonstrate that TET family proteins are important epigenetic players involved in the regulation of pluripotency and reprogramming of DNA methylation, and are thus crucial for normal embryo development. The findings in the dissertation will improve our understanding of epigenetic events occurring in mammalian embryos, and have the potential to overcome epigenetic defects that are observed in pluripotent stem cells and in-vitro derived embryos.

Maintaining Proper Levels of DNA Methylation Marks Through the TET Family is Critical for Normal Embryo Development in Pigs

Kyungjun Uh

GENERAL AUDIENCE ABSTRACT

Epigenetic modifications are heritable changes affecting the level of gene expression without changing the sequence of the genome. DNA methylation, one of the biggest epigenetic marks in mammalian genome, is often correlated to gene repression. In mammals, DNA methylation patterns are dramatically changed during preimplantation development to acquire embryonic developmental potential. Understanding of the epigenetic changes occurring in preimplantation embryos is necessary for producing healthy domestic animals in agriculture and for developing strategies for the treatment of epigenetic defects in human. Ten-eleven translocation (TET) family enzymes, TET1, TET2, and TET3, are known to function as a DNA methylation modifier by converting 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC). My dissertation research was performed to elucidate the role of TET family during preimplantation development using porcine embryos as a model.

Pluripotency refers to the ability of cells to differentiate into all cell types of a mature organism. Pluripotent cells emerge in embryos as embryonic cells acquire lineage-specific characteristics. The first set of studies focused on the role of TET enzymes in regulating the pluripotency of porcine embryos. The impacts of inhibited activities of TET enzymes on the expression of pluripotency related genes were examined. We found that the inhibition of all TET enzymes leads to a decreased expression of pluripotency related genes, an altered DNA methylation level on a gene segment controlling pluripotency, and the impaired formation of pluripotent cell lineage in porcine embryos. This study demonstrates that the TET family is critical for the acquisition of pluripotency in porcine embryos. In the following sets of studies, the function of TET3 protein in the demethylation process occurring in preimplantation embryos was investigated. Fertilized mammalian embryos undergo genome-wide demethylation process to reset germ cell specific epigenetic marks into the embryonic epigenome. Previous studies indicate that TET3 is responsible for the demethylation process in mammalian embryos, although detailed mechanistic action of TET3 is still elusive. Here, we identified a predominant expression of a specific *TET3* gene in porcine oocytes. The *TET3* gene contained a CXXC domain, a potential DNA binding module, suggesting that TET3 may mediate DNA demethylation through its DNA binding property. To examine the function of the CXXC domain in TET3-mediated DNA demethylation, isolated CXXC domain was injected into porcine oocytes. The injection of CXXC domain facilitated DNA demethylation in embryos, demonstrating that the DNA binding property of TET3 is important for its functionality. In the last study, we investigated the importance of genes known to interact with TET enzymes in porcine embryos. Methyl-CpG-binding domain proteins (MBDs) have the ability to bind methylated region on the genome and play a critical role in mediating DNA methylation and gene repression. Our hypothesis was that a competitive binding of MBD and TET

proteins to methylated regions was critical for proper DNA methylation levels in embryos. We identified that porcine *MBD* sequences were very similar to other species in terms of gene structure, indicating that the genes could also possess gene repressing activity by competing with TET enzymes during porcine embryo development. Injection of *MBD1* mRNA to oocytes impaired normal embryo development, suggesting that the injected *MBD1* may have negatively affected early embryo development in pigs by disrupting the proper maintenance of DNA methylation levels.

My dissertation researches demonstrate that maintaining proper DNA methylation levels through the TET family is critical for normal embryo development in pigs. This research assists in improving our understating of epigenetic dynamics occurring in mammalian embryos and offers a potential solution to the epigenetic defects frequently observed in mammalian embryos produced through artificial reproductive technologies and pluripotent stem cells reprogrammed from somatic cells.

ACKNOWLEDGEMENTS

I wish to express my deepest gratitude to my supervisor, Dr. Kiho Lee, who convincingly guided me and encouraged me to be professional and be on the right track even when I get bogged down. Without his persistent help, my PhD studies would not have been successful. I hope to someday tell my own students what I heard and learned from you.

I would like to thank the members of the Lee lab, past and present. Without your help, my projects would not have been completed. And all of you were the best English teachers to me. Thank you for your patience in my poor English proficiency.

To my committee members, Dr. Alan Ealy, Dr. Will Eyestone, Dr. Michelle Rhoads, and Dr. Hehuang Xie, thank you for your guidance. Your advice expanded my perspective and pushed me to be a better scientist.

I also would like to thank my family. My parents, thank you for supporting me and believing in me throughout my academic carrier; I could not have done it without them. I also thank my brother, Dr. Kyungchan Uh, for encouraging me whenever I'm stressed out. Finally, to my loving and supportive wife, Junghwa, I would like to express my deepest gratitude. When the times got rough, your encouragement and consideration pulled me up out of it. My heartfelt thanks.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AID: activation-induced deaminase

BER: base pair excision repair

CDS: coding sequence

COC: cumulus oocyte complex

DMOG: dimethyloxalylglycine

DNMT: DNA methyltransferase

DSBH: double-stranded β-helix

ICC: immunocytochemistry

ICM: inner cell mass

ESC: embryonic stem cell

GV: germinal vesicle

HSC: hematopoietic stem cell

iPSC: induced pluripotent stem cell

IVF: *in vitro* fertilization

KO: knockout

MBD protein: methyl-CpG-binding domain protein

MII: metaphase II

NLS: nuclear localization sequence

PGC: primordial germ cell

PN: pronuclear

SCNT: somatic cell nuclear transfer

TDG: thymine-DNA glycosylase

TE: trophectoderm

TET family: ten-eleven translocation family

UTR: untranslated region

ZGA: zygotic genome activation

2OG: 2-oxoglutarate

5mC: 5-methylcytosine

5hmC: 5-hydroxymethylcytosine

5fC: 5-formylcytosine

5caC: 5-carboxycytosine

Chapter 1. Literature Review

Preimplantation embryo development is initiated by the fusion of two gametes, the sperm and oocyte. Once the haploid genomes of each gamete are combined, embryos go through cell cleavage and morula compaction to reach the blastocyst stage. During preimplantation embryo development, the epigenetic landscape of embryos is dynamically changed through a series of well-orchestrated epigenetic changes that are essential for embryo survival and successful development. The epigenetic changes that preimplantation embryos undergo include DNA methylation, histone modification, and chromatin remodeling. Epigenetic reprogramming factors originating in oocytes play a crucial role during early embryo development by remodeling the epigenetic status of the zygotic genome. Although proper DNA methylation reprogramming is key for development, the detailed mechanism governing the reprogramming process has remained elusive. Furthermore, different mammals present unique epigenetic dynamics during embryo development. Understanding key molecules governing the differences as a comparative study can expand our knowledge on epigenetic regulations during embryogenesis. Recent studies indicate that the Ten-eleven translocation (TET) family is a key epigenetic modulator regulating DNA methylation in preimplantation embryos. Specifically, TET family is suggested to initiate the DNA demethylation process after fertilization and maintain proper levels of DNA methylation during embryogenesis by catalyzing the oxidation of 5-methylcytosine (5mC). Although the role of TET family as an initiator of DNA demethylation has been reported, detailed mechanistic action of TET-mediated epigenetic control is still under investigation. In this review, I illustrate the importance of DNA methylation for normal embryo development and current knowledge on the function of TET family in controlling DNA methylation in mammalian embryos.

1. Developmental consequences of epigenetic reprogramming failure

Genome-wide epigenetic reprogramming, including imprinting, DNA methylation, histone modification, and chromatin remodeling, occurs during germ cell and preimplantation embryo development. Epigenetic reprogramming is involved in embryo development, fetal development, and birthweight; thus, disrupted reprogramming can lead to embryonic defects, mortality, or postnatal disorders. DNA methylation is established and maintained by DNA methyltransferases, DNMT1 and DNMT3A/B. In mice, the loss of DNMTs results in midgestational or postnatal lethality (1). In humans, mutations in the DNMT3B gene lead to immunodeficiency, facial anomalies, mental retardation and developmental delay in patients (2). Methyl-CpG binding proteins, such as MECP2, are recruited to methylated genomic regions and induce changes in chromatin configuration and transcription (3). The methyl-CpG binding proteins are crucial for embryo viability during peri-implantation development, thus mouse embryos lacking the methyl-CpG binding proteins cannot survive beyond 12.5 days postcoitum (4). In humans, mutations in the methyl-CpG binding proteins prevent transcriptional repression of methylated DNA, ultimately leading to Rett syndrome or X-linked syndrome of psychosis (5, 6). DNA methylation is also crucial for regulating the expression of imprinted genes.

Unlike most autosomal genes, imprinted genes only express one allele of a gene pair and expression occurs in a parent-of-origin-specific manner, which is mostly determined by the methylation level in imprinting control regions (7, 8). Genomic imprinting is erased in primordial germ cells (PGCs), allowing germ cell specific imprints to be established in gametes and maintained during post-fertilization development (9). The allele-specific expression of imprinted genes, controlled by DNA methylation, is crucial for embryonic and postnatal growth, and placental functions (10). For example, an oocyte specific variant of DNMT1 (DNMT1o) is

responsible for maintenance of genomic imprints in preimplantation embryos. Mouse embryos lacking DNMT1o lose genomic imprints, and subsequently die before birth (1, 11, 12). Mutations in imprinted genes cause a number of human imprinting diseases. For instance, genetic or epigenetic abnormalities at the *H19*/*IGF2* locus results in Beckwith-Wiedemann syndrome or Russell-Silver syndrome, and failure of allele-specific expression of *SNRPN* gene causes Prader-Willi syndrome or Angelman syndrome (13). Cancers in various tissues are also caused by mutations in epigenetic modifiers, such as DNMTs, TET enzymes, and histone modifiers (14). Hence, studies of epigenetic regulation of gene expression and chromatin configuration during development will be beneficial to reveal the underlying causes of epigenetic reprogramming related human diseases.

2. DNA methylation dynamics during preimplantation development

In 1975, it was reported that methylation of cytosine residues in the genome can serve as an epigenetic mark in vertebrates (15, 16). Methylation on the fifth carbon of cytosine produces 5-methylcytosine (5mC), which carries epigenetic information, primarily in the context of CpG dinucleotides in mammals (17, 18). Non-CpG methylation has recently been identified in oocytes, embryonic stem cells (ESCs), and neurons (19-22); however, the epigenetic role of such methylation marks still remain to be identified. Generally, high level of DNA methylation is linked to gene repression because high CpG density in promoter regions generally codes for the repression of gene transcription (23-25). The formation of DNA methylation is mediated by DNA methyltransferases (DNMTs), which catalyze the transfer of methyl groups from S-adenosyl-l-methionine to the fifth carbon of cytosine residues (26).

Although differences in methylation levels between the two gametes exist, the genomes of mature sperm and oocytes are highly methylated compared to that of somatic cells (27). Following fertilization, global epigenetic reprogramming occurs during early embryo development by erasing germ cell-specific epigenetic marks and replacing them with embryo-specific markers. This global epigenetic reprogramming is crucial for establishing pluripotency of developing embryos. In the mouse, highly methylated genomes, inherited from germ cells, are dramatically demethylated after fertilization until the blastocyst stage (28-31); the demethylation excludes imprinted genes and repetitive elements. The paternal and maternal genome exhibits different rates of demethylation during preimplantation development and involvement of distinct demethylation pathways in the parental genomes has been reported (28-31). Overall, the level of CpG methylation of the mature sperm genome is 80 – 90%, while the methylation level of a mature oocyte is comparatively lower (about 40 %) in mice (32). The highly methylated paternal genome is completely demethylated shortly after fertilization, whereas the maternal genome is gradually demethylated as embryos develop, implicating the presence of distinct demethylation mechanisms in the parental genomes (33). The gradual loss of methylation in the maternal genome indicates that methylation marks are presumably diluted following successive DNA replications in the absence of active DNMT1 (34-36), as DNMT1 adds methyl groups to the nascent strand of hemimethylated DNA after DNA replication (18, 37). While the maternal genome is passively demethylated following DNA replication, the complete loss of methylation in the paternal genome suggests the existence of an active demethylation process, as methylation marks are erased before the onset of DNA replication at the pronuclear stage 3 (PN3) (33, 38). Although immunocytochemistry assays for 5mC have demonstrated the rapid loss of methylation marks in the paternal pronucleus, incomplete demethylation has been detected in bisulfite

sequencing analysis at the same embryonic stage (39, 40), raising a possibility of an unknown mechanism that mediates demethylation of the paternal genome. A leading mechanistic explanation is deamination of 5mC to thymidine by the activation-induced deaminase (AID) followed by thymine-DNA glycosylase (TDG) recognition and base pair excision repair (BER) (41). However, weak catalytic activity of TDG on the demethylation of 5mC:G dinucleotides (42) suggests that 5mC removal can be induced by unknown secondary modification of 5mC that induce TDG activity. Indeed, the discovery of novel modifications of cytosine in mammalian cells led to the identification of different demethylation mechanism in zygotes (43).

3. 5-hydroxymethycytosine (5hmC) and Ten-eleven translocation (TET) family

For more than six decades, 5mC has been recognized as the fifth base of mammalian DNA because of its important role as an epigenetic mark. On the other hand, the hydroxylated form of 5mC, 5-hydroxymethylcytosine (5hmC), received only little attention although the existence in mammalian DNA was first reported in the early 1970s by Penn et al. (44). In the study, they reported that 5hmC account for ~15% of total cytosines in DNA isolated from brain tissues of rat, mouse, and frog; however, the presence of 5hmC in mammalian DNA could not be confirmed by other studies until it was robustly detected in the mouse cerebellum (45) and ESCs (43) by two research groups in 2009. 5hmC accounted for 0.6% of the total nucleotides in mouse Purkinje neurons (45) and 0.03% of the total nucleotides in mouse ESCs (43). Later studies found 5hmC in various mouse and human tissues with high levels in the central nervous system (46). These findings led to the hypothesis that 5hmC may be an intermediate base pair during the process of 5mC removal.

Discovery of ten-eleven translocation (TET) proteins was inspired by the presence of base J (β-D-glucosyl hydroxymethyluracil) in trypanosomes, a modified thymine associated with gene silencing similar to 5mC (47). The production of base J is initiated by hydroxylation of the methyl group of thymine (47). JBP1 and JBP2, enzymes of the 2-oxoglutarate (2OG)- and Fe(II)-dependent oxygenase superfamily, have been predicted to catalyze the hydroxylation of the methyl group of thymine (48, 49). In 2009, Tahiliani et al. identified the ten-eleven translocation (TET) proteins as mammalian homologs of the trypanosome proteins JBP1 and JBP2 (43). TET enzymes are 2OG- and Fe(II)-dependent enzymes that catalyze the conversion of 5mC to 5hmC. Subsequent studies revealed that 5hmC can be further oxidized to 5-formylcytosien (5fC) and 5-carboxycytosine (5caC) by the TET enzymes (50, 51). The discovery of the three 5mC derivatives and TET enzymes suggested a new demethylation mechanism. Recent studies found that TDG, an enzyme mediating base excision repair of DNA, has direct activity with 5fC and 5caC, implying that the two 5mC derivatives are intermediates of the active demethylation process (50, 52). In addition, conversion of 5mC to 5hmC by TET enzymes may aid in the acceleration of passive demethylation even in the presence of active DNMT1 because the affinity of DNMT1 for 5hmC is much lower than for 5mC in hemi-methylated (5mC or 5hmC) DNA strands (53, 54).

In mammals, the TET family consists of three enzymes, TET1/2/3, and they are differentially expressed during embryo development and in adult tissues (43). TET1 and TET2 are highly expressed in the inner cell mass (ICM) of blastocysts, embryonic stem cells (ESCs), and primordial germ cells (PGCs) (55, 56). TET2 is also highly expressed in hematopoietic stem cells (HSCs) and essential in hematopoiesis, including HSCs self-renewal and lineage commitment (57). TET1 knockout (KO) resulted in reduced birth weight and subfertility in both

male and female mice; however, the modification did not lead to embryo lethality (58, 59). The mild impact of TET1 KO on development likely is due to the functional redundancy with TET1 and TET2, as they possess overlapping expression patterns in ESCs and HSCs (60, 61). Indeed, TET1/2 double KO led to more obvious developmental defects than the TET1 single KO such as perinatal lethality, although some of the double KO mice were viable and fertile (62). TET3 is detected in oocytes and early stage preimplantation embryos including zygotes, and also highly expressed in neurons (63-66). Homozygous TET3 KO resulted in neonatal lethality, and the loss of maternal TET3 caused embryonic sublethality (67).

3.1. Active DNA demethylation in the paternal genome of zygotes

As noted above, different demethylation statuses in the paternal genome of zygotes from immunocytochemistry assays and bisulfite sequencing analyses (39, 40) imply the involvement of unknown DNA demethylation pathways. The discovery of TET3 protein and 5hmC in oocytes and zygotes could provide an explanation for a mechanism mediating demethylation of the paternal genome in fertilized oocytes (67). Indeed, studies have demonstrated that the TET3 protein mediates demethylation of the paternal genome in zygotes by converting 5mC to 5hmC (63, 64, 67). The different results of methylation status from immunocytochemistry and bisulfite sequencing was due to the presence of 5hmC in the paternal pronucleus. Specifically, 5mC and its oxidative derivative 5hmC cannot be distinguished by conventional bisulfite sequencing analysis, whereas 5mC antibody specifically recognizes 5mC through immunocytochemistry. TET3 is enriched in the paternal pronucleus, where it catalyzes the oxidation of 5mC to 5hmC (63, 67). In the absence of TET3, 5mC remained constant in the paternal genome in zygotes (67). The deficiency of TET3 also impeded the demethylation of paternal alleles of genes essential for

embryo development and pluripotency (e.g. *OCT4* and *NANOG*), and delayed activation of paternally derived genes (67). These findings demonstrate that demethylation of the paternal genome in zygotes is initiated by TET3-mediated conversion of 5mC to 5hmC and is critical for proper activation of genes related to embryo development and pluripotency. In addition to 5hmC, 5fC and 5caC appear in zygotes concurrently with the loss of 5mC, suggesting that 5hmC is further oxidized potentially by TET3 before cleavage (68). The three 5mC derivatives, 5hmC, 5fC, and 5caC, can be direct targets for TDG/BER pathways (50, 69), thus ultimately can be converted into unmethylated cytosines. Further studies are necessary to clarify the subsequent demethylation process that reverts oxidized 5mC to unmethylated cytosine because TDG is not consistently detectable in zygotes (70).

3.2. Absence of active DNA demethylation in maternal pronucleus

In contrast to the paternal genome, 5mC in the maternal pronucleus is not targeted by TET3, thus protected from active demethylation in zygotes (63). STELLA protein, a maternal factor, is known to play a protective role against TET3-mediated 5mC oxidation in methylation status of the maternal pronucleus (63, 71). STELLA (also known as PGC7 and DPPA3) is a nuclear polypeptide that is highly expressed in PGCs, oocytes, and pluripotent cells, and is essential for embryo viability (72, 73). While mating heterozygous STELLA mutant mice resulted in the birth of STELLA-null offspring without developmental defects, the development of oocytes from STELLA-deficient females arrested during early cleavage (mostly arrested at 4-cell stage) following fertilization due to the lack of maternally inherited STELLA in the oocytes (74, 75). Prior to the discovery of 5hmC and TET proteins in mammals, in 2007, Nakamura and colleagues suggested that STELLA likely protects the maternal genome from demethylation

shortly after localizing to the maternal pronucleus in zygotes (76). Indeed, STELLA deficiency resulted in TET3-mediated 5hmC accumulation in the maternal pronucleus, demonstrating that STELLA is an important maternal factor for protecting the genome of maternal pronucleus against active TET3-mediated demethylation (63, 71). It has been demonstrated that the protective role of STELLA is determined by its interaction with dimethylated histone H3 Lys9 (H3K9me2), a histone methylation mark enriched only in the maternal pronucleus (71). STELLA preferentially binds to the maternal genome harboring H3K9me2 marks in zygotes and alters chromatin configuration, thus preventing TET3 binding and inhibiting TET3-mediated 5mC oxidation (71, 77). Interestingly, a large portion of the maternal genome remained methylated in the absence of STELLA (76) and, in the most recent study, global 5mC level was increased, rather than decreased in the STELLA-deficient zygotes (78). Further studies are needed to clarify the molecular mechanism by which STELLA suppresses TET3-mediated demethylation in the maternal pronucleus.

3.3. TET family and pluripotency

Pluripotency refers to the ability of cells to differentiate into all cell types of a mature organism. DNA methylation is closely associated with pluripotency because it can regulate activation or repression of genes that determine cell fate (79). Although DNA methylation on the gene promoter mostly represses transcription (23), the effects can vary depending on the regions. For instance, a positive correlation between DNA methylation and gene transcription activity has been reported (80, 81). DNA methylation status is determined and balanced by the interconnected activities of methylation writers and erasers: establishment of *de novo* methylation marks and maintenance of methylation patterns are accomplished by DNMT3A/B

and DNMT1 respectively, and TET proteins remove methylation marks by oxidizing 5mC to 5hmC. Among the TET family members, TET1 and TET2 proteins are detected in the ICM of blastocysts and ESCs, implicating that the two TET proteins could be related to the establishment or regulation of pluripotency (55). Indeed, skewed commitment to trophectoderm lineage in TET1 knockdown embryos demonstrates that TET1 is important for ICM specification during preimplantation embryo development (55). TET1 knockdown also impaired the self-renewal of mouse ESCs via downregulation of *NANOG* expression, a key pluripotency marker (55). Along with the reduced *NANOG* expression, an increased DNA methylation level in the *NANOG* promoter supported the role of TET1 as a regulator of DNA methylation (55). In contrast, subsequent studies found that TET1-depleted mouse ESCs are morphologically indistinguishable from control ESCs and possessed self-renewal capacity without changes in *NANOG* expression (58, 82-85). Expression of other key pluripotency genes, including *OCT4* (*POU5F1*) and *SOX2*, was also unaffected by the loss of TET1 in mouse ESCs (82, 85). The contrasting results of TET1-depleted mouse ESCs could be caused by different knockdown efficiency or off-target RNA interference. Although the loss of TET1 showed little impact on the ESCs maintenance, differentiation of TET1-depleted ESCs skewed into certain lineages, such as trophectoderm or endoderm, *in vitro* (58, 82), suggesting that TET1 is critical in lineage differentiation of ESCs.

Although TET1 is responsible for the formation of 5hmC, the loss of TET1 results in partial reduction of 5hmC (~35%) and a very subtle increase of 5mC (~1%) in mouse ESCs (58, 82). This result implies that other TET proteins have an overlapping function with TET1 or compensate for the loss of TET1 in ESCs. Similar to TET1-depletion, loss of TET2 does not affect the maintenance of pluripotency and has little impact on the global methylation level,

including 5hmC in mouse ESCs (62, 82). Unlike TET1 or TET2 single mutant ESCs, TET1/2 double KO ESCs were depleted of 5hmC and exhibited a greater increase in the global 5mC level, indicating that both TET1 and TET2 are responsible for regulating the methylation level in mouse ESCs. Despite the dramatic change in the global methylation level, the TET1/2 double KO ESCs were pluripotent and had skewed differentiation capacity, similar to TET1 single KO ESCs (62). Interestingly, TET3 expression was increased and a significant level of 5hmC was detected in the TET1/2 double KO ESCs and in various tissues of the postnatal mouse, raising the possibility that TET3 compensates for the loss of TET1/2 (62). The reduction of 5hmC was enhanced in TET1/2/3 triple KO ESCs compared to that in TET1/2 double KO ESCs; there was no detectable 5hmC in TET1/2/3 triple KO ESCs while there was ~ 5% 5hmC in TET1/2 double KO ESCs (86, 87). The triple KO ESCs also possessed a higher global 5mC level (about 6 %) than the double KO ESCs (about 4.5%) (87). The triple KO ESCs were largely normal under the stem cell culture conditions, but their differentiation potential was restricted to endoderm and other specified lineages were absent in the teratoma (87). Collectively, according to the studies of mouse ESCs, TET proteins are likely to be involved in regulating lineage differentiation rather than the maintenance of pluripotency, and have overlapping functions.

The pluripotent state of ESCs is controlled by a network of key transcription factors, such as OCT4, NANOG, and SOX2. Although the pluripotency of mouse ESCs can be maintained in the absence of TET proteins, the expression of key pluripotency factors is highly associated with TET proteins. For instance, OCT4 can promote TET2 expression by directly binding to its proximal promoter region (82, 88). NANOG and TET1 physically interact and co-occupy genome loci of genes related to maintenance of pluripotency (60). The transcription of TET1 is also dependent upon OCT4 and SOX2 (82, 89). Furthermore, the close relationship between TET

proteins and pluripotency factors was also identified in the establishment of induced pluripotent stem cells (iPSCs). In general, iPSCs are derived via reprogramming of differentiated somatic cells by ectopic expression of the key pluripotency factors including OCT4, SOX2, KLF4, and c-MYC (OKSM) (90-92). At the early stage of somatic cell reprogramming, TET2 is recruited to the endogenous NANOG locus and establishes epigenetic marks that direct the subsequent transcriptional induction of pluripotency genes (93). In addition, TET1 facilitates derivation of iPSCs by promoting demethylation and reactivation of endogenous OCT4 (94). Moreover, TET1 can replace OCT4 and initiate the reprogramming of somatic cells when combined with the other reprogramming factors, SOX2, KLF2, and c-MYC (94). Although there was no obvious effect of TET protein deficiency on pluripotency maintenance in mouse ESCs, these data demonstrate the involvement of TET proteins in the pluripotency circuits.

4. Epigenetic reprogramming of mammalian embryos and pluripotent cells

Dysregulation of epigenetic reprogramming during germ cell and embryo development has been implicated in human disease and infertility. Intensive investigations using mouse models have allowed unprecedented mapping of epigenetic modifications in germ cells and embryos. However, the epigenetic dynamics in human germ cells and embryos has not been fully defined due to the limitations of obtaining a large number of samples. Recent advances in low input sequencing technologies improved the understanding of epigenetic dynamics including DNA methylation in human germ cells and embryos (95-97) and eventually, are expected to aid in the development of strategies for resolving issues regarding human disease and infertility.

To date, human studies have shown that DNA methylation patterns and reprogramming events are relatively conserved between mouse and human. Human PGCs undergo global

demethylation and mature gametes acquire germ cell specific methylation patterns (98-100). The genome of human mature sperm is hypermethylated (~75%), similar to mouse sperm (~80%) (97, 101). However, fully grown human oocytes have a higher average methylation level (~54%) than mouse oocytes (~40%) (28, 102). After fertilization, global reprogramming occurs and the methylation level reaches the lowest level at the blastocyst stage (95, 97, 102). A distinct feature of the demethylation process is a dramatic decrease of methylation that occurs between fertilization and 2-cell stage in human and at the zygote stage in mouse (97). The difference could be attributed to the discrepancy in the timing of zygotic genome activation (ZGA) between human and mouse embryos because human embryos have a prolonged period before the onset of ZGA; while mouse embryos undergo major ZGA at 2-cell stage, it occurs at 8-cell stage in human embryos (103, 104).

Although the functions of TET proteins and 5hmC marks in developing human embryos have not been fully investigated due to the scarcity of samples, studies using ESCs proposed that TET proteins may be crucial players in maintaining pluripotency networks in human cells. In human ESCs, 5hmC marks can be found in enhancers and gene-bodies and enriched in transcription factor binding sites, suggesting a potential role of TET proteins in gene regulation (105, 106). Indeed, all three TET genes were expressed in human ESCs and were responsible for the appearance of 5hmC (107). TET2 protein could bind to the NANOG promoter and the methylation level was correlated with decreased expression of NANOG when TET2 was knocked down in human ESCs (108). Furthermore, a recent study showed that TET proteins are crucial for maintaining bivalent promoters in a hypomethylated state in human ESCs by preventing *de novo* methylation (107). Similar to mouse ESCs, human ESCs can be maintained under stem cell culture conditions, regardless of TET proteins. TET2-deficient human ESCs

presented normal morphology and self-renewal capacity; however, their differentiation skewed to the neuroectoderm lineage both *in vitro* and *in vivo* (108). Human ESCs lacking all three TET genes were not distinguishable from wild-type cells in morphology, self-renewal capacity, or pluripotency marker expression (107). Nonetheless, the TET1/2/3 triple KO human ESCs were not able to form a teratoma and unable to fully induce key differentiation genes during spontaneous *in vitro* differentiation (107). These data suggest that the TET proteins are also crucial in regulating lineage differentiation in human ESCs, supporting a conserved action of TET family among species.

5. Epigenetic reprogramming in domestic animal species

Domestic animals, such as cow, pig, and sheep, are valuable resources in agriculture, as well as in biomedicine as a model of human disease. For agricultural or biomedical purposes, genetically engineered animals have been generated by the use of somatic cell nuclear transfer (SCNT). Despite the generation of cloned animals in multiple species (109-111), the overall efficiency of obtaining live cloned animals remains extremely low (2-5%) due to high losses during embryonic and fetal development (112, 113). Incomplete reprogramming of the somatic genome has been suggested as a major cause of the developmental failure of cloned animals. Still, it is challenging to understand the molecular and epigenetic events occurring in cloned embryos because of the lack of knowledge in the dynamics of epigenetic reprogramming process during embryo development in domestic species. In addition, the limited knowledge of reprogramming is associated with the difficulty in deriving stable iPSCs in domestic species. Although pluripotent cells are great sources for genetic engineering and the study of developmental biology, establishment of stable iPSC lines in domestic species have been

plagued by incomplete reprogramming of the somatic cell genome. Therefore, it is essential to understand the epigenetic events, including DNA methylation dynamics, during embryo development to overcome the difficulties in the production of SCNT animals and derivation of iPSCs in domestic animal species.

5.1. Bovine

The global DNA methylation dynamics of bovine embryos have been investigated by immunostaining of 5mC in studies (114-116). These studies found that active demethylation of the paternal genome and passive demethylation of the maternal genome are also observed in bovine zygotes. The global methylation level is reduced from 2-cell to 8-cell stage through DNA replication; embryos between 8-cell and 16-cell stage acquire *de novo* methylation (115). In addition to the immunostaining results, recent studies provided more detailed and sequence-specific methylation profiles of bovine embryos using genome-scale bisulfite sequencing tools (117, 118). The average global CpG methylation levels of sperm and oocytes are about 70% and 30% respectively, and subsequently decrease to 25% in 2-cell stage embryos (117). As development progresses, further demethylation occurs and the level reaches the lowest point (15%) at the 8-cell stage (117), which coincides with the onset of zygotic genome activation in cattle (119). At the 16-cell stage, the methylation level increases to 32% (117, 118), which is consistent with elevated expression of DNMT3A/B at that stage (120), suggesting that *de novo* methylation initiates at 16-cell stage in bovine embryos. TET3 expression is detected in bovine oocytes and early stage embryos and TET1 is highly expressed in blastocysts (121, 122). Intriguingly, TET2 follows a similar expression pattern to TET3, which is distinct from mouse and human embryos where TET2 is highly expressed in the morula and blastocyst stage (121,

122). TET3 knockdown in oocytes resulted in an elevated 5mC level and a reduced 5hmC level in early cleaving embryos, indicating TET3 is responsible for active demethylation in bovine embryos (121). Similar to mouse, STELLA is also likely to be involved in the protection of the maternal genome from TET3-mediated active demethylation in bovine zygotes because STELLA knockdown leads to an increased 5hmC level in the maternal pronucleus (122). The relationship between TET proteins and pluripotency in bovine embryos has been poorly investigated. A recent study reported that bovine embryos treated with an inhibitor of TET proteins exhibited increased DNA methylation level in the *NANOG* promoter and reduced expression of *NANOG* (123). This result suggests a potential role of TET proteins in regulating pluripotency in bovine embryos, although further investigation is needed.

5.2. Porcine

As observed in other mammals, the differential demethylation process has been reported, (115, 124) implying conserved active and passive demethylation processes. Specifically, in the immunostaining assay, 5mC intensity decreases and 5hmC intensity increases in the paternal pronucleus during the transition from early to late zygote stage, whereas the 5mC level is not changed and the 5hmC signal remains very low in the maternal pronucleus during the same time period (125). Consistent with the global demethylation results from the immunostaining assay, locus specific analysis showed an increase of 5hmC and a decrease of 5mC in the repetitive element and OCT4 gene during the progression from early to late zygote stage (126). These results indicate potential involvement of TET proteins in demethylation of the paternal pronucleus in porcine zygotes. Indeed, TET3 is highly expressed in oocytes and zygotes and mainly localized to the male pronucleus (125-127). Knockdown of maternal TET3 interferes

with the formation of 5hmC in the paternal pronucleus, demonstrating that TET3 mediates demethylation of the paternal genome by catalyzing the conversion of 5mC to 5hmC in porcine zygotes (127). The methylation level continuously decreases from 2-cell to 8-cell stage, and methylation marks are reestablished at the blastocyst stage (124). The implication of TET proteins in regulation of porcine pluripotency is unclear; however, several studies provided evidence supporting the involvement of TET proteins in pluripotency maintenance of porcine embryos and iPSCs. Post-fertilization demethylation by TET3 is likely important for proper expression of pluripotency markers in porcine blastocysts because the knockdown of maternal *TET3* resulted in decreased *NANOG* expression in blastocysts (127). In porcine iPSCs, TET1 knockdown directly affects the pluripotency maintenance; TET1-deficient iPSCs are morphologically altered and differentiated with decreased expression of pluripotency-related genes and upregulation of differentiation markers (128). This result is contrary to the mouse and human pluripotent stem cells, which maintain pluripotency in the absence of all three TET proteins. Overexpression of TET1 was complementary to the incomplete reprogramming of the somatic cell genome during iPSC derivation, leading to increased expression of pluripotency genes and enhanced differentiation capacity (128). These data suggest that TET1 is an important player in the regulation of pluripotency in porcine pluripotent cells. However, the molecular mechanism of TET1-mediated reprogramming and pluripotency maintenance is still ambiguous. For example, it remains unknown whether 5hmC formation by TET1 is necessary for reducing the global DNA methylation level or related to maintaining pluripotency.

5.3. Ovine

Contrary to other mammalian species, ovine embryos undergo a very limited demethylation process rather than global demethylation during preimplantation development. Several studies reported that active demethylation does not occur in the paternal pronucleus of ovine zygotes (114, 116). However, the absence of active demethylation is controversial because the paternal pronucleus is less methylated than the maternal pronucleus in other studies, although methylation level of paternal pronucleus is higher compared to that in mouse zygotes (129, 130). Furthermore, recent findings of the restricted accumulation of 5hmC on the paternal pronucleus and expression of TET3 in oocytes and zygotes raise the possibility of active demethylation of the paternal genome (130, 131). The global methylation level is partially reduced between 2- and 8-cell stage, but then stabilizes through the morula stage (114, 132). The decrease in methylation level between the 2-cell and 8-cell stage could be related to the preparation for ZGA, which occurs at 8-16 cell stage in ovine embryos (133). However, in contrast to other species such as mouse and bovine, dramatic genome-wide demethylation does not occur before the onset of ZGA. At the blastocyst stage, trophectoderm cells are partially demethylated compared to morula embryos, while the ICM is still substantially methylated (114, 132). Interestingly, *de novo* methylation is not obviously detected in fertilized ovine embryos before the blastocyst stage (114, 132). Taken together, these findings indicate that ovine embryos likely do not undergo dynamic reprogramming of DNA methylation during preimplantation development compared to other mammalian embryos. The distinct epigenetic features of ovine embryos can provide meaningful information to understand the importance of DNA methylation reprogramming in developmental competence of mammalian embryos.

6. Porcine embryos as a model

Mouse models carrying gene modifications have been used to understand epigenetic reprogramming events occurring in human germ cells and embryos; however, differences in the epigenetic status and mechanisms between the human and mouse provides a challenge to interpret data generated using mouse models. For example, DNMTs, which are critical for mouse spermatogenesis, are not expressed during human spermatogenesis. Human oocytes have a higher average DNA methylation level than mouse oocytes (28, 102). After fertilization, the maternal pronucleus of a human zygote undergoes less demethylation than the maternal pronucleus of a mouse zygote (95, 97, 102). These differences indicate that mouse models may not be representative of the reprogramming events of human embryos. Despite the differences between mouse and human, mouse models are still utilized as a key model because direct examination of human embryos using variable genetic tools, such as gene editing, is restricted due to the scarcity of human embryos and ethical issues. Therefore, use of other model organisms can contribute to the elucidation of mechanisms of epigenetic reprogramming in early human embryos. Porcine embryos are suitable to model embryogenesis in humans because porcine embryos share key developmental events with humans. First, both human and porcine embryos present similar demethylation patterns after fertilization. In human and porcine embryos, post-fertilization demethylation progresses until the 4- to 8-cell stage (95, 134), whereas the major demethylation wave is completed before 2-cell stage in mouse (135). Second, the timing of zygotic genome activation is similar between human (103) and porcine (136). The zygotic genome activation occurs at 4-cell stage and 8-cell stage in porcine and human embryos, respectively. Third, unlike mouse, *OCT4*, the key pluripotent gene, is initially expressed in both the ICM and trophectoderm in porcine (137) and human (138) blastocysts. Lastly, human and

porcine embryos share similar amino acid metabolism, especially during morula to blastocyst transition (139). Given these similarities between human and porcine embryos, it is beneficial to use porcine embryos, instead of mouse embryos, for a better understanding of reprogramming events and underlying mechanisms in human embryos.

In addition to the beneficial use as a model for human embryo studies, basic research of porcine embryos in the areas of molecular and developmental biology is necessary to achieve biomedical and agricultural advances. Domestic pigs are an important meat source commonly consumed worldwide. Traditionally, the desired characteristics of pigs for agricultural purposes have been acquired through selective breeding, which is time consuming. However, the efficiency of the process could be revolutionarily improved by genetic engineering, including gene knockout and knockin. Economically valuable traits, such as increased meat production (140-142) and enhanced disease resistance, (143-145) could be efficiently obtained through genetic engineering in pigs. In biomedical research, pigs serve as an important model for human disease because of high similarities with humans in body/organ size, physiology, metabolism, and lifespan, compared to small rodents. For example, the pig genome can be modified to carry the same gene mutation found in human disease (146), or to provide organs with reduced immune rejection for xenotransplantation (147). In order to exploit the full potential of pigs for agricultural and biomedical purposes, understanding the molecular events occurring during embryo development or reprogramming processes, such as SCNT and iPSC derivation, is essential. Although porcine embryos appear to share similarities with the embryos of other species, they possess unique features with respect to significant aspects of developmental biology. The difficulty in establishing well-characterized pluripotent stem cell lines in pigs is one example that illustrates the different features of porcine embryos. Furthermore, molecular events

during embryo development, including cell signaling pathways and epigenetic changes, are less defined in porcine embryos than in other species. Research on molecular mechanisms controlling initial development of porcine embryos is necessary to facilitate agricultural and biomedical advances.

7. Conclusion

This literature reviewed here suggests that there is a lack of understanding of the epigenetic reprogramming processes occurring during mammalian preimplantation embryo development. Expanding our knowledge of the reprogramming process is critical to address the current problem of incomplete reprogramming of the somatic cell genome present in cloned embryos and iPSCs, and provide strategies to cure related human diseases. Topics in this dissertation will focus on elucidating the role of TET proteins in reprogramming of the embryonic genome in mammals. Although TET proteins are essential in genome-wide demethylation of fertilized embryos, the mechanism of TET-mediated demethylation remains unclear. Furthermore, the functions of TET proteins as regulators of pluripotency is controversial, as shown in the inconsistent results from the previous reports. In these studies, the mechanistic role of TET proteins in regulating the DNA methylation level and pluripotency in mammals will be explored using pig embryos. This work is expected to contribute to the understanding of how the reprogramming process resets the embryonic epigenome and affects the acquisition of pluripotency.

Chapter 2. TET family regulates the embryonic pluripotency of porcine preimplantation embryos by maintaining the DNA methylation level of *NANOG*

Abstract

The TET family (TET1/2/3) initiates conversion of 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC), thereby orchestrating the DNA demethylation process and changes in epigenetic marks during early embryogenesis. In this study, CRISPR/Cas9 technology and a TET specific inhibitor were applied to elucidate the role of TET family in regulating pluripotency in preimplantation embryos using porcine embryos as a model. Disruption of *TET1* unexpectedly resulted in the upregulation of *NANOG* and *ESRRB* transcripts, although there was no change to the level of DNA methylation in the promoter of *NANOG*. Surprisingly, a three-fold increase in the transcript level of *TET3* was observed in blastocysts carrying modified TET1, which may explain the upregulation of *NANOG* and *ESRRB*. When the activity of TET enzymes was inhibited by dimethyloxalylglycine (DMOG) treatment, a dioxygenase inhibitor, to investigate the role of TET1 while eliminating the potential compensatory activation of TET3, reduced level of pluripotency genes including *NANOG* and *ESRRB*, and increased level of DNA methylation in the *NANOG* promoter were detected. Blastocysts treated with DMOG also presented a lower ICM/TE ratio, implying the involvement of TET family in lineage specification in blastocysts. Our results indicate that the TET family modulates proper expression of *NANOG*, a key pluripotency marker, by controlling its DNA methylation profile in the promoter during embryogenesis. This study suggests that TET family is a critical component in pluripotency network of porcine embryos by regulating gene expression involved in pluripotency and early lineage specification.

Introduction

DNA methylation at CpG dinucleotides is an important epigenetic mark that regulates gene expression and its proper maintenance is essential for normal embryo development (17, 148, 149). Although DNA methylation patterns are stably preserved without dynamic changes in somatic cells, genome-wide reprogramming of DNA methylation occurs after fertilization and during germ cell development (150, 151). Asymmetric global DNA demethylation in paternal and maternal genomes of fertilized zygotes has been reported in multiple species (115, 116, 152). Specifically, the paternal genome undergoes an active demethylation process in zygotes, while DNA methylation level is gradually decreased following embryo cleavage in the maternal genome through DNA replication dependent manner in the absence of maintenance methyltransferase DNMT1 (33, 38). More recent studies challenges the theory and suggest that the presence of active and passive demethylation processes in both maternal and paternal genomes of zygotes (153, 154), indicating that the mechanism of the global reprogramming of DNA methylation marks upon fertilization is a complicated process. De novo methylation near or at the blastocyst stage establishes hypermethylation of inner cell mass (ICM) compared to trophectoderm (135, 155), which is essential for establishing lineage-specific pluripotency of embryos.

Ten-eleven translocation (TET) proteins are 2-oxoglutarate (2OG)- and Fe(II)-dependent enzymes that catalyze successive conversion of 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC) (43, 50, 55). These 5mC oxidation products are implicated as intermediates of the DNA demethylation process since the oxidized 5mC derivatives are ultimately transformed to unmodified cytosine

through thymine DNA glycosylase (TDG)-mediated base excision repair (50, 69). The TET gene family includes three members, TET1, TET2, and TET3 (43, 55), and their expression patterns are individually distinct during preimplantation development. TET3 is highly expressed in oocytes and fertilized zygotes, and known to be responsible for active DNA demethylation in zygote genomes (64, 67, 153, 154). In contrast, the expression level of *TET1* and *TET2* are very low in oocytes or zygotes, but increases following preimplantation development and both are highly expressed in blastocysts (55, 82).

To understand the physiological relevance of TET enzymes and 5hmC in development, mouse models carrying knockout (KO) alleles for *TET* genes have been generated. Mouse with an individual KO of *TET1* (58) or *TET2* (156, 157) gene develops to adults without critical developmental defects, whereas *TET3* depletion leads to neonatal lethality for unknown reasons (67). In contrast, embryos deficient in both *TET1* and *TET2* present substantial perinatal lethality, while a fraction of the double KO embryos survives and develops normally in mice (62). Triple KO (*TET1*, *TET2*, and *TET3*) mouse ES cells fail to support embryonic development at E9.5 in tetraploid complementation assay (87). These suggests that the effects of single or dual KO of *TET* genes on development can be partially offset by the involvement of redundancy in the function among the TET enzymes.

Both TET1 and 5hmC are highly detected in mouse embryonic stem (ES) cells and the ICM of blastocyst. The role of TET1 in maintaining pluripotency and development has been studied using a TET1 knockdown model (43, 55) and knockout mouse ES cells (58). Using shRNA to knockdown *TET1* gene in mouse ES cells, one study reported that TET1 is involved in maintaining *NANOG* gene expression and important for self-renewal of ES cells (55). However, in another study, TET1 KO lead to reduction of the 5hmC level, but did not affect the

pluripotency of mouse ES cells (58). Moreover, pluripotency is maintained in TET1 and TET2 double KO mouse ES cells, albeit with some defects in differentiation (62). These reports raise the possibility that activities of more than one TET family are involved in regulation of pluripotency, despite of the higher abundance of TET1 than other TET proteins in ES cells.

Although the subtle effect of TET1 depletion on pluripotency and full-term development has been demonstrated using mouse ES cells, the role of TET1 during preimplantation embryo development, especially in blastocysts where embryonic blastomeres differentiate into ICM and trophectoderm, has not been well characterized. Here, we investigated the role of TET1 in regulating gene expression involved in pluripotency and preimplantation development using porcine embryos to expand our understanding of TET1-mediated regulation of pluripotency. To define the necessity of combined activities of TET family in pluripotency of developing embryos, gene expression profiles, as well as global and locus-specific DNA methylation (5hmC and 5mC) status were examined in *TET1* KO blastocysts and blastocysts impaired in overall activities of TET enzymes. Our results identify that a proper level of *TET1* is critical for establishing proper expression of pluripotency genes in embryos by delineating the level of DNA methylation.

Materials and Methods

Chemicals

All chemicals used in the experiments were purchased from Sigma Aldrich Chemical Company (St. Louise, MU, USA) unless indicated otherwise.

In vitro embryo production

Porcine oocytes were collected from ovaries obtained from an abattoir or purchased from Desoto Biosciences LLC (Seymour, TN, USA). Collected cumulus oocyte complexes (COCs) were placed in 4-well dishes containing maturation medium, which was TCM-199 (Invitrogen) supplemented with 3.05-mM glucose, 0.91-mM sodium pyruvate, 0.57-mM cysteine, 10-ng/mL EGF, 0.5-mg/mL LH, 0.5-mg/mL FSH, 10-ng/mL gentamicin, and 0.1% polyvinyl alcohol (PVA) at 38.5 ℃, 5% CO_2 in humidified air. After 42 – 44 h of maturation, cumulus cells were removed by vortexing in the presence of 0.03% hyaluronidase. Oocytes with a polar body were collected in manipulation medium (TCM-199 supplemented with 0.6 mM NaHCO3, 2.9 mM Hepes, 30 mM NaCl, 10 ng/ml gentamicin, and 3 mg/ml bovine serum albumin (BSA)) and placed in 50 µl droplets of fertilization medium (modified Tris-buffered medium with 113. 1mM NaCl, 3mM KCl, 7.5mM $CaCl_2$, 11mM glucose, 20mM Tris, 2mM caffeine,5mM sodium pyruvate, and 2 mg/ml BSA) in a group of 25 – 30 oocytes. Fresh semen obtained from boars was diluted with semen extender (MOFA, Verona, WI, USA) and stored at 17 ℃ for one week. To wash semen, 1 ml of diluted semen was added to 9 ml of DPBS supplemented with 0.1 % BSA and washed at 720 × g for 3 min by centrifugation. After 3 washing steps, the semen pellet was resuspended with fertilization medium to 0.5×10^6/ml. Fifty microliters of diluted semen was added to the droplets with oocytes. Oocytes and sperms were co-incubated in fertilization medium at 38.5 ℃, 5% CO_2 in humidified air for 5 h. After fertilization, embryos were cultured in PZM3 medium (158) at 38.5 ℃, 5% CO_2, 5% O_2 in humidified air.

Dimethyloxallyl glycine (DMOG) treatment

To test the ability of DMOG blocking catalytic activity of TET3 enzyme, oocytes were incubated in fertilization medium in the presence of 1mM DMOG for 30 min before fertilization. Then, oocytes and sperms were co-incubated in fertilization medium with 1mM DMOG for 5 h. Subsequently, oocytes were cultured in PZM3 medium with 1mM DMOG for 16 h and collected for measurement of 5hmC level by immune-staining. For inhibition of TET family in blastocysts, oocytes and sperms were fertilized and cultured until day 4 without DMOG. At day 4, embryos were moved to PZM3 medium with 1mM DMOG and cultured until day 7. Blastocysts were collected at day 7 for further examinations.

Microinjection

Three sgRNAs targeting different regions of *TET1* gene were designed using the Zhang laboratory CRISPR design tool (http://crispr.mit.edu) (Fig. 2-1A and Table 2-1). *In vitro*-synthesized Cas9 mRNA (20ng/µl) and three sgRNAs (10ng/µl each) were injected into the cytoplasm of presumable zygotes after *in vitro* fertilization (IVF) using the FemtoJet microinjector (Eppendorf) as previously described (159, 160). Embryos were microinjected in manipulation medium on a heated stage of a Nikon inverted microscope. After the microinjection, the zygotes were washed and then cultured in PZM3 media for 7 days.

Quantification of TET enzyme activity

Nuclear proteins were extracted from control and *TET1* KO blastocysts (30 each) at day 7 using the Nuclear Extraction Kit (Abcam). The TET enzyme activity was quantified in triplicate using the TET Hydroxylase Activity Quantification Kit (Abcam) from the extracted nuclear

proteins according to the manufacturer's instruction. The activity was measured on a Tecan Infinite M200 Pro plate reader (Tecan) by measuring the fluorescence intensity at the excitation wavelength of 530 nm and the emission wavelength of 590 nm. The relative fluorescence units of control and *TET1* KO samples were subtracted by that of blank and then compared each other.

Immunocytochemistry and blastomere counting

Zona-free embryos were fixed in 4% paraformaldehyde for 15 min at room temperature. The fixed embryos were washed and permeabilized in PBS containing 0.25% TritonX-100 for 1h. Then, they were treated with 2N HCl for 30 min and neutralized in Tris–HCl pH 8.5 for 10 min. The samples were incubated in PBS containing 0.1% Tween- 20 and 2% bovine serum albumin for 1h at room temperature. After blocking, embryos were incubated in blocking solution together with 5hmC (dilution 1:100; Active Motif) or 5mC (dilution 1:100; Active Motif) or CDX2 (dilution 1:20; Biogenex) antibody overnight at 4℃. The next day, the samples were washed in blocking solution and stained with FITC-conjugated secondary antibodies (dilution 1:200; Santa Cruz Biotechnology or dilution 1:500; Thermo Fisher Scientific) for 1h at room temperature. DNA was stained with 10 µg/ml propidium iodide (PI) or 1 µg/ml DAPI. The total cell number was calculated by counting PI/DAPI positive cells in day 7 blastocysts stained with 5hmC antibody, then the average numbers were compared between control and TET-modified blastocysts. The number of ICM and TE cells were calculated by counting the numbers of CDX2 negative and positive cells in DAPI-stained blastocysts.

Quantitative RT-PCR (RT-qPCR)

Nine to ten blastocysts per group were collected at day 7 to analyze gene expression patterns of pluripotency-related and extra-embryonic lineage genes using RT-qPCR. mRNA was immediately isolated from the pooled embryos using Dynabeads mRNA Direct Kit (Thermo Fisher Scientific), followed by cDNA synthesis using Maxima H Minus First Strand cDNA Synthesis Kit (Thermo Fisher Scientific) according to the manufacturer's instructions. Amplification and detection were conducted with the ABI 7500 Real-Time PCR System (Applied Biosystems) using PerfeCTa SYBR Green SuperMix (Quantabio) under the following conditions: 95℃ for 3 min, 40 cycles of denaturation at 95℃ for 10 sec, and annealing at 60℃ for 60 sec. Primers used in RT-qPCR analysis are listed in Table 2-2. All of the threshold cycle (CT) values of the tested genes were normalized to *GAPDH* level, and relative ratios were calculated using the $2^{-\Delta\Delta}$Ct method. Four biological and three experimental replications were used. Differences in the gene expression were evaluated by student's t-test. $P<0.05$ was considered as statistically significant.

Bisulfite DNA sequencing

For analysis of DNA methylation status of repetitive elements and *NANOG* gene, 15 to 18 blastocysts at day 7 were collected and their DNA was treated with bisulfite using EZ DNA methylation kit (Zymo Research) following the manufacturer's instruction. Subsequently, the bisulfite-treated DNA was PCR amplified using specific primer sets. Details of primer information and PCR conditions are described in Table 2-3 and Table 2-4. The PCR products were purified using GeneJet gel extraction kit (Thermo Fisher Scientific) and were then ligated into pCR 2.1 TA cloning vector (Invitrogen). Twelve to sixteen colonies of each cloned samples were sequenced and evaluated.

Detection of locus-specific 5hmC

Detection of locus-specific 5hmC was done using Quest 5-hmC detection kit (Zymo Research). This kit enables detection of 5hmC in DNA sequence through glycosylation of 5hmC and treatment of restriction endonucleases; 5hmC in DNA is specifically tagged with a glucose moiety by glucosyltransferase yielding a modified base, glucosyl-5hmC, which is not digested with a glucosyl-5hmC sensitive restriction endonuclease (MspI) whereas 5C and 5mC are digested with the enzyme. To analyze 5hmC level in promoter and gene-body regions of *NANOG* gene, genomic DNA was isolated from 15 to 18 day 7 blastocysts using Purelink genomic DNA mini kit (Invitrogen), and then equal amounts of DNA from samples were divided into two groups; one group was treated with a glucosyltransferase and the other group did not receive the treatment. An internal control sample was included for reference. Glycosylation reaction was performed at 37 ℃ for 2 h, and then both groups were digested with 15 units of MspI for 4 h following manufacturer's instruction. The enzyme-digested DNA was purified with DNA clean & concentrator kit (Zymo Research) and used for qPCR. Test DNA that had not been processed was used as an internal control. *NANOG* promoter and gene-body regions were amplified from DNA using the following primers: 5'-ACAGACCAATGGAACAGAATAG-3' (forward) and 5'-CACTCATGTTGAGTTGAAGAG-3' (reverse) for promoter region and 5'-AGGACAGCCCTGATTCTTCCACAA-3' (forward) and 5'-GTTGCTCCATGATGGGTTAT-3' (reverse) for gene-body region. PCR amplification was performed with the ABI 7500 Real-Time PCR System (Applied Biosystems) using PerfeCTa SYBR Green SuperMix (Quantabio) under the following conditions: 95℃ for 1 min, 40 cycles of denaturation at 95℃ for 30 sec, annealing at 60℃ for 15 sec and extension at 72℃ for 25 sec. Percentage of hydroxymethylation

was obtained by the following formula: % = ((–control) – (+5hmC)/(–control) – (no treatment)) × 100, where "–control" is unglycosylated but MspI digested, "+5hmC" is glycosylated and MspI digested, and no treatment is unglycosylated and not digested with MspI

Statistical analysis

Differences in the frequency of blastocyst formation were determined by the chi-square test. Average total cell numbers in IVF control, injection control, and TET1 KO blastocysts were compared using one-way ANOVA, and Student t-test was used to determine the difference in the average total cell numbers in control and DMOG-treated blastocysts. To determine the difference in the ICM/TE ratio between control and TET1 KO or DMOG treatment embryos, Student t–test was used. Statistical calculations of TET enzyme activity, RT-qPCR and locus-specific 5hmC levels were performed using Student t-test. For calculations, the statistical software GraphPad Prism was used. Differences with $P < 0.05$ were considered significant.

Results

Disruption of *TET1* leads to an abnormal level of global 5hmC and 5mC in blastocysts

To disrupt TET1 function in blastocysts, *TET1*-specific CRISPR/Cas9 system containing three sgRNAs targeting different regions of *TET1* was injected into one-cell stage embryos (Fig. 2-1A). Efficiency of the targeted disruption was verified by genotyping ten blastocysts that were injected by the CRISPR/Cas9 system. All of blastocysts tested (n=10) possessed mutations on both alleles, demonstrating efficacy of the approach (Table 2-5). The total activity of TET enzymes was significantly decreased in *TET1* KO blastocysts compared to that in controls,

indicating successful disruption of *TET1* by the CRISPR/Cas9 system (Fig. 2-2A). The frequency of blastocyst formation was reduced by the *TET1* disruption; however, the average total cell number of blastocysts at day 7 was not affected (Table 2-6 ad Fig. 2-2B), suggesting that the loss of TET1 does not lead to developmental stall.

The global level of 5mC and 5hmC was monitored in *TET1* KO blastocysts as TET family is known to oxidize 5mc to initiate DNA demethylation process. The disruption of *TET1* dramatically reduced the level of global 5hmC in blastocysts (Fig. 2-1B); however, distinct differences in the 5mC level between *TET1* KO and control blastocysts were not observed in immunocytochemistry analysis (Fig. 2-1C). To investigate the global level of 5mC at the nucleotide level, DNA methylation status of repetitive elements was quantified by using bisulfite sequencing analysis. Specifically, the DNA methylation level of PRE-1 (porcine specific SINE retrotransposon) and centromeric satellites regions can capture the overview of methylation status on euchromatin and heterochromatin, respectively (161, 162). In the bisulfite sequencing analysis, DNA methylation on the PRE-1 region was increased in *TET1* KO blastocysts compared to that in control blastocysts, while centromeric satellites were hypomethylated in both *TET1* KO and control blastocysts (Fig. 2-1D). Our findings from the immunocytochemistry and bisulfite sequencing suggest that TET1 is responsible for the formation of 5hmC in blastocysts and important for maintaining global DNA methylation levels, especially in the euchromatic region of the genome.

<u>Disruption of *TET1* leads to abnormal expression of pluripotency genes without converting lineage specification</u>

Altered global 5hmC level and DNA methylation pattern by the *TET1* KO implied that the KO embryos potentially carried abnormal gene expression patterns. To verify, transcript abundance of pluripotency-related genes and extra-embryonic lineage markers were quantified in *TET1* KO blastocysts by RT-qPCR. The lack of functional TET1 did not alter the transcript level of pluripotency-related genes except for *NANOG* and *ESRRB* (Fig. 2-3A). Unexpectedly, an elevated level of *NANOG* and *ESRRB* transcript was observed in *TET1* KO blastocysts. There was no change in the transcript level of selected extra-embryonic lineage genes after *TET1* KO except for *GATA4* (Fig. 2-3B); the transcript level of *GATA4* was higher in *TET1* KO blastocysts compared to that in controls. RT-qPCR was also performed to monitor the level of *TET* family (Fig. 2-3C) as previous knockout experiments in the mouse suggest functional redundancy of the TET enzymes. The expression level of *TET1* transcript was significantly downregulated in *TET1* KO blastocyst, indicating that the mutations introduced by CRISPR/Cas9 system interfered with transcription, as well as protein function via generation of premature stop codons. Interestingly, the disruption of *TET1* induced activation of *TET3* as indicated by the 3-fold increase in its transcript level compared to the control blastocysts. Expression level of *TET2* was also numerically increased (1.5-fold), although the change was not statistically significant ($p = 0.16$). The overexpression of *TET3* suggests that disruption of *TET1* activated *TET3* as compensatory action and could potentially contribute to the limited changes in gene expression profile of *TET1* KO blastocysts.

To examine the impact of *TET1* on early lineage commitment, day 7 blastocysts derived from CRISPR/Cas9 injected and control group were stained with CDX2, a trophectoderm marker (Fig. 2-3D). The ICM and trophectoderm (TE) lineages were distinguished by counting CDX2 negative and positive cells, respectively. In both control and *TET1* KO blastocysts, a cluster of

CDX2 negative cells was identified, indicating ICM lineage formation was not affected by the disruption of *TET1*. Analysis of the ICM/TE ratio between control and *TET1* KO embryos suggested that early lineage commitment in blastocysts was not affected by the lack of *TET1* (Fig. 2-3E).

Inhibition of catalytic activity of TET family and its impact on DNA methylation

To investigate whether the findings from the *TET1* KO blastocysts was influenced by the compensatory increase in other *TET* genes, overall inhibition of TET family was performed. Specifically, catalytic activity of TET enzymes was blocked by treating embryos with dimethyloxallyl glycine (DMOG), a small-molecule inhibitor of 2OG-dependent oxygenases. To confirm the effectiveness of DMOG in blocking oxygenase activity of TET, especially TET3, fertilized zygotes were incubated with 1mM DMOG, then patterns of DNA demethylation after fertilization were monitored by uncovering the appearance of 5hmC after fertilization (Fig. 2-4A). Because the DNA demethylation process is orchestrated by TET3 (67, 153, 154), the inhibitor would alter the level of 5hmc, if effective. The DMOG treatment diminished the appearance of 5hmC in both paternal and maternal pronuclei compared to the level of 5hmC in control zygotes, demonstrating the capability of DMOG to inhibit the conversion from 5mC to 5hmC by TET (Fig. 2-4B). To inhibit the activity of TET enzymes effectively at the blastocyst stage, embryos were incubated with 1mM DMOG from morula (day 4) to blastocyst stage (day 7) (Fig. 2-5A). The treatment would keep the activity of TET3 on initiating DNA demethylation process after fertilization, but prevent activity of all TET family in blastocysts. Consistent with the *TET1* KO experiment, frequency of blastocyst formation on day 7 was decreased in DMOG-treated embryos compared to that in control embryos (Table 2-7) while total cell number in

blastocyst was not affected by the inhibition of TET activity (Fig. 2-2C). Similar to the findings in *TET1* KO blastocysts, the global level of 5hmC was dramatically reduced in blastocysts treated with DMOG compared to the control blastocysts, indicating that active TET family is necessary for the formation of 5hmC at the blastocyst stage (Fig. 2-5B). Dramatic changes of global level of 5mC were not observed by immunocytochemistry analysis in blastocysts treated with DMOG (Fig. 2-5C). The DNA methylation pattern of repetitive elements in DMOG-treated blastocysts was also similar to that in *TET1* KO blastocysts; methylation level of PRE-1 region increased when embryos were incubated with DMOG, while hypomethylation status of centromeric satellites was observed in both control and DMOG-treated blastocysts (Fig. 2-5D). These data align with conclusions from the *TET1* KO experiment that TET1 is responsible for the formation of 5hmC and maintenance of global DNA methylation level in the euchromatic region at the blastocyst stage.

Inhibition of overall TET family activity impairs expression of pluripotency and extra-embryonic lineage genes

To further investigate whether the abnormal level of *NANOG* and *ESRRB* in *TET1* KO blastocysts was influenced by the compensatory increase of other *TET* genes, gene expression patterns in DMOG-treated blastocysts were compared to that in control blastocysts. A significant number of pluripotency-related genes was downregulated in DMOG-treated blastocysts; the transcript abundance of *POU5F1*, *NANOG*, *ESRRB*, *PRDM14*, *ZFP42*, and *DPPA3* was lower in DMOG-treated blastocysts (Fig. 2-6A). However, there was no significant change in the level of *SOX2*, *KLF2*, and *TCL1A* after the DMOG treatment. The transcript abundance of trophoblast markers (*CDX2*, *TEAD4*, and *GATA3*) and primitive endoderm marker (*SOX17*, *GATA6*, and

GATA4) was downregulated in DMOG-treated blastocysts (Fig. 2-6B). The DMOG treatment also affected the level of *TET* family genes; the level of *TET1*, *2*, and *3* transcripts was downregulated by the DMOG treatment in blastocysts (Fig. 2-6C). The gene expression analysis indicates that TET enzymes are critical for proper expression of pluripotency and extraembryonic lineage markers. In addition, the dynamic changes in transcripts after the DMOG treatment but not after *TET1* KO suggests that compensatory action of TET family (TET2 or TET3) in the absence of TET1 assisted in maintaining the normal level of selected markers.

Intriguingly, capturing the number of CDX2 positive and negative cells through immunohistochemistry revealed that the DMOG treatment reduced the number of CDX2 negative cells, i.e. ICM, in blastocysts compared to that in control embryos (Fig. 2-6D). The ICM/TE ratio was also decreased in DMOG-treated blastocysts (Fig. 2-6E), demonstrating that TET family is essential for normal lineage commitment in blastocysts.

TET enzymes regulate DNA methylation level in the promoter and gene-body regions of *NANOG*

Contradicting effect between DMOG-treated and *TET1* KO blastocysts on the level of *NANOG* implied a differential level of DNA methylation on the promoter of *NANOG* from the two treatments. To investigate that possibility, the level of DNA methylation (5mC and/or 5hmC) levels in promoter and gene-body regions of *NANOG* gene were analyzed (Fig. 2-7A). The 5hmC level in *NANOG* promoter was not different between control and *TET1* KO blastocysts (Fig. 2-7B). However, the 5hmC level in gene-body regions was increased in *TET1* KO blastocysts compared to that in controls. Bisulfite sequencing analysis revealed that there

was no difference in the promoter methylation level between *TET1* KO and control blastocysts, while higher methylation level in gene-body region was observed in *TET1* KO blastocysts compared to that in controls (Fig. 2-7C). When embryos were incubated with DMOG, 5hmC level of blastocysts was not changed in the promoter or gene-body regions (Fig. 2-7D). However, the DMOG treatment increased DNA methylation level in both promoter and gene-body regions (Fig. 2-7E). Bisulfite sequencing analysis cannot distinguish between 5mC and 5hmC; therefore, it is difficult to conclude whether the increase is due to the conversion of 5mC into 5hmC. Considering no alteration in 5hmC level by the DMOG treatment, it is arguable that the increase in DNA methylation level seen from bisulfite sequencing is due to 5mC and the elevated 5mC level in promoter and gene-body regions led to the downregulation of *NANOG* expression in DMOG-treated blastocysts.

Discussion

Recent studies have expanded biological significance of the TET family on early development. Here, we demonstrated that the loss of TET activities alters the expression pattern of genes related to pluripotency and early lineage specification in porcine blastocysts; specific changes differed when activity of TET1 or overall TET family was interfered. Global reduction in the 5hmC level after *TET1* KO suggests that TET1 is the main 5-methyl-dioxygenase responsible for establishing the 5hmC marks in porcine blastocysts. Studies using mouse ES cells, indicates that while Tet1 is important for establishing 5hmC marks, Tet1 deletion alone has a relatively limited influence on the level of 5hmC whereas the lack of *Tet1* and *Tet2* leads to a complete loss of 5hmC (43, 58, 62, 82). Furthermore, Tet2 depletion resulted in a greater decrease in 5hmC level than Tet1 depletion in the mouse ES cells (82, 163). Concomitant with

the decrease in global 5hmC level, DNA methylation level including 5mC and 5hmC was increased in euchromatic SINE, PRE-1, in both *TET1* KO and DMOG-treated blastocysts, consistent with a previous report indicating that TET proteins play a protective role against de novo DNA methylation (164). A similar level of increase in DNA methylation level (about 30 %) in PRE-1 regions between TET1 KO and DMOG-treated blastocysts suggests that TET1 is likely to be the main TET enzyme responsible for maintaining global methylation level in blastocysts among the TET family members (TET1 – 3). Hypomethylation of the centromeric region is common in germ cells and preimplantation embryos (165) and disruption of TET1 or overall TET family did not change the level, indicating that TET proteins are not involved in methylation maintenance in heterochromatic regions. The result is consistent with previous studies in the mouse where 5hmC modifications are rarely detected in heterochromatic regions, whereas euchromatic regions are enriched with Tet1 binding-sites and a wide distribution of 5hmC in the mouse ES cells (84, 166, 167).

The role of TET1 in cellular pluripotency and differentiation has been extensively studied in ES cells and blastocysts because its level is highly enriched in the cells and embryos compared to other TET family genes. Impaired or skewed differentiation of ES cells caused by TET1 depletion has been demonstrated (58, 82, 84, 87); however the impact of TET1 loss on the pluripotency of ES cells has not been consistent in the previous reports. Using siRNA or shRNA approach, multiple studies reported downregulation of pluripotency-associated genes and impaired self-renewal in ES cells after targeted disruption of TET1 (55, 84, 167, 168). However, other studies argue that TET1 depletion did not alter the expression of pluripotency markers (58, 82, 85). This variance could be caused by the compensatory role of other TET family members under the absence of functional *TET1*. Similarly, we observed a compensatory and overlapping

function of TET family. The disruption of *TET1* did not lead to total repression of pluripotency genes, potentially due to the compensatory activation of TET3. Transcript levels of pluripotency genes (*SOX2*, *KLF2*, *PRDM14*, *TCL1A*) were numerically increased in *TET1* KO blastocysts. However, the differences were not statistically significant due to high variations among biological replications, which may be explained by a mouse study in which Tet1/Tet3 double KO mouse embryos presented a significant transcriptome variability among embryos (169). Blocking overall TET activity using DMOG repressed the expression of pluripotency related genes, confirming an overlapping function between TET1 and TET3. The finding is consistent with the result from Tet1/Tet3 double KO mouse blastocysts which displayed severe loss in Nanog expression and dysregulation of extra-embryonic lineage markers (169). Another possibility that cannot be fully excluded is that the DMOG is an inhibitor of 2OG-dependent oxygenases and decreased activity of other 2OG-dependent oxygenases such as Jumonji-domain-containing enzymes could also contribute to the repressed expression of pluripotency related genes.

Although the majority of TET1 binding sites are located around transcription start sites of CpG-rich promoters and gene-bodies, TET1 also binds to CpG-poor gene promoters, such as *NANOG*, *TCL1*, and *ESRRB*, important genes for maintaining pluripotency of ES cells (85, 167). In contrast to a mouse study where the lack of Tet1 reduced *Nanog* expression in ES cells (167), the expression of *NANOG* was upregulated and proximal promoter regions were hypomethylated after the removal of *TET1* in porcine blastocysts. Previous reports in the mouse suggest that Tet1 can act as a repressor to genes related to development, but not to pluripotency related genes (167, 170). Therefore, it is unlikely that the upregulation of *NANOG* in *TET1* KO porcine blastocysts was the result of lifting TET1-mediated repression. Interestingly, the inhibition of overall TET by DMOG repressed *NANOG* expression and increased the methylation level of its promoter

region, suggesting that the expression of *NANOG* is directly controlled by the level of DNA methylation in the promoter region, similar to transcriptionally active genes (171). The contrasting results between the *TET1* KO and DMOG treatment suggests that other TET enzymes, such as TET2 and TET3, are involved in the regulation of *NANOG* expression. For instance, despite its low expression level, Tet3 contributes to a small portion of 5hmC in mouse ES cells, suggesting its action in pluripotent cells (86, 87). Maternal Tet3 is required for demethylation of the promoter region of paternal *Nanog* in mouse zygotes (67), and knockdown of maternal *TET3* leads to downregulation of *NANOG* expression in porcine blastocysts (127). In addition, TET3 activated by a hypoxia microenvironment regulates *NANOG* transcription by direct binding to promoter region in brain tumor cells (172). The previous publications and findings from current study imply that *TET3* possesses a role in the transcriptional regulation of *NANOG* gene in blastocysts by controlling the level of DNA methylation on its promoter. Findings in this study lead us to a model of how TET family regulates global DNA methylation and *NANOG* expression in porcine blastocysts (Fig. 2-8). Under the presence of functional TET1, the DNA methylation level of the genome, including the promoter regions of *NANOG*, is regulated by TET1, thus maintaining the level of *NANOG* in the blastocysts. In the absence of functional TET1, a compensatory increase in TET3 potentially reprograms the promoter region of *NANOG* to ensure its expression; however, TET3 may have little effect on the genome-wide DNA methylation level, i.e. gene-specific regulation of DNA methylation rather than global level. Suppression of overall TET family led to the downregulation of *NANOG* and increased the level of DNA methylation on *NANOG* promoter and the euchromatic genome, indicating the involvement of TET family in establishing pluripotency through regulating the level of DNA methylation.

The expression pattern of *ESRRB* was also aligned with *NANOG*: upregulation after *TET1* KO and downregulation with DMOG treatment. The changes in the abundance of *ESRRB* is likely to be a secondary effect by the altered *NANOG* expression rather than the direct effect from the shift in TET family because *ESRRB* is known as a direct downstream target gene of *NANOG* in pluripotent cells (173). For example, depletion of Tet1 in mouse ES cells led to the downregulation of *Esrrb* and *Nanog*; however, the expression of *Esrrb* was restored by the overexpression of exogenous *Nanog* (167), indicating that the expression of *Esrrb* is regulated by *Nanog*.

The transcription factors *OCT4* (*POU5F1*) and *NANOG* are essential for the regulation of early embryo development and pluripotency in mammalian embryos. In porcine, similar to human blastocysts, OCT4 expression is maintained in both ICM and trophectoderm cells in blastocysts (174, 175). However, unlike mice and human, unique expression patterns of NANOG can be found in porcine blastocysts. For instance, although *NANOG* transcripts can be detected in early blastocyst stage, NANOG proteins are not detectable in early porcine blastocyst, but begin to appear after epiblast formation when embryos arrive in uterus (176). The unique NANOG expression implies distinct pluripotency network in pig blastocysts and may render difficult derivation of ES cells from porcine embryos under the conventional ES cell culture condition. Therefore, an understanding of expressional regulation of *NANOG* gene is likely to unveil pluripotency regulation in porcine embryos and contribute to the establishment of porcine ES cells.

The CRISPR/Cas9 technology was used to disrupt *TET1* in porcine embryos. Efficacy of the approach was high enough that no wild-type allele was identified in two target loci. The *TET1* transcript level was also reduced in blastocysts carrying mutated *TET1*. One potential

explanation could be that TET1 has a positive feedback to its transcription activity, thus the lack of functional TET1 reduced the amount of *TET1* transcripts. Another explanation could be that mutations on *TET1,* introduced by CRISPR/Cas9 system, could interfere with proper secondary structure formation of *TET1* mRNA, which would lead to the degradation of *TET1* transcript. The direct injection of CRISPR/Cas9 system has been successfully utilized in pigs to introduce targeted modifications at the high level (159, 160, 177) and efficiency of the current study aligns with the previous reports. Pig embryos are suitable to model embryogenesis in humans because pig embryos mirror developmental events with human embryos. For example, dynamic changes in the level of DNA methylation on paternal DNA is observed in humans (178) as well as pigs (124, 179). In addition, the timing of zygotic genome activation is similar between human (103) and pig (136). Despite the similarities, pig models are rarely used for comparative studies due to the lack of tools available to induce genetic modifications. Unlike rodent models, the production of KO embryos in pigs through breeding founder animals is an unrealistic option considering cost associated with producing KO pigs and gestation period (114 days). As demonstrated in this study, utilization of CRISPR/Cas9 system allowed us to generate KO embryos without having to produce founder KO pigs. These technical advancements may expand the use of pig models to understand developmental events in embryos as a comparative study.

In summary, we demonstrate that the TET enzymes are closely involved in maintaining proper level of pluripotency genes in preimplantation embryos. Although TET1 has a major role in genome-wide 5hmC formation at the blastocyst stage, *TET1* disruption has a minor impact on the pluripotency genes of preimplantation embryos. On the other hand, overall inhibition of TET family results in a defective expression of pluripotency genes, abnormal lineage specification, and methylation increase in *NANOG* promoter. Our data indicate *NANOG,* a key gene in

pluripotency, is delicately regulated by TET family through promoter methylation. Our findings propose that the TET1/3-mediated demethylation at promoter region is a key element for transcriptional activation of genes related to pluripotency and early lineage specification in mammalian blastocysts.

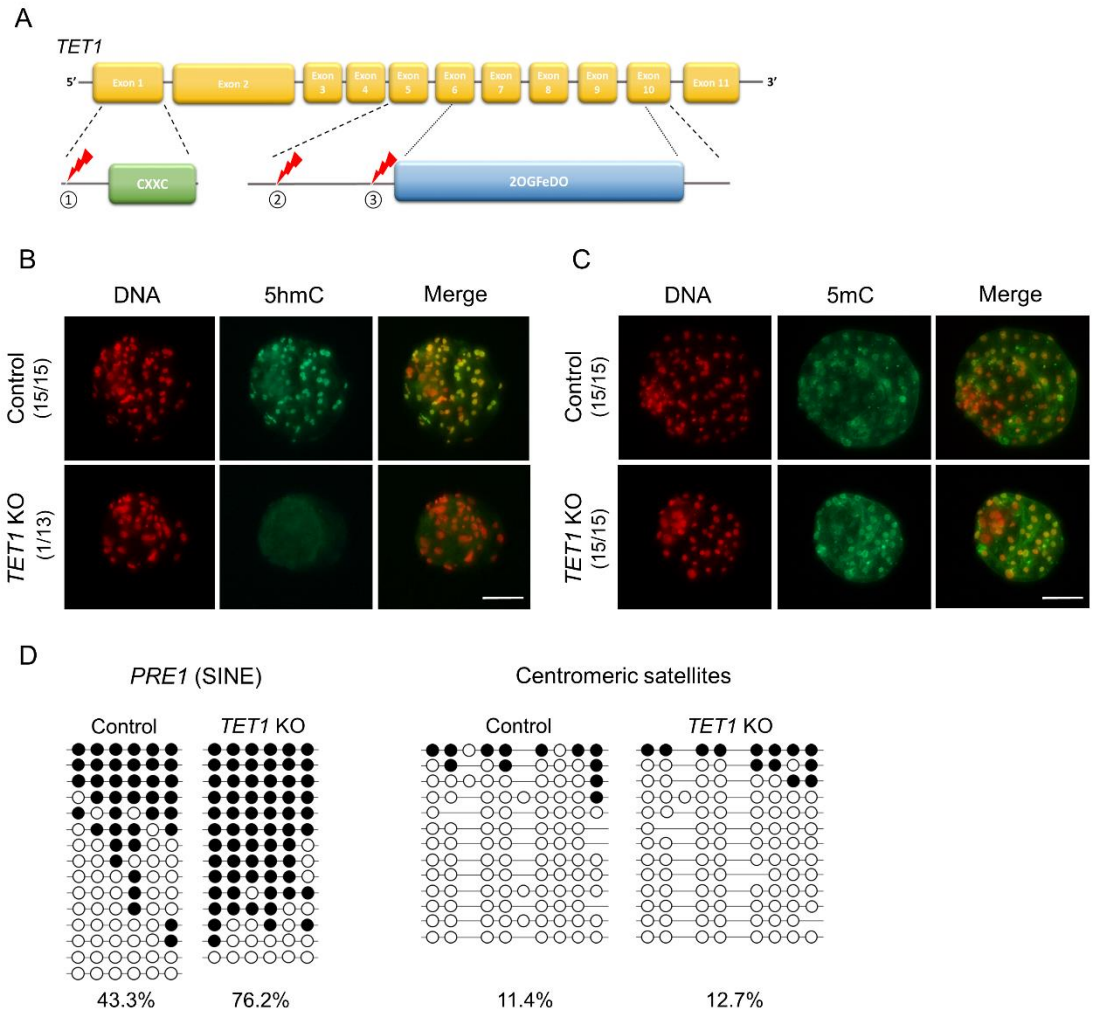


Fig. 2-1. Global levels of 5hmC and 5mC in *TET1* KO blastocysts. (A) Strategy to disrupt *TET1* gene in porcine embryos. Three sgRNAs targets different regions of *TET1* gene; one on the immediate downstream of the presumable translation start site and two on the 5' side of the 2-oxoglutarate-Fe(II)-oxygenase domain. (B) Global 5hmC level of blastocysts was dramatically decreased by *TET1* KO. (C) No differences in the global 5mC level between *TET1* KO and control blastocysts were detected by immune-staining. Numbers in parentheses indicate the number of embryos displaying a positive signal for 5hmC or 5mC out of the total number of embryos examined. (D) Bisulfite sequencing analysis of DNA methylation pattern in repetitive

elements. Methylation level of PRE-1 was increased by *TET1* KO whereas centromeric satellites were hypomethylated in both *TET1* KO (n = 17) and control blastocysts (n = 17). Scale bar indicates 100 μm. Methylated and unmethylated CpG dinucleotides are indicated by filled circle and open circle, respectively.

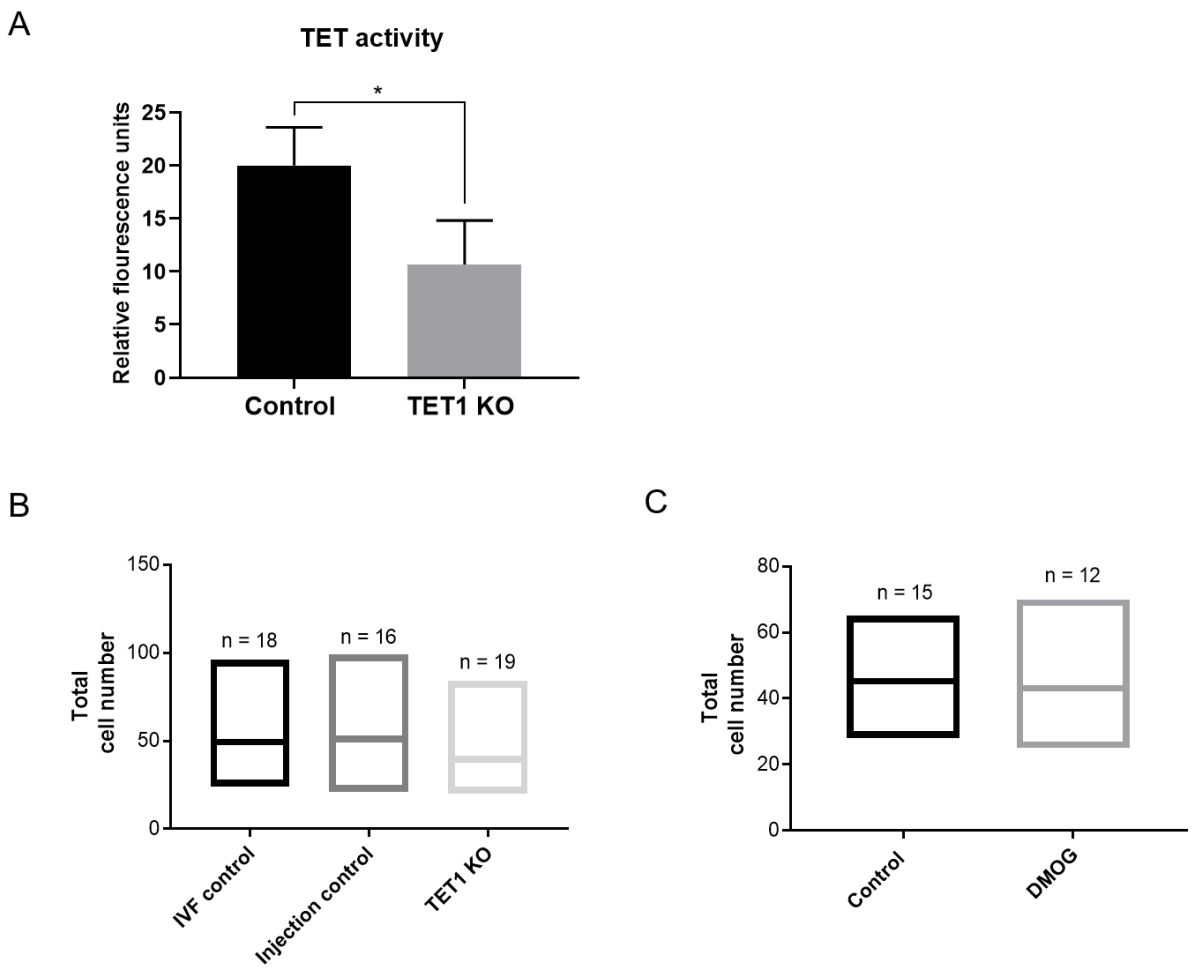


Fig. 2-2. TET activity and total cell number in TET1-disrupted and DMOG-treated blastocysts. (A) Measurement of TET enzyme activity in *TET1* KO blastocysts. Total TET enzyme activity in *TET1* KO blastocysts was decreased compared to that in control blastocysts. Error bars represent standard deviation. A p-value < 0.05 was considered statistically significant. (B) There was no significant difference in the total cell number between IVF control, injection control (Cas9 only), and *TET1* KO blastocysts. (C) Total cell number of blastocysts incubated with DMOG was not different compared to that of control blastocysts. Top horizontal line and bottom horizontal line of the box indicate maximum and minimum values, respectively. Middle horizontal line indicates average value. A p-value < 0.05 was considered statistically significant.

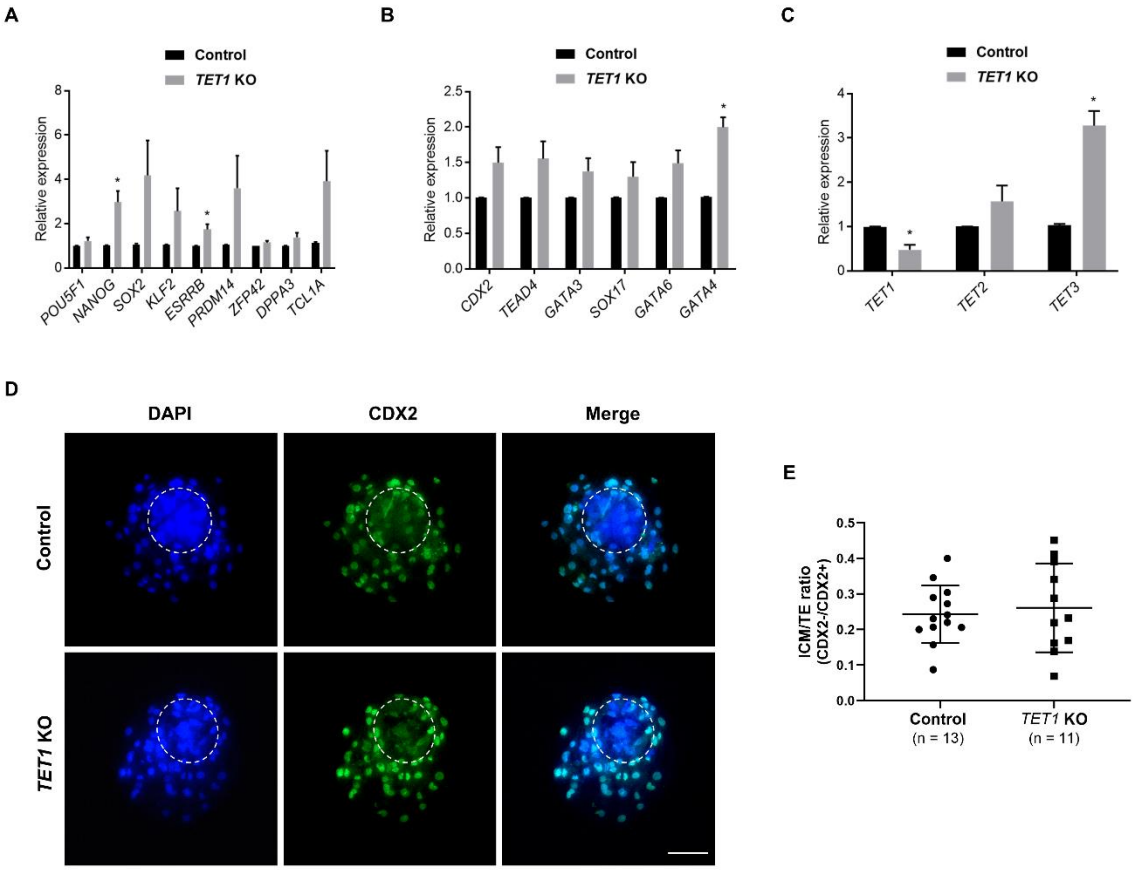


Fig. 2-3. Impact of *TET1* KO on gene expression profile and lineage specification in blastocysts. (A) Expression levels of *NANOG* and *ESRRB* were increased by *TET1* KO, however levels of other pluripotency genes were not significantly changed. (B) Transcript levels of extra-embryonic lineage genes were not changed by *TET1* KO, except *GATA4*. (C) The level of *TET1* transcript was decreased, however, *TET3* levels were increased approximately threefold. Relative expression levels were normalized to the *GAPDH* level. Error bars represent S.E.M. A p-value < 0.05 was considered statistically significant. (D) CDX2 staining revealed that *TET1* KO did not alter lineage commitment in blastocysts. The dotted white circle indicates ICM lineage. Scale bar indicates 100 µm. (E) ICM/TE cell number ratio was calculated by counting CDX2 negative and CDX2 positive cells. The ICM/trophectoderm ratio was not different between control and *TET1*

KO blastocysts. Error bars indicate SD. A p-value < 0.05 was considered statistically significant. TE: trophectoderm.

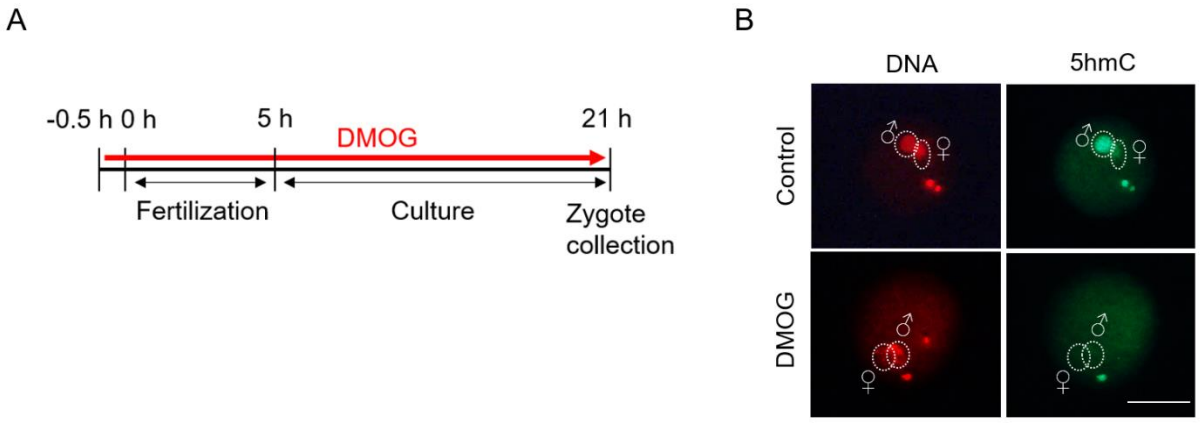


Fig. 2-4. The effect of DMOG treatment on TET3 inhibition in zygotes. (A) Timeline of DMOG treatment to inhibit TET3 activity in zygotes. Oocytes were incubated with 1mM DMOG for 30 min before fertilization and the treatment was maintained during fertilization and embryo culture for further 21 h, then zygotes were collected for immunocytochemistry analysis. (B) Inhibition of TET3 activity by DMOG treatment. In control zygotes, formation of 5hmC was detected in both the paternal and maternal pronuclei. 5hmC disappeared in pronuclei when zygotes were incubated with DMOG.

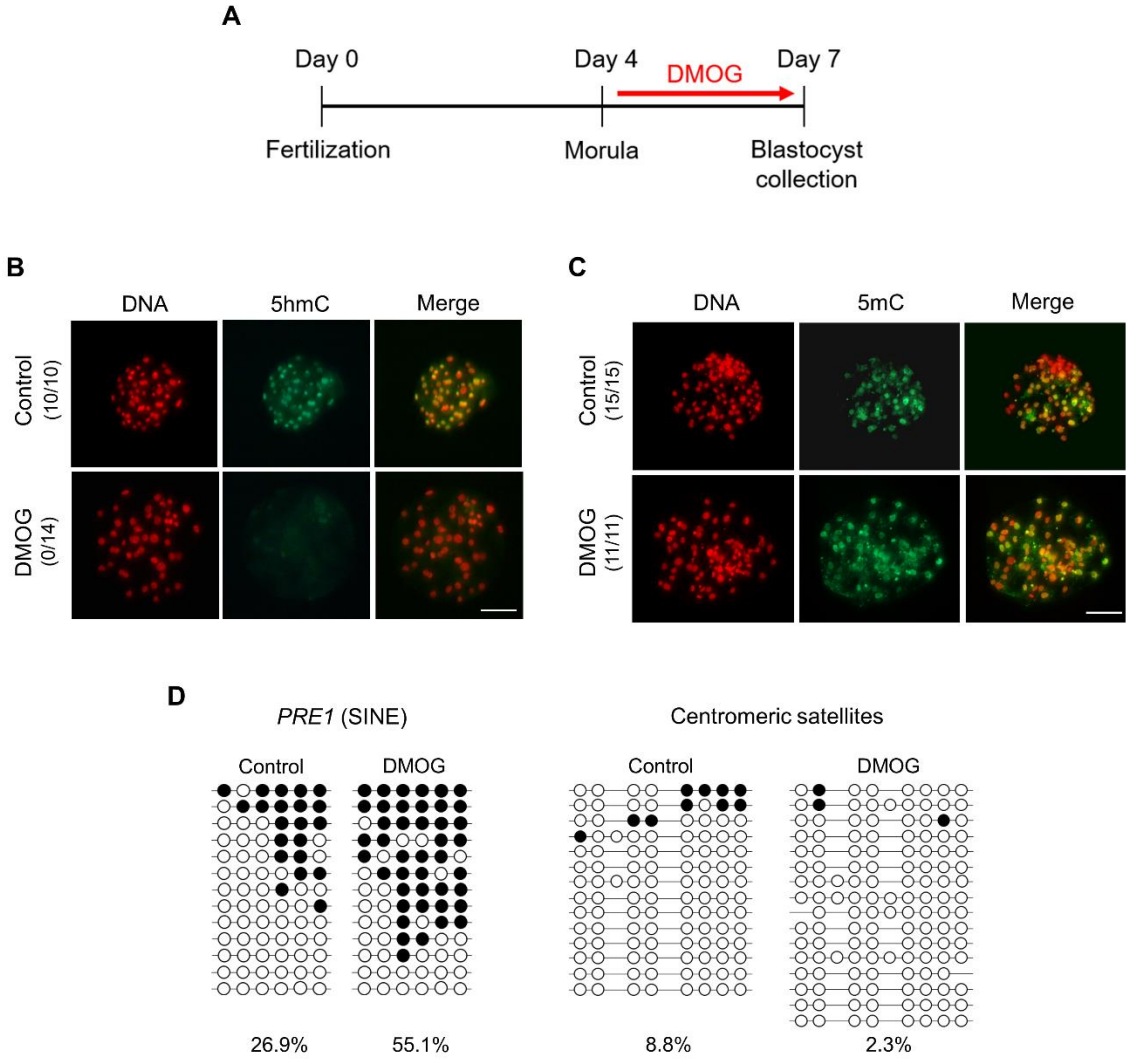


Fig. 2-5. Global levels of 5hmC and 5mC in DMOG-treated blastocysts. (A) Timeline of DMOG treatment to inhibit TET activities in blastocysts. To block TET activities at blastocyst stage, embryos were incubated with 1mM DMOG from morula stage (day 4) to blastocyst stage (day 7), then blastocysts were collected for further analysis. (B) Global 5hmC level of blastocysts was dramatically decreased by DMOG treatment. (C) No differences in the global 5mC level between DMOG-treated and control blastocysts were detected by immunocytochemistry. Numbers in parentheses indicate the number of embryos displaying a positive signal for 5hmC or 5mC out of

the total number of embryos examined. (D) Bisulfite sequencing analysis of DNA methylation pattern in repetitive elements. Methylation level of PRE-1 increased after DMOG treatment whereas centromeric satellites were hypomethylated in both DMOG-treated (n = 15) and control blastocysts (n =16). Scale bar indicates 100 µm. Methylated and unmethylated CpG dinucleotides are indicated by filled circle and open circle, respectively.

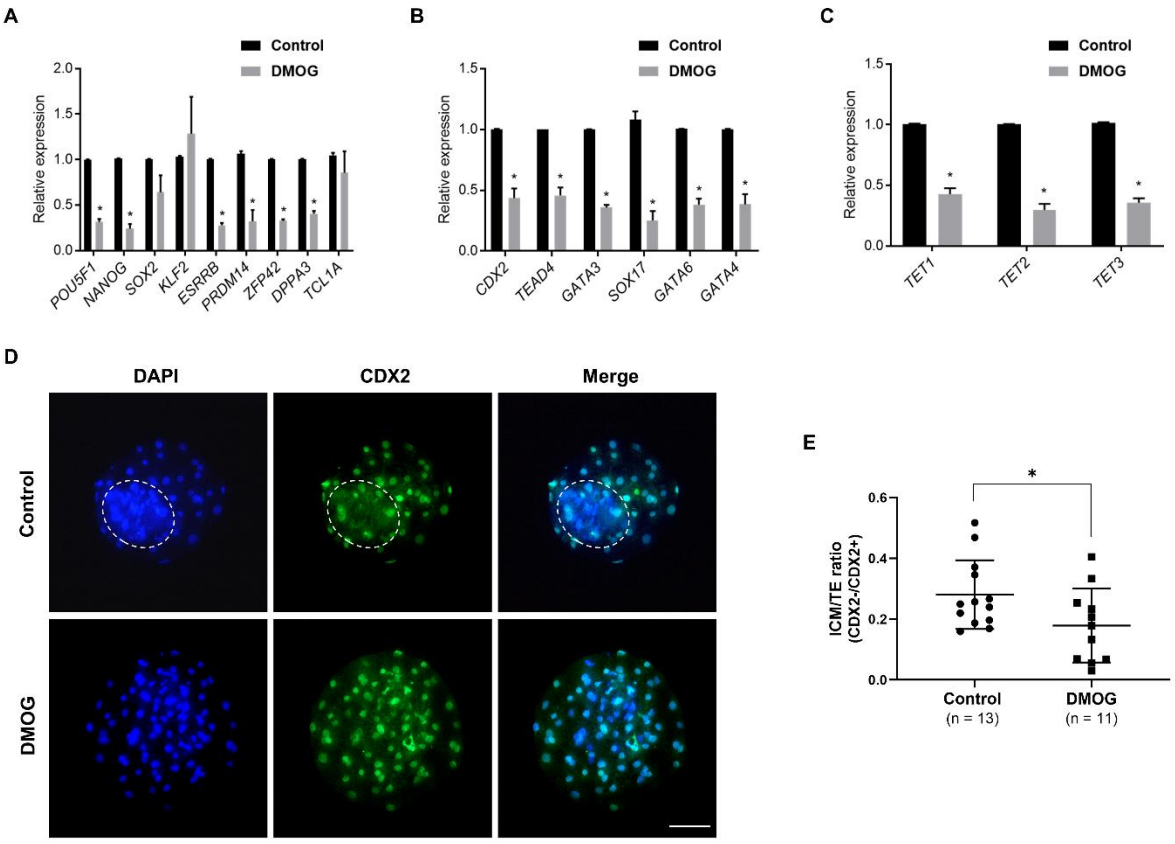


Fig. 2-6. Relative mRNA levels of pluripotency genes, extra-embryonic lineage genes, and *TET* family genes in DMOG-treated blastocysts. (A) Expression levels of pluripotency genes were downregulated in DMOG-treated blastocysts, except for *SOX2*, *KLF2*, and *TCL1A*. (B) Transcript levels of extra-embryonic lineage genes were reduced in DMOG-treated blastocysts. (C) Expression level of *TET* family genes was reduced in DMOG-treated blastocysts. Relative expression levels were normalized to the *GAPDH* level. Error bars represent S.E.M. A p-value < 0.05 was considered statistically significant. (D) The number of CDX2 negative cells (ICM) was reduced in DMOG-treated blastocysts compared to that in control blastocysts. The dotted white circle indicates ICM lineage. Scale bar indicates 100 μm. (E) DMOG treatment lowered ICM/TE ratio. Error bars indicate SD. A p-value < 0.05 was considered statistically significant. TE: trophectoderm.

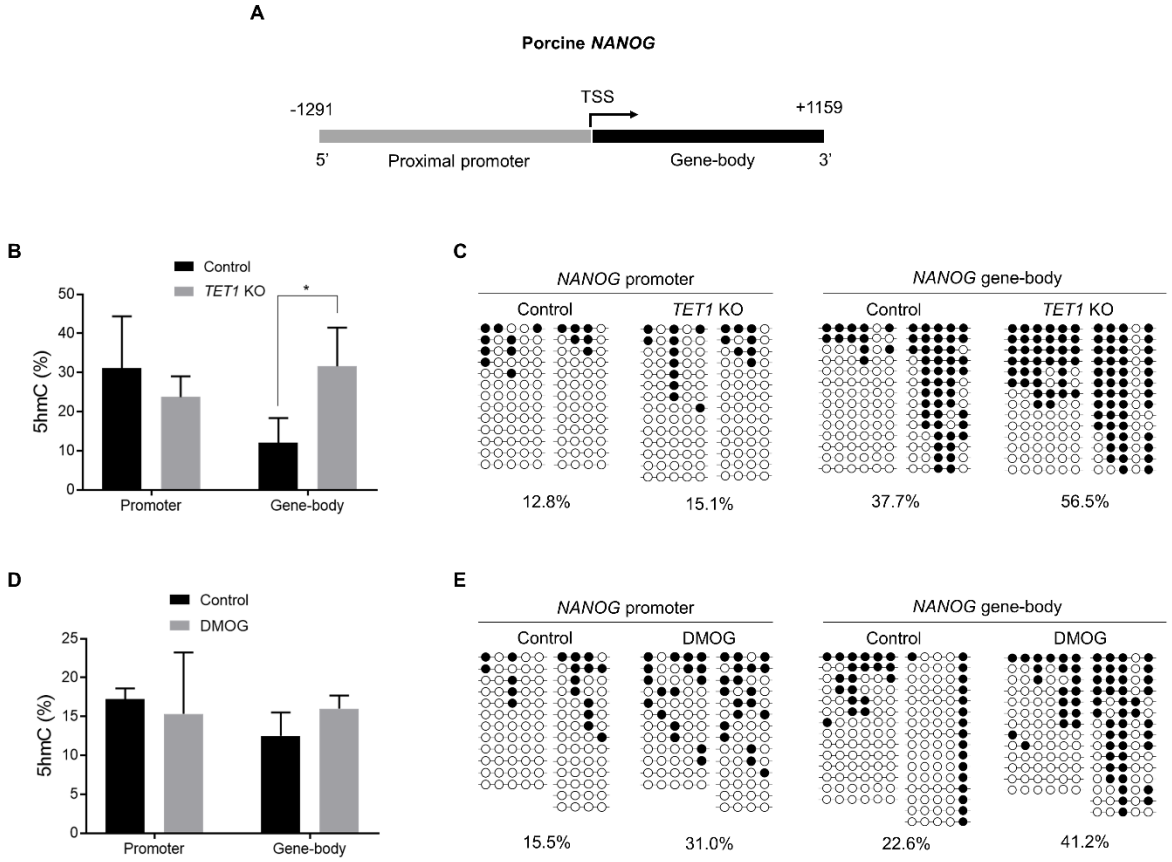


Fig. 2-7. DNA methylation (5hmC and 5mC) status of *NANOG* gene in *TET1* KO and DMOG-treated blastocysts. (A) Analysis of methylation status of porcine *NANOG* gene. The proximal promoter region (1291bp upstream the TSS) and gene-body region (1159bp downstream the TSS) of *NANOG* gene were separately examined. (B) 5hmC level in promoter and gene-body regions of *NANOG* gene in *TET1* KO blastocysts. (C) Bisulfite sequencing analysis of promoter and gene-body regions of *NANOG* gene in *TET1* KO (n = 15) and control (n = 15) blastocysts. (D) 5hmC level in promoter and gene-body regions of *NANOG* gene in DMOG-treated blastocysts. (E) Bisulfite sequencing analysis of promoter and gene-body regions of *NANOG* gene in DMOG-treated (n = 15) and control (n = 15) blastocysts. 5hmC data were analyzed with student's t-test for three replicates and a p-value < 0.05 was considered statistically significant.

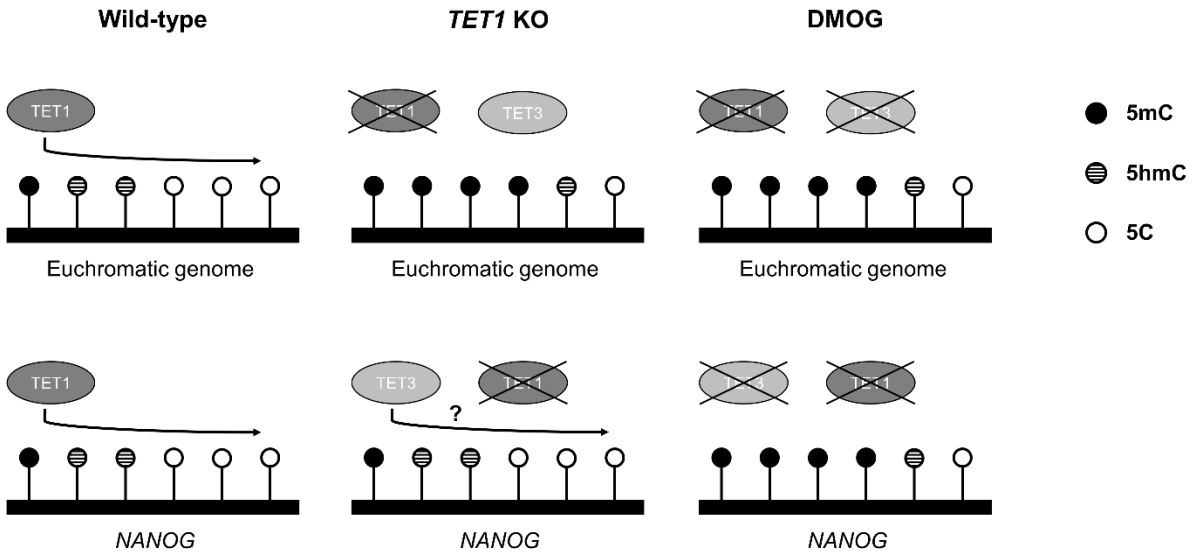


Fig. 2-8. Graphical summary and proposed mechanism of TET actions on DNA methylation in porcine blastocysts under the *TET1* KO or TET inhibition.

Table 2-1. Information of sgRNAs used for depletion of *TET1* gene.

sgRNAs	Sequence (5' → 3')
*TET1*_sgRNA1	TGTCTCGATCTCGCCATGCA
*TET1*_sgRNA2	GTGCTCATCATGGTATGGGA
*TET1*_sgRNA3	AGTCGAACCTGTACATGTCA

Table 2-2. Primers used for RT-qPCR analysis.

Primers	Sequence (5' → 3')
*GAPDH*_forward	ATGACATCAAGAAGGTGGTGAAGC
*GAPDH*_reverse	CCAGCATCAAAAGTGGAAGAGTGA
*OCT4*_forward	TTTGGGAAGGTGTTCAGCCAAACG
*OCT4*_reverse	TCGGTTCTCGATACTTGTCCGCTT
*NANOG*_forward	AGGACAGCCCTGATTCTTCCACAA
*NANOG*_reverse	AAAGTTCTTGCATCTGCTGGAGGC
*SOX2*_forward	TGTCGGAGACGGAGAAGCG
*SOX2*_reverse	CGGGGCCGGTATTTATAATCC
*KLF2*_forward	CGATCCTCCTTGACGAATTT
*KLF2*_reverse	CAAGCCTCGATCCTCTAGT
*ESRRB*_forward	CTGCAAGGCCTTCTTCAA
*ESRRB*_reverse	CGTTTGGTGATCTCACACTC
*ZFP42*_forward	GGATTCCTTCTCTGACTGTTAC
*ZFP42*_reverse	GCTCTTGTTCTGATCCTTCTT
*DPPA3*_forward	CTCAGGCTTGTCCTCAAATG
*DPPA3*_reverse	CGTCAAGTTACTGAGGTTCTG
*PRDM14*_forward	GAGCCTGCAGGTCATAAAG
*PRDM14*_reverse	CTTGAGATGCTTGTCTCTGTAA
*TCL1A*_forward	GGCAAAGGCTGTGTATGT
*TCL1A*_reverse	CCTGACGCATGAGTACTTG

*SOX17*_forward	CTTCATGGTGTGGGCTAAGG
*SOX17*_reverse	CGGCCGGTACTTGTAGTTG
*GATA6*_forward	GCTGCACAGTCTACAGAGTC
*GATA6*_reverse	AGCGGTTGCACAAGTAGT
*GATA4*_forward	AAGAGATGCGTCCCATCAAG
*GATA4*_reverse	GACTGGCTGACCGAAGATG
*CDX2*_forward	AACCTGTGCGAGTGGATG
*CDX2*_reverse	CCTTTCTCCGAATGGTGATGTA
*TEAD4*_forward	TGTGAGTACATGGTCAACTTCAT
*TEAD4*_reverse	GCTGACACCTCGAAGACATAC
*GATA3*_forward	TACTACGGAAACTCGGTGAGG
*GATA3*_reverse	TGGATGGACGTCTTGGAGAA
*TET1*_forward	TGTCGGCTTGGCAAGAAAGA
*TET1*_reverse	AGACCACTGTGCTGCCATTA
*TET2*_forward	GTGAGATCACTCACCCATCGCATA
*TET2*_reverse	TACTGGCACTATCAGCATCACAGG
*TET3*_forward	TCTTCCGTCGTTCAGCTACTACAG
*TET3*_reverse	GTGGAGGTCTGGCTTCTTAAA

Table 2-3. Primers used for PCR amplification of bisulfite converted DNA.

Primers	Sequence (5' → 3')
NANOG promoter1_forward	AAAATTAGGTAGAGATATTATTAAAAA
NANOG promoter1_reverse	AAATATTCCCTCTATACCCACTTAAC
NANOG promoter2_forward	CTTATATAGGAAGAGAAGAGATTAAATTG
NANOG promoter2_reverse	CCCAACAATACTTACTAAATAAACTTTCC
NANOG gene-body1_forward	CTAATTTAATATGAGTGTGGA
NANOG gene-body1_reverse	GAATATTAAAAAATTCTTACATCTACTAAAA
NANOG gene-body2_forward	GAGAGGTAGAAGTATTTTAGTTTTTAGTAGA
NANOG gene-bopy2_reverse	GTAAAATAATTTAAAAATAAATCCATAATTT

Table 2-4. PCR conditions for amplification of bisulfite converted DNA.

Amplified regions	Primers	PCR conditions
NANOG promoter part 1	*NANOG* promoter1_forward	Initial denaturation for 2 min at 95℃ and 40 cycles of denaturation for 30 sec at 95℃, annealing for 30 sec at 50℃, extension for 30 sec min at 72℃, followed by a final extension for 2 min at 72℃.
	NANOG promoter1_reverse	
NANOG promoter part 2	*NANOG* promoter2_forward	Initial denaturation for 2 min at 95℃ and 40 cycles of denaturation for 30 sec at 95℃, annealing for 30 sec at 50℃, extension for 45 sec min at 72℃, followed by a final extension for 2 min at 72℃.
	NANOG promoter2_reverse	
NANOG gene-body part 1	*NANOG* gene-body1_forward	Initial denaturation for 2 min at 95℃ and 40 cycles of denaturation for 30 sec at 95℃, annealing for 30 sec at 45℃, extension for 30 sec min at 72℃, followed by a final extension for 2 min at 72℃.
	NANOG gene-body1_reverse	
NANOG gene-body part 2	*NANOG* gene-body2_forward	Initial denaturation for 2 min at 95℃ and 40 cycles of denaturation for 30 sec at 95℃, annealing for 30 sec at 50℃, extension for 30 sec min at 72℃, followed by a final extension for 2 min at 72℃.
	NANOG gene-bopy2_reverse	

Table 2-5. Validation of targeting efficiency by CRISPR/Cas9 at *TET1* gene locus. All blastocysts genotyped carried mutations at the three target sites indicating 100% targeting efficiency.

Embryos genotyped (n=10)	Target site 1	Target site 2	Target site 3
Embryo #1	Homozygous	Homozygous	Mosaic w/ wildtype
Embryo #2	Homozygous	Homozygous	Mosaic w/ wildtype
Embryo #3	Homozygous	Biallelic	Mosaic w/ wildtype
Embryo #4	Biallelic	Biallelic	Mosaic w/o wildtype
Embryo #5	Biallelic	Biallelic	Mosaic w/o wildtype
Embryo #6	Biallelic	Homozygous	Mosaic w/o wildtype
Embryo #7	Biallelic	Homozygous	Mosaic w/ wildtype
Embryo #8	Biallelic	Biallelic	Mosaic w/o wildtype
Embryo #9	Homozygous	Biallelic	Mosaic w/ wildtype
Embryo #10	Biallelic	Biallelic	Mosaic w/ wildtype

Table 2-6. Development of *TET1* KO embryos to blastocyst stage. Frequency of blastocyst formation at day 7 decreased in *TET1* KO embryos compared to that in IVF control and injection control (Cas9 only) embryos. Different letters indicate statistical difference.

	Number of blastocysts at day 7 (%)	Number of oocytes used
IVF control	111 (28.8)[a]	385
Injection control (Cas9)	164 (18.5)[b]	885
TET1 KO	119 (12.4)[c]	958

Table 2-7. Development of embryos treated with DMOG. Frequency of blastocyst formation at day 7 decreased in embryos treated with DMOG compared to that in control embryos. Different letters indicate statistical difference.

	Number of blastocysts at day 7 (%)	Number of oocytes used
Control	304 (21.3)[a]	1429
DMOG	261 (17.6)[b]	1484

Chapter 3. Presence of porcine *TET3L* isoform in oocytes: potential involvement in the DNA demethylation process

Abstract

Fertilized oocytes undergo genome-wide DNA demethylation with the exception of imprinted genes and certain repetitive elements. Ten-eleven translocation 3 (TET3) protein is known to be responsible for the DNA demethylation process by catalyzing the oxidation of 5-methylcytosine (5mC) into 5-hydroxymethylcytosine (5hmC). Recent studies in the mouse indicate that multiple *Tet3* isoforms exist in oocytes, implying differential actions of the isoforms. Although dissecting gene structures of *TET3* isoforms is crucial to understand the mechanism of the TET3-driven DNA demethylation process, major isoforms present in oocytes or early stage embryos have not been identified beyond the mouse model. Here, we investigated the presence of *TET3* isoforms in porcine oocytes and cumulus cells, and followed changes in the level of *TET3* isoforms during oocyte maturation to further understand the mechanism of *TET3* action in the DNA demethylation process. The 5'RACE revealed three different *TET3* isoforms present in GV and MII cumulus cells. Among the isoforms, the longest variant (*pTET3L*) contained sequences for the CXXC domain, known to carry DNA binding property. Interestingly, the expression of the *pTET3L* isoform was only verified in oocytes, suggesting that the isoform may be the predominant isoform in porcine oocytes. To characterize transcript abundance of *pTET3L*, RNA was isolated from different cells and/or tissues, including cumulus cells, oocytes, brain, spleen, and lung. The expression level of *pTET3L* was much greater in MII oocytes compared to that in somatic tissues. In addition, the expression level of *pTET3L* increased over 3-fold during oocyte maturation, i.e. from GV to MII stage oocytes, indicating

that *pTET3L* may have a significant role in DNA demethylation after fertilization. In conclusion, the *TET3* isoform containing CXXC domain is predominantly expressed in mature porcine oocytes, suggesting an important role of the *TET3* CXXC domain in DNA demethylation in zygotes.

Introduction

Epigenetic marks on the genome, such as DNA methylation and histone modifications, are stably maintained in somatic cells; however, genome-wide reprogramming of epigenetic marks can be seen during mammalian germ cell development and preimplantation embryogenesis (150, 151). During primordial germ cell (PGC) development, genome-wide DNA demethylation occurs to erase and reestablish parental genomic imprints (180). Fertilized oocytes also undergo genome-wide reprogramming of DNA methylation to erase germ cell-specific epigenetic marks in order to acquire totipotent developmental potential (181). The genome of fertilized oocytes is demethylated by active (replication independent) and passive (replication dependent) pathways and *de novo* methylated by DNA methyltransferase 3A and B (DNMT3A/B) to reestablish cell type-specific DNA methylation levels as embryos develop (182). Upon fertilization, distinct demethylation processes are observed in the two parental genomes of zygotes. The DNA methylation level of the paternal genome is dramatically reduced shortly after fertilization by the active pathway, whereas the level of maternal DNA methylation is gradually decreased following embryo cleavage through the passive pathway (33, 38).

The discovery of ten-eleven translocation (TET) family as the main dioxygenases of 5mC led to the understanding of novel active demethylation process. Active demethylation is initiated by the oxidation of 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC) via TET family

enzymes in a 2-oxoglutarate- and Fe (II)-dependent manner (43). All three TET family proteins, TET1, TET2, and TET3 proteins, possess the 5mC oxidase activity. The 5hmC form can be further oxidized to 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) by the TET proteins (50, 51). These 5mC derivatives can be recognized by thymine DNA glycosylase, then subsequently converted to unmethylated cytosine through the base excision repair pathway (50, 52). In addition, because DNMT1 has a low affinity to the DNA template containing 5hmC, TET-mediated 5mC oxidation could facilitate replication-dependent passive demethylation (53, 54).

The three TET proteins are differentially expressed during preimplantation embryo development; TET1 and TET2 are highly expressed in blastocysts, while TET3 is detected in oocytes and early stage embryos (55, 82). Among the three TET proteins, TET3 is responsible for the demethylation of the paternal genome in zygotes (63, 64, 67). The maternal DNA is not affected by the TET3-mediated demethylation process due to the binding of STELLA (also known as PGC7 or DPPA3) to maternal chromatin, thus inhibiting the access of TET3 (63, 71, 77). Conditional knockout of *TET3* in early stage mouse embryos resulted in the impairment of 5hmC formation and caused abnormal expression of *Oct4* and *Nanog* due to incomplete demethylation of promoter regions of these genes (67). TET3-mediated demethylation is the first step in the process of epigenetic transition from gametes to embryonic status; however, little is known about the molecular mechanism of TET3 by which a methylated region is recognized and oxidized into 5hmC.

All TET family proteins possess a conserved C-terminal catalytic domain, which includes a Cys-rich insert and a double-stranded β-helix (DSBH) domain. However, only TET1 and TET3 possess the N-terminal CXXC domain, a potential DNA binding module. The CXXC domains

found in other proteins, such as MLL and DNMT1, specifically bind to unmethylated CpG dinucleotides (183, 184). However, TET1 CXXC domain is known to bind CpG sequences regardless of cytosine modification types in the mouse (83, 185), and TET3 CXXC is reported to bind unmethylated cytosine instead of CpG dinucleotide in Xenopus (186). Although the DNA binding nature of TET3 CXXC domain is still elusive, the coexistence of CXXC domain and catalytic domain can confer DNA demethylase property to TET3. Curiously, different *TET3* isoforms lacking the N-terminal CXXC domain are also identified in mouse and human (187). These isoforms are speculated to be recruited to a target genome through the association with other proteins carrying CXXC domain or DNA binding factors (66, 188). To better understand the demethylation mechanism of TET3 in fertilized oocytes, dissection of the structure and function of TET3 isoforms expressed in oocytes is critical. Recently, the identification of an oocyte-specific *Tet3* isoform lacking the CXXC domain was reported in a mouse study (189), raising a possibility that the recruitment of TET3 to the methylated paternal genome of zygotes is independent of the intrinsic CXXC domain. However, further evidence for the existence of an oocyte-specific *TET3* isoform has not been demonstrated in other species. In this study, three *TET3* isoforms were identified in germinal vesicle (GV) or metaphase II (MII) stage porcine cumulus cells. We demonstrate that a *TET3* isoform containing CXXC domain is only expressed in MII stage porcine oocytes among the three isoforms, implying a predominant role of the isoform in demethylation of the paternal pronucleus in porcine zygotes.

Materials and Methods

In vitro maturation of porcine oocytes

Sow ovaries were obtained from a local abattoir and transported to the laboratory. Immature oocytes were aspirated from the follicles using an 18-gauge needle attached to a syringe. Aspirates were washed with TL-Hepes medium, and then cumulus-oocyte complexes (COCs) with evenly granulated cytoplasm and intact surrounding cumulus cells were collected using a finely drawn glass pipet under a microscope. Approximately 50 COCs were placed in a single well of a 4-well plate containing maturation medium (medium 199 supplemented with 3.05 mM glucose, 0.91 mM sodium pyruvate, 0.57 mM cysteine, 10 ng/ml epidermal growth factor (EGF), 0.5 μg/ml luteinizing hormone (LH), 0.5 μg/ml follicle stimulating hormone (FSH), 10 ng/ml gentamicin, and 0.1% polyvinyl alcohol (PVA); pH 7.4), and then incubated for 42–44 h at 38.5 ℃, 5% CO2, and 100% humidity.

Rapid amplification cDNA ends (RACE)

To obtain oocytes and cumulus cells separately, approximately 60 immature or mature COCs were placed in denuding medium (0.3 M mannitol, 0.001% BSA, 0.03% hyaluronidase, 5% TL-Hepes medium in distilled water; pH 7.4) and vortexed. Subsequently, polyadenylated RNAs were isolated from the oocytes and cumulus cells using the Dynabeads® mRNA DIRECT™ Micro Kit (ThermoFisher Scientific, USA). To identify sequence variants at the 5' end of *TET3* mRNA, 5'RACE was performed using the SMARTer® RACE 5'/3' kit (Takara Bio, USA) following manufacturer's instructions. Briefly, the first-strand cDNA was synthesized from the polyadenylated RNAs, and then 5' ends of *TET3* mRNA were amplified from the cDNA using a gene-specific primer (5' - GATTACGCCAAGCTTATCACGGCATTCTGGCAGTGATGGCCC - 3') at the touchdown PCR condition shown in Table 3-1. The gene-specific primer was designed based on the

previously cloned partial porcine *TET3* sequences (GenBank accession number: KC137685.1). The amplicons were gel extracted and ligated into the pRACE vector, then subsequently transformed into TOP10 competent cells. Plasmids were isolated from the randomly selected 3 – 5 clones using the GeneJET Plasmid Miniprep Kit (ThermoFisher Scientific, USA), and sent out for Sanger sequencing.

RT-PCR

Polyadenylated RNAs were isolated from the pooled oocytes (n=30) and cumulus cells at both GV and MII stages using Dynabeads mRNA Direct Kit (Thermo Fisher Scientific). cDNAs were synthesized from isolated mRNAs isolated using Maxima H Minus First Strand cDNA Synthesis Kit (Thermo Fisher Scientific). cDNAs were PCR amplified with isoform-specific primers designed based on 5'RACE results and EST databases using the DreamTaq DNA Polymerases (ThermoFisher Scientific) under the following PCR condition: initial denaturation at 95℃ for 2min, 40 cycles of denaturation at 95℃ for 30s, annealing at 60℃ for 30s, and extension at 72℃ for 60s, and final extension at 72℃ for 5min. Primer sequences are described in Table 3-2.

Quantitative RT-PCR

Polyadenylated RNAs were isolated from the pooled oocytes (n=30) and cumulus cells using Dynabeads mRNA Direct Kit (Thermo Fisher Scientific). RNA was isolated from porcine ovary, spleen, lung, and brain tissues using RNeasy Plus Mini Kit (QIAGEN). cDNAs were synthesized using Maxima H Minus First Strand cDNA Synthesis Kit (Thermo Fisher Scientific)

according to the manufacturer's instructions. Two primer sets spanning CXXC domain or oxygenase domain (Table 3-2) were used for quantitative RT-PCR to characterize the expression pattern of the *TET3L* isoform. PCR amplification and detection of the amplicons were conducted with the ABI 7500 Real-Time PCR System (Applied Biosystems) using PerfeCTa SYBR Green SuperMix (Quantabio) under the following conditions: 95℃ for 3min, 40 cycles of denaturation at 95℃ for 10s, and annealing at 60℃ for 60s. All of the threshold cycle (CT) values of the tested genes were normalized to *GAPDH* level, and relative ratios were calculated using the $2^{-\Delta\Delta}$Ct method. Three biological and three experimental replications were used. Differences in gene expression were evaluated by Student's t-test. $P<0.05$ was considered as statistically significant.

Results

Three *TET3* isoforms in porcine cumulus cells

Previous studies in mouse and human demonstrate that the presence of the N-terminal CXXC domain varies among *TET3* isoforms; C-terminal Fe(II)- and 2-ketoglutarate-dependent catalytic domain is conserved in all isoforms. Previously, we cloned the *TET3* mRNA sequence lacking CXXC domain from MII stage porcine oocytes through RT-PCR. Due to the lack of complete porcine EST database, it was uncertain whether the cloned *TET3* sequence represented the entire coding of *TET3* mRNAs in porcine oocytes. Therefore, to identify the full length of *TET3* mRNA sequences and its isoforms in porcine oocytes, 5'RACE was conducted using a primer designed based on the previously cloned *TET3* and computationally predicted TET3 sequences (Fig. 3-1A). Polyadenylated mRNAs were isolated from oocytes and cumulus cells at both GV and MII stages, and 5' ends of TET3 cDNA were amplified using the *TET3*-specific

primer. In the 5'RACE assay, one PCR amplicon from GV cumulus cells and two amplicons from MII cumulus cells were detected (Fig. 3-1B). However, there was no obvious PCR product in both GV and MII oocytes (Fig. 3-1B). Sequencing of the PCR products from cumulus cells revealed that one *TET3* isoform and two *TET3* isoforms were present in GV and MII stage cumulus cells, respectively. The three isoforms possessed distinct exons at the 5' end, whereas exons at the 3' side (7 exons), constituting Fe(II)- and 2-ketoglutarate-dependent catalytic domain, were conserved between the isoforms (Fig. 3-2). Among the three isoforms, only the longest form, identified in MII cumulus cells (referred hereafter to as *pTET3L*), contained sequences representing the CXXC domain at N-terminal. The coding sequence (CDS) of *pTET3L* was identical with that of the computationally predicted TET3; however, a unique 5' untranslated region (5'UTR) was identified in the first two exons of the *pTET3L* (Fig. 3-2). Another isoform detected in MII cumulus cells (referred hereafter to as *pTET3M*) was shorter than the *pTET3L* isoform and lacked the N-terminal CXXC domain (Fig. 3-2). The shortest isoform detected in GV cumulus cells (referred hereafter to as *pTET3S*) also did not contain CXXC domain (Fig. 3-2).

pTET3L as a predominant isoform of *TET3* in porcine oocytes

The absence of amplicons from *TET3* 5'RACE in oocytes could be attributed to the amount of template originated from oocytes. Since we were able to clone a TET3 isoform from porcine oocytes during our previous study, we were certain that at least one TET3 isoform was present in mature porcine oocytes. Therefore, presence of the three *TET3* isoforms detected in cumulus cells was explored in MII oocytes using RT-PCR. The isoform-specific primers were designed based on sequences obtained from the 5'RACE results and EST databases (Fig. 3-3A

and 3-4A). Interestingly, only *pTET3L* isoform was detected from MII oocyte-derived cDNA, while MII cumulus cells possessed both *pTET3L* and *pTET3M* (Fig. 3-3B, 3-3C, 3-4B, and 3-4C). The expression of *pTET3S* isoform was identified in GV cumulus cells, but not detected in MII oocytes (data not shown). The RT-PCR data suggests that *pTET3L* may be the predominant isoform in porcine MII oocytes. At the amino acid sequence level, the pTET3L isoform containing CXXC domain was homologous to the full-length TET3 isoforms of mouse (mTET3FL) and human (hTET3FL); amino acid sequence identities of pTET3L with mTET3FL and hTET3FL were 87% and 91%, respectively (Fig. 3-5). The *pTET3M* isoform, which was not expressed in porcine MII oocytes, was homologous to the short *TET3* isoforms of mouse (mTET3S) and human (hTET3S) lacking CXXC domain (Fig. 3-5).

To analyze the transcript abundance of the *pTET3L* isoform in MII oocytes, quantitative RT-PCR was performed using two primer sets. One primer set was designed to amplify *pTET3L*-specific CXXC domain, and the other primer set spanned a catalytic domain, which is conserved among the *pTET3* isoforms (Fig. 3-6A). Intriguingly, the expression level of *pTET3L* increased over 3-fold during oocyte maturation, i.e. from GV to MII stage oocytes, implying that oocytes are preparing to demethylate the paternal pronucleus by elevating *pTET3L* transcription. A similar increase (approximately 3-fold) was found in the overall *TET3* transcripts during maturation (Fig. 3-6B). The similar transcriptional increases between the *pTET3L* and the overall *pTET3* isoforms indicate that *pTET3L* is the major isoform present in porcine oocytes. The transcript abundance of *pTET3L* in oocytes was compared to that in different cells and/or tissues including ovary, brain, spleen, and lung. The expression level of *pTET3L* isoform was much higher in MII oocytes compared to that in somatic tissues; MII oocytes expressed the *pTET3* isoform over 180-fold higher than brain cells and 20-fold higher than lung cells (Fig. 3-6C). A

similar expression pattern was observed when the abundance of the overall *TET3* transcripts were measured using the primers spanning oxygenase domain (Fig. 3-6C). The PCR results suggest that *pTET3L*, the predominant isoform expressed in oocytes, may have a significant role in DNA demethylation after fertilization.

Discussion

Since the report of the actions of TET proteins in mammalian embryos and pluripotent stem cells, TET3 has been considered as one of the major factors responsible for epigenetic reprogramming of preimplantation stage embryos. However, little is known about how TET3 targets specific genome loci because of limited information on the predominant *TET3* isoform in oocytes and the role of the presumable DNA binding domain of TET3, the CXXC domain. Although two *TET3* isoforms, CXXC-containing *TET3FL* and CXXC-lacking *TET3S*, have been identified in mouse and human, the major isoform of *TET3* in oocytes and embryos has not been experimentally determined. Interestingly, a recent study identified oocyte-specific *Tet3* isoform (*mTet3O*) in mouse, which is structurally different from the other two isoforms. The *mTet3O* transcript is created by an alternative promoter located approximately 5kb upstream of the start codon of *mTet3FL* isoform, and lacks CXXC domain via alternative splicing (189). As the isoform is not detected in embryonic stem cells or somatic tissues, including neuronal cells, the study argues that the isoform is predicted to be involved in preferential global oxidation of the paternal genome in mouse zygotes. In our study, we cloned three different *TET3* isoforms in porcine cumulus cells and detected an enriched expression of a single isoform in mature porcine oocytes. Consistent with the mouse study (189), the short *TET3* isoforms, lacking CXXC domain, were not expressed in porcine oocytes. In a previous study in the mouse, the relative

expression level of CXXC-containing *mTet3FL* is much weaker than that of *mTet3O* in oocytes (189), while *mTet3FL* is considered as a predominant form expressed in neuronal cells (190, 191). Contrary to mouse oocytes, we observed a high expression level of CXXC-containing *TET3* isoform (*pTET3L*) in porcine oocytes compared to somatic tissues, including brain tissue. Furthermore, approximately a 3-fold increase in transcript amount was observed in both *pTET3L* and overall *pTET3* during oocyte maturation. These results suggest that the *TET3* isoform possessing CXXC domain may play an important role in global demethylation of the paternal genome in porcine zygotes. Although the possibility of the presence of another oocyte-specific *TET3* isoform in porcine oocytes cannot be excluded, we could not find an isoform homologous to *mTet3O* isoform via *in silico* searches of porcine EST databases and genomic sequences.

Although 5'- and 3'-untranlated regions (UTRs) do not contribute to amino acid sequence of proteins, the regions contribute to the regulation of protein translation by forming specific secondary structure or providing binding sites for RNA binding proteins (192). Translation of *TET* family proteins are also regulated by interactions between the UTRs of TET family genes and miRNAs (193, 194). In 5'RACE assay, we found 5'UTR sequences (786bp) of the *pTET3L*, highly homologous to that of *mTet3FL* and *hTET3FL*; nucleotide sequence identity with *mTet3FL* and *hTET3FL* was 81% and 88%, respectively. The 5'UTR sequence of the *pTET3L* was longer than the reported sequences of *mTet3FL* (532bp) and *hTET3FL* (672bp), suggesting that the full 5'UTR sequences of *mTet3FL* and *hTET3FL* may be longer than previously reported. In addition, the 5'UTR sequence of *pTET3L* included a number of CpG sites (72 CpG sites). Particularly, the newly identified sequences that do not overlap with reported 5'UTR sequences in mouse and human possessed extremely high CpG density (24 CpG sites within

118bp). The high CpG content implies an important role of the 5'UTR in transcriptional regulation of *pTET3L* as a part of the CpG islands overlap with the promoter region of *TET3*.

Identification of *TET3* isoform types located in oocyte and zygote is essential to understand the mechanistic action of *TET3* in the demethylation process. Unlike CXXC domain of other proteins, such as MLL and CFP1, which require unmethylated CpG dinucleotides for DNA binding (183, 195), the CXXC domain of *TET3* has unique DNA binding characteristics. Specifically, CpG dinucleotide context is not essential, but the presence of unmodified cytosine base is required for DNA binding of the CXXC domains of Xenopus and human *TET3* (186). Similarly, CpG dinucleotides are not absolutely required for DNA binding of the CXXC domain of mouse *Tet3*, although CpG dinucleotides are slightly preferred compared to other bases (188). These reports suggest that the specific DNA binding property of *TET3* CXXC domain could influence *pTET3L* targeting loci on the genome of porcine preimplantation embryos. Interestingly, *mTet3S* isoform lacking CXXC domain displays similar DNA binding characteristics to the *mTet3FL* isoform; *mTet3S* also has a slight preference for CpG dinucleotides, especially for substrates with methylated CpG sites (188). However, the CXXC domain alone possesses lower binding affinity toward methylated CpG substrates compared to *mTet3S* and *mTet3FL* forms (188), indicating a possibility that sequences or structures outside of CXXC domain could affect the binding property of TET3 protein. The Cys-rich region, located in the catalytic domain of *Tet3*, is postulated to be involved in target recognition as a part of the DNA binding surface (196). Furthermore, *mTet3S* lacking intrinsic CXXC appears to associate with an extrinsic CXXC domain, CXXC4 protein, which may directly bind *mTetS* to DNA (188). Taken together, the oocyte specific isoform lacking CXXC domain may also have the potential to bind to DNA for global demethylation by utilizing other intrinsic domains, such as Cys-rich

region, or via interaction with an extrinsic CXXC domain. The possibility is only arguable in mouse because the isoform lacking CXXC domain has only been detected in mouse oocytes.

In this study, we identified multiple *TET3* isoforms in porcine oocytes and cumulus cells and predominant expression of *pTET3L* isoform containing CXXC domain in porcine oocytes. The findings imply an important role of CXXC domain in targeting DNA for genome-wide demethylation in zygotes. The differences in major *TET3* isoform between mouse and porcine oocytes suggest a possibility of species-specific reprogramming of DNA methylation marks during embryo development. To clarify the role of TET3 protein in reprogramming the genome of porcine embryos, further studies should focus on the mechanistic action of CXXC domain during early post-fertilization stages.

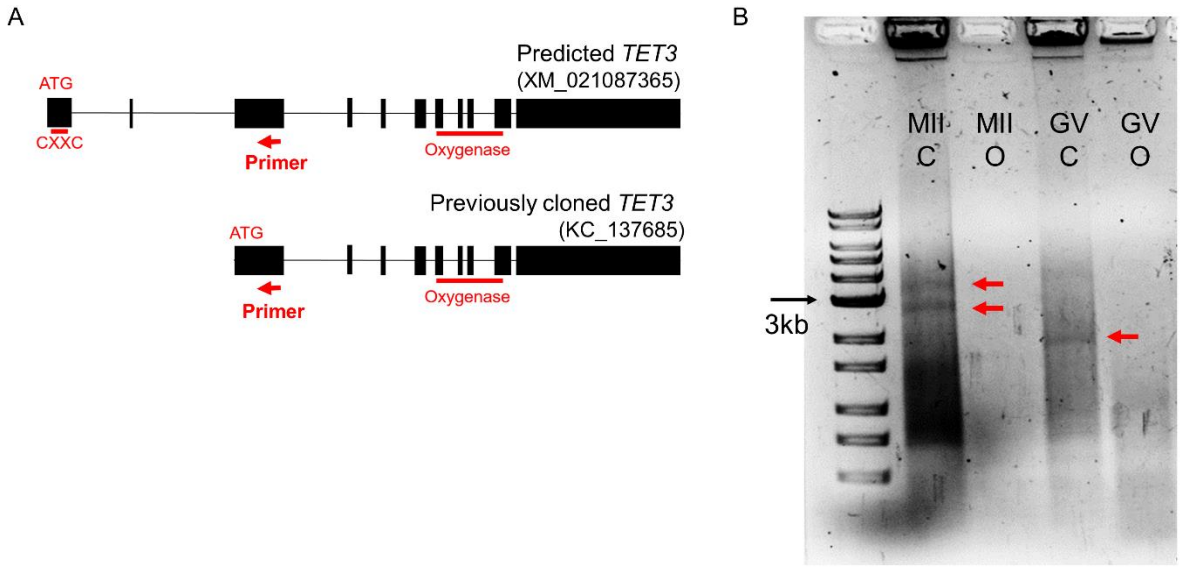


Fig. 3-1. *TET3* transcripts detected in porcine cumulus cells and oocytes by 5'RACE assay. (A) Structures of predicted and cloned porcine *TET3* genes in the previous study. Gene-specific primer (red arrows) for 5'RACE was designed to target the conserved exon between the predicted and the previously cloned sequences. (B) Gel electrophoresis image of *TET3* 5'RACE PCR. Two and single TET3 amplicon(s) (red arrows) were detected in GV (GV C) and MII stage (MII C) cumulus cells, respectively. There were no PCR amplicons in either GV (GV O) or MII (MII O) stage oocytes.

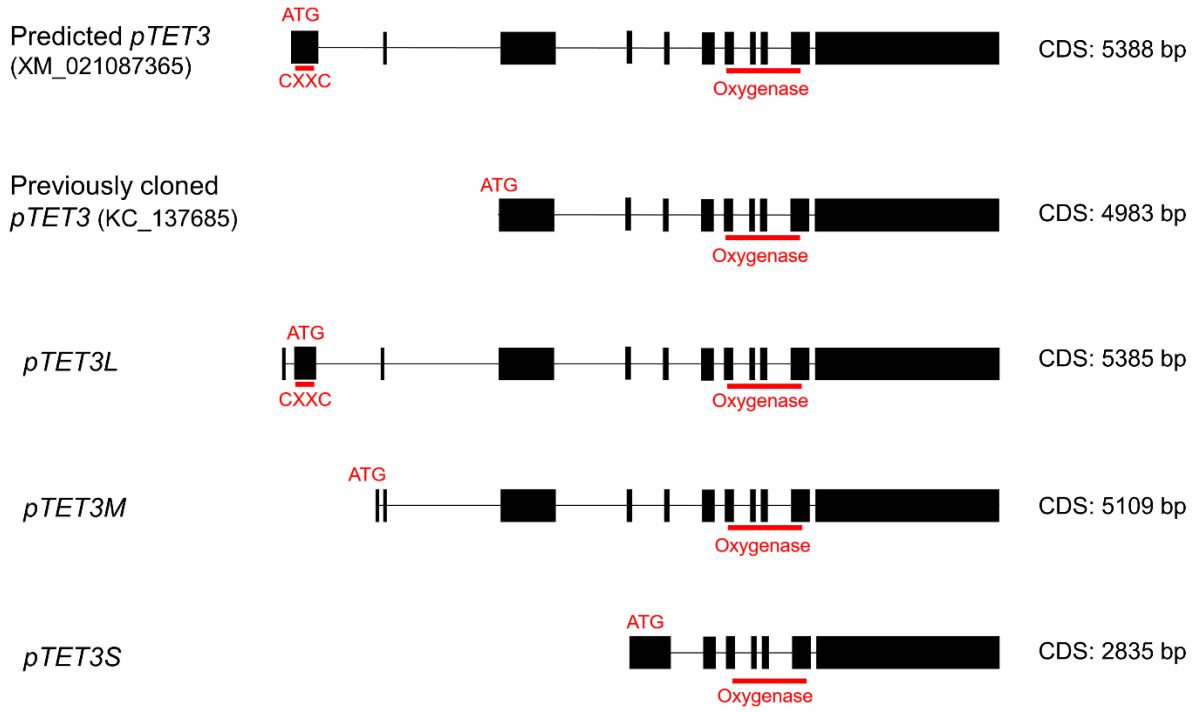


Fig. 3-2. Gene structure of and *TET3* isoforms depicted based on 5'RACE results and EST databases. *pTET3L* isoform containing N-terminal CXXC domain was structurally similar to the predicted *pTET3* isoform. *pTET3M* and *pTET3S* isoforms possessed the conserved C-terminal catalytic domain, but lacked CXXC domain.

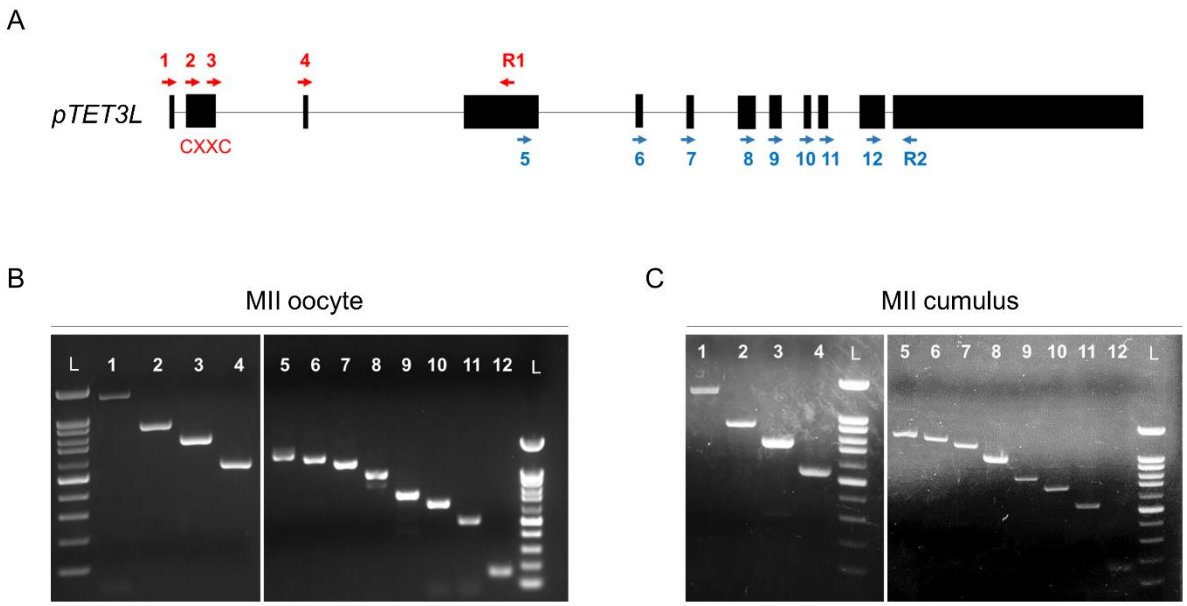


Fig. 3-3. RT-PCR of *pTET3L* isoform in MII oocytes and MII cumulus cells. (A) Primers used in RT-PCR analysis; four forward primers (1 – 4) were coupled with reverse primer 1 (R1) and eight forward primers (5 – 12) were coupled with reverse primer 2 (R2). *pTET3L* isoform was identified in both MII oocytes (B) and MII cumulus cells (C).

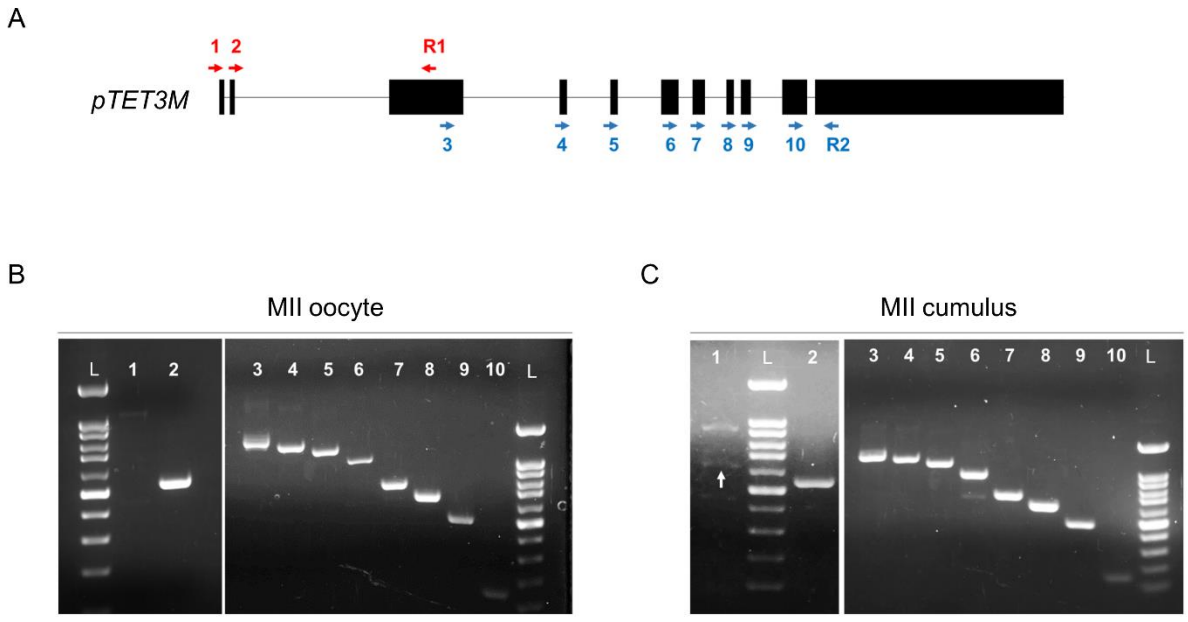


Fig. 3-4. RT-PCR of *pTET3M* isoform in MII oocytes and MII cumulus cells. (A) Primers used in RT-PCR analysis; two forward primers (1 – 2) were coupled with reverse primer 1 (R1) and eight forward primers (3 – 10) were coupled with reverse primer 2 (R2). (B) *pTET3M* isoform was not detected in MII oocytes; exon1, which is distinct in *pTET3M* isoform, was not detected, while other conserved regions were amplified. (C) Presence of *pTET3M* isoform was confirmed in MII cumulus cells by detecting the distinct exon1 (white arrow).

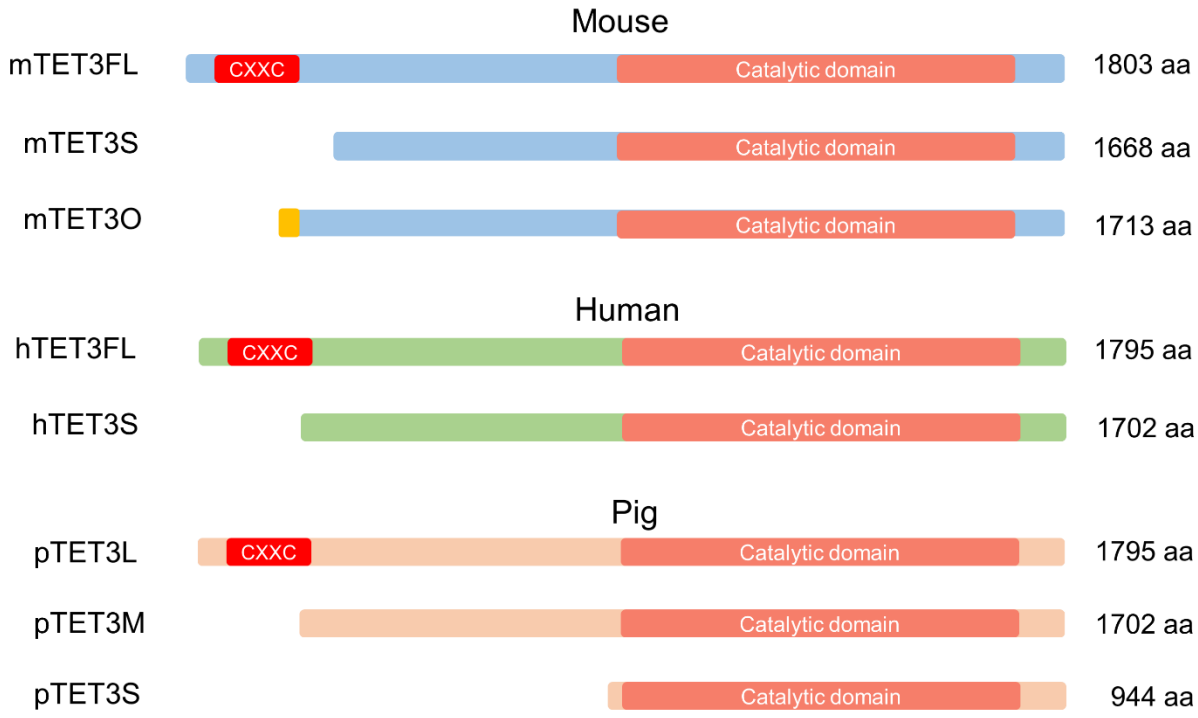


Fig. 3-5. Homology of porcine TET3 isoforms to mouse and human TET3 major isoforms. pTET3L isoform was homologous to the mouse and human full-length TET3 isoforms (mTET3FL and hTET3FL) containing CXXC domain; amino acid sequence identities of pTET3L with mTET3FL and hTET3FL were 87% and 91%, respectively. pTET3M isoform lacking CXXC domain was homologous to mouse and human short TET3 isoforms (mTET3S and hTET3S); sequence identities of pTET3M with mTET3S and hTET3s were also 87% and 91%, respectively. Yellow box in mTET3O isoform indicates an additional N-terminal exon distinct from mTET3FL and mTET3S isoforms.

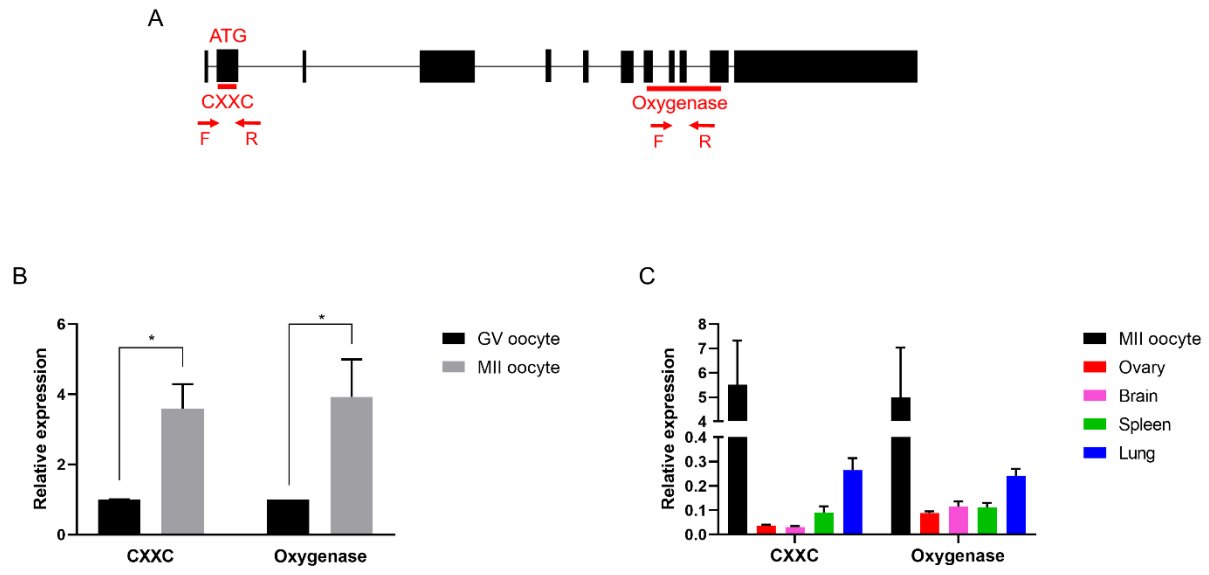


Fig. 3-6. Relative transcript abundance of *pTET3L* isoforms across cells and tissues. (A) Expression levels of *pTET3L* isoform and overall isoforms were measured using primer sets spanning CXXC domain and oxygenase domain, respectively. (B) The *pTET3L* isoform and overall isoforms showed similar increase (approximately 3-fold) during oocyte maturation. (C) The transcript amount of *pTET3L* isoform was much higher in oocytes than in somatic tissues (e.g. 20-fold higher than lung tissue), similar to the expression pattern of overall *TET3*. *GAPDH* was used as a normalizer. Results are shown as mean ± SEM (n = 3).

Table 3-1. PCR condition for amplification of 5'ends of *TET3* mRNA

Cycles	**Temperature**	**Time**
5	94℃	30 sec
	72℃	3 min 30 sec
5	94℃	30 sec
	70℃	30 sec
	72℃	3 min 30 sec
30	94℃	30 sec
	68℃	30 sec
	72℃	3 min 30 sec

Table 3-2. Primers used for RT-PCR and quantitative RT-PCR

Usage	Primer	Sequence (5' → 3')
RT-PCR (TET3L)	TET3L_Exon1_F1	GGAAACTTTGCCCCTTTGTGC
	TET3L_Exon2_F1	AGCATCATGAGTCAGTTTCAG
	TET3L_Exon2_F2	GGGCGGAAGAAACGGAAACGGT
	TET3L_Exon3_F1	GACAAGGGGCAGCTGTCAA
	TET3L_Exon4_R1	CACTCGAGGTAGTCAGGGCA
	TET3L_Exon4_F1	GTTCCCCACTTGTGATTGCG
	TET3L_Exon5_F1	CACTCATCTGGGATCTGGCC
	TET3L_Exon6_F1	CATCCGGATCGAGAAGGTCA
	TET3L_Exon7_F1	GTGATTGTCATCCTCATCCTGG
	TET3L_Exon8_F1	CAAGACTCCTCGCAAGTTCC
	TET3L_Exon9_F1	GCTCCCCTCTATAAGCGGCT
	TET3L_Exon10_F1	CCACAAGGACCAGCATAACCT
	TET3L_Exon11_F1	GAGAAGATCAAGCAGGAGGCC
	TET3L_Exon12_R1	GCGTTGCCACTGTATTTGAAGG
RT-PCR (TET3M)	TET3M_Exon1-F1	GATAATGGGAGGCCCGCC
	TET3M_Exon2-F1	GACAAGGGGCAGCTGTCAA
	TET3M_Exon3-R1	CACTCGAGGTAGTCAGGGCA
	TET3M_Exon3-F1	GTTCCCCACTTGTGATTGCG
	TET3M_Exon4-F1	CACTCATCTGGGATCTGGCC
	TET3M_Exon5-F1	CATCCGGATCGAGAAGGTCA

	TET3M_Exon6-F1	GTGATTGTCATCCTCATCCTGG
	TET3M_Exon7-F1	CAAGACTCCTCGCAAGTTCC
	TET3M_Exon8-F1	GCTCCCCTCTATAAGCGGCT
	TET3M_Exon9-F1	CCACAAGGACCAGCATAACCT
	TET3M_Exon10-F1	GAGAAGATCAAGCAGGAGGCC
	TET3M_Exon11-R1	GCGTTGCCACTGTATTTGAAGG
qRT-PCR	TET3_CXXC_F1	CAATGGCCGGGAGTGAAT
	TET3_CXXC_R1	GCTCCTTCAGCAGCCTTTAT
	TET3_Oxygenase_F1	TATCAGAACCAGGTGACCAATG
	TET3_Oxygenase_R1	GTAGAGGTTATGCTGGTCCTTG

Chapter 4. *TET3* CXXC domain is critical for post-fertilization demethylation and expression of pluripotency genes in pig embryos

Abstract

Ten-eleven translocation (TET) methylcytosine dioxygenases are considered to play an important role in the regulation of DNA methylation patterns by converting 5-methylacytosine (5mC) to 5-hydroxymethylcytosine (5hmC). TET3 protein, enriched in mature oocytes and early stage embryos, is known to initiate DNA demethylation of the paternal genome in zygotes. Previous studies in mouse cells indicate that the N-terminal CXXC domain of TET3 is important in catalyzing the oxidation of 5mC through its potential DNA binding ability; however, it is not clear whether the DNA binding capacity of CXXC domain is required for the 5hmC conversion in mammalian embryos. Here, we investigated the role of the *TET3* CXXC domain in controlling post-fertilization demethylation in porcine embryos by injecting *TET3* GFP-CXXC domain into mature porcine oocytes. The injected CXXC was exclusively localized in the pronuclei, indicating that the CXXC domain may act as a nuclear localization signal to TET3. The CXXC overexpression reduced the 5mC level in zygotes and enhanced the DNA demethylation of the *NANOG* promoter in 2-cell stage embryos. Furthermore, there was an increase in transcript abundance of *NANOG* and *ESRRB* in blastocysts developed from GFP-CXXC injected oocytes compared to control blastocysts. In this study, we found that the CXXC domain of *TET3* is critical for post-fertilization demethylation of porcine embryos and proper expression of pluripotency related genes (*NANOG* and *ESRRB*) in blastocysts.

Introduction

5-methylcytosine (5mC) is a key epigenetic mark implicated in orchestrating gene expression levels (171). 5mC has been regarded as the only modified cytosine form playing important roles in transcriptional regulation in mammalian DNA (197, 198). Recently, it has been demonstrated that 5mC can be oxidized into 5-hydroxymethylcytosine (5hmC) by ten-eleven translocation (TET) family enzymes (43). 5hmC can also be further oxidized to 5-formylcytosine (5fC) and 5-carboxylcytosine (5CaC) by TET enzymes (50, 51). 5hmC is thought to be an intermediate base pair of the demethylation process (199-201), as well as a stable epigenetic mark possessing functions distinct from 5mC (202, 203).

All three TET enzymes, *TET1*, *TET2*, and *TET3*, possess 5mC oxidase activity; however, their expression is often tissue specific and gene structure consists of different domains. *TET1* and *TET2* are abundant in pluripotent stem cells and primordial germ cells (55, 201), while *TET3* is only substantially expressed in oocytes, zygotes, and neuronal cells (63-65, 67). All the three TET enzymes contain C-terminal Fe(II)- and 2-oxoglutarate (2OG)-dependent catalytic domain. The 5mC hydroxylation is directly mediated by the domain, and thus mutation on the iron binding site (43) or inhibitor of 2OG-dependent dioxygenases impairs the catalytic activity of TET enzymes (204). *TET1* and *TET3* have CXXC domain, a potential DNA binding module at their N-terminus, whereas *TET2* lacks the CXXC domain.

The mechanistic action of TET enzymes to oxidize 5mC is relatively well characterized; however, molecular mechanisms for their DNA recognition and targeting are still elusive. N-terminal CXXC domain has been considered as a potential DNA binding module of *TET1* and *TET3*. The CXXC domain of other proteins, such as DNMT1 and MLL, specifically binds to unmethylated CpG dinucleotides (183, 184). However, different DNA binding properties of

CXXC domain of TET enzymes have been reported. For example, *TET1* CXXC domain is known to bind CpG sequences regardless of cytosine modification types in mouse (83, 185), and *TET3* CXXC is reported to bind unmethylated cytosine instead of CpG dinucleotides in Xenopus (186). *TET2* lacks CXXC domain, but it can bind DNA sequences containing unmethylated CpG dinucleotides via association with an extrinsic CXXC domain (IDAX) (205). A similar interaction between TET3 and extrinsic CXXC domain was also reported in mouse (188). Furthermore, the Cys-rich region located in the catalytic domain of TET3 was postulated to be involved in target recognition as a part of the DNA binding surface (196). Despite efforts to characterize DNA binding properties of CXXC domain, it is not clear how TET proteins target to specific DNA loci in distinct cell types and developmental stages.

TET3 is an essential epigenetic factor for genome-wide demethylation of the genome in fertilized oocytes in mammals (67). Two major isoforms have been identified in mouse and human; *TET3* full-length isoform (*TET3FL*) containing CXXC domain, and *TET3* short isoform (*TET3S*) lacking CXXC domain (187). Although there are differences in the domain architecture between the two major isoforms, it has not been clear which isoform is responsible for the post-fertilization demethylation. Recently, a mouse study found an oocyte-specific *TET3* isoform distinct from the two major isoforms (189); the oocyte-specific *TET3* isoform lacks N-terminal CXXC domain. In the previous chapter, we identified *TET3* isoform containing CXXC domain (*pTET3L*) as a predominant form in porcine oocytes. Here, we provide evidence that global demethylation is mediated by *pTET3L* using the intrinsic CXXC domain in pig zygotes, and that the transcription of *NANOG* in blastocyst is affected by *TET3* CXXC targeting *NANOG* promoter region in early stage. Our data suggest that CXXC domain is responsible for the activity of *pTET3L* and subsequent genome-wide demethylation process.

Materials and Methods

Cloning of pig *TET3* CXXC

CXXC domain of pig *TET3* was identified through bioinformatics comparison of *TET3* sequences among different species. mRNAs were isolated from *in vitro* matured pig oocytes (n=40) using Dynabeads mRNA Direct Kit (Thermo Fisher Scientific), and subsequently cDNAs were synthesized using Maxima H Minus First Strand cDNA Synthesis Kit (Thermo Fisher Scientific). CXXC domain region was amplified from the cDNAs with a primer set (forward 5'-GGGCGGAAGAAACGGAAACGGT-3' and reverse 5'-ACACTTGCGCAGTTTGCAGATCTGG-3') under the following PCR condition: initial denaturation at 98℃ for 30s, 40 cycles of denaturation at 98℃ for 10s, annealing at 60℃ for 30s, and extension at 72℃ for 30s, and final extension at 72℃ for 5min, using Phusion DNA Polymerases (Thermo Fisher Scientific). To construct GFP-CXXC fusion protein, the amplified CXXC sequence was subcloned into the N-terminal GFP fusion vector (Thermo Fisher Scientific) and the insertion of the CXXC sequence was confirmed by Sanger sequencing. Subsequently, GFP-CXXC mRNA was synthesized from the linearized vector through *in vitro* transcription using the mMESSAGE mMACHINE T7 Ultra Kit following the manufacturer's instruction.

Microinjection

Sow ovaries were obtained from a local abattoir and gilt oocytes were purchased from Desoto Biosciences Inc. Oocytes were aspirated from the follicles with a needle attached to a

syringe, and *in vitro* matured in maturation medium (medium 199 supplemented with 3.05 mM glucose, 0.91 mM sodium pyruvate, 0.57 mM cysteine, 10 ng/ml epidermal growth factor (EGF), 0.5 μg/ml luteinizing hormone (LH), 0.5 μg/ml follicle stimulating hormone (FSH), 10 ng/ml gentamicin, and 0.1% polyvinyl alcohol (PVA); pH 7.4) for 30–32 h at 38.5 ℃, 5% CO_2, and 100% humidity. Cumulus cells were removed by vortexing in denuding medium (0.3 M mannitol, 0.001% BSA, 0.03% hyaluronidase, 5% TL-Hepes medium in distilled water; pH 7.4). *In vitro*-synthesized *TET3* CXXC mRNA (100ng/μl) was injected into the collected oocytes at 30-32h post the start of *in vitro* maturation using the FemtoJet microinjector (Eppendorf) in manipulation medium (medium 199 supplemented with 0.6mM $NaHCO_3$, 2.9mM Hepes, 30mM NaCl, 10ng/ml gentamicin, and 3mg/ml BSA; pH 7.4) on a heated stage of a Nikon inverted microscope. To confirm intracellular localization of TET3 CXXC, GFP mRNA (100ng/μl) injected oocytes used as a control. The GFP mRNA injected embryos were also used as a control for immunocytochemistry (ICC) and water injected oocytes served as a control for quantitative RT-PCR analysis. The injected oocytes were placed in maturation medium and incubated for 7–8 h to complete maturation. After *in vitro* maturation, oocytes were fertilized in mTBM medium (modified Tris-buffered medium with 113.1mM NaCl, 3mM KCl, 7.5mM $CaCl_2$, 11mM glucose, 20mM Tris, 2mM caffeine, 5mM sodium pyruvate, and 2mg/ml BSA; pH 7.4) for 5 h at 38.5 ℃, 5% CO_2, and 100% humidity, and then placed in PZM3 medium (158) for embryo culture.

Immunocytochemistry (ICC)

Zona-free embryos were fixed in 4% paraformaldehyde for 15 min at room temperature. The fixed embryos were washed and permeabilized in PBS containing 0.25% TritonX-100 for 1h. Then, the embryos were treated with 2N HCl for 30min and neutralized in Tris–HCl pH 8.5

for 10min. The samples were incubated in PBS containing 0.1% Tween- 20 and 2% bovine serum albumin for 1h at room temperature. After blocking, embryos were incubated in blocking solution together with 5mC (dilution 1:100; Active Motif) antibody overnight at 4℃. The next day, the samples were washed in blocking solution and stained with secondary antibodies conjugated with CF^{TM} 555 (dilution 1:500) for 1h at room temperature. DNA was stained with 1 μg/ml DAPI.

Quantitative RT-PCR

Seven to nine blastocysts per group were collected on day 7 to analyze gene expression patterns of *TET* family and pluripotency-related genes using quantitative RT-PCR. mRNA was immediately isolated from the pooled embryos using Dynabeads mRNA Direct Kit (Thermo Fisher Scientific), followed by cDNA synthesis using Maxima H Minus First Strand cDNA Synthesis Kit (Thermo Fisher Scientific) according to the manufacturers' instructions. Amplification and detection of PCR products were conducted with the ABI 7500 Real-Time PCR System (Applied Biosystems) using PerfeCTa SYBR Green SuperMix (Quantabio) under the following conditions: 95℃ for 3min, 40 cycles of denaturation at 95℃ for 10s, and annealing at 60℃ for 60s. Primer information can be seen in Table 4-1. All of the threshold cycle (CT) values of the tested genes were normalized to *GAPDH* level, and relative ratios were calculated using the $2^{-\Delta\Delta}$Ct method. Three biological and three experimental replications were used. Differences in the gene expression were evaluated by Student's t-test. $P<0.05$ was considered as statistically significant.

Bisulfite DNA sequencing

For analysis of DNA methylation status of repetitive elements and *NANOG* promoter regions, 25 2-cell embryos were pooled and their DNA was treated with bisulphite using EZ DNA methylation kit (Zymo Research) following the manufacturer's instruction. Subsequently, the bisulphite-treated DNA was PCR-amplified using specific primer sets under the following PCR conditions: initial denaturation for 2min at 95℃ and 40 cycles of denaturation for 30s at 95℃, annealing for 30s at 50℃, extension for 30s min at 72℃, followed by a final extension for 2min at 72℃. Primers used for amplification of *PRE-1* are forward 5'-GAGTATGTAGTTAGAAAAATAAAAATTAAA-3' and reverse 5'-CACAAAAATATTTCAACCCAAATAAAAT-3'. Primers used for amplification of *NANOG* promoter are forward 5'-AAAATTAGGTAGAGATATTATTAAAAA-3' and reverse 5'-AAATATTCCCTCTATACCCACTTAAC-3'. The PCR products were purified using GeneJet gel extraction kit (Thermo Fisher Scientific) and were then ligated into pCR 2.1 TA cloning vector (Invitrogen). Fourteen to 16 colonies of each cloned sample were sequenced and evaluated.

Results

TET3 CXXC as nuclear localization signal

In the previous chapter, we found that *pTET3L* isoform containing N-terminal CXXC domain is predominantly present in porcine MII oocytes. Sequences representing porcine *TET3* CXXC domain, consisting of 40 amino acids, were successfully cloned from the mRNA of MII oocytes (Fig. 4-1A) and fused with GFP gene at N-terminal region to track localization of the

CXXC in embryos (Fig. 4-1B). The GFP-CXXC was exclusively expressed in the (pro)nuclei of zygotes and 2-cell embryos, while control GFP was identified in both nucleic and cytoplasm (Fig. 4-2). These findings indicate that CXXC domain may be responsible for nuclear localization of *pTET3L* in early stage embryos. Although the nuclear localization signal (NLS) of *TET3* short form is located at the C-terminus (206, 207) (Fig. 4-3A), sequences acting as the NLS for *TET3* long form, containing CXXC domain, has not been experimentally determined. The nuclear localization of injected GFP-CXXC implies sequences within CXXC domain may lead *TET3* long form to nucleus. Therefore, a web-based computer software (208) was used to search for NLS sequences within the CXXC domain. One predicted NLS region (GRKKRK) was identified within the CXXC domain sequences which is conserved in mouse, human and porcine *TET3* long form (Fig. 4-3B). This result indicates that the NLS within CXXC domain may transport *TET3* long form into the nucleus.

Overexpression of *TET3* CXXC facilitates global demethylation at zygote stage

As *pTET3L* is the predominant isoform in porcine oocytes and the isolated CXXC domain is localized to the pronucleus of zygotes, we assumed that CXXC domain is important for DNA recognition of *pTET3L* and subsequent post-fertilization demethylation process. To examine the function of *TET3* CXXC domain in the demethylation process, embryos overexpressing CXXC domain was analyzed for the level of 5mC in the embryos. In ICC assay, zygotes injected with CXXC exhibited decreased 5mC level compared to control embryos (Fig. 4-4A). However, there was no detectable difference in 5mC levels between the CXXC overexpression and the control group at 2-cell stage (Fig. 4-4A). When the intensity of fluorescence was quantified, the 5mC level of zygotes overexpressing CXXC was significantly

lower than that of control zygotes (Fig. 4-4B upper-left). However, the 5mC level of 2-cell embryos injected with CXXC did not differ from that of control 2-cell embryos (Fig. 4-4B upper-right). While the level of 5mC gradually decreased from zygotes to 2-cell stage in control embryos (Fig. 4-4B bottom-left), a similar level of 5mC was detected between zygotes and 2-cell embryos overexpressing TET3 CXXC (Fig. 4-4B bottom-right). This finding suggests that injected CXXC facilitates global demethylation at the zygote stage, resulting in the low DNA methylation status of zygotes corresponding to the level that is normally attained at 2-cell stage.

CXXC-triggered hyper-demethylation increases the expression of *NANOG*

Although overexpression of CXXC facilitated demethylation at zygote stage, significant changes in the global methylation level were not detected at 2-cell stage. Because subtle methylation changes are not detectable by ICC, bisulfite genomic sequencing was conducted to observe locus-specific methylation status in 2-cell embryos. The methylation level of repetitive elements, PRE-1, in 2-cell embryos was not changed by CXXC injection compared to that of control embryos, consistent with the ICC results (Fig. 4-5A).

Because deficiency in maternal *Tet3* impeded initial demethylation of *Nanog* promoter in mouse embryos (67), we assessed the methylation level of *NANOG* promoter in porcine embryos to examine the effect of CXXC overexpression on demethylation of the region. Interestingly, 2-cell embryos injected with CXXC displayed a lower DNA methylation level than control embryos, indicating demethylation of *NANOG* promoter was enhanced by the injected CXXC domain (Fig. 4-5B). Next, expression levels of pluripotency genes, including *NANOG,* were measured by quantitative RT-PCR to examine whether the altered methylation level by CXXC

overexpression affects the transcription of pluripotency genes in blastocysts. Frequency of blastocyst formation at day 7 in embryos injected with CXXC was not different from that in injection control embryos, although the blastocyst formation was lower compared to non-injected IVF control embryos (Table 4-2). The transcript amounts of *NANOG* and *ESRRB* were significantly increased (approximately 3- and 2-fold, respectively) by CXXC overexpression (Fig. 4-6A). Expression levels of other pluripotency genes, *OCT4* (*POU5F1*) and *SOX2,* were not changed (Fig. 4-6A). Also, CXXC overexpression did not impact the transcription of *TET* family genes (Fig. 4-6B). Along with the bisulfite sequencing data, the RT-qPCR result indicates that enhanced demethylation of *NANOG* promoter in early stage embryos by CXXC overexpression leads to the increased *NANOG* expression at blastocyst stage.

Discussion

Mammalian preimplantation embryos undergo a series of dynamic epigenetic reprogramming and the genome-wide demethylation is the first step of the reprogramming process. Although *TET3* is an essential player for the demethylation, especially in the paternal pronucleus, very limited information is available on possible mechanisms of *TET3* action in the recognition of methylated genomic regions and target specificity. A previous study using Xenopus embryos suggested that *TET3* CXXC domain possesses unique DNA binding properties critical for *TET3* targeting (186). However, the role of CXXC domain in *TET3*-driven demethylation has remained elusive in mammalian preimplantation embryos. In this study, we provide an evidence that the global DNA demethylation after fertilization is mediated by *pTET3L* isoform potentially by utilizing its CXXC domain in porcine zygotes. In the subcellular localization examination, we confirmed exclusive localization of isolated *TET3* CXXC domain

into the (pro)nuclei of zygotes and 2-cell embryos, suggesting that CXXC domain lead to nuclear localization of *pTET3L* isoform in early porcine embryos. Furthermore, the identification of putative NLS within CXXC of *pTET3L* supports the role of CXXC domain in nuclear localization of *pTET3L*. In Xenopus study, CXXC-deleted *Tet3* mutant could be localized into the nucleus of somatic cells (186). However, the nuclear localization of CXXC-deleted *Tet3* mutant seems to be enabled by using an alternative NLS located at C-terminus, which is responsible for the nuclear localization of *TET3* short form lacking CXXC domain (206, 207). Meanwhile, *TET3* has been known to preferentially demethylate the paternal pronucleus of zygotes because catalytic activity of *TET3* is protected by *STELLA* in the maternal pronucleus (71, 77). Although the activity of *TET3* is limited or vague in the maternal pronucleus, recent studies argued that the maternal genome is demethylated by *TET3* to some extent (153, 154). In this study, injected CXXC domains were detected in both pronuclei in zygotes, even though parental origin of the pronucleus could not be distinguished. Our data support the notion of *TET3*-mediated active demethylation in both parental pronuclei in zygotes (153, 154).

The *TET3*-driven active demethylation in the paternal pronucleus has been demonstrated by an increase in the 5mC level in *Tet3*-deficient mouse zygotes (63, 67, 153). The *TET3*-mediated 5mC oxidation is also reported in preimplantation embryos of other species, such as porcine (209) and bovine (121), indicating that the demethylation process is conserved across species. In the studies mentioned above, overall *TET3* was ablated, regardless of isoforms, by targeting conserved regions between the isoforms to observe the effect of overall *TET3* deficiency on 5mC oxidation. However, the presence of two isoforms with different domain architecture in mouse oocytes (189) raises a question about distinct roles of the isoforms in the demethylation process. Indeed, *mTet3FL* isoform containing CXXC domain was less effective in

5mC oxidation compared to other isoforms lacking CXXC domain in dot blot analysis and luciferase reporter assays (189). The predominant expression of *pTET3L* isoform containing CXXC domain in porcine oocytes (see Chapter 3) implies an important role of CXXC domain in the demethylation process occurring in early preimplantation embryos. In ICC assay, the global 5mC level was significantly decreased in porcine zygotes overexpressing CXXC domain, indicating that CXXC domain facilitated the demethylation process. On the other hand, the enhanced demethylation by CXXC overexpression was not detected at 2-cell stage porcine embryos. A plausible explanation for this result is that 5mC oxidation is mostly achieved at zygote stage by *pTET3L* isoform via DNA binding of CXXC, and then oxidation derivatives, such as 5hmC, 5fC, and 5caC, are removed in a DNA replication-dependent manner from 2-cell stage porcine embryos, as suggested by a mouse study (210). Further studies on dynamics of the isoform-specific expression of *TET3* and the level of 5mC derivatives during the preimplantation development of porcine embryos will be helpful to validate the speculation.

Consistent with the ICC results, the methylation level of the repetitive element (PRE-1) in 2-cell embryos was not changed by CXXC overexpression in bisulfite sequencing analysis. This result is in agreeance with a previous mouse study where the absence of *Tet3* had a limited impact on the methylation level of repetitive elements (211). However, CXXC overexpression induced decreases in the methylation level of *NANOG* promoter region in 2-cell embryos. This finding is an opposing effect of the outcomes from *Tet3* depletion demonstrated in mouse study, where *Tet3* deficiency interfered with demethylation of *Nanog* and *Oct4* in zygotes (67). The hypermethylation of *Oct4* promoter by *Tet3* deficiency resulted in a delayed activation of *Oct4* in mouse embryos (67). Similarly, our previous study demonstrates that *TET3* knockdown leads to a decreased *NANOG* expression in porcine blastocysts (209). Contrary to the effect of *TET3*

knockdown on *NANOG* expression, the CXXC overexpression in zygotes resulted in an increased *NANOG* expression in blastocysts. The enhanced demethylation on *NANOG* promoter and the hyperactivation of *NANOG* in blastocysts after *TET3* CXXC overexpression suggests that the DNA binding of CXXC domain is important for *TET3* action on the reprogramming of *NANOG* gene in porcine embryos. The findings lead us to a model in which *pTET3L* mediates DNA demethylation utilizing CXXC domain critical for *NANOG* expression in porcine embryos (Fig. 4-7). Specifically, we hypothesize that *pTET3L* is not involved in the demethylation of repetitive elements because *TET3* CXXC domain lacks binding affinity towards the regions. The CXXC overexpression facilitates demethylation of *NANOG* by recruiting more *pTET3L* to the CXXC binding sites, thus upregulating *NANOG* expression in blastocysts.

Meanwhile, along with *NANOG*, expression of *ESRRB* was also increased in blastocysts by CXXC overexpression. The expression pattern of *ESRRB* was also aligned with *NANOG* in *TET1* KO or DMOG-treated blastocysts; expression of *NANOG* and *ESRRB* was increased in *TET1* KO blastocysts, while their expression was decreased in DMOG-treated blastocysts (see Chapter 2). This result reinforces the notion that transcription of *ESRRB* is regulated by *NANOG* in porcine embryos (see Chapter 2) similar to the pluripotency gene networks of mouse pluripotent stem cells in which Nanog directly binds to *Esrrb* and stimulates *Esrrb* transcription (173).

In summary, we demonstrate that the CXXC domain of *pTET3L* isoform is closely involved in post-fertilization demethylation of porcine zygotes. Although the global methylation level of zygotes was decreased by overexpression of CXXC domain, it had a limited impact on repetitive elements, suggesting that *pTET3L* did not target these regions. On the other hand, overexpression of the CXXC domain resulted in enhanced demethylation of *NANOG* promoter in

early embryos and increased *NANOG* expression in blastocysts. Our findings suggest that DNA targeting of CXXC domain is critical for *TET3*-mediated global demethylation in zygotes and proper expression of *NANOG* in blastocysts.

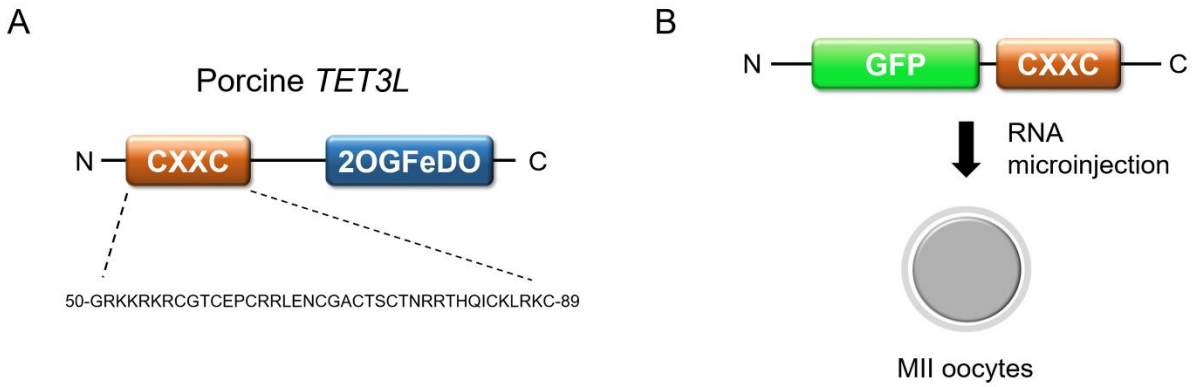


Fig. 4-1. Strategy of cloning porcine *TET3* CXXC domain and expression of GFP-CXXC via microinjection. (A) N-terminal porcine *TET3* sequences reflecting CXXC domain (40 amino acids) was identified through NCBI's conserved domain database. (B) CXXC domain was subcloned into N-terminal GFP fusion vector and mRNA was *in vitro* transcribed to inject into metaphase II (MII) oocytes.

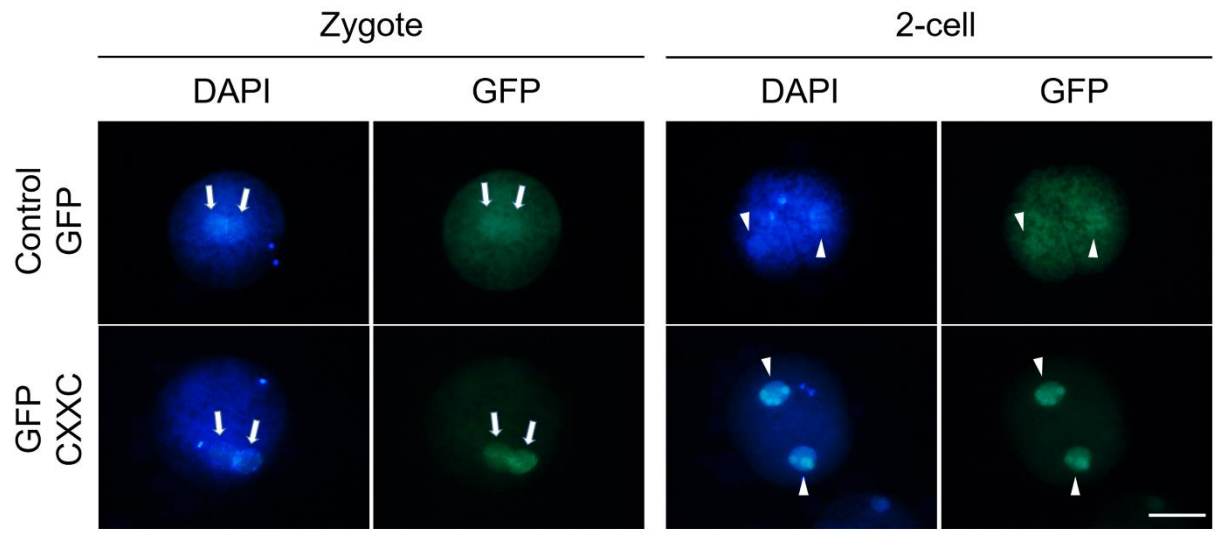


Fig. 4-2. Nuclear localization of *TET3* CXXC domain. The GFP-CXXC domain was exclusively localized in pronuclei (zygotes) and nuclei (2-cell embryos) whereas expression of control GFP was detected in both (pro)nuclei and cytosol. Arrows and arrow heads indicate pronuclei and nuclei, respectively. Scale bar = 100μm.

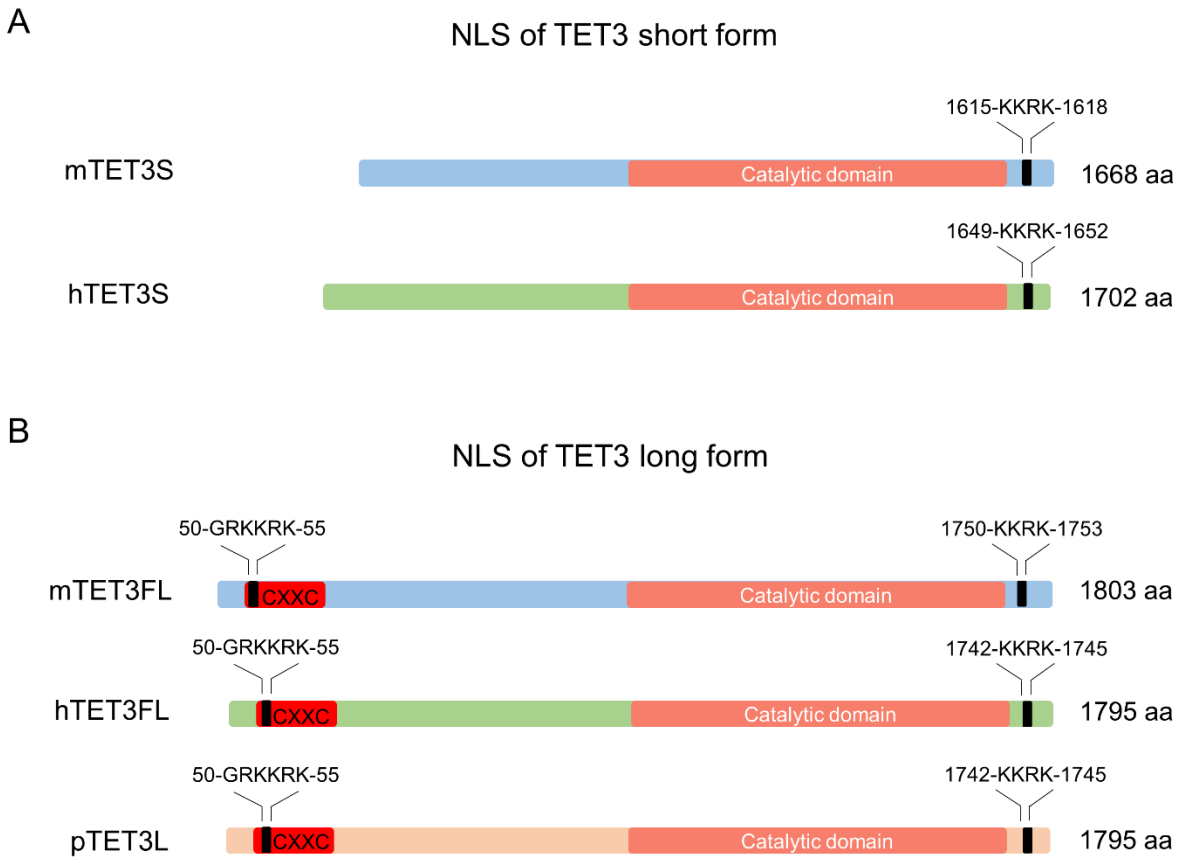


Fig. 4-3. Nuclear localization sequences (NLS) of TET3 isoforms. (A) NLS located in C-terminus of mouse and human TET3 short forms identified in previous studies. (B) TET3 long form contains the C-terminal NLS as well as a potential N-terminal NLS within CXXC domain. mTET3S, mouse TET short form; hTET3S, human TET3 short form; mTET3FL, mouse TET3 full length; hTET3FL, human TET3 full length; and pTET3L, porcine TET3 long form.

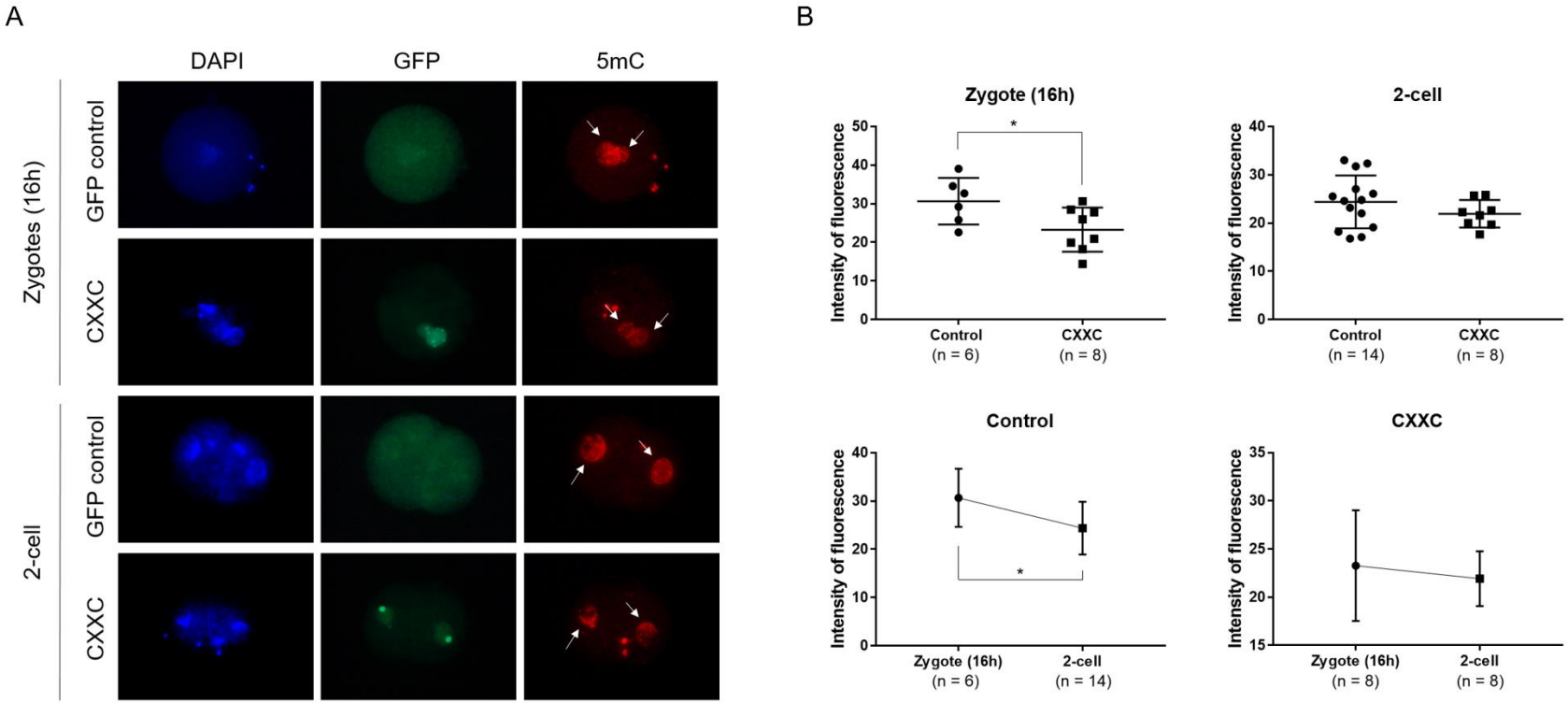


Fig. 4-4. Changes of 5mC levels in embryos by CXXC overexpression. (A) ICC assay with 5mC antibody showed decrease of 5mC level by CXXC overexpression in zygotes. Arrows indicate pronuclei or nuclei of embryos. (B) Intensity of 5mC fluorescence was quantified using the Image J software. 5mC level of CXXC-injected zygotes was lower than that of control zygotes, while there were

no differences in 5mC levels between CXXC-injected and control embryos at 2-cell stage. In control embryo, 5mC level was decreased following embryo cleavage, however, the low 5mC level was maintained during the transition from zygotes to 2-cell embryos.

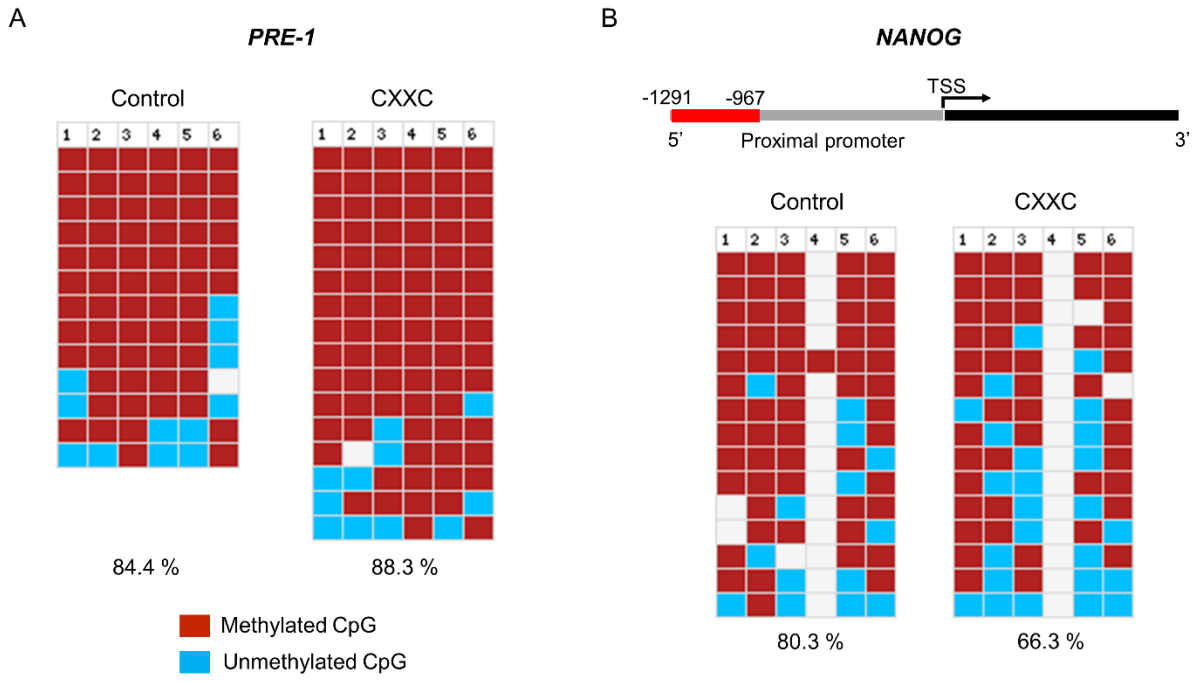


Fig. 4-5 Bisulfite sequencing analysis of DNA methylation level in 2-cell embryos. (A) Methylation level of *PRE-1*, the repetitive elements, was not altered by CXXC overexpression. (B) Methylation level of *NANOG* promoter region was reduced in embryos injected with CXXC compared to that in control embryos. Red color in the diagram of *NANOG* gene indicates the promoter region that methylation level was measured.

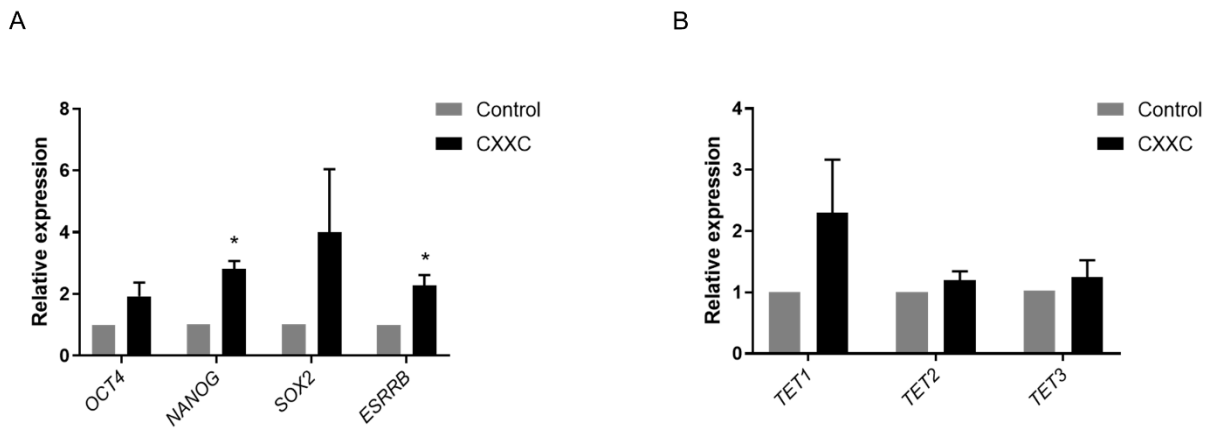


Fig. 4-6. Transcript abundance of pluripotency related genes and TET family genes in blastocysts derived from CXXC-injected zygotes. The expression level of *NANOG* and *ESRRB* was increased in CXXC-injected blastocysts compared to that of control blastocysts. No statistical difference in expression of TET family genes was observed between CXXC-injected blastocysts and control blastocysts. *GAPDH* was used as a normalizer. Results are shown as mean ± SEM (n = 3). Statistical significance is indicated by asterisks presenting P-values (* $P < 0.05$).

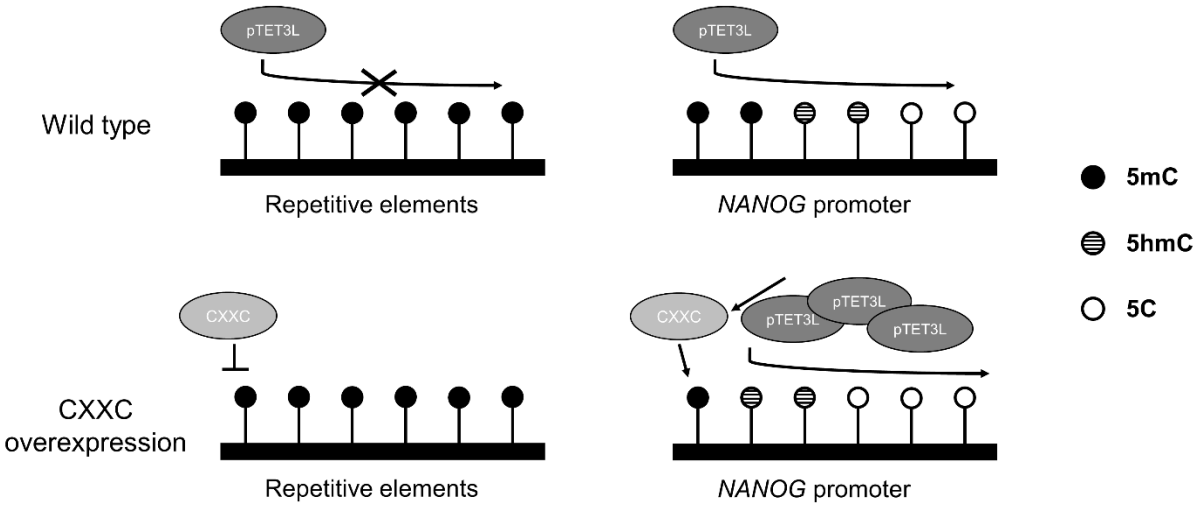


Fig. 4-7. Graphical summary and proposed mechanism of *pTET3L* actions on DNA demethylation in porcine embryos under the CXXC overexpression.

Table 4-1. Primers used for RT-qPCR analysis.

Primers	Sequence (5' → 3')
*GAPDH*_forward	ATGACATCAAGAAGGTGGTGAAGC
*GAPDH*_reverse	CCAGCATCAAAAGTGGAAGAGTGA
*OCT4*_forward	TTTGGGAAGGTGTTCAGCCAAACG
*OCT4*_reverse	TCGGTTCTCGATACTTGTCCGCTT
*NANOG*_forward	AGGACAGCCCTGATTCTTCCACAA
*NANOG*_reverse	AAAGTTCTTGCATCTGCTGGAGGC
*SOX2*_forward	TGTCGGAGACGGAGAAGCG
*SOX2*_reverse	CGGGGCCGGTATTTATAATCC
*ESRRB*_forward	CTGCAAGGCCTTCTTCAA
*ESRRB*_reverse	CGTTTGGTGATCTCACACTC
*TET1*_forward	TGTCGGCTTGGCAAGAAAGA
*TET1*_reverse	AGACCACTGTGCTGCCATTA
*TET2*_forward	GTGAGATCACTCACCCATCGCATA
*TET2*_reverse	TACTGGCACTATCAGCATCACAGG
*TET3*_forward	TCTTCCGTCGTTCAGCTACTACAG
*TET3*_reverse	GTGGAGGTCTGGCTTCTTAAA

Table 4-2. Development of embryos injected with *TET3* CXXC. Frequency of blastocyst formation at day 7 decreased in embryos injected with CXXC compared to that in IVF control embryos, but did not differ from that in injection control embryos. Different letters indicate statistical difference.

	Number of blastocysts at day 7 (%)	Number of oocytes used
IVF control	17 (20.5)[a]	83
Injection control (water)	22 (12.3)[ab]	179
CXXC injection	35 (9.6)[b]	364

Chapter 5. Overexpression of methyl-binding domain protein 1 (MBD1) disrupts preimplantation embryo development in pigs

Abstract

The conversion of 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC) by TET enzymes initiates DNA demethylation following fertilization in mammals. The 5hmC also serves as a stable epigenetic marker that regulates the level of transcription. The presence of CXXC domain in TET1 and TET3 indicates that direct interaction with the genome may be critical for their function; however, the binding requirement of TET family for their function still remains elusive. Methyl-CpG-binding domain proteins (MBDs) have the ability to bind 5mC and play a critical role in mediating DNA methylation, histone modification, and act as transcriptional repressors. A recent study demonstrated the MBD proteins can inhibit the function of TET1 by occupying TET1 target loci in mouse cells. The conversion of 5mC to 5hmC by TET family can keep genes in an activated state because 5hmC is not typically recognized by MBD. This competitive action between TET and MBD family may be involved in maintaining a proper level of DNA methylation during embryogenesis. The aim of this study was to clone the full coding sequence (CDS) of maternal porcine MBD members (*MBD1*, *MBD3*, and *MBD4*) from MII stage oocytes to study the impact of MBD family during porcine embryo development. Cloning of porcine *MBD1*, *MBD3*, and *MBD4* revealed that the genes are highly conserved among different species, indicating that the genes could be involved in maintaining the proper level of DNA methylation during porcine embryo development. When *MBD1* mRNA was overexpressed in embryos, the frequency of blastocyst formation was significantly lower compared to that from *IVF* control and water injected groups. This result indicates that the overexpression of *MBD1*

may negatively affect early embryo development in pigs by disrupting proper maintenance of DNA methylation levels. Specifically, we speculate that proper competitive binding of MBD1 and TET3 to 5mC is critical for normal development in pigs.

Introduction

In mammals, the fifth carbon of cytosine residue at CpG dinucleotides can be modified to 5-methylcytosine (5mC) by DNA methyltransferases (DNMTs) (212). Generally, vertebrate DNA carry low number of CpG dinucleotides, except for some regions with high frequencies of CpG sites, termed CpG islands (213, 214). CpG islands are commonly found in promoter regions and related to gene expression; methylated CpG islands in promoter or enhancer regions are typically associated with gene repression (81, 215). Non-CpG methylation marks have also been recently identified in oocytes, embryonic stem cells (ESCs), and neurons (19-22); however, the epigenetic role of such methylation still remains elusive. Although technological developments allow for the genome-scale profiling of CpG distribution in cells, interpretation of DNA methylation patterns and their role in epigenetic plasticity have not been fully identified.

DNA methylation patterns are modulated and interpreted by groups of proteins: DNA methylation 'writers', 'readers', and 'editors'. While methylation marks are added to DNA by DNMTs, the methylation 'writers', 5mC is specifically recognized and translated by methylation 'readers'. Methyl-CpG binding domain (MBD) proteins represent one family of methylation 'readers' (216, 217). Five members of MBD protein family have been well characterized in mammals, including MBD1, MBD2, MBD3, MBD4, and MeCP2. These proteins often bind to methylated CpG dinucleotides with high affinity (218, 219). In addition, MBD proteins play an

important role as a transcriptional regulator by coordinating crosstalk between DNA methylation and histone modification (220, 221).

Ten-elven translocation (TET) proteins are recently defined DNA methylation 'editors' and carry out hydroxylation of 5mC to generate 5-hydroxymehtylcytosine (5hmC) (43). Initially, the 5hmC bases are known as an intermediate during the DNA demethylation process because 5hmC can be further oxidized by TET proteins to 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC), and ultimately demethylated through the thymine DNA glycosylase (TDG)-dependent base excision repair (222). Interestingly, recent studies indicate that 5hmC serves as a stable epigenetic marker involved in transcriptional regulation (167), cellular reprogramming (223), and stem cell pluripotency (84). The TET proteins can regulate gene expression by binding to different transcription regulation sites. Specifically, direct binding of TET proteins to the genome can regulate gene expression by controlling binding activity of specific transcription factors at various promoter regions (83, 167). TET1 can also act as a gene repressor through association with a transcriptional repressor complex (85). Furthermore, TET proteins can affect gene expression by regulating the level of histone O-GlcNAcylation (224, 225). These results suggest that the role of TET family members is not limited to a key factor in DNA demethylation as a hydroxylase, but can orchestrate the fine tuning of transcription.

Genomic distribution of 5hmC shows a unique pattern in different cells, tissues, and development stages (226), implying that 5hmC formation by TET proteins is involved in the identity of cells and tissues. Since typically, MBD proteins repress gene expression by specific binding to 5mC (227), conversion of 5mC to 5hmC by TET proteins could release the MBD-mediated gene repression. Meanwhile, since MBD and TET proteins both have an affinity toward 5mC, competitive binding could occur. Indeed, a recent study demonstrated that binding

of MBD proteins to DNA segments carrying 5mC protects the region from TET1-mediated oxidation in mouse cells (228). Fertilized mammalian embryos undergo genome-wide epigenetic reprogramming; 5mC oxidation by TET proteins is a critical step of the reprogramming process for embryo development and pluripotency acquisition (229). In the previous chapters we confirmed that TET family play important roles in methylation reprogramming and regulation of pluripotency in porcine preimplantation embryos. We hypothesize that exogenous expression of MBD proteins in embryos negatively impact embryo development because the overexpression may interfere with the activity of TET proteins if both TET and MBD proteins possess overlapping 5mC target sites. In this study, we cloned *MBD* sequences from porcine oocytes and induced overexpression of *MBD1* in embryos. Our data show that overexpression of MBD1 causes a developmental stall to preimplantation porcine embryos.

Materials and Methods

In vitro maturation of porcine oocytes

Sow ovaries were obtained from a local abattoir and transported to the laboratory. Oocytes were aspirated from the follicles with a needle attached to a syringe. Oocytes were washed with TL-Hepes and *in vitro* matured in maturation medium (medium 199 supplemented with 3.05 mM glucose, 0.91 mM sodium pyruvate, 0.57 mM cysteine, 10 ng/ml epidermal growth factor (EGF), 0.5 μg/ml luteinizing hormone (LH), 0.5 μg/ml follicle stimulating hormone (FSH), 10 ng/ml gentamicin, and 0.1% polyvinyl alcohol (PVA); pH 7.4) for 30–32 h at 38.5 ℃, 5% CO_2, and 100% humidity. After maturation, cumulus cells were removed by vortexing in denuding medium (0.3 M mannitol, 0.001% BSA, 0.03% hyaluronidase, 5% TL-

Hepes medium in distilled water; pH 7.4). The cumulus-free oocytes were used for MBD cloning or microinjection.

Cloning of porcine MBDs

To clone the CDS of porcine MBD (MBD1, MBD3, and MBD4), primers were designed from porcine ESTs that matched predicted porcine MBD sequences (Genbank accession numbers: FS668599, EW639347, HX233243, DN124476, DN111806, and DN122093). Primer information is presented in Table 5-1. mRNAs were isolated from *in vitro* matured pig oocytes (n=30) using Dynabeads mRNA Direct Kit (Thermo Fisher Scientific), and subsequently cDNAs were synthesized using Maxima H Minus First Strand cDNA Synthesis Kit (Thermo Fisher Scientific). Coding sequences of porcine MBDs were amplified from the cDNAs. Phusion High-Fidelity DNA Polymerase (Thermo Fisher Scientific) was used for the PCR. PCR conditions follow; initial denature at 98℃ for 30s, denature at 98℃ for 15s, annealing at 60℃ for 30s, extension at 72℃ for 90s (*MBD1* and *MBD4*) or 60s (*MBD3*) for 35 cycles. Potential porcine MBD PCR products were cloned into TOPO vector (Thermo Fisher Scientific) and the Sanger sequencing was performed to identify the amplicons. The porcine MBD sequences (both nucleotide and predicted peptide) were then compared against human and mouse MBD for comparative analysis.

Microinjection of *MBD1* mRNA

To overexpress *MBD1* in preimplantation porcine embryos, *MBD1* mRNA was synthesized from the cloned *MBD1* CDS using the mMESSAGE mMACHINE T7 Ultra Kit

(Thermo Fisher Scientific). Then, the *MBD1* mRNA (100ng/µl) was introduced into oocytes that were *in vitro* matured for 36 h in manipulation medium (medium 199 supplemented with 0.6mM $NaHCO_3$, 2.9mM Hepes, 30mM NaCl, 10ng/ml gentamicin, and 3mg/ml BSA; pH 7.4) on a heated stage of a Nikon inverted microscope. The *MBD1* mRNA injected oocytes were placed in maturation medium and incubated for 2 h. After the incubation, live oocytes were fertilized and then cultured in PZM3 media (158) for 7 days. Blastocyst formation of the embryos were recorded. Non-injected and water injected embryos served as a control.

Results

Molecular cloning of porcine *MBDs*

In mammals, five MBD proteins are known to possess a homologous methyl-CpG binding domain; MeCP2, MBD1, MBD2, MBD3, and MBD4 (220). Previous deep sequencing results showed expression of the five major MBD proteins in porcine oocytes and blastocysts (230, 231). Based on the deep sequencing results, we cloned the full CDS of three porcine *MBDs*, *MBD1*, *MBD3*, and *MBD4* from mature oocytes to investigate whether their overexpression would have an impact on embryo development. *MBD2* and *MeCP2* were excluded in this study because a mutation in *MBD2* did not have an impact on development in mice (232), and *MeCP2* can bind to both 5mC and 5hmC at high affinity and activate transcription in the brain (233), thus reducing its possibility as the main MBD recognizing 5mC during post-fertilization reprogramming process. The CDS length of *MBD1*, *MBD3*, and *MBD4* were 1812 bp, 780 bp, and 1704 bp, respectively (Fig. 5-1A, B, and C). Sequencing results revealed that the sequences of porcine *MBDs* were conserved among various species (Table 5-2). For instance, when the amino acid sequence of porcine *MBD1* was compared to that of human,

the porcine *MBD1* had 84% identity with human *MBD1*; the identity was only 56% compared to the mouse. On the other hand, the amino acid sequence of porcine *MBD3* showed high similarities with both human and mouse *MBD3*, 94% and 96% of sequence identity, respectively. The porcine *MBD4* had a higher identity with human *MBD4* compared to mouse: 76% and 66%, respectively.

Overexpression of *MBD1* causes developmental stall to preimplantation porcine embryos

A recent study reported that binding of MBD1 to DNA protects 5mC from TET-mediated oxidation in somatic cells (228). To examine the effect of exogenous expression of MBD1 on preimplantation embryo development, porcine mature oocytes were injected with *MBD1* mRNA (100ng/µl) and fertilized. When *MBD1* mRNA was overexpressed in embryos, the frequency of blastocyst formation was only 2.4%, whereas *IVF* control and water injected groups showed a significantly higher blastocyst formation frequency: 33.3% and 22.8%, respectively (Table 5-3). This result indicates that overexpression of *MBD1* may negatively affect early embryo development in pigs by disrupting the proper maintenance of DNA methylation level.

Discussion

The MBD proteins provide a critical link between DNA methylation and the establishment of transcriptional repression (234) and the proper expression of MBD is important for normal development. For example, deletion of the *MBD3* gene causes post-implantation embryo lethality in mice (232) and *MBD1* KO mouse shows abnormal neurogenesis (235). Although the importance of MBD proteins in transcriptional regulation and embryo development

is apparent, mRNA sequences of MBDs have not been annotated in porcine species. In this study we cloned the full coding sequences of porcine *MBDs* (*MBD1*, *MBD3*, and *MBD4*) from mature oocytes and confirmed that MBD proteins are conserved across mouse, human, and porcine species. Especially, porcine MBDs showed higher homology to human than mouse at the amino acid level. This high identity suggests that porcine and human MBDs could possess similar functions. Considering the similarities in MBD sequences between porcine and human, porcine embryos could be a more suitable model than mouse for understanding molecular and biological functions of human MBDs in developing embryos. MBD proteins typically recognize methylated CpGs using the methyl-CpG binding domain (218). Through NCBI's conserved domain database, we confirmed that all porcine MBDs also contain methyl-CpG binding domains. Interestingly, in addition to the methyl-CpG binding domain, the full length MBD1 isoform also contains three CXXC domains that bind to unmethylated CpGs in human (236, 237). We also found that the cloned porcine *MBD1* also possesses the CXXC domains (Fig. 5-2). Although the CXXC domains are characterized by two conserved cysteine-rich clusters, they can be segregated into three subtypes according to the sequence similarities (238). Among the three subtypes, only type 1 CXXC domain, such as the CXXC domains of DNMT1 and MLL, can specifically recognize unmethylated CpGs. Type 2 and 3 domains are known to have less or no specificity for unmethylated CpGs. One of the three CXXC domains found in porcine *MBD1* can be categorized into type 1 and other two domains correspond to type 2. This structural information of *MBD1* suggests that porcine MBD1 may be able to recognize both methylated and unmethylated CpGs using the methyl-CpG binding domain and CXXC domain.

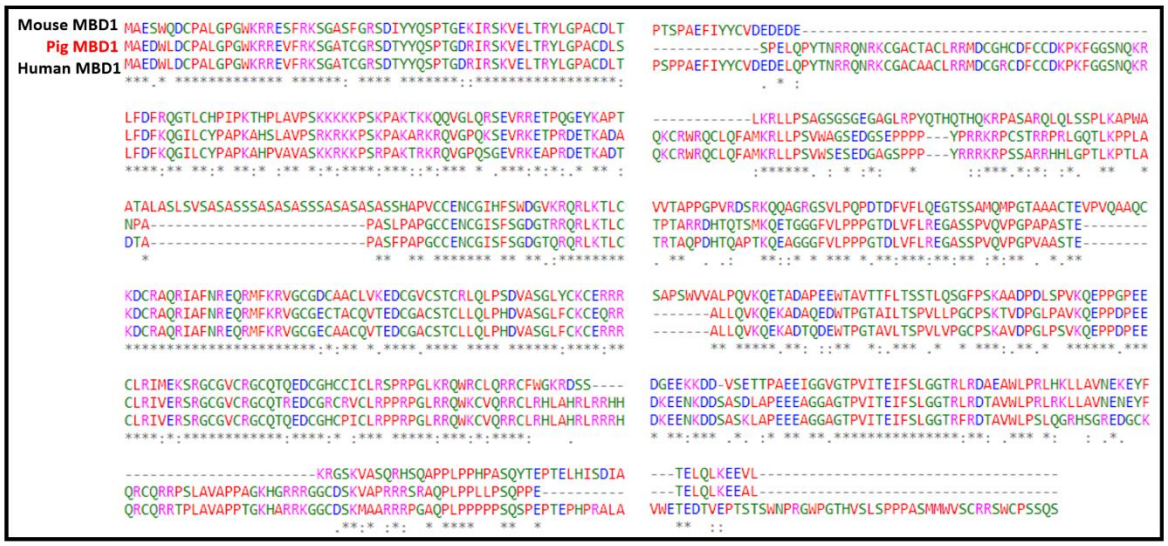
In this study, overexpression of *MBD1* was induced in porcine oocytes and it led to failed preimplantation development, suggesting that improper maintenance of DNA methylation levels

by the MBD1 overexpression may negatively affect embryo development. A recent study demonstrated that binding of MBD2 and MeCP2 to DNA protects 5mC from TET1-mediated oxidation by directly restricting access of TET1 to DNA (228). However, the stalled porcine embryo development by MBD1 overexpression did not seem to be caused by interference from TET1 activity because depletion of *TET1* does not impact preimplantation embryo development (58) (see Chapter 2) and MBD1 enhances TET1-mediated 5hmC formation rather than blocking the 5mC oxidation in mouse cells (239). Instead, overexpressing MBD1 might negatively affect the 5mC oxidation activity of TET3. In the previous chapters, we detected predominant expression of *pTET3L* isoform containing CXXC domain in porcine mature oocytes (see Chapter 3) and provided evidence suggesting that the *pTET3L* isoform mediates the initial demethylation in zygotes through the binding of CXXC domain to the genome (see Chapter 4). The type 3 CXXC domain found in TET3 protein has low specificity for unmethylated CpGs and differs from the three CXXC domains (type 1 and type 2) of MBD1 protein (238). Interestingly, a previous study showed that the type 1 CXXC domain of MBD1 localizes TET1 to heterochromatic DNA and enhances TET1-mediated 5hmC formation in somatic cells (239). Contrarily, MBD1 blocks the TET1-mediated 5hmC formation when the CXXC domain of MBD1 is removed. Although there are ambivalent impacts of MBD1 on TET1 activity, the CXXC domain of MBD1 does not facilitate the localization of TET3 to heterochromatic DNA (239), suggesting that MDB1 can only inhibit the oxidation activity of TET3. The previous findings and our data lead to a proposed model in which, competitive action between MBD1 and TET3 modulates proper methylation levels in preimplantation embryos (Fig. 5-3). In this model, the post-fertilization demethylation process is controlled by the binding of MBD1 and TET3 to methylated CpGs under physiological conditions. However, overexpression of MBD1 interferes

with TET3 accessibility by occupying CpGs, and thus leads to improper maintenance of methylation levels.

In summary, we cloned MBD proteins (MBD1, MBD3, and MBD4) from mature porcine oocytes and identified that overexpression of *MBD1* disrupts preimplantation development of porcine embryos. Further studies on the changes of methylation levels (5mC and 5hmC) by exogenous expression or depletion of MBD proteins can provide further insight into the interactive roles of MBD and TET proteins in the regulation of methylation levels in embryos. Based on previous studies and our findings, we propose that competitive binding of MBD1 and TET3 to methylated CpGs orchestrates fine tuning of transcription by maintaining proper methylation levels in preimplantation embryos.

A



B

C

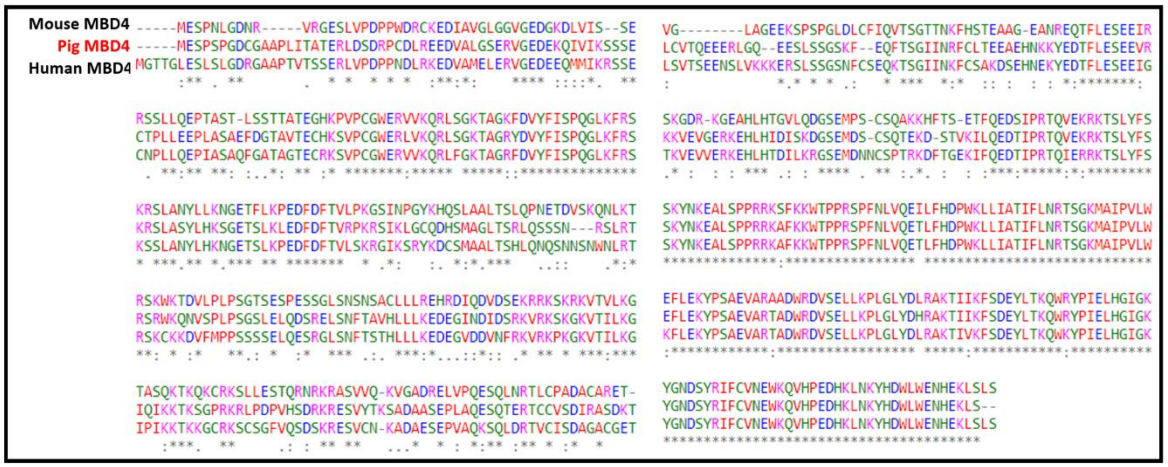


Fig. 5-1. Amino acid sequence of the cloned porcine MBD proteins and homology with other species. (A) Porcine MBD1 and homology with mouse and human MBD1 (B) Porcine MBD3

and homology with mouse and human MBD3. (C) Porcine MBD4 and homology with mouse and human MBD4.

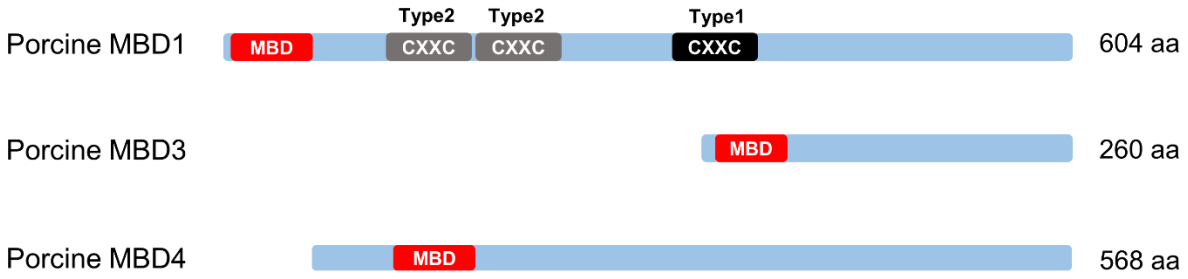


Fig. 5-2. Domain structure of the cloned porcine MBD proteins. All cloned MBD proteins (MBD1, MBD3, and MBD4) contain a methyl-CpG binding domain (MBD) at N-terminus (red box). MBD1 contains three CXXC domains; one type 1 CXXC domain (black box) and two type 2 CXXC domains (gray box).

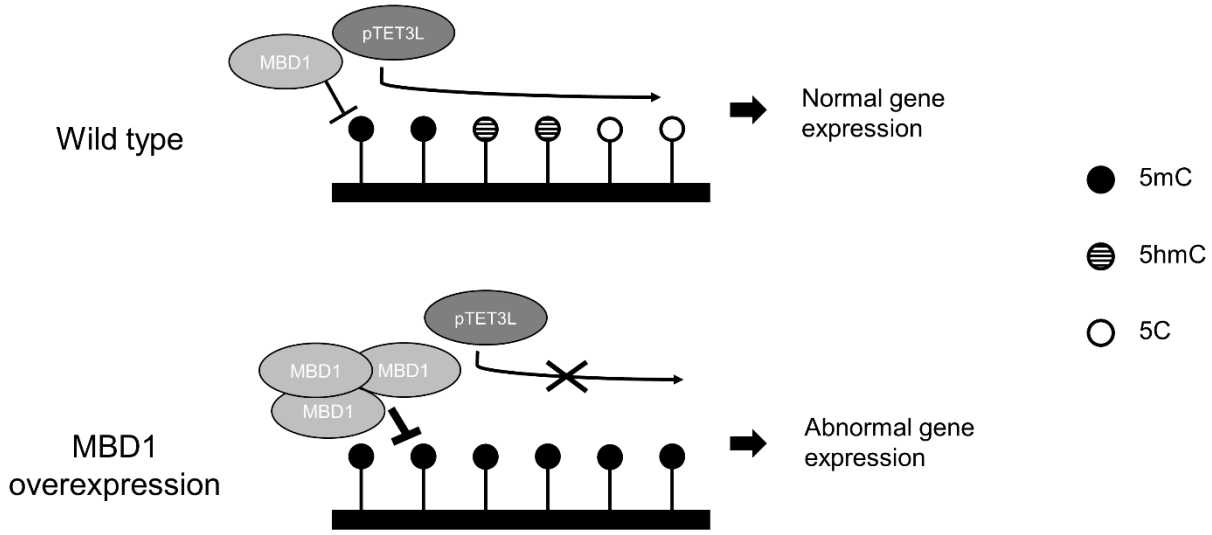


Fig. 5-3. Proposed mechanism for regulation of DNA methylation level in porcine embryos by competitive actions between MBD1 and pTET3L.

Table 5-1. Primers used for the RT-PCR of porcine MBD proteins.

Primers	Sequences (5' → 3')
MBD1_forward	GAGTCAAAACCGGGACCCTCGTCA
MBD1_reverse	CCGCCTACTTCAGATTTTGTATGCGGTACG
MBD3_forward	AACAGCCGCGGCAAGTGG
MBD3_reverse	GTGTGGCAAGGGCTGAGTCCG
MBD4_forward	CGTTGCTACAGCTGGATCTCGC
MBD4_reverse	TTAAGATAGACTTAACTTTTCATGATTTTCCCAAAGCCA

Table 5-2. Amino acid identity of cloned porcine MBDs against human and mouse.

Porcine MBDs	Homology against human MBDs	Homology against mouse MBDs
MBD1	84%	56%
MBD3	94%	96%
MBD4	76%	66%

Table 5-3. Overexpression of MBD1 causes developmental stall to preimplantation porcine embryos. Different letters indicate a statistical difference ($P < 0.01$).

Embryo groups	Day 7 blastocysts, %	Total number of embryos
IVF control	33.3[a]	75
Injection control (water)	22.8[a]	202
MBD1 injection	2.4[b]	253

Chapter 6. Conclusions and perspectives

The overall aim of this dissertation was to improve our understanding of mechanisms underlying the control of epigenetic modifications during porcine embryo development. Since the discovery of converting 5mC to 5hmC by the TET family proteins, efforts have been made to elucidate the roles of TET proteins and 5hmC in mammalian embryos and ESCs. Our studies demonstrate that TET proteins are important epigenetic modulators involved in the dynamic reprogramming of DNA methylation and acquisition of pluripotency in porcine embryos. Our studies support findings in previous research using mouse models and therefore, highlights conserved actions of TET proteins, and contributes to defining mechanistic actions of TET family-mediated reprogramming through novel findings in porcine embryos. Investigating epigenetic events in human embryos have a restriction due to the scarcity of available embryos and ethical issues. Considering similarities between porcine and human embryos in physiology, genome information, and molecular events, our studies can provide valuable information for understanding epigenetic events in human embryos.

The first objective of this dissertation was to investigate the role of *TET* family in regulating pluripotency in porcine preimplantation embryos. We demonstrated that *TET* family modulates *NANOG* expression by controlling promoter methylation levels in porcine embryos. *TET1* KO blastocysts generated by the CRISPR/Cas9 system unexpectedly displayed upregulated *NANOG* expression without any change in DNA methylation levels in the promoter region. Interestingly, the amount of *TET3* transcript was increased approximately 3-fold in *TET1* KO blastocysts, suggesting a compensatory increase of *TET3* transcription may be involved in the upregulation of *NANOG* expression. Indeed, overall inhibition of *TET* proteins by *TET* inhibitor downregulated expression of pluripotency genes, including *NANOG,* and resulted in

increased methylation levels of the *NANOG* promoter. Although functional redundancy between *TET1* and *TET3* has been reported, the specific molecular functions overlapping between the two proteins are not clear. For example, locus-specific activity of TET proteins for 5hmC formation and transcriptional regulation has not been characterized, although genome-wide impacts of the loss of TET proteins on DNA methylation patterns (5mC and 5hmC) and pluripotency have been reported (55, 85, 87, 167). To define distinctive roles of *TET1* and *TET3* in regulation of pluripotency genes, further studies, such as profiling of methylation patterns and gene expression in *TET1*/*2*/*3*-null blastocysts with the exogenous expression of *TET3,* are necessary. In addition, analysis of the global demethylation process occurring in early stage embryos where the expression of *TET3* is replaced with *TET1* will be beneficial to clarify the functional redundancy between the two proteins. The bisulfite sequencing technique used in this study for evaluating DNA methylation cannot distinguish 5mC and 5hmC. Taking a simultaneous measurement of locus-specific 5mC and 5hmC levels through advanced technologies, such as single cell 5hmC sequencing, will provide more accurate information about the oxidation activity of *TET* family in embryos. It would also be interesting to determine whether *TET1* and *TET3* have different 5mC oxidation activities by profiling further oxidized forms of 5mC, such as 5fC and 5caC, in embryos expressing mutated *TET1* or *TET3*.

Impaired ICM formation followed by the inhibition of TET proteins implies the involvement of *TET* family in early lineage specification in porcine blastocysts. The segregation of ICM and trophectoderm lineages follows a different pattern in porcine embryos compared to that in other species (174, 240). The signaling pathways for the formation of ICM in porcine embryos is also different from that in other species (241). The unique features of porcine ICM formation confers resistance to the establishment of stable porcine ESC lines. Impaired ICM

formation by TET family inhibition implies that genes targeted by TET proteins are involved in the regulation of early lineage segregation of porcine blastocysts. Investigations to identify target genes of TET proteins during lineage segregation can contribute to the elucidation of mechanisms regulating ICM formation and the establishment of authentic porcine pluripotent stem cells. In future studies, the effect of enhanced expression of *TET* family on the derivation or maintenance of porcine pluripotent stem cells can be a topic of interest. For example, treatment of TET protein activator, such as Vitamin C (242, 243), to ESCs or iPSCs during derivation of the cells will be a simple approach to verify the relationship between TET family and porcine pluripotency.

Distinctive domain structures of TET isoforms could reveal molecular functions of TET proteins (187). The second study focused on identifying *TET3* isoforms present in oocytes to determine the major isoform of *TET3* involved in the post-fertilization demethylation process. We demonstrated that *pTET3L* isoform containing N-terminal CXXC domain is predominantly expressed in MII stage porcine oocytes. Among the three different *TET3* isoforms identified from GV or MII stage cumulus cells, only *pTET3L* isoform expression was detected in MII oocytes. The RT-qPCR analysis confirmed the predominant expression of *pTET3L* isoform in MII oocytes, suggesting the potential involvement of CXXC domain of the *pTET3L* isoform in DNA targeting for genome-wide DNA demethylation. Although the role of *TET3* in 5mC hydroxylation in mammalian zygotes is conserved (67, 121, 209), our data is not consistent with a previous mouse study (189), which argues that oocyte-specific *Tet3* isoform lacking CXXC domain is predicted to play a major role in DNA demethylation. Similar catalytic activities for converting 5mC to 5hmC among different *TET3* isoforms containing or lacking CXXC domain in non-embryonic cells (186, 188) supports the notion that the oocyte-specific *Tet3* isoform can

be involved in global 5mC demethylation in zygotes (189). However, this isoform could not be found in porcine oocytes through *in silico* analysis and has not been reported in other species.

Identification of major *TET3* isoforms in oocytes of other species, such as human and bovine, will provide clues to identify the necessity of DNA binding features of *TET3* in mammalian embryos. If the expression of oocyte-specific *TET3* isoform is not universal among mammalian oocytes, different DNA recognition properties of *TET3* isoforms may represent species-specific oocyte-mediated reprogramming. For example, because DNA methylation of paternal genome continues until 4-cell stage in porcine embryos, *pTET3L* isoform can gradually demethylate the porcine embryo genome through relatively low CpG binding activity of CXXC domain, whereas oocyte-specific *Tet3* isoform lacking CXXC domain can lead to complete genome-wide demethylation shortly after fertilization. Analysis of DNA methylation patterns in porcine embryos, where *pTET3L* isoform is replaced with the oocyte-specific isoform of mouse *Tet3*, will provide valuable information concerning different activities of *TET3* isoforms.

Based on the presence of *TET3* isoforms in oocytes, the role of *TET3* CXXC domain in maintaining DNA methylation during preimplantation development was evaluated in the following study. We found that CXXC domain of *TET3* is critical for post-fertilization demethylation of porcine embryos and proper expression of pluripotency related genes in blastocysts. Exogenous *TET3* CXXC protein was localized to the pronucleus of zygotes, indicating that CXXC domain may be responsible for nuclear localization of *TET3* in early stage embryos. Although the CXXC domain of *TET3* is localized in the nucleus, functionality of the NLS needs to be examined in further studies because *pTET3L* still contains a NLS sequence at its C-terminus, identical to the NLS of *TET3* short form (206, 207). The amino acid sequence of the NLS located at C-terminus is KKRK. The KKRK peptide is known to be responsible for nuclear

localization of CUL4B, FMIP, δB-CaM kinase, and DNA helicase Q1 (244-247). These proteins are transported into the nucleus by importin-α1, -α2, -α3, -α5, and -β1, implying that nuclear localization of *TET3* can also be meditated by importin-α and importin-β1. The sequence of the predicted NLS in the CXXC domain is similar to that present in the C-terminus, but not identical. Although the precise NLS peptide needs to be experimentally identified, the different NLSs implies that nuclear transport can be mediated by proteins other than importin-α and importin-β1. If so, the TET3 protein expression pattern in the nucleus can be different by the presence of CXXC domain. Investigations on the subcellular localization of *TET3* isoforms carrying mutated N-terminal and/or C-terminal NLSs will characterize role of the NLS for the function of pTET3L. In the experiment in which we induced overexpression of *TET3* CXXC, we detected promoted global DNA demethylation in zygotes and enhanced demethylation of the *NANOG* promoter region in 2-cell stage embryos. Interestingly, *NANOG* expression in blastocysts was increased by the CXXC overexpression in zygotes, suggesting that enhanced demethylation of the promoter leads to the upregulation of *NANOG* transcription. In this study, the impact of CXXC overexpression on the methylation levels of *NANOG* promoter was evaluated solely at the 2-cell stage. Profiling the methylation levels of this region at later stages, such as 4-cell, 8-cell, and blastocyst, will be beneficial to clarify the effect of CXXC overexpression on the programming of *NANOG* promoter for its transcription. In addition, isoform specific knockdown studies are necessary to further confirm the role of the CXXC domain in DNA methylation reprogramming and proper expression of pluripotency related genes in preimplantation embryos. Although the negative impact of *TET3* deficiency on DNA demethylation and expression of pluripotency related genes has been reported in previous studies (67, 209), the results were obtained by the elimination of overall *TET3*. Elimination of specific *TET3* isoforms by targeting

distinctive regions, such as CXXC domain, will clarify the functions of the *TET3* isoforms with different domain architectures in DNA targeting and methylation reprogramming.

The aim of the last study was to investigate the importance of MBD proteins for porcine embryo development and explore potential interaction between MBD proteins and TET family proteins during preimplantation development. We identified full coding sequences of porcine MBD proteins and showed that exogenous expression of *MBD1* stalls preimplantation development. Based on the results and previous literature, we speculate that the occupation of 5mC sites by MBD1 interferes with 5mC oxidation activity of TET3 protein. Unfortunately, detailed evidence for the interaction of TET proteins with MBD1 in the maintenance of DNA methylation have not been explored in this study. Further experiments are needed to confirm the competitive action between MBD1 and TET3. First, profiling of DNA methylation (5mC and 5hmC) patterns and gene expression in embryos overexpressing *MBD1* is necessary to confirm the cause of developmental failure. Second, the effect of *TET3* overexpression on DNA methylation, gene expression and embryo development should be examined; *TET3* overexpression should relieve the impact of *MBD1* overexpression on epigenetic modifications and embryo development if they bind to 5mC competitively. These experiments will verify the hypothesis that interactive actions of MBD1 and TET3 orchestrate the fine tuning of transcription through the maintenance of DNA methylation. In addition to *MBD1*, expression of other *MBD* proteins (*MBD3* and *MBD4*) were detected in porcine oocytes. Additional studies on the association of TET proteins with MBD3 or MBD4 in porcine embryos should be conducted. Interestingly, the cloned porcine MBD1 possessed a methyl-CpG binding domain as well as CXXC domains, indicating ambivalent DNA binding properties of MBD1. Indeed, 5hmC formation by *Tet1* was facilitated by interacting with CXXC domain of MBD1 in mouse cells

(239). It would be interesting to investigate the DNA binding property of MBD1 in mammalian embryos via mutations on either or both of the domains. Furthermore, such study will be beneficial to understand how DNA binding of MBD1 affects the 5mC oxidation activity of TET proteins in embryos. In addition to MBD proteins, other proteins involved in chromatin remodeling have been known to interact with TET proteins to regulate transcription. For example, TET1 represses transcription by recruiting the SIN3A co-repressor complex to promoter in mouse ESCs (164, 248), and TET proteins associate with OGT (O-linked β-D-N-acetylglucosamine (O-GlcNAc) transferase), which regulates gene expression by adding GlcNAc sugar to histone H2B and other chromatin modifiers (224, 248, 249). Further studies on the association between TET proteins and other chromatin remodeling complexes in embryos will improve our understanding of transcriptional regulation during preimplantation development.

Collectively, these studies emphasize that TET family-mediated epigenetic modifications are critical for proper gene expression and normal embryo development in pigs. Although pigs serve as importance agricultural resources and animal models for biomedical studies, little has been defined on the epigenetic dynamics of porcine embryos during preimplantation development compared to mouse. Our studies provide new insights into the epigenetic events occurring in porcine preimplantation embryos. Continued research in this area will unveil the mechanisms underlying epigenetic reprogramming and pluripotency in porcine embryos, and eventually contribute to improving the efficiency of pig production for agricultural and biomedical purposes.

References

1. Howell CY, Bestor TH, Ding F, Latham KE, Mertineit C, Trasler JM, et al. Genomic imprinting disrupted by a maternal effect mutation in the Dnmt1 gene. Cell. 2001;104(6):829-38.

2. Hansen RS, Wijmenga C, Luo P, Stanek AM, Canfield TK, Weemaes CM, et al. The DNMT3B DNA methyltransferase gene is mutated in the ICF immunodeficiency syndrome. Proc Natl Acad Sci U S A. 1999;96(25):14412-7.

3. Wan M, Zhao K, Lee SS, Francke U. MECP2 truncating mutations cause histone H4 hyperacetylation in Rett syndrome. Hum Mol Genet. 2001;10(10):1085-92.

4. Carlone DL, Skalnik DG. CpG binding protein is crucial for early embryonic development. Mol Cell Biol. 2001;21(22):7601-6.

5. Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. Nature genetics. 1999;23(2):185-8.

6. Oberle I, Rousseau F, Heitz D, Kretz C, Devys D, Hanauer A, et al. Instability of a 550-base pair DNA segment and abnormal methylation in fragile X syndrome. Science. 1991;252(5009):1097-102.

7. Spielman M, Vinkenoog R, Dickinson HG, Scott RJ. The epigenetic basis of gender in flowering plants and mammals. Trends Genet. 2001;17(12):705-11.

8. Reinhart B, Eljanne M, Chaillet JR. Shared role for differentially methylated domains of imprinted genes. Mol Cell Biol. 2002;22(7):2089-98.

9. Barlow DP, Bartolomei MS. Genomic imprinting in mammals. Cold Spring Harb Perspect Biol. 2014;6(2).

10. Falls JG, Pulford DJ, Wylie AA, Jirtle RL. Genomic imprinting: implications for human disease. Am J Pathol. 1999;154(3):635-47.

11. Cirio MC, Martel J, Mann M, Toppings M, Bartolomei M, Trasler J, et al. DNA methyltransferase 1o functions during preimplantation development to preclude a profound level of epigenetic variation. Dev Biol. 2008;324(1):139-50.

12. Toppings M, Castro C, Mills PH, Reinhart B, Schatten G, Ahrens ET, et al. Profound phenotypic variation among mice deficient in the maintenance of genomic imprints. Human reproduction (Oxford, England). 2008;23(4):807-18.

13. Kalish JM, Jiang C, Bartolomei MS. Epigenetics and imprinting in human disease. Int J Dev Biol. 2014;58(2-4):291-8.

14. Brookes E, Shi Y. Diverse epigenetic mechanisms of human disease. Annu Rev Genet. 2014;48:237-68.

15. Holliday R, Pugh JE. DNA modification mechanisms and gene activity during development. Science. 1975;187(4173):226-32.

16. Riggs AD. X inactivation, differentiation, and DNA methylation. Cytogenet Cell Genet. 1975;14(1):9-25.

17. Bird A. DNA methylation patterns and epigenetic memory. Genes Dev. 2002;16(1):6-21.

18. Law JA, Jacobsen SE. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nature reviews Genetics. 2010;11(3):204-20.

19. Lister R, Mukamel EA, Nery JR, Urich M, Puddifoot CA, Johnson ND, et al. Global epigenomic reconfiguration during mammalian brain development. Science. 2013;341(6146):1237905.

20. Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. Nature. 2009;462(7271):315-22.

21. Shirane K, Toh H, Kobayashi H, Miura F, Chiba H, Ito T, et al. Mouse oocyte methylomes at base resolution reveal genome-wide accumulation of non-CpG methylation and role of DNA methyltransferases. PLoS genetics. 2013;9(4):e1003439.

22. Xie W, Barr CL, Kim A, Yue F, Lee AY, Eubanks J, et al. Base-resolution analyses of sequence and parent-of-origin dependent DNA methylation in the mouse genome. Cell. 2012;148(4):816-31.

23. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nature reviews Genetics. 2012;13(7):484-92.

24. Weber M, Hellmann I, Stadler MB, Ramos L, Paabo S, Rebhan M, et al. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. Nature genetics. 2007;39(4):457-66.

25. Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, et al. Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature. 2008;454(7205):766-70.

26. Hermann A, Gowher H, Jeltsch A. Biochemistry and biology of mammalian DNA methyltransferases. Cell Mol Life Sci. 2004;61(19-20):2571-87.

27. Bestor TH. The DNA methyltransferases of mammals. Hum Mol Genet. 2000;9(16):2395-402.

28. Kobayashi H, Sakurai T, Imai M, Takahashi N, Fukuda A, Yayoi O, et al. Contribution of intragenic DNA methylation in mouse gametic DNA methylomes to establish oocyte-specific heritable marks. PLoS genetics. 2012;8(1):e1002440.

29. Smith ZD, Chan MM, Mikkelsen TS, Gu H, Gnirke A, Regev A, et al. A unique regulatory phase of DNA methylation in the early mammalian embryo. Nature. 2012;484(7394):339-44.

30. Borgel J, Guibert S, Li Y, Chiba H, Schubeler D, Sasaki H, et al. Targets and dynamics of promoter DNA methylation during early mouse development. Nature genetics. 2010;42(12):1093-100.

31. Smallwood SA, Tomizawa S, Krueger F, Ruf N, Carli N, Segonds-Pichon A, et al. Dynamic CpG island methylation landscape in oocytes and preimplantation embryos. Nature genetics. 2011;43(8):811-4.

32. Popp C, Dean W, Feng S, Cokus SJ, Andrews S, Pellegrini M, et al. Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. Nature. 2010;463(7284):1101-5.

33. Oswald J, Engemann S, Lane N, Mayer W, Olek A, Fundele R, et al. Active demethylation of the paternal genome in the mouse zygote. Current biology : CB. 2000;10(8):475-8.

34. Monk M, Adams RL, Rinaldi A. Decrease in DNA methylase activity during preimplantation development in the mouse. Development. 1991;112(1):189-92.

35. Rougier N, Bourc'his D, Gomes DM, Niveleau A, Plachot M, Paldi A, et al. Chromosome methylation patterns during mammalian preimplantation development. Genes Dev. 1998;12(14):2108-13.

36. Haaf T. Methylation dynamics in the early mammalian embryo: implications of genome reprogramming defects for development. Curr Top Microbiol Immunol. 2006;310:13-22.

37. Song J, Teplova M, Ishibe-Murakami S, Patel DJ. Structure-based mechanistic insights into DNMT1-mediated maintenance DNA methylation. Science. 2012;335(6069):709-12.

38. Mayer W, Niveleau A, Walter J, Fundele R, Haaf T. Demethylation of the zygotic paternal genome. Nature. 2000;403(6769):501-2.

39. Hajkova P, Ancelin K, Waldmann T, Lacoste N, Lange UC, Cesari F, et al. Chromatin dynamics during epigenetic reprogramming in the mouse germ line. Nature. 2008;452(7189):877-81.

40. Wossidlo M, Arand J, Sebastiano V, Lepikhov K, Boiani M, Reinhardt R, et al. Dynamic link of DNA demethylation, DNA strand breaks and repair in mouse zygotes. Embo j. 2010;29(11):1877-88.

41. Morgan HD, Dean W, Coker HA, Reik W, Petersen-Mahrt SK. Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming. The Journal of biological chemistry. 2004;279(50):52353-60.

42. Hardeland U, Bentele M, Jiricny J, Schär P. The versatile thymine DNA-glycosylase: a comparative characterization of the human, Drosophila and fission yeast orthologs. Nucleic Acids Res. 2003;31(9):2261-71.

43. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science (New York, NY). 2009;324(5929):930-5.

44. Penn NW, Suwalski R, O'Riley C, Bojanowski K, Yura R. The presence of 5-hydroxymethylcytosine in animal deoxyribonucleic acid. The Biochemical journal. 1972;126(4):781-90.

45. Kriaucionis S, Heintz N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. Science. 2009;324(5929):929-30.

46. Song CX, Szulwach KE, Fu Y, Dai Q, Yi C, Li X, et al. Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. Nat Biotechnol. 2011;29(1):68-72.

47. Borst P, Sabatini R. Base J: discovery, biosynthesis, and possible functions. Annu Rev Microbiol. 2008;62:235-51.

48. Yu Z, Genest PA, ter Riet B, Sweeney K, DiPaolo C, Kieft R, et al. The protein that binds to DNA base J in trypanosomatids has features of a thymidine hydroxylase. Nucleic Acids Res. 2007;35(7):2107-15.

49. Cliffe LJ, Kieft R, Southern T, Birkeland SR, Marshall M, Sweeney K, et al. JBP1 and JBP2 are two distinct thymidine hydroxylases involved in J biosynthesis in genomic DNA of African trypanosomes. Nucleic Acids Res. 2009;37(5):1452-62.

50. He YF, Li BZ, Li Z, Liu P, Wang Y, Tang Q, et al. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. Science. 2011;333(6047):1303-7.

51. Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, et al. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. Science. 2011;333(6047):1300-3.

52. Shen L, Wu H, Diep D, Yamaguchi S, D'Alessio AC, Fung HL, et al. Genome-wide analysis reveals TET- and TDG-dependent 5-methylcytosine oxidation dynamics. Cell. 2013;153(3):692-706.

53. Hashimoto H, Liu Y, Upadhyay AK, Chang Y, Howerton SB, Vertino PM, et al. Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation. Nucleic Acids Res. 2012;40(11):4841-9.

54. Ji D, Lin K, Song J, Wang Y. Effects of Tet-induced oxidation products of 5-methylcytosine on Dnmt1- and DNMT3a-mediated cytosine methylation. Mol Biosyst. 2014;10(7):1749-52.

55. Ito S, D'Alessio AC, Taranova OV, Hong K, Sowers LC, Zhang Y. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. Nature. 2010;466(7310):1129-33.

56. Vincent JJ, Huang Y, Chen PY, Feng S, Calvopina JH, Nee K, et al. Stage-specific roles for tet1 and tet2 in DNA demethylation in primordial germ cells. Cell Stem Cell. 2013;12(4):470-8.

57. Solary E, Bernard OA, Tefferi A, Fuks F, Vainchenker W. The Ten-Eleven Translocation-2 (TET2) gene in hematopoiesis and hematopoietic diseases. Leukemia. 2014;28(3):485-96.

58. Dawlaty MM, Ganz K, Powell BE, Hu YC, Markoulaki S, Cheng AW, et al. Tet1 is dispensable for maintaining pluripotency and its loss is compatible with embryonic and postnatal development. Cell Stem Cell. 2011;9(2):166-75.

59. Yamaguchi S, Hong K, Liu R, Shen L, Inoue A, Diep D, et al. Tet1 controls meiosis by regulating meiotic gene expression. Nature. 2012;492(7429):443-7.

60. Costa Y, Ding J, Theunissen TW, Faiola F, Hore TA, Shliaha PV, et al. NANOG-dependent function of TET1 and TET2 in establishment of pluripotency. Nature. 2013;495(7441):370-4.

61. Zhao Z, Chen L, Dawlaty MM, Pan F, Weeks O, Zhou Y, et al. Combined Loss of Tet1 and Tet2 Promotes B Cell, but Not Myeloid Malignancies, in Mice. Cell Rep. 2015;13(8):1692-704.

62. Dawlaty MM, Breiling A, Le T, Raddatz G, Barrasa MI, Cheng AW, et al. Combined deficiency of Tet1 and Tet2 causes epigenetic abnormalities but is compatible with postnatal development. Dev Cell. 2013;24(3):310-23.

63. Wossidlo M, Nakamura T, Lepikhov K, Marques CJ, Zakhartchenko V, Boiani M, et al. 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. Nat Commun. 2011;2:241.

64. Iqbal K, Jin SG, Pfeifer GP, Szabo PE. Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine. Proc Natl Acad Sci U S A. 2011;108(9):3642-7.

65. Yu H, Su Y, Shin J, Zhong C, Guo JU, Weng YL, et al. Tet3 regulates synaptic transmission and homeostatic plasticity via DNA oxidation and repair. Nat Neurosci. 2015;18(6):836-43.

66. Perera A, Eisen D, Wagner M, Laube SK, Kunzel AF, Koch S, et al. TET3 is recruited by REST for context-specific hydroxymethylation and induction of gene expression. Cell Rep. 2015;11(2):283-94.

67. Gu TP, Guo F, Yang H, Wu HP, Xu GF, Liu W, et al. The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. Nature. 2011;477(7366):606-10.

68. Inoue A, Shen L, Dai Q, He C, Zhang Y. Generation and replication-dependent dilution of 5fC and 5caC during mouse preimplantation development. Cell Res. 2011;21(12):1670-6.

69. Maiti A, Drohat AC. Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites. The Journal of biological chemistry. 2011;286(41):35334-8.

70. Hajkova P, Jeffries SJ, Lee C, Miller N, Jackson SP, Surani MA. Genome-wide reprogramming in the mouse germ line entails the base excision repair pathway. Science. 2010;329(5987):78-82.

71. Nakamura T, Liu YJ, Nakashima H, Umehara H, Inoue K, Matoba S, et al. PGC7 binds histone H3K9me2 to protect against conversion of 5mC to 5hmC in early embryos. Nature. 2012;486(7403):415-9.

72. Sato M, Kimura T, Kurokawa K, Fujita Y, Abe K, Masuhara M, et al. Identification of PGC7, a new gene expressed specifically in preimplantation embryos and germ cells. Mech Dev. 2002;113(1):91-4.

73. Bowles J, Teasdale RP, James K, Koopman P. Dppa3 is a marker of pluripotency and has a human homologue that is expressed in germ cell tumours. Cytogenet Genome Res. 2003;101(3-4):261-5.

74. Payer B, Saitou M, Barton SC, Thresher R, Dixon JP, Zahn D, et al. Stella is a maternal effect gene required for normal early development in mice. Current biology : CB. 2003;13(23):2110-7.

75. Bortvin A, Goodheart M, Liao M, Page DC. Dppa3 / Pgc7 / stella is a maternal factor and is not required for germ cell specification in mice. BMC Dev Biol. 2004;4:2.

76. Nakamura T, Arai Y, Umehara H, Masuhara M, Kimura T, Taniguchi H, et al. PGC7/Stella protects against DNA demethylation in early embryogenesis. Nature cell biology. 2007;9(1):64-71.

77. Bian C, Yu X. PGC7 suppresses TET3 for protecting DNA methylation. Nucleic Acids Res. 2014;42(5):2893-905.

78. Li Y, Zhang Z, Chen J, Liu W, Lai W, Liu B, et al. Stella safeguards the oocyte methylome by preventing de novo methylation mediated by DNMT1. Nature. 2018;564(7734):136-40.

79. Wang R, Li T. DNA Methylation is Correlated with Pluripotency of Stem Cells. Curr Stem Cell Res Ther. 2017;12(6):442-6.

80. Hellman A, Chess A. Gene body-specific methylation on the active X chromosome. Science. 2007;315(5815):1141-3.

81. Ball MP, Li JB, Gao Y, Lee JH, LeProust EM, Park IH, et al. Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. Nat Biotechnol. 2009;27(4):361-8.

82. Koh KP, Yabuuchi A, Rao S, Huang Y, Cunniff K, Nardone J, et al. Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. Cell Stem Cell. 2011;8(2):200-13.

83. Xu Y, Wu F, Tan L, Kong L, Xiong L, Deng J, et al. Genome-wide regulation of 5hmC, 5mC, and gene expression by Tet1 hydroxylase in mouse embryonic stem cells. Mol Cell. 2011;42(4):451-64.

84. Ficz G, Branco MR, Seisenberger S, Santos F, Krueger F, Hore TA, et al. Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. Nature. 2011;473(7347):398-402.

85. Williams K, Christensen J, Pedersen MT, Johansen JV, Cloos PA, Rappsilber J, et al. TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. Nature. 2011;473(7347):343-8.

86. Lu F, Liu Y, Jiang L, Yamaguchi S, Zhang Y. Role of Tet proteins in enhancer activity and telomere elongation. Genes Dev. 2014;28(19):2103-19.

87. Dawlaty MM, Breiling A, Le T, Barrasa MI, Raddatz G, Gao Q, et al. Loss of tet enzymes compromises proper differentiation of embryonic stem cells. Developmental cell. 2014;29(1):102-11.

88. Wu Y, Guo Z, Liu Y, Tang B, Wang Y, Yang L, et al. Oct4 and the small molecule inhibitor, SC1, regulates Tet2 expression in mouse embryonic stem cells. Mol Biol Rep. 2013;40(4):2897-906.

89. Sohni A, Bartoccetti M, Khoueiry R, Spans L, Vande Velde J, De Troyer L, et al. Dynamic switching of active promoter and enhancer domains regulates Tet1 and Tet2 expression during cell state transitions between pluripotency and differentiation. Mol Cell Biol. 2015;35(6):1026-42.

90. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell. 2006;126(4):663-76.

91. Wernig M, Meissner A, Foreman R, Brambrink T, Ku M, Hochedlinger K, et al. In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. Nature. 2007;448(7151):318-24.

92. Okita K, Ichisaka T, Yamanaka S. Generation of germline-competent induced pluripotent stem cells. Nature. 2007;448(7151):313-7.

93. Doege CA, Inoue K, Yamashita T, Rhee DB, Travis S, Fujita R, et al. Early-stage epigenetic modification during somatic cell reprogramming by Parp1 and Tet2. Nature. 2012;488(7413):652-5.

94. Gao Y, Chen J, Li K, Wu T, Huang B, Liu W, et al. Replacement of Oct4 by Tet1 during iPSC induction reveals an important role of DNA methylation and hydroxymethylation in reprogramming. Cell Stem Cell. 2013;12(4):453-69.

95. Zhu P, Guo H, Ren Y, Hou Y, Dong J, Li R, et al. Single-cell DNA methylome sequencing of human preimplantation embryos. Nature genetics. 2018;50(1):12-9.

96. Yu B, Dong X, Gravina S, Kartal Ö , Schimmel T, Cohen J, et al. Genome-wide, Single-Cell DNA Methylomics Reveals Increased Non-CpG Methylation during Human Oocyte Maturation. Stem Cell Reports. 2017;9(1):397-407.

97. Guo H, Zhu P, Yan L, Li R, Hu B, Lian Y, et al. The DNA methylation landscape of human early embryos. Nature. 2014;511(7511):606-10.

98. Guo F, Yan L, Guo H, Li L, Hu B, Zhao Y, et al. The Transcriptome and DNA Methylome Landscapes of Human Primordial Germ Cells. Cell. 2015;161(6):1437-52.

99. Guo H, Hu B, Yan L, Yong J, Wu Y, Gao Y, et al. DNA methylation and chromatin accessibility profiling of mouse and human fetal germ cells. Cell Res. 2017;27(2):165-83.

100. Gkountela S, Zhang KX, Shafiq TA, Liao WW, Hargan-Calvopina J, Chen PY, et al. DNA Demethylation Dynamics in the Human Prenatal Germline. Cell. 2015;161(6):1425-36.

101. Molaro A, Hodges E, Fang F, Song Q, McCombie WR, Hannon GJ, et al. Sperm methylation profiles reveal features of epigenetic inheritance and evolution in primates. Cell. 2011;146(6):1029-41.

102. Okae H, Chiba H, Hiura H, Hamada H, Sato A, Utsunomiya T, et al. Genome-wide analysis of DNA methylation dynamics during early human development. PLoS genetics. 2014;10(12):e1004868.

103. Braude P, Bolton V, Moore S. Human gene expression first occurs between the four- and eight-cell stages of preimplantation development. Nature. 1988;332(6163):459-61.

104. Aoki F, Worrad DM, Schultz RM. Regulation of transcriptional activity during the first and second cell cycles in the preimplantation mouse embryo. Dev Biol. 1997;181(2):296-307.

105. Stroud H, Feng S, Morey Kinney S, Pradhan S, Jacobsen SE. 5-Hydroxymethylcytosine is associated with enhancers and gene bodies in human embryonic stem cells. Genome biology. 2011;12(6):R54.

106. Szulwach KE, Li X, Li Y, Song CX, Han JW, Kim S, et al. Integrating 5-hydroxymethylcytosine into the epigenomic landscape of human embryonic stem cells. PLoS genetics. 2011;7(6):e1002154.

107. Verma N, Pan H, Dore LC, Shukla A, Li QV, Pelham-Webb B, et al. TET proteins safeguard bivalent promoters from de novo methylation in human embryonic stem cells. Nature genetics. 2018;50(1):83-95.

108. Langlois T, da Costa Reis Monte-Mor B, Lenglet G, Droin N, Marty C, Le Couedic JP, et al. TET2 deficiency inhibits mesoderm and hematopoietic differentiation in human embryonic stem cells. Stem cells (Dayton, Ohio). 2014;32(8):2084-97.

109. Campbell KH, McWhir J, Ritchie WA, Wilmut I. Sheep cloned by nuclear transfer from a cultured cell line. Nature. 1996;380(6569):64-6.

110. Kato Y, Tani T, Sotomaru Y, Kurokawa K, Kato J, Doguchi H, et al. Eight calves cloned from somatic cells of a single adult. Science. 1998;282(5396):2095-8.

111. Onishi A, Iwamoto M, Akita T, Mikawa S, Takeda K, Awata T, et al. Pig cloning by microinjection of fetal fibroblast nuclei. Science. 2000;289(5482):1188-90.

112. Wakayama T, Yanagimachi R. Cloning the laboratory mouse. Semin Cell Dev Biol. 1999;10(3):253-8.

113. Solter D. Mammalian cloning: advances and limitations. Nature reviews Genetics. 2000;1(3):199-207.

114. Beaujean N, Hartshorne G, Cavilla J, Taylor J, Gardner J, Wilmut I, et al. Non-conservation of mammalian preimplantation methylation dynamics. Current biology : CB. 2004;14(7):R266-7.

115. Dean W, Santos F, Stojkovic M, Zakhartchenko V, Walter J, Wolf E, et al. Conservation of methylation reprogramming in mammalian development: aberrant reprogramming in cloned embryos. Proc Natl Acad Sci U S A. 2001;98(24):13734-8.

116. Park JS, Jeong YS, Shin ST, Lee KK, Kang YK. Dynamic DNA methylation reprogramming: active demethylation and immediate remethylation in the male pronucleus of bovine zygotes. Developmental dynamics : an official publication of the American Association of Anatomists. 2007;236(9):2523-33.

117. Duan JE, Jiang ZC, Alqahtani F, Mandoiu I, Dong H, Zheng X, et al. Methylome Dynamics of Bovine Gametes and in vivo Early Embryos. Front Genet. 2019;10:512.

118. Jiang Z, Lin J, Dong H, Zheng X, Marjani SL, Duan J, et al. DNA methylomes of bovine gametes and in vivo produced preimplantation embryos. Biol Reprod. 2018;99(5):949-59.

119. Misirlioglu M, Page GP, Sagirkaya H, Kaya A, Parrish JJ, First NL, et al. Dynamics of global transcriptome in bovine matured oocytes and preimplantation embryos. Proc Natl Acad Sci U S A. 2006;103(50):18905-10.

120. Jiang Z, Sun J, Dong H, Luo O, Zheng X, Obergfell C, et al. Transcriptional profiles of bovine in vivo pre-implantation development. BMC Genomics. 2014;15:756.

121. Cheng H, Zhang J, Zhang S, Zhai Y, Jiang Y, An X, et al. Tet3 is required for normal in vitro fertilization preimplantation embryos development of bovine. Molecular reproduction and development. 2019;86(3):298-307.

122. Bakhtari A, Ross PJ. DPPA3 prevents cytosine hydroxymethylation of the maternal pronucleus and is required for normal development in bovine embryos. Epigenetics. 2014;9(9):1271-9.

123. Zhang J, Zhang S, Wang Y, Cheng H, Hao L, Zhai Y, et al. Effect of TET inhibitor on bovine parthenogenetic embryo development. PloS one. 2017;12(12):e0189542.

124. Deshmukh RS, Ostrup O, Ostrup E, Vejlsted M, Niemann H, Lucas-Hahn A, et al. DNA methylation in porcine preimplantation embryos developed in vivo and produced by in vitro fertilization, parthenogenetic activation and somatic cell nuclear transfer. Epigenetics. 2011;6(2):177-87.

125. Cao Z, Zhou N, Zhang Y, Zhang Y, Wu R, Li Y, et al. Dynamic reprogramming of 5-hydroxymethylcytosine during early porcine embryogenesis. Theriogenology. 2014;81(3):496-508.

126. Nie X, Liu Q, Wang R, Sheng W, Li X, Zhang M, et al. DNA demethylation pattern of in-vitro fertilized and cloned porcine pronuclear stage embryos. Clin Chim Acta. 2017;473:45-50.

127. Lee K, Hamm J, Whitworth K, Spate L, Park KW, Murphy CN, et al. Dynamics of TET family expression in porcine preimplantation embryos is related to zygotic genome activation and required for the maintenance of NANOG. Dev Biol. 2014;386(1):86-95.

128. Fan A, Ma K, An X, Ding Y, An P, Song G, et al. Effects of TET1 knockdown on gene expression and DNA methylation in porcine induced pluripotent stem cells. Reproduction. 2013;146(6):569-79.

129. Hou J, Liu L, Zhang J, Cui XH, Yan FX, Guan H, et al. Epigenetic modification of histone 3 at lysine 9 in sheep zygotes and its relationship with DNA methylation. BMC Dev Biol. 2008;8:60.

130. Masala L, Burrai GP, Bellu E, Ariu F, Bogliolo L, Ledda S, et al. Methylation dynamics during folliculogenesis and early embryo development in sheep. Reproduction. 2017;153(5):605-19.

131. Jafarpour F, Hosseini SM, Ostadhosseini S, Abbasi H, Dalman A, Nasr-Esfahani MH. Comparative dynamics of 5-methylcytosine reprogramming and TET family expression during preimplantation mammalian development in mouse and sheep. Theriogenology. 2017;89:86-96.

132. Young LE, Beaujean N. DNA methylation in the preimplantation embryo: the differing stories of the mouse and sheep. Animal reproduction science. 2004;82-83:61-78.

133. Telford NA, Watson AJ, Schultz GA. Transition from maternal to embryonic control in early mammalian development: a comparison of several species. Mol Reprod Dev. 1990;26(1):90-100.

134. Deshmukh RS, Ø strup O, Ø strup E, Vejlsted M, Niemann H, Lucas-Hahn A, et al. DNA methylation in porcine preimplantation embryos developed in vivo and produced by in vitro fertilization, parthenogenetic activation and somatic cell nuclear transfer. Epigenetics. 2011;6(2):177-87.

135. Santos F, Hendrich B, Reik W, Dean W. Dynamic reprogramming of DNA methylation in the early mouse embryo. Dev Biol. 2002;241(1):172-82.

136. Prather RS. Nuclear control of early embryonic development in domestic pigs. J Reprod Fertil Suppl. 1993;48:17-29.

137. Oestrup O, Hall V, Petkov SG, Wolf XA, Hyldig S, Hyttel P. From zygote to implantation: morphological and molecular dynamics during embryo development in the pig. Reprod Domest Anim. 2009;44 Suppl 3:39-49.

138. Hansis C, Grifo JA, Krey LC. Oct-4 expression in inner cell mass and trophectoderm of human blastocysts. Mol Hum Reprod. 2000;6(11):999-1004.

139. Humpherson PG, Leese HJ, Sturmey RG. Amino acid metabolism of the porcine blastocyst. Theriogenology. 2005;64(8):1852-66.

140. Qian L, Tang M, Yang J, Wang Q, Cai C, Jiang S, et al. Targeted mutations in myostatin by zinc-finger nucleases result in double-muscled phenotype in Meishan pigs. Sci Rep. 2015;5:14435.

141. Wang K, Ouyang H, Xie Z, Yao C, Guo N, Li M, et al. Efficient Generation of Myostatin Mutations in Pigs Using the CRISPR/Cas9 System. Sci Rep. 2015;5:16623.

142. Rao S, Fujimura T, Matsunari H, Sakuma T, Nakano K, Watanabe M, et al. Efficient modification of the myostatin gene in porcine somatic cells and generation of knockout piglets. Mol Reprod Dev. 2016;83(1):61-70.

143. Whitworth KM, Rowland RR, Ewen CL, Trible BR, Kerrigan MA, Cino-Ozuna AG, et al. Gene-edited pigs are protected from porcine reproductive and respiratory syndrome virus. Nat Biotechnol. 2016;34(1):20-2.

144. Yang H, Zhang J, Zhang X, Shi J, Pan Y, Zhou R, et al. CD163 knockout pigs are fully resistant to highly pathogenic porcine reproductive and respiratory syndrome virus. Antiviral Res. 2018;151:63-70.

145. Burkard C, Lillico SG, Reid E, Jackson B, Mileham AJ, Ait-Ali T, et al. Precision engineering for PRRSV resistance in pigs: Macrophages from genome edited pigs lacking CD163 SRCR5 domain are fully resistant to both PRRSV genotypes while maintaining biological function. PLoS Pathog. 2017;13(2):e1006206.

146. Perleberg C, Kind A, Schnieke A. Genetically engineered pigs as models for human disease. Dis Model Mech. 2018;11(1).

147. Hryhorowicz M, Zeyland J, Słomski R, Lipiński D. Genetically Modified Pigs as Organ Donors for Xenotransplantation. Mol Biotechnol. 2017;59(9-10):435-44.

148. Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nature genetics. 2003;33 Suppl:245-54.

149. Robertson KD, Wolffe AP. DNA methylation in health and disease. Nature reviews Genetics. 2000;1(1):11-9.

150. Reik W. Stability and flexibility of epigenetic gene regulation in mammalian development. Nature. 2007;447(7143):425-32.

151. Surani MA, Hayashi K, Hajkova P. Genetic and epigenetic regulators of pluripotency. Cell. 2007;128(4):747-62.

152. Zaitseva I, Zaitsev S, Alenina N, Bader M, Krivokharchenko A. Dynamics of DNA-demethylation in early mouse and rat embryos developed in vivo and in vitro. Mol Reprod Dev. 2007;74(10):1255-61.

153. Shen L, Inoue A, He J, Liu Y, Lu F, Zhang Y. Tet3 and DNA replication mediate demethylation of both the maternal and paternal genomes in mouse zygotes. Cell Stem Cell. 2014;15(4):459-71.

154. Guo F, Li X, Liang D, Li T, Zhu P, Guo H, et al. Active and passive demethylation of male and female pronuclear DNA in the mammalian zygote. Cell Stem Cell. 2014;15(4):447-59.

155. Santos F, Dean W. Epigenetic reprogramming during early development in mammals. Reproduction. 2004;127(6):643-51.

156. Quivoron C, Couronne L, Della Valle V, Lopez CK, Plo I, Wagner-Ballon O, et al. TET2 inactivation results in pleiotropic hematopoietic abnormalities in mouse and is a recurrent event during human lymphomagenesis. Cancer cell. 2011;20(1):25-38.

157. Ko M, Bandukwala HS, An J, Lamperti ED, Thompson EC, Hastie R, et al. Ten-Eleven-Translocation 2 (TET2) negatively regulates homeostasis and differentiation of hematopoietic stem cells in mice. Proc Natl Acad Sci U S A. 2011;108(35):14566-71.

158. Yoshioka K, Suzuki C, Tanaka A, Anas IM, Iwamura S. Birth of piglets derived from porcine zygotes cultured in a chemically defined medium. Biology of reproduction. 2002;66(1):112-9.

159. Yugo DM, Heffron CL, Ryu J, Uh K, Subramaniam S, Matzinger SR, et al. Infection Dynamics of Hepatitis E Virus in Wild-Type and Immunoglobulin Heavy Chain Knockout JH (-/-) Gnotobiotic Piglets. Journal of virology. 2018;92(21).

160. Lei S, Ryu J, Wen K, Twitchell E, Bui T, Ramesh A, et al. Increased and prolonged human norovirus infection in RAG2/IL2RG deficient gnotobiotic pigs with severe combined immunodeficiency. Sci Rep. 2016;6:25222.

161. Kang YK, Koo DB, Park JS, Choi YH, Kim HN, Chang WK, et al. Typical demethylation events in cloned pig embryos. Clues on species-specific differences in epigenetic reprogramming of a cloned donor genome. The Journal of biological chemistry. 2001;276(43):39980-4.

162. Archer GS, Dindot S, Friend TH, Walker S, Zaunbrecher G, Lawhorn B, et al. Hierarchical phenotypic and epigenetic variation in cloned swine. Biol Reprod. 2003;69(2):430-6.

163. Huang Y, Chavez L, Chang X, Wang X, Pastor WA, Kang J, et al. Distinct roles of the methylcytosine oxidases Tet1 and Tet2 in mouse embryonic stem cells. Proc Natl Acad Sci U S A. 2014;111(4):1361-6.

164. Williams K, Christensen J, Helin K. DNA methylation: TET proteins-guardians of CpG islands? EMBO Rep. 2011;13(1):28-35.

165. Yamagata K, Yamazaki T, Miki H, Ogonuki N, Inoue K, Ogura A, et al. Centromeric DNA hypomethylation as an epigenetic signature discriminates between germ and somatic cell lineages. Dev Biol. 2007;312(1):419-26.

166. Pastor WA, Pape UJ, Huang Y, Henderson HR, Lister R, Ko M, et al. Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. Nature. 2011;473(7347):394-7.

167. Wu H, D'Alessio AC, Ito S, Xia K, Wang Z, Cui K, et al. Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. Nature. 2011;473(7347):389-93.

168. Freudenberg JM, Ghosh S, Lackford BL, Yellaboina S, Zheng X, Li R, et al. Acute depletion of Tet1-dependent 5-hydroxymethylcytosine levels impairs LIF/Stat3 signaling and results in loss of embryonic stem cell identity. Nucleic Acids Res. 2012;40(8):3364-77.

169. Kang J, Lienhard M, Pastor WA, Chawla A, Novotny M, Tsagaratou A, et al. Simultaneous deletion of the methylcytosine oxidases Tet1 and Tet3 increases transcriptome variability in early embryogenesis. Proc Natl Acad Sci U S A. 2015;112(31):E4236-45.

170. Shi FT, Kim H, Lu W, He Q, Liu D, Goodell MA, et al. Ten-eleven translocation 1 (Tet1) is regulated by O-linked N-acetylglucosamine transferase (Ogt) for target gene repression in mouse embryonic stem cells. The Journal of biological chemistry. 2013;288(29):20776-84.

171. Klose RJ, Bird AP. Genomic DNA methylation: the mark and its mediators. Trends in biochemical sciences. 2006;31(2):89-97.

172. Prasad P, Mittal SA, Chongtham J, Mohanty S, Srivastava T. Hypoxia-Mediated Epigenetic Regulation of Stemness in Brain Tumor Cells. Stem cells (Dayton, Ohio). 2017;35(6):1468-78.

173. Festuccia N, Osorno R, Halbritter F, Karwacki-Neisius V, Navarro P, Colby D, et al. Esrrb is a direct Nanog target gene that can substitute for Nanog function in pluripotent cells. Cell Stem Cell. 2012;11(4):477-90.

174. Kuijk EW, Du Puy L, Van Tol HT, Oei CH, Haagsman HP, Colenbrander B, et al. Differences in early lineage segregation between mammals. Developmental dynamics : an official publication of the American Association of Anatomists. 2008;237(4):918-27.

175. Hall VJ, Christensen J, Gao Y, Schmidt MH, Hyttel P. Porcine pluripotency cell signaling develops from the inner cell mass to the epiblast during early development.

Developmental dynamics : an official publication of the American Association of Anatomists. 2009;238(8):2014-24.

176. du Puy L, Lopes SM, Haagsman HP, Roelen BA. Analysis of co-expression of OCT4, NANOG and SOX2 in pluripotent cells of the porcine embryo, in vivo and in vitro. Theriogenology. 2011;75(3):513-26.

177. Whitworth KM, Lee K, Benne JA, Beaton BP, Spate LD, Murphy SL, et al. Use of the CRISPR/Cas9 system to produce genetically engineered pigs from in vitro-derived oocytes and embryos. Biol Reprod. 2014;91(3):78.

178. Santos F, Hyslop L, Stojkovic P, Leary C, Murdoch A, Reik W, et al. Evaluation of epigenetic marks in human embryos derived from IVF and ICSI. Human reproduction (Oxford, England). 2010;25(9):2387-95.

179. Fulka J, Fulka H, Slavik T, Okada K, Fulka J, Jr. DNA methylation pattern in pig in vivo produced embryos. Histochemistry and cell biology. 2006;126(2):213-7.

180. Sasaki H, Matsui Y. Epigenetic events in mammalian germ-cell development: reprogramming and beyond. Nature reviews Genetics. 2008;9(2):129-40.

181. Zhou LQ, Dean J. Reprogramming the genome to totipotency in mouse embryos. Trends Cell Biol. 2015;25(2):82-91.

182. Messerschmidt DM, Knowles BB, Solter D. DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos. Genes Dev. 2014;28(8):812-28.

183. Allen MD, Grummitt CG, Hilcenko C, Min SY, Tonkin LM, Johnson CM, et al. Solution structure of the nonmethyl-CpG-binding CXXC domain of the leukaemia-associated MLL histone methyltransferase. Embo j. 2006;25(19):4503-12.

184. Pradhan M, Esteve PO, Chin HG, Samaranayke M, Kim GD, Pradhan S. CXXC domain of human DNMT1 is essential for enzymatic activity. Biochemistry. 2008;47(38):10000-9.

185. Zhang H, Zhang X, Clark E, Mulcahey M, Huang S, Shi YG. TET1 is a DNA-binding protein that modulates DNA methylation and gene transcription via hydroxylation of 5-methylcytosine. Cell Res. 2010;20(12):1390-3.

186. Xu Y, Xu C, Kato A, Tempel W, Abreu JG, Bian C, et al. Tet3 CXXC domain and dioxygenase activity cooperatively regulate key genes for Xenopus eye and neural development. Cell. 2012;151(6):1200-13.

187. Melamed P, Yosefzon Y, David C, Tsukerman A, Pnueli L. Tet Enzymes, Variants, and Differential Effects on Function. Front Cell Dev Biol. 2018;6:22.

188. Liu N, Wang M, Deng W, Schmidt CS, Qin W, Leonhardt H, et al. Intrinsic and extrinsic connections of Tet3 dioxygenase with CXXC zinc finger modules. PloS one. 2013;8(5):e62755.

189. Jin SG, Zhang ZM, Dunwell TL, Harter MR, Wu X, Johnson J, et al. Tet3 Reads 5-Carboxylcytosine through Its CXXC Domain and Is a Potential Guardian against Neurodegeneration. Cell Rep. 2016;14(3):493-505.

190. Hahn MA, Qiu R, Wu X, Li AX, Zhang H, Wang J, et al. Dynamics of 5-hydroxymethylcytosine and chromatin marks in Mammalian neurogenesis. Cell Rep. 2013;3(2):291-300.

191. Colquitt BM, Allen WE, Barnea G, Lomvardas S. Alteration of genic 5-hydroxymethylcytosine patterning in olfactory neurons correlates with changes in gene expression and cell identity. Proc Natl Acad Sci U S A. 2013;110(36):14682-7.

192. Araujo PR, Yoon K, Ko D, Smith AD, Qiao M, Suresh U, et al. Before It Gets Started: Regulating Translation at the 5' UTR. Comp Funct Genomics. 2012;2012:475731.

193. Lv X, Jiang H, Liu Y, Lei X, Jiao J. MicroRNA-15b promotes neurogenesis and inhibits neural progenitor proliferation by directly repressing TET3 during early neocortical development. EMBO Rep. 2014;15(12):1305-14.

194. Zhang P, Huang B, Xu X, Sessa WC. Ten-eleven translocation (Tet) and thymine DNA glycosylase (TDG), components of the demethylation pathway, are direct targets of miRNA-29a. Biochem Biophys Res Commun. 2013;437(3):368-73.

195. Xu C, Bian C, Lam R, Dong A, Min J. The structural basis for selective binding of non-methylated CpG islands by the CFP1 CXXC domain. Nat Commun. 2011;2:227.

196. Iyer LM, Tahiliani M, Rao A, Aravind L. Prediction of novel families of enzymes involved in oxidative and other complex modifications of bases in nucleic acids. Cell Cycle. 2009;8(11):1698-710.

197. Bestor TH, Coxon A. Cytosine methylation: the pros and cons of DNA methylation. Current biology : CB. 1993;3(6):384-6.

198. Bird AP. CpG-rich islands and the function of DNA methylation. Nature. 1986;321(6067):209-13.

199. Guo JU, Su Y, Zhong C, Ming GL, Song H. Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. Cell. 2011;145(3):423-34.

200. Guo JU, Su Y, Zhong C, Ming GL, Song H. Emerging roles of TET proteins and 5-hydroxymethylcytosines in active DNA demethylation and beyond. Cell Cycle. 2011;10(16):2662-8.

201. Hackett JA, Sengupta R, Zylicz JJ, Murakami K, Lee C, Down TA, et al. Germline DNA demethylation dynamics and imprint erasure through 5-hydroxymethylcytosine. Science. 2013;339(6118):448-52.

202. Koh KP, Rao A. DNA methylation and methylcytosine oxidation in cell fate decisions. Curr Opin Cell Biol. 2013;25(2):152-61.

203. Pfeifer GP, Kadam S, Jin SG. 5-hydroxymethylcytosine and its potential roles in development and cancer. Epigenetics Chromatin. 2013;6(1):10.

204. Xu W, Yang H, Liu Y, Yang Y, Wang P, Kim SH, et al. Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of α-ketoglutarate-dependent dioxygenases. Cancer cell. 2011;19(1):17-30.

205. Ko M, An J, Bandukwala HS, Chavez L, Aijö T, Pastor WA, et al. Modulation of TET2 expression and 5-methylcytosine oxidation by the CXXC domain protein IDAX. Nature. 2013;497(7447):122-6.

206. Xiao P, Zhou XL, Zhang HX, Xiong K, Teng Y, Huang XJ, et al. Characterization of the nuclear localization signal of the mouse TET3 protein. Biochem Biophys Res Commun. 2013;439(3):373-7.

207. Zhang Q, Liu X, Gao W, Li P, Hou J, Li J, et al. Differential regulation of the ten-eleven translocation (TET) family of dioxygenases by O-linked β-N-acetylglucosamine transferase (OGT). The Journal of biological chemistry. 2014;289(9):5986-96.

208. Nguyen Ba AN, Pogoutse A, Provart N, Moses AM. NLStradamus: a simple Hidden Markov Model for nuclear localization signal prediction. BMC Bioinformatics. 2009;10:202.

209. Lee K, Hamm J, Whitworth K, Spate L, Park K-w, Murphy CN, et al. Dynamics of TET family expression in porcine preimplantation embryos is related to zygotic genome activation and required for the maintenance of NANOG. Developmental biology. 2014;386(1):86-95.

210. Inoue A, Zhang Y. Replication-dependent loss of 5-hydroxymethylcytosine in mouse preimplantation embryos. Science. 2011;334(6053):194.

211. Peat JR, Dean W, Clark SJ, Krueger F, Smallwood SA, Ficz G, et al. Genome-wide bisulfite sequencing in zygotes identifies demethylation targets and maps the contribution of TET3 oxidation. Cell Rep. 2014;9(6):1990-2000.

212. Bestor T, Laudano A, Mattaliano R, Ingram V. Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. J Mol Biol. 1988;203(4):971-83.

213. Stirzaker C, Taberlay PC, Statham AL, Clark SJ. Mining cancer methylomes: prospects and challenges. Trends Genet. 2014;30(2):75-84.

214. Gardiner-Garden M, Frommer M. CpG islands in vertebrate genomes. J Mol Biol. 1987;196(2):261-82.

215. Ogoshi K, Hashimoto S, Nakatani Y, Qu W, Oshima K, Tokunaga K, et al. Genome-wide profiling of DNA methylation in human cancer cells. Genomics. 2011;98(4):280-7.

216. Filion GJ, Zhenilo S, Salozhin S, Yamada D, Prokhortchouk E, Defossez PA. A family of human zinc finger proteins that bind methylated DNA and repress transcription. Mol Cell Biol. 2006;26(1):169-81.

217. Hendrich B, Bird A. Identification and characterization of a family of mammalian methyl-CpG binding proteins. Mol Cell Biol. 1998;18(11):6538-47.

218. Ludwig AK, Zhang P, Cardoso MC. Modifiers and Readers of DNA Modifications and Their Impact on Genome Structure, Expression, and Stability in Disease. Front Genet. 2016;7:115.

219. Brero A, Leonhardt H, Cardoso MC. Replication and translation of epigenetic information. Curr Top Microbiol Immunol. 2006;301:21-44.

220. Bogdanović O, Veenstra GJ. DNA methylation and methyl-CpG binding proteins: developmental requirements and function. Chromosoma. 2009;118(5):549-65.

221. Hashimoto H, Vertino PM, Cheng X. Molecular coupling of DNA methylation and histone methylation. Epigenomics. 2010;2(5):657-69.

222. Wu X, Zhang Y. TET-mediated active DNA demethylation: mechanism, function and beyond. Nature reviews Genetics. 2017;18(9):517-34.

223. Wang T, Wu H, Li Y, Szulwach KE, Lin L, Li X, et al. Subtelomeric hotspots of aberrant 5-hydroxymethylcytosine-mediated epigenetic modifications during reprogramming to pluripotency. Nature cell biology. 2013;15(6):700-11.

224. Chen Q, Chen Y, Bian C, Fujiki R, Yu X. TET2 promotes histone O-GlcNAcylation during gene transcription. Nature. 2013;493(7433):561-4.

225. Ito R, Katsura S, Shimada H, Tsuchiya H, Hada M, Okumura T, et al. TET3-OGT interaction increases the stability and the presence of OGT in chromatin. Genes to cells : devoted to molecular & cellular mechanisms. 2014;19(1):52-65.

226. Nestor CE, Ottaviano R, Reddington J, Sproul D, Reinhardt D, Dunican D, et al. Tissue type is a major modifier of the 5-hydroxymethylcytosine content of human genes. Genome research. 2012;22(3):467-77.

227. Valinluck V, Tsai HH, Rogstad DK, Burdzy A, Bird A, Sowers LC. Oxidative damage to methyl-CpG sequences inhibits the binding of the methyl-CpG binding domain (MBD) of methyl-CpG binding protein 2 (MeCP2). Nucleic Acids Res. 2004;32(14):4100-8.

228. Ludwig AK, Zhang P, Hastert FD, Meyer S, Rausch C, Herce HD, et al. Binding of MBD proteins to DNA blocks Tet1 function thereby modulating transcriptional noise. Nucleic Acids Res. 2017;45(5):2438-57.

229. Rasmussen KD, Helin K. Role of TET enzymes in DNA methylation, development, and cancer. Genes Dev. 2016;30(7):733-50.

230. Redel BK, Brown AN, Spate LD, Whitworth KM, Green JA, Prather RS. Glycolysis in preimplantation development is partially controlled by the Warburg Effect. Mol Reprod Dev. 2012;79(4):262-71.

231. Spate LD, Brown AN, Redel BK, Whitworth KM, Murphy CN, Prather RS. Dickkopf-related protein 1 inhibits the WNT signaling pathway and improves pig oocyte maturation. PloS one. 2014;9(4):e95114.

232. Hendrich B, Guy J, Ramsahoye B, Wilson VA, Bird A. Closely related proteins MBD2 and MBD3 play distinctive but interacting roles in mouse development. Genes Dev. 2001;15(6):710-23.

233. Mellén M, Ayata P, Dewell S, Kriaucionis S, Heintz N. MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. Cell. 2012;151(7):1417-30.

234. Wade PA. Methyl CpG binding proteins: coupling chromatin architecture to gene regulation. Oncogene. 2001;20(24):3166-73.

235. Zhao X, Ueba T, Christie BR, Barkho B, McConnell MJ, Nakashima K, et al. Mice lacking methyl-CpG binding protein 1 have deficits in adult neurogenesis and hippocampal function. Proc Natl Acad Sci U S A. 2003;100(11):6777-82.

236. Fujita N, Takebayashi S, Okumura K, Kudo S, Chiba T, Saya H, et al. Methylation-mediated transcriptional silencing in euchromatin by methyl-CpG binding protein MBD1 isoforms. Mol Cell Biol. 1999;19(9):6415-26.

237. Fujita N, Shimotake N, Ohki I, Chiba T, Saya H, Shirakawa M, et al. Mechanism of transcriptional regulation by methyl-CpG binding protein MBD1. Mol Cell Biol. 2000;20(14):5107-18.

238. Long HK, Blackledge NP, Klose RJ. ZF-CxxC domain-containing proteins, CpG islands and the chromatin connection. Biochem Soc Trans. 2013;41(3):727-40.

239. Zhang P, Rausch C, Hastert FD, Boneva B, Filatova A, Patil SJ, et al. Methyl-CpG binding domain protein 1 regulates localization and activity of Tet1 in a CXXC3 domain-dependent manner. Nucleic Acids Res. 2017;45(12):7118-36.

240. Fujii T, Sakurai N, Osaki T, Iwagami G, Hirayama H, Minamihashi A, et al. Changes in the expression patterns of the genes involved in the segregation and function of inner cell mass and trophectoderm lineages during porcine preimplantation development. The Journal of reproduction and development. 2013;59(2):151-8.

241. Rodríguez A, Allegrucci C, Alberio R. Modulation of pluripotency in the porcine embryo and iPS cells. PloS one. 2012;7(11):e49079.

242. Blaschke K, Ebata KT, Karimi MM, Zepeda-Martínez JA, Goyal P, Mahapatra S, et al. Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. Nature. 2013;500(7461):222-6.

243. Chen J, Guo L, Zhang L, Wu H, Yang J, Liu H, et al. Vitamin C modulates TET1 function during somatic cell reprogramming. Nature genetics. 2013;45(12):1504-9.

244. Zou Y, Mi J, Cui J, Lu D, Zhang X, Guo C, et al. Characterization of nuclear localization signal in the N terminus of CUL4B and its essential role in cyclin E degradation and cell cycle progression. The Journal of biological chemistry. 2009;284(48):33320-32.

245. Srinivasan M, Edman CF, Schulman H. Alternative splicing introduces a nuclear localization signal that targets multifunctional CaM kinase to the nucleus. J Cell Biol. 1994;126(4):839-52.

246. Mancini A, Koch A, Whetton AD, Tamura T. The M-CSF receptor substrate and interacting protein FMIP is governed in its subcellular localization by protein kinase C-mediated phosphorylation, and thereby potentiates M-CSF-mediated differentiation. Oncogene. 2004;23(39):6581-9.

247. Miyamoto Y, Imamoto N, Sekimoto T, Tachibana T, Seki T, Tada S, et al. Differential modes of nuclear localization signal (NLS) recognition by three distinct classes of NLS receptors. The Journal of biological chemistry. 1997;272(42):26375-81.

248. Deplus R, Delatte B, Schwinn MK, Defrance M, Méndez J, Murphy N, et al. TET2 and TET3 regulate GlcNAcylation and H3K4 methylation through OGT and SET1/COMPASS. Embo j. 2013;32(5):645-55.

249. Vella P, Scelfo A, Jammula S, Chiacchiera F, Williams K, Cuomo A, et al. Tet proteins connect the O-linked N-acetylglucosamine transferase Ogt to chromatin in embryonic stem cells. Mol Cell. 2013;49(4):645-56.