8-2020

Comparative Analysis of Small Non-Coding RNA and Messenger RNA Expression in Somatic Cell Nuclear Transfer and In Vitro-Fertilized Bovine Embryos During Early Development Through the Maternal-to-Embryonic Transition

Jocelyn Marie Cuthbert
Utah State University

COMPARATIVE ANALYSIS OF SMALL NON-CODING RNA AND MESSENGER RNA

EXPRESSION IN SOMATIC CELL NUCLEAR TRANSFER AND IN VITRO-FERTILIZED

BOVINE EMBRYOS DURING EARLY DEVELOPMENT THROUGH THE

MATERNAL-TO-EMBRYONIC TRANSITION

by

Jocelyn Marie Cuthbert

A dissertation submitted in partial fulfillment

of the requirements for the degree

of

DOCTOR OF PHILOSOPHY

in

Animal, Dairy and Veterinary Sciences

Approved:

_____
Abby D. Benninghoff, Ph.D.
Major Professor

_____
Kenneth L. White, Ph.D.
Major Professor

_____
S. Clay Isom, Ph.D.
Committee Member

_____
Irina A. Polejaeva, Ph.D.
Committee Member

_____
Gregory J. Podgorski, Ph.D.
Committee Member

_____
Richard S. Inouye, Ph.D.
Vice Provost for Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2020

ABSTRACT

Profiling of sncRNA and mapping transcriptional changes in bovine somatic cell nuclear transfer

embryos during early development through the maternal-to-embryonic transition

by

Jocelyn M. Cuthbert, Doctor of Philosophy

Utah State University, 2020

Major Professor: Dr. Abby Benninghoff and Dr. Kenneth White

Department: Animal, Dairy, and Veterinary Sciences

The efficiency of embryo production by somatic cell nuclear transfer (scNT) is far below that of *in vitro* fertilized embryos (IVF), likely due to the accumulation of errors in genome reprogramming, which may encompass the coding genes for small non-coding RNAs (sncRNAs). While sncRNAs have been reported to be important in mammalian embryo development, their potential functions have not been assessed through the critical maternal-to-embryonic transition (MET) in cattle embryos. The objectives of this study were to examine changes in expression of sncRNA during the MET in bovine IVF embryos, to assess the dysregulation of sncRNA and transcriptome profiles in scNT bovine embryos compared to IVF embryos, and to correlate transcript expression profiles with micro-RNA expression profiles. First, RNA sequencing was utilized to profile sncRNAs in bovine IVF oocytes and 8-cell- and blastocyst-staged bovine embryos. The analyses revealed a strong dynamic shift in relative abundance of differentially expressed sncRNAs, suggesting important roles for sncRNAs during the MET. Next, small and large RNA fractions were isolated from the same samples of cattle embryos produced via IVF or scNT at the 2-cell, 8-cell, morula and blastocyst developmental stages, and from MII oocytes, donor cell fibroblasts and either IVF or scNT blastocyst-derived cells (BDCs). RNA sequencing

showed that few sncRNAs were differentially expressed between scNT and IVF embryos, with all significant differences occurring at the morula stage. Differential expression of sncRNAs was apparent between stages of development within type (scNT or IVF), with changes in sncRNA populations appearing to be based on the relative differentiation status of developmental stage. Last, transcriptome data analysis revealed large-scale differences between scNT and IVF embryos at each developmental stage examined. Interestingly, altered transcripts at the 8-cell stage were associated with biological functions critical for the MET. The abundance of two miRNAs differentially expressed in scNT morulae, miR-34a and miR-345, was negatively correlated with the expression of some predicted mRNA targets. However, broad changes in mRNA expression were not consistently correlated with aberrations in miRNA expression, suggesting that other mechanisms leading to altered expression of mRNA in scNT embryos may be at play.

(382 pages)

PUBLIC ABSTRACT

Profiling of sncRNA and mapping transcriptional changes in bovine somatic cell nuclear transfer

embryos during early development through the maternal-to-embryonic transition

Jocelyn M. Cuthbert

Major Professor: Drs. Abby D. Benninghoff and Kenneth L. White

Department: Animal, Dairy, and Veterinary Sciences

Cloning animals using somatic cell nuclear transfer (scNT) was first successfully demonstrated with the birth of Dolly the sheep, but the process of cloning remains highly inefficient. By improving our understanding of the errors that may occur during cloned cattle embryo development, we could obtain a greater understanding of how specific molecular events contribute to successful development. The central dogma of biology refers to the process of DNA being transcribed into messenger RNA (mRNA) and the translation of mRNA into proteins, which ultimately carry out the functions encoded by genes. The epigenetic code is defined as the array of chemical modifications, or "marks", to DNA molecules that do not change the genome sequence but do allow for control of gene expression. During early development, genome reprogramming involves the removal of epigenetic marks from the sperm and egg and re-establishment of marks for the embryonic genome that code for proper gene expression to support embryo development. The point during this process at which the embryo's genes are turned on is known as embryonic genome activation (EGA). Small non-coding RNAs (sncRNAs), including microRNAs (miRNAs), may also contribute to the this process. For example, miRNA molecules do not code for proteins themselves, but rather bind to mRNAs and effectively block their translation into protein. We hypothesized that aberrant expression of sncRNAs in cloned embryos may lead to anomalous abundance of mRNA molecules, thus explaining poor

development of cloned embryos. First, we used RNA sequencing to examine the total population of sncRNAs in cattle embryos produced by *in vitro* fertilization (IVF) and found a dramatic shift in populations at the EGA. Next, we collected both sncRNA and mRNA from scNT cattle embryos, and again performed sequencing of both RNA fractions. We found that few sncRNAs were abnormally expressed in scNT embryos, with all differences appearing after EGA at the morula developmental stage. However, notable differences in the populations of sncRNAs were evident when comparing embryos by developmental stage. For populations of mRNA, we observed dramatic differences when comparing scNT and IVF cattle embryos, with the highest number of changes occurring at the EGA (8-cell stage) and after (morula stage). While changes in specific miRNA molecules (miR-34a and miR-345) were negatively correlated with some of their predicted target mRNAs, this pattern was not widespread as would be expected if these sncRNAs are functionally binding to all of the predicted mRNA targets. Collectively, our observations suggest that other mechanisms leading to altered expression of mRNA in cloned embryos may be responsible for their relatively poor development.

ACKNOWLEDGMENTS

None of this research would have been possible without the support of my major advisors. Dr. Ken White provided invaluable expertise and guidance in cloning, embryo culture, and research methods. Dr. Benninghoff mentored and supported me through my program, in both in the laboratory and professionally. I would not have been able to develop and pursue my interest in teaching without her support and example of passion and drive. I would like to thank everyone in my committee for their guidance and support including Dr. Irina Polejaeva who gave valuable advice on embryo and embryonic stem cell culture, Dr. Clay Isom for always being willing to talk through not only laboratory issues but bigger picture questions in the scientific field, and Dr. Greg Podgorski for giving me opportunities and guidance in front of a classroom.

This work would not have happened without the help of many others as well. I would like to thank Drs. Qinggang Meng and Ying Liu for their expertise and advice in embryo micromanipulation and cloning, Misha Regouski for her expertise and hard work to organize animal management and transfers, Jacob Keim for his assistance in embryo fusion, and Porter Nielson and Kurtis Jewkes for their support in the lab. I would also like to thank Dr. Rusty Stott and Dr. Alexis Sweat in their efforts in cattle transfers and many others in the White and Benninghoff laboratories who gave support over the years.

I would not have achieved any of this without my family or friends, and I would like to give special thanks to my dad, Brian Cuthbert, for his continued support through the entire

educational process, and Mike Vakula his patience and support. The help of my friends and the community in Logan meant the world, and I could not have done this without you.

<div align="right">Jocelyn M. Cuthbert</div>

CONTENTS

xii

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND SYMBOLS

Abbreviations

5mC	5-Methylcytosine
ADAR	Adenosine deaminases acting on RNA
AGO2	Argonaute protein
ARE	AU-rich elements
ART	Artificial reproductive technology
ATP	Adenosine triphosphate
BDC	Blastocyst derived cell
BER	Base excision repair
COC	Cumulus oocyte complex
CpG	Cytosine-phosphate-guanine
Ct	Cycle threshold
DMR	Differentially methylated region
DNA	Deoxyribonucleic acid
DNMT	DNA methyltransferase
EGA	Embryonic genome activation
ESC	Embryonic stem cell
HAT	Histone acetyltransferase
HDAC	Histone deacetylase
iPSC	Induced pluripotent stem cell
ICM	Inner cell mass
IVF	*In vitro* fertilization
IVP	*In vivo*-produced
LINE	Long interspersed nuclear repeat element
LH	Luteinizing hormone
lncRNA	Long non-coding RNA
LTR	Long terminal repeat transposons
MaLR	Mammalian long terminal repeat elements
miRNA	MicroRNA
MZT	Maternal-to-zygotic transition
MET	Maternal-to-embryonic transition
ncRNA	Non-coding RNA
NPC	Nuclear pore complex
nt	Nucleotide
PABP	Poly A binding proteins
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PGC	Primordial germ cell
piRNA	PIWI-interacting RNA
pilRNA	PIWI-interaction-like RNA
pre-miRNA	Precursor microRNA
pri-miRNA	Primary microRNA
PTM	Post-translational modification
RISC	miRNA-induced silencing complex
RNA	Ribonucleic acid
RNAi	RNA interference

RNP	Ribonucleoprotein complex
scNT	Somatic cell nuclear transfer
SINE	Short interspersed nuclear repeat element
sncRNA	Small non-coding RNA
snoRNA	Small nucleolar RNA
snRNA	Small nuclear RNA
snRNP	Small nuclear ribonucleoprotein particle
TE	Transposable element
tRF	tRNA fragment
UTR	Untranslated region

Symbols

C	Celsius
cm	Centimeter
DC	Direct current
g	Grams
hr	Hour
mg	Milligram
min	Minute
ml	Milliliter
µg	Microgram
µsec	Microsecond
µl	Microliter
µM	Micromolar
µm	Micrometer
ng	Nanogram
U	Units

CHAPTER 1

REVIEW OF THE LITERATURE

Classes of Ribonucleic Acids

The field of molecular biology has long been dominated by the central dogma that reading the DNA code to generate the protein machinery of the cell is a two-step process, first requiring the copying of the DNA code into messenger RNA via transcription and then converting that RNA code into functional protein via translation. However, scientists now understand that this central dogma is inadequate to fully explain the highly complex process of copying and translating DNA code to generate the necessary molecular machinery necessary for cellular function. Indeed, researchers have determined that multiple types of RNA molecules, including some types that do not code for proteins, are involved in regulating this flow of genetic information within cells. Small non-coding RNA (sncRNA) includes all RNA under 200 nt in length that are functional despite not coding for protein. Alternatively, large RNA includes molecules more than 200 nt in length, both protein coding (messenger or mRNA) and non-protein coding (long non-coding or lncRNA). RNA interference (RNAi) is the main pathway by which sncRNAs regulate the flow of genetic information, through the binding of transcripts based on sequence complementarity. Once bound to the target transcript, sncRNAs cause repression of translation and degradation. The field of sncRNA is rapidly expanding, as RNAi control of gene expression was once thought limited to microRNA (miRNA) and small interfering RNA (siRNA), but new functions in RNAi have been discovered for piwi-interacting RNAs (piRNAs), transfer RNA fragments (tRFs), and small nucleolar RNA (snoRNAs) fragments. The synthesis, function, and regulation of these classes of sncRNA, along with the classes of large RNA, share some biological processes, while differing in some key ways that allow for functional differentiation between classes.

Messenger RNA

The fundamentals of messenger RNA (mRNA) synthesis, transcription, and translation are very well understood and covered in any basic molecular biology textbook. However, there are nuances in the biosynthesis and modification of transcripts that change the regulation of mRNA molecules by sncRNA. Localization, RNA-bound proteins, *cis*-regulatory elements, and other chemical changes, especially those that impact the 3' untranslated region (UTR) of the mRNA, are important in regulating the transcript's availability for binding by miRNAs and other sncRNAs that participate in RNAi. The poly-adenylation of transcripts also has specialized function in early development, as embryos have unique pathways of degradation to protect maternal transcripts until embryonic genome activation (EGA). Due to the specialized nature of these pathways, background on traditional polyadenylation and degradation pathways will be thoroughly covered as well.

The control of protein-coding gene expression is vital to many biological processes and is highly regulated at the level of transcription. DNA transcription results in in the production of pre-mRNA molecules, which undergo splicing and process events to produce mRNA. Many of these mRNA molecules undergo the process of translation to ultimately produce proteins. Initiation of transcription begins with the sigma subunit of RNA polymerase binding to the TATA box, which is located at 35 and 25 bases upstream of the initiation site. A group of transcription factors form the pre-initiation complex, which binds to the promoter region and recruits RNA polymerase. Once bound, the RNA polymerase is released from the other transcription factors and elongation begins as pre-mRNA is synthesized from the 5' to 3' direction as the polymerase travels on the template DNA in the 3' to 5' direction. The RNA polymerase unwinds the double-stranded template DNA, and template DNA behind rewinds in a transcription bubble of about 25 unwound DNA base pairs. Because eukaryotic DNA is packaged into chromatin and nucleosomes when the cells are not dividing, a special protein complex facilitates chromatic

transcription (FACT), pulling away histones from the DNA template and replacing and recreating the nucleosomes once transcription is complete. Termination signals are different for each polymerase, with RNA polymerase I requiring a specific termination sequence that causes a termination protein to bind and block further transcription. RNA polymerase II transcribes the majority of protein-encoding RNAs, structural RNAs, and regulatory RNA genes. These genes lack a termination signal, resulting in elongation continuing beyond the gene body; the resulting tail is removed during mRNA processing, thus releasing the pre-mRNA from the polymerase as it continues to transcribe, with the continued transcription product degraded by 5' exonuclease, which disengages the polymerase from the template DNA. RNA polymerase III requires a hairpin structure termination signal to terminate transcription that is not completely understood [1, 2].

The primary transcript originating from the DNA is referred to as "pre-mRNA", which includes sequence regions destined to be translated into amino acids, located in exons – and sections that are excised and not translated – the introns. Removal of introns from pre-mRNA sequences occurs via the spliceosome, through trans-esterification reactions that recognize the splicing signals in the intron. Alternative splicing of these sequences can give rise to multiple mature mRNA sequences for a particular gene, thus greatly amplifying the message communicated via the genomic code; this process occurs in over 90% of genes in mammals. Alternative splicing takes place in the nucleus and can occur via intron retention, exon inclusion or skipping, alternative 3' splicing, or alternative 5' splicing [3]. Additional processing includes modification of the 5' end of the mRNA molecule to include a specialized structure referred to as a 7-methylguanylate cap (m^7G), a modification that facilitates transport into the cytoplasm, translation, and protection from degradation. The poly-A tail is created through polyadenylation of the 3' end, which also allows for nuclear export of the mRNA molecule, stabilizes the structure in the cytoplasm, and is involved in the initiation of translation [1].

In addition to the coding information embedded within its structure, mRNA also contains information that communicates how frequently the molecule should be translated, how long it will persist, and where in the cell it will be translated. The lifespan of an mRNA transcript greatly impacts its cellular function. As a case in point, so-called "housekeeping" genes are those that are typically ubiquitously and consistently expressed in all cell types as they are required for basic cell function and maintenance. Thus, housekeeping genes typically have longer half-lives compared to those encoding regulatory proteins [4]. A transcript's half-life is encoded in the form of non-coding genomic sequences called *cis*-acting sequences that are typically functionally relevant when located in the mRNA 5' or 3' untranslated region [1]. These *cis*-acting sequences serve as binding sites for *trans*-acting factors, RNA binding proteins, and sncRNA and can impact translation and degradation of the transcript [5]. Messenger RNA has also been found to be the subject of base modifications via chemical alterations, and it is likely that these changes impact RNA stability, base pairing, and RNA confirmation [6].

The functional forms of pre-mRNAs and mRNAs in cells are bound by proteins as ribonucleoprotein complexes (RNPs). Changes in mRNA structure can occur through modification of the nucleotide sequence through splicing [7]. These changes can dictate whether proteins dissociate and rebind later in the mRNA's lifespan or remain stably bound. The early processing of the mRNA occurs in the nucleus, after which the mRNA is exported to the cytoplasm. The nuclear pore complex (NPC) is a supramolecular assembly that acts as a gatekeeper to prevent the movement of proteins and protein RNA complexes between the nucleus and cytoplasm. Several mRNA binding proteins act as export adaptors, and they bridge the mRNA to the receptor proteins, which in turn contracts components of the NPC [8]. Once the mRNAs are in the cytoplasm, many of them are immediately translationally active. These changes in localization, bound proteins, and structure of mRNAs can significantly change the regulation of mRNA by sncRNAs, especially those that impact the 3'UTR.

Gene expression is highly regulated in a multitude of pathways, allowing for the fine-tuning of expression during both transcription and translation, as well as during RNA processing. RNA transport and localization, mRNA degradation, translational control, and protein activity all regulate gene expression. Proteins with DNA binding motifs can recognize DNA sequences and bind to base pairs in the major or minor groove. These proteins include a DNA binding domain, which guide the protein to complex with DNA at specific sequences, and a transcriptional effect domain, which impacts gene transcription [9]. More than 5% of the human genome codes for transcriptional regulatory machinery, which often function in a combinational, synergistic fashion [10]. Transcriptional activators recruit basal (general) transcription factors, the transcription-initiation complex, or factors that loosen chromatin. Transcriptional repressors can prevent the binding of activators, mask activation domains, block transcription factors, or recruit factors that tighten chromatin. Some of these activators or repressors also have an impact on the splicing of pre-mRNA, due to the close relationship between transcription and RNA processing. Regulation of transcriptional activity allows a broad range of expression levels, leading to the fine-tuning of transcription depending on the factors present [9]. Some epigenetic mechanisms that affect the level of transcription that occur are discussed in more detail below. Highly unstable species of mRNA (half-lives of 10-15 min) are present at low, steady levels. While low transcription would conserve resources, using a high rate of turnover to control levels of mRNAs allows for rapid induction of gene expression when needed and rapid loss when transcripts are no longer necessary. Transcripts with these types of rapid controls are often critical for regulatory function during developmental processes, during which period they are needed only briefly [11].

Messenger RNAs are relatively unstable, as the cell contains a multitude of RNA degrading proteins called ribonucleases. Ribonucleases cleave the phosphodiester linkage connecting the ribonucleotides, and are either endoribonucleases, which cleave RNA at internal linkages, or exoribonucleases, which remove nucleotides from a terminus in a directional manner.

Five-prime (5') exonucleases hydrolyze RNA from the 5' end, and 3' exonucleases hydrolyze

RNA from the 3' end [12]. The half-life of mRNA is a characteristic of each individual mRNA

molecule [11]. The abundance of a particular mRNA within a cell is often more dependent on the

control of decay processes than on the generation of the mRNA through transcription and RNA

processing [8]. Helicases are an ATP-dependent protein family that participate both in almost all

pathways of RNA processing and degradation. Helicases function with both exonucleases and

endonucleases, unwind secondary structure or remove RNA bound proteins, and may remain in

position while recruiting degradation machinery [13].

The degradation of the majority of mRNAs depends on deadenylation. The poly(A) tail

in mammals is about 200 nt long and is coated by Poly A binding proteins (PABPs). The poly(A)

tail is subject to gradual shortening in the cytoplasm by poly(A) nucleases. The PABPs protect

the poly(A) tail from degradation through interaction with translation factors that bridge to the 5'

cap. Eukaryotic translation initiation factor 4 (eIF4G), functions as a multifunctional adaptor

bringing together translational machinery through interaction with the 5'cap and poly(A) tail to

create a pseudo circular structure. The pseudo circular structure depends on the PABP, which

stimulates translation [2]. The bridging circularizes the mRNA molecule, which increases

translation and prevents degradation [14]. When degraded, the poly(A) tail is initially shortened

by the PAN2/3 complex, which leaves a 60-80 nt long tail. The rest of the poly(A) tail is

degraded rapidly by the CCR4-NOT complex, which contains exonucleases. Another normal

pathway for degradation is for the poly(A) tail to be digested down to 10-12 nt in length, which

then triggers decapping of the 5' end [14]. During translation, the cap is resistant to decapping

because the cytoplasmic cap binding protein is bound. In addition, non-coding RNA (ncRNA)

will target and degrade mRNA in the cytoplasm, with the specific ncRNA and the method of

function in degradation and translation discussed in depth in the next section. The constellation

of RNA bound proteins, ncRNA, and sequence signals dictate the mRNAs degradation rates, acting as a powerful regulator of gene expression [15].

In somatic cells, mRNA decay often occurs in P-bodies, which are discrete cytoplasmic loci. P-bodies are also the site for miRNA suppression, mRNA storage, and translational repression. P-bodies lack translational initiation factors or ribosomes, creating an environment of translational silence, and the number of untranslated mRNAs is correlated with the P-body number and size [16]. While P-bodies are not essential for mRNA degradation, they form in response to mRNA degradation. P-bodies contain three classes of mRNA decapping components: decapping enzymes, activators of decapping, and 5'-3' exonucleases[17]. The ratio of messenger ribonuclearproteins and the polysomes influence the rates of translation and decay. Not all mRNAs that enter the P body are degraded, and silenced mRNAs can exit the P body and reenter a translational state [16]. Messenger RNA-specific regulators control translation and degradation by recruiting the decay machinery. Deadenylation of mRNA can take place outside of P bodies as well, which induces degradation of the transcript [2].

Small Non-Coding RNA

MicroRNA

MicroRNAs (miRNAs) canonically function by binding to the 3' untranslated region (UTR) of transcripts, repressing translation; often multiple miRNAs will target a single mRNA for maximum repression. The majority of miRNAs investigated experimentally inhibit gene expression through binding interactions that block translation and decrease stability of the target transcript, but there are miRNAs known to activate translation and gene expression [18]. Facilitation of mRNA degradation is typically based on the recognition of a short sequence at the 3' region, which corresponds to a matching seed sequence within the first 8 nucleotides on the 5' end of the miRNA. If there is a high degree of complementarity within this seed sequence, there

does not need to be strong binding within other areas of the miRNA. Less commonly (<5%), strong interactions in the 3' end of miRNA can compensate for weaker interactions in the 5' end of the miRNA, which means that weak complementarity in the 3' end can still result in mRNA degradation [19]. In cases of perfect complementary binding, miRNAs will cause the degradation of their target mRNAs, mediated by the Argonaute protein AGO2 [20]. In cases of imperfect complementary binding, bulges in the double stranded mRNA/miRNA duplex prevent endonucleolytic cleavage of the RNAs [21], but result in repression of translation by inhibition of translation initiation [22]. The mechanism of translational repression involves AGO2, which binds the m^7 cap using a domain that is structurally similar to one contained in the translation initiation factor eIF4E. AGO2 may then hinder recognition of the m^7 cap in the initiation of translation [23]. This mechanism of action is supported by the finding that mRNAs that lack a 5' cap are refractory to miRNA mediated repression [22]. However, some miRNA binding is observed in actively translating polysomes, which may mean that miRNA-mediated translational inhibition results either from rapid protein product degradation or a high drop rate in ribosomes during elongation, resulting in incomplete protein products [24]. However, it appears that even non-complementary binding promotes degradation after translationally repressing the mRNA [25]. It appears that miRNA that do not bind with perfect complementarity to their mRNA targets cause degradation of mRNA, as they expedite poly(A) tail removal and cause degradation of their RNA targets to reduce translation [26]. Regardless of the specific miRNA molecule, the cell type or cell growth condition, or translational state of the mRNA, the destabilization of mRNA explains 66% to more than 90% of miRNA repression [27].

MicroRNAs promote destabilization of target mRNAs by promoting mRNA decapping and de-adenylation, leading to the localization of the target mRNA to cytoplasmic degradative processing bodies [22]. Importantly for embryonic development, non-polyadenylated mRNAs can also undergo miRNA-mediated repression, although this repression typically is not as strong

as for mRNA with poly(A) tails [22]. It also appears that the mRNA target's translational efficiency impacts miRNA binding, with high translation efficiency linked to a more robust repression by the miRNA [28]. Beyond their well-established cytoplasmic actions, miRNAs also have some nuclear functions that have not been as well investigated [29].

MicroRNAs can be transcribed from intergenic regions and from either the intronic or exonic portions of protein coding genes. The first step in miRNA biogenesis involves a nuclear protein complex, made up of Drosha and the DiGeorge syndrome chromosomal region 8 (DGCR8) microprocessor complex subunit in mammals. Drosha contains two RNase III domains for cleavage, and DGCR8 contains two double-strand RNA binding domains for recognition of the cleavage site [30]. The Drosha/DGCR8 microprocessor cleaves primary-miRNA (pri-miRNA) to generate precursor miRNA (pre-miRNA) molecules (Figure 1). Drosha is susceptible
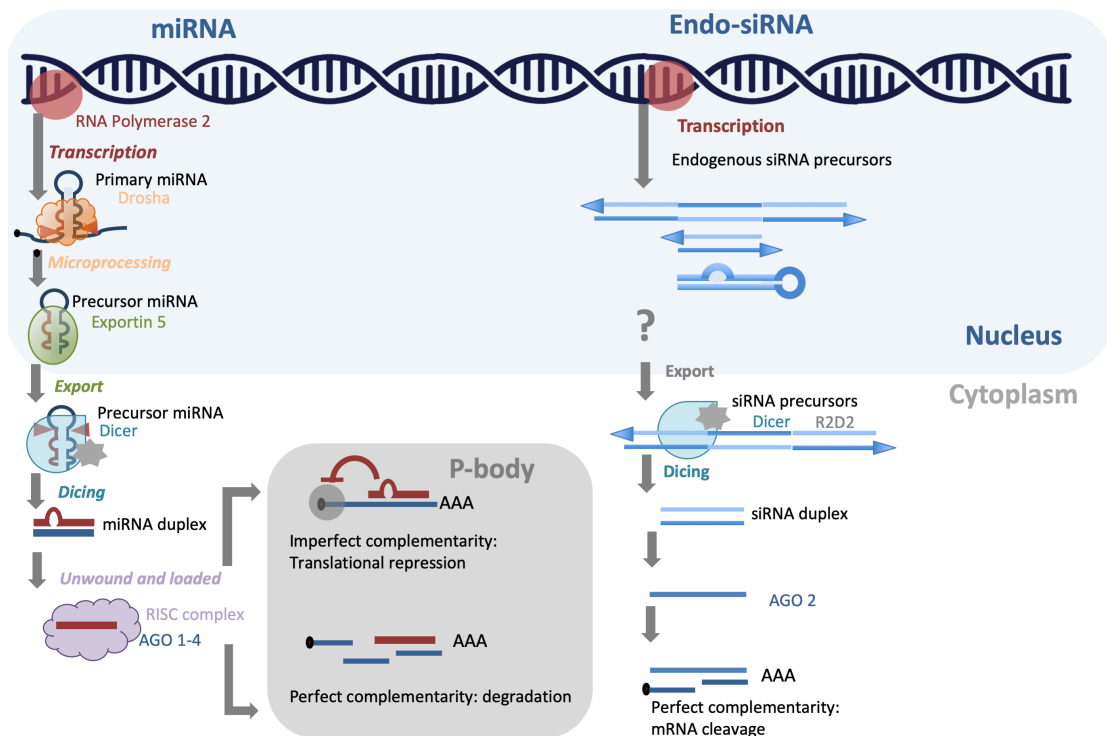
Figure 1.1. Biogenesis and actions of miRNA and siRNA. The details of the biogenesis pathways for miRNA and siRNA are discussed in the text.

to several modifications, including phosphorylation and acetylation, which leads to nuclear

localization and stabilization of the protein, respectively. Drosha binding to the protein TDP43

also causes stabilization and increased processing power. Phosphorylation of DGCR8 stabilizes

this protein, while its acetylation increases affinity for pri-miRNA. Alternatively, binding to

MECP2 can inhibit DGCR8 binding to primary miRNA (pri-miRNA). After processing,

Exportin 5 transports pre-miRNAs to the cytoplasm, where they are ultimately cleaved into

double-stranded, non-hairpin structures about 22 nt in length by the Dicer complex. Exportin

recognizes and interacts with the stem-loop and 3' 2 nt overhang on the pre-miRNA that is

created by Drosha. Unlike mRNAs, miRNAs do not have a universal recognition sequence, such

as the poly(A) tail. Dicer associates with TRBP, an RNA binding protein, as well as PKR, a

protein activator that is not required but serves to stabilize the protein-miRNA complex that

targets mRNA, the RNA-induced silencing complex (RISC) [31]. Dicer binds to the 5'

phosphorylated end, and will then cleave the pre-miRNA at specific locations, a fixed distance

from the 3' end that spans from 21-25 nt in length. The RISC complex is made of associated

proteins, most importantly Dicer and Argonaute (AGO) proteins, although the specific protein

members of RISC differ by species. The specific AGO protein that binds to RISC may be

dependent on the 5' U or C of the incorporated miRNA. AGO proteins are stabilized by miRNA

binding, and can be ubiquitylated and degraded. The AGO protein unwinds the guide strand,

which is determined by thermodynamic stability, and loads it into the RISC complex [30]. The

processed miRNAs then assemble with RISC, which directs the miRNA to its target messenger

RNA sequence. The unused passenger strand is then degraded. MicroRNA-dependent

degradation or silencing of mRNA can be prevented by mRNA binding proteins such as DND1,

which blocks miRNA binding, or DAZL, which promotes poly(A) tail lengthening. A minority

of miRNAs can also be processed independently of both Drosha and Dicer using alternative

biogenesis pathways [32]. For example, RNases such as RNase Z and Integrator can function as

Drosha, splicing in mirtrons make Drosha unnecessary, and Ago2 can replace the function of Dicer [33].

While the canonical biogenesis pathway for miRNAs includes both Drosha and Dicer processing, non-canonical pathways have also been identified. Generally, these non-canonical pathways can be grouped into Drosha/DGCR8-independent, and Dicer-independent pathways. Mirtrons are an example of a Drosha/DGCR8 independent sncRNA, and are produced from mRNA introns at splicing [34, 35]. 7-methylguanosine (m^7G) capped pre-miRNA are also able to bypass Drosha/DGCR8 processing through direct export to the cytoplasm [36]. Drosha can process Dicer-independent miRNAs, but Dicer processing requires AGO2 for complete maturation [37]. Whereas knockout of Drosha expression in mice abolishes canonical miRNA expression, many miRNAs are still present when Dicer expression is blocked [38]. Interestingly, knockout of exportin showed that this protein is dispensable for a substantial proportion of miRNA maturation [39]. While canonical miRNA pathways are generally constitutive, alternate pathways exist that may allow miRNA expression when core biogenesis proteins are reduced.

P-bodies are cytoplasmic foci that are enriched with mRNA degradation machinery, such as mRNA decapping proteins and CCR4-NOT complexes [40]. While P-bodies are not necessary for mRNA degradation, they appear to form in response to activation of these pathways. Because of their role in mRNA degradation, they were identified early on as potential sites involved with miRNA-mediation mRNA degradation and silencing pathways. The RISC complex and miRNAs localize target transcripts to the P-body for degradation. P-body remodeling after miRNA recruitment modifies the translational complex and facilitates access of the decapping complex to the cap structure to induce degradation [41]. In addition to facilitating target degradation, P-body components have also been implicated in the translational repression of targets. The P-body components GW182 and RCK/p54, when knocked down, relieve miRNA-mediated translational repression, although the components that contribute to this process appear to differ for individual

miRNAs and their targets [42]. The target mRNAs can be stored in P-bodies and released and returned to the translational machinery with the right signals [41]. P-bodies are not required for miRNA silencing as the mRNA degradation machinery exists diffusely in the cytoplasm. Thus, RNA interference pathways are still active in the absence of P-bodies, and P-bodies appear to arise as a consequence of the silencing pathways [43].

Many studies have identified miRNAs located outside of cells, with one population found in vesicles including exosomes, microvesicles and apoptotic bodies, and the other population associated with proteins such as AGO2. Like cellular miRNAs that are bound to the RISC complex, these extracellular miRNAs appear to be highly stable and are able to survive at room temp for more than four days. These extracellular miRNAs have been found in plasma, serum, cerebrospinal fluid, saliva, breast milk, seminal fluid, and ovarian follicular fluid [44]. The extracellular miRNA populations show promise as biomarkers of disease, such as ovarian cancer [45]. Extracellular miRNAs may be a byproduct of normal cellular activity, as it appears that the release of miRNAs is a highly regulated process, and exosomal miRNAs are functional in recipient cells [46]. MiRNAs are not randomly incorporated into exosomes, and a subset of miRNAs preferentially entered exosomes, with several models for sorting suggested [47]. These models include an nSMase2-dependent pathway, a miRNA motif/sumoylated hnRNP dependent pathway, 3' miRNA sequence dependent pathway, and a miRISC related pathway [48]. Support for these models includes evidence that preferentially included miRNAs have a GGAG motif in the 3' miRNA sequence, or more poly U than poly A at the 3' end of the miRNAs, and that miRISC complexes co-localize with the sites of exosomes biogenesis [48]. The exact functions of these extracellular miRNAs are an expanding area of research. While extracellular miRNAs may be of keen interest for identifying biomarkers and regulators of disease states such as cancer, these molecules have also been implicated in cross-species regulation through diet. While one such study suggested that miRNAs that are ingested could transfer to the blood, accumulate in

tissues, and exert regulation on host genes [49], follow up work has not been able to repeat these results [50]. However, recent work from Li et al. found plant miRNAs in human amniotic fluid and umbilical cord blood and demonstrated that oral gavage of pregnant mice with exogenous miRNA from the influenza virus was sufficient to elevate levels in serum and fetal liver, and the miRNA down-regulated mRNA levels of targets in these tissues [51]. The ability of miRNA to pass from cell to cell and regulate gene expression, including cross-species, opens up new possibilities for miRNA functional pathways.

When miRNAs are bound in the RISC complex, they are highly stable and can have a half-life extending out to days, allowing for accumulation at high amounts. However, specific miRNAs decay rapidly in certain environments [52]. Several sequences have been identified that regulate miRNA decay rate, including 7 nt at the 3' end [53], the seed sequence [52], and three uridines at nt positions 9-11 [54]. Interestingly, it appears that with high complementarity binding and extensive pairing through the 3' region of a miRNA with uracil tails, the mRNA target can evade silencing and cause degradation of the miRNA [55]. Four separate enzymes have been identified that degrade miRNAs, including the small RNA degrading nuclease 1 (SDN1), which destroys miRNA with high efficiency using a catalytic sequence-independent reaction proceeding in a 3'-5' direction. XRN2 is an exoribonuclease that cuts miRNA in a 5'-3' direction after facilitating miRNAs release from the RISC complex, and also acts in a sequence-independent manner. XRN2's close relative XRN1 is involved in the degradation of both the passenger and guide strand of the duplex miRNA. PNPT1 is a PNPase, acting as a 3'-5' exoribonuclease and exosome subunit that catalyzes phosphorolysis of miRNAs [52]. Modulation of miRNA levels can be controlled by degradation of the pre-miRNAs as well as mature miRNAs. The ribonuclease complex Translin/Trax (TN/TX) is a degrading enzyme that competes with pre-miRNA processing to degrade pre-miRNA in Dicer-deficient contexts [56]. Select pre-miRNAs can also be cleaved by the mammalian endoribonucleases MCP-induced

protein 1 (MCPIP1), as well as Ser/Thr protein kinase/endoribonucleases IREI-alpha [57]. Recent evidence has suggested that there are different pathways of degradation for specific populations of miRNA. For example, the RNA binding protein LIN28 initiates degradation of let-7 precursors through recruitment of terminal uridylyltransferases (TUTases) and causing oligouridylation [58]. However, LIN18 also binds to miRNA-9 and destabilizes its precursor in a poly U independent manner, and constitutive expression of LIN18 reduces let-7 but not miRNA-9. It appears that the 3'-5' exoribonuclease Dis312 also seems to contribute highly to control miRNA-9 production, and there may be multiple degradation pathways working together for one miRNA population. LIN28 may bind to many miRNAs, but unlike its destabilization action for Let-7 and miRNA-9, there are several miRNA up-regulated by LIN18 overexpression [59]. It is unknown whether there is conserved machinery for miRNA decay. Complex systems for degradation of miRNA that may act differently in different populations of miRNA make elucidating pathways of degradation a challenge, and more work is needed to understand the breadth of degradation pathways in different environments.

MiRNAs are subject to RNA tailing, which can result in multiple isoforms with different 3' ends. The 5' ends of miRNA are relatively invariable, as they contain the seed sequence and need to be able to bind to the target mRNA without interruption. The 3' end of the miRNA does not always need to bind to the target mRNA, unless the seed sequence has less sequence complementary. Therefore, the 3' end of miRNAs tends to be heterogeneous in its modifications [60]. Changes to an individual miRNA length or sequence can have implications for that miRNA isoform's target selection, miRNA stability, and RISC loading [61]. Adenosine deaminases acting on RNA (ADAR) can convert adenosine to inosine in segments and target many pri- and pre-miRNAs. The changes in sequence affects Drosha and Dicer processing and can prevent miRNA export [62]. Because tailing is rare with as few as 5% of miRNAs have these modifications, tailing may not be a piece of normal miRNA functionality. However, specific miRNA species

can be highly modified leading to functional consequences under certain conditions. In mammals, U/A tailing of miRNAs is induced through high complementary binding to targets, which induces trimming or decay of both the target and miRNA through several enzymes, one of which may be the 3'-5' associated exonuclease DIS3L2 [63]. Viruses have been shown to hijack this degradation pathway in order to degrade host miRNA during infections. Mammalian GLD2 functions in polyadenylation of mRNA as well as adenylating miRNA. In contrast to A tailing induced by binding, this adenylation appears to stabilize the miRNA [64]. Also, mono-uridylation of pre-miRNA promotes processing via DICER and increases populations of mature miRNA, while oligouridylation blocks DICER processing and facilitating decay of pre-miRNA, thus reducing populations of mature miRNA [65]. Several species of pre-miRNA require 3' end mono-uridylation for Dicer processing due to a short 1 nt 3' overhang from Drosha processing [66]. In activated Drosophila eggs, over 30% of miRNAs carry a non-template A addition, and these levels drop in the hours after egg laying. Comparatively, in sea urchin embryos up to the 32-cell stage, adenylation reached levels over 30% in miRNAs [67]. Also, in mouse metaphase II stage oocytes, adenylation of miRNAs reaches 30% and drops as the embryo develops [67]. These modifications included mono-, di- and tri-adenylation and were carried out by the Wispy protein. Because adenylation takes place on both the 5' and 3' ends of miRNA, the modification must take place downstream of DICER processing. It also appears that there is no sequence motif conserved among adenylated miRNAs, and therefore adenylation may not be driven by intrinsic sequences. It was demonstrated that Wispy associated transiently with AGO1, which makes up part of the RISC complex that aids in miRNA binding to targets [67]. Therefore, the adenylation process may be triggered by miRNA activity, as opposed to a particular sequence. When a miRNA binds to a target with high complementarity, the PAZ domain releases the 3' end of the guide miRNA. The release of the 3' end may make that miRNA susceptible to adenylation via Wispy, and physical separation from Wispy may protect miRNAs from adenylation. Adenylation

may act as a molecular basis for the clearance of maternal miRNAs through destabilization and degradation during early embryogenesis, discussed in more depth below [67]. To complicate this theory, polyadenylation has a stabilizing effect and we do not have a clear understanding of how these similar mechanisms have such context-dependent outcomes. When considering miRNA adenylation, the miRNA molecules that are expressed in early embryogenesis and decline after the EGA have the highest levels of adenylation as compared to those expressed at or after the EGA [67]. The conservation between such distinct model species would predict that it might be a mechanism of clearing maternal miRNA, although timing varies.

Endogenous small interfering RNA

Endogenous small interfering RNAs (endo-siRNAs) have been found in plants and worms, and generally are produced by RNA-dependent RNA polymerases (RdRPs), which are lacking in mammalian genomes. However, recently diverse substrates that generate siRNA were found in *Drosophila* and mouse oocytes [68]. Small interfering RNAs exert a repressive effect on mRNA, similar to that of miRNA, through the RISC complex and AGO proteins. Small interfering RNAs bind with perfectly complementary to target sequences, which then induces endonucleolytic degradation [68]. SiRNA molecules are exactly 21 nt long, present in both sense and antisense orientations, with a 3' modified end. One defining factor that distinguishes siRNA from miRNA and piRNA is that siRNA molecules are not biased towards uracil in the first nucleotide position on the 5' end[69]. In the *Drosophila*, the genomic sequence coding for siRNA transcripts are derived from transposons, heterochromatic sequences, intergenic regions, long RNA transcripts, and mRNAs. In mice, siRNAs have also been found to originate from psuedogenes. Ghildiya and coworkers showed that pseudogene transcripts can generate siRNA by annealing to their functional transcripts, and therefore pseudogenes may suppress their functional counterparts through siRNA degradation [69]. Endo-siRNA has been studied in depth

in plants and *C. Elegans,* but the majority of that information does not translate to mammals due to difference in sncRNA bioprocessing pathways. In mammals, endo-siRNA has only been found in mouse oocytes, which may be a non-conserved function due to a rodent-specific Dicer isoform that lacks the N-terminal helicase domain and efficiently generates siRNA [70]. Likely due in part to the limited known expression of siRNA, little information is currently available with respect to the persistence or degradation of siRNAs.

The production of endo-siRNA precursors can occur by sense-antisense pairs, which are derived from transposons [31]. siRNAs are processed from fully complementary, long double stranded RNA. In both mice and *Drosophila*, a siRNA precursor can also be a single stranded RNA that self-hybridizes in a similar fashion to miRNA but with an extended stem length. It does not appear that these precursor molecules require processing in the nucleus, and the protein responsible for exportation from the nucleus to the cytoplasm is unknown [31]. The biogenesis pathway for endogenous siRNA is similar to that for miRNA, but evidence from the studies in *Drosophila* indicate that production of siRNA requires Dicer 2 (DCR-2) as opposed to Dicer 1 (DCR-1). In DCR-2 knockout studies, a subpopulation of siRNA persists, which may be due to the activity of a non-canonical pathway similar to what has been found in miRNA [69]. These siRNAs produced via a non-canonical pathway require DCR-2 to bind with Loquacious (LOQS). In a more normal setting, R2D2 will bind to DCR-2, and LOQS will bind to DCR-1. When siRNA is introduced to the cell from an exogenous source, such as via viral dsRNA, similar processing occurs save for DCR-2 binding to R2D2 [31]. Also, for processing of exogenous siRNA, R2D2 is required for loading the siRNA onto AGO2. However, while it is known that endo-siRNA acts through AGO2, the extent to which R2D2 is needed for loading is still unknown. Due to the assumption that mammals did not express endo-siRNA due to the lack of RdRP, and the more recent discovery of these alternative biogenesis pathways mammals, less is known about the biogenesis of endo-siRNAs than many other sncRNA species reviewed herein.

Piwi interacting RNA

Piwi-interacting RNAs (piRNA) are single-stranded RNA molecules associated with the PIWI subfamily of proteins. Most of these proteins are specifically expressed in the germ cell line, but not in somatic tissues. PiRNAs are highly conserved class of sncRNA about 21-34 bp in length. In germ cells, piRNAs act as a type of immune system to protect the germ cell genomes from transposons, but this mechanism is silent in somatic cells [71]. Piwi-interacting RNAs bind RNA targets based on sequence complementarity and repress targets via slicing or chromatin mark recruitment. PiRNAs repress transposable elements (TEs) that pose a threat to genomic stability during genomic remodeling periods in both a transcriptional and post-transcriptional manner [72]. Piwi-interacting RNAs were originally found to function in gametogenesis and silence TEs during reprogramming events throughout primordial germline cell development [73]. During pachytene spermatogenesis, piRNAs target not only TEs, but also target mRNA for degradation, which suggests that piRNAs may function to suppress gene expression in other cellular functions as well [74]. It has been shown that transcripts derived from TEs are expressed in bovine embryos at the point at which the embryonic genome is activated, or the maternal-to-embryonic transition (MET) [75]. Therefore, piRNAs may be targeting and suppressing TEs during early development, but not maintain this silencing during the EGA. In mammals, piRNAs have also been implicated in directing transposon methylation in primordial germ cells (PGCs) during germ cell development. Knockout studies of piRNA have correlated loss of piRNAs with elevated transposon activity, compromised histone methylation, diminished *de novo* DNA methylation, and impaired control of translation [76, 77]. While piRNAs generally are present only in early embryos and germline cells, they have been found to be present in cumulus cells, which are somatic differentiated cells that are highly associated with the oocyte [78]. In *Drosophila*, piRNAs also function in genome maintenance, telomere protection [79], maternal mRNA deadenylation and decay [80] and suppress phenotype variation via transposon silencing.

PiRNAs could be responsible for the deposition of DNA methylation over a large proportion of the genome through directing methylation to unmethylated transposons they target, although the mechanism behind this DNA methylation via piRNA is still unknown [81, 82]. PiRNAs are expressed during periods of epigenetic reprogramming when methylation and other silencing marks are removed, such as during the reprogramming of PGCs. With low methylation, active repeats are transiently reactivated and converted into a primary pool of piRNAs. When the wave of *de novo* methylation is engaged, the relative abundance of these transcripts is reduced, but a fraction of retrotransposon copies are able to evade this silencing and stay transcriptionally active by mirroring the activity of protein coding genes [83]. These transcripts then engage in secondary piRNA amplification, which can contribute to the second wave of specificity in the second wave of *de novo* methylation in PGC reprogramming [81].

Different piRNAs specifically target different types of TEs, which are mobile genomic elements that can integrate into the genome at different sites. DNA transposons are full-length autonomous elements that can encode the protein transposase, which allows an element to be removed from one position in the genome and inserted at another [84]. Transposons generally have short inverted repeats at each end and can move if activated. This capacity for mobility poses a genomic threat; for example, long interspersed nuclear elements (LINEs) encode for a reverse transcriptase and comprise up to 20% of the human genome [85]. An old lineage of LINES, LINE2, stopped relocation in the genome before the evolutionary radiation of mammals. However, a newer LINE family, LINE1, is still active with evidence for genomic insertion in mammals [86]. Short interspersed nuclear repeat elements (SINEs) comprise up to 15% of the human genome and rely on the reverse transcriptase from LINEs for movement [87, 88]. Different mammalian lineages have evolved SINE elements independently, with primate genomes containing Alu elements and mouse genomes containing B1 and B2 type elements. The insertion process of LINEs and SINEs into the genome causes a short (7-21 nt) sequence to be

repeated, with one copy at each end of the inserted sequence. Alu elements accumulate

preferentially in GC rich areas, and L1 accumulate preferentially in GC poor regions [89]. Long

terminal repeat (LTR) transposons are characterized by the long terminal repeat on each end of

the transposon that is several hundred base-pairs long [90]. Some such elements are endogenous

retroviruses, which are cousins of retroviruses, such as HIV, and are unable to survive outside of

the cell. In human genomes, none are currently known to still be active. However, active

mobility of non-autonomous repeats are still seen in the mouse genome, referred to as

mammalian LTR elements (MaLR) [91].

While piRNAs may act similarly to miRNAs, they are produced via a distinct biogenesis

pathway (Figure 2). The genomic source of piRNAs and the precise transcript marks required for

piRNA production are unsolved questions in the field. Genomic loci-producing piRNA are

termed piRNA clusters, due to the high density of piRNA sequences mapping to those locations.

In mammals, piRNA precursors are indistinguishable from canonical euchromatic RNA pol 2

transcription units, and they are often produced from primary transcripts derived from active

transposable elements or repetitive loci [71]. PiRNA biogenesis differs from miRNA and siRNA

in that the process employs single-stranded RNA (ssRNA) as a precursor molecule, that piRNA

molecules do not have secondary structures, and DICER is dispensable to piRNA processing.

Precursor piRNAs require endonucleolytic cleavage of their 7-methylguanosine cap to yield a 5'

monophosphate. Precursor piRNA molecules must be long, ssRNA with a 5' monophosphate cap

in order to enter the piRNA biogenesis pathway [71]. There are two pathways of biogenesis for

piRNAs with the correct 5' ends. The first pathway is primary piRNA biogenesis, from which

piRNAs originate from the genomic sequence and then bind with complementarity to TE

transcripts, thus acting as guides for PIWI proteins to target and degrade those TE transcripts. In

Drosophila, the nuclear primary piRNAs are processed into mature cytoplasmic primary piRNAs.

The RNA helicase Armitage resolves any secondary structures, and then the piRNA are cleaved

by Zucchini to generate pre-piRNAs with the 5' monophosphate. The 3'-5' exonucleases Nibbler

trims the pre-piRNAs to a final length once they are loaded onto PIWI proteins. Hen1 increases

PIWI binding affinity and piRNA stability, and methylates the 2' hydroxyl group at the 3' end.

The piRISC complex migrates to the nucleus, and represses transcription of targets by

establishing a H3K9me3 chromatin state in the DNA producing the transposon targets [92].

These primary piRNAs may be one initiator of the ping-pong cycle, but maternally inherited

piRNA populations also appear to be important [93]. The details and exact proteins involved

differ in somatic and germ cells [94]. The second pathway, the ping-pong cycle, begins with the
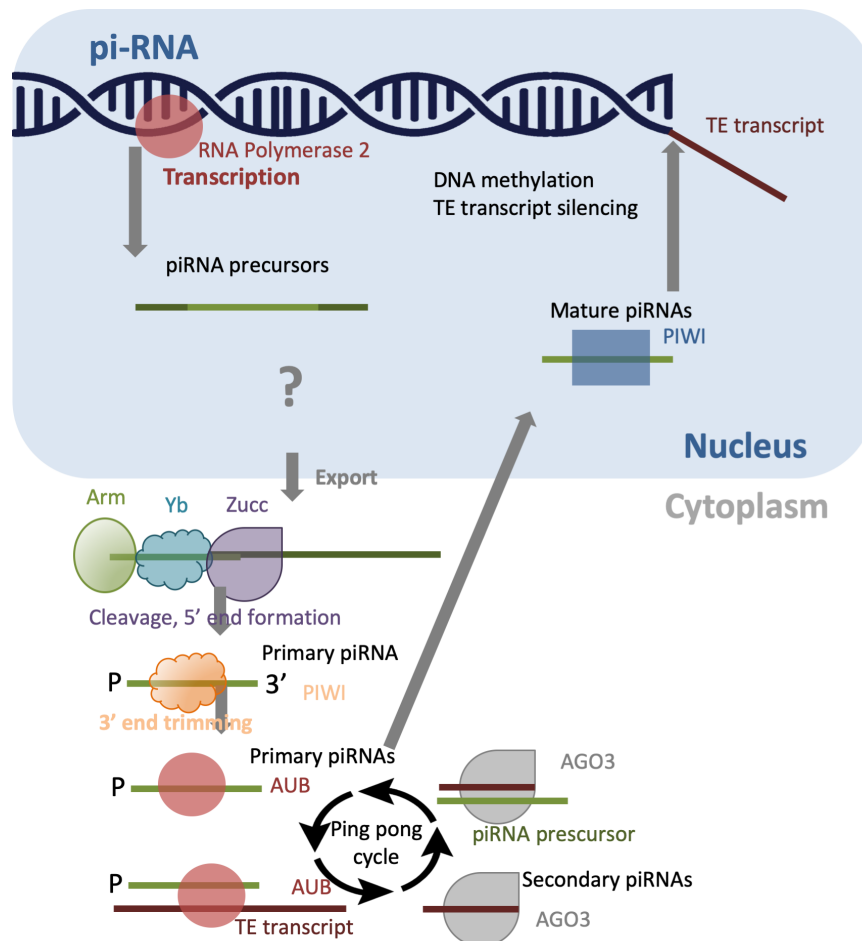
Figure 1.2. Biogenesis and actions of piRNA. The details of the biogenesis pathways for piRNA are discussed in the text.

PIWI protein Aub guided by an initiator piRNA (antisense), which cleaves a complementary

transposon transcript. The cleavage produces a pre-pre-piRNA, and once the PIWI protein

AGO3 binds that product, it is committed to produce a responder piRNA at the 5' end (sense).

The pre-piRNA intermediate product of this process is too long and requires cleavage by AGO at

the first 10 nt that are complementary between the responder piRNA and the initiator piRNA.

The resulting piRNA can itself then begin to act as an initiator piRNA that produces new

responder piRNA identical to the original initiator piRNA [71]. The primary pathway processes

many piRNAs from a single precursor transcript, and the secondary pathway utilizes

amplification of a single piRNA, which ensures the piRNA quantity is related to the expression of

their actual target transposons that are active in a cell at any given time [95].

Transfer RNA

Transfer RNA (tRNA) is a key player in the deciphering of the genetic code. In the

process of translating the genetic code, tRNA functions as the carrier of incoming amino acids for

the growing polypeptide chain. Specific transfer RNAs chemically link to the particular

individual amino acids, are then recognized by an amino acyl-tRNA synthetase, and base pairs

with the correct corresponding codon sequence in the mRNA [96, 97]. Transfer RNA molecules

are 70-80 nt in length and form precise three-dimensional structures resembling a clover leaf in

their stem-loop arrangements [97]. These structures dictate the function of tRNAs. Four stems

form by Watson Crick base pairing, creating helices with loops on the end of 7-8 bases. The

anticodon, where the tRNA binds to the mRNA, is made up of three nucleotides on the end on

one of the loops; this anticodon can bind to more than one mRNA codon through nonstandard

base pairing [98]. Transfer RNAs can undergo structural modifications, with methylation of

tRNA acting to stabilize and prevent cleavage [99]. There is growing evidence that tRNA

abundance is actively regulated through transcription and degradation pathways, enabling the cell

to coordinate translation efficiency and gene expression. The correlation between tRNA availability, specific codon used, and efficiency of translation have demonstrated that codons recognized by abundant tRNAs are most efficiently translated. Tumor tissues express aberrant tRNA profiles, which correlated with the codon usage of genes during proliferation. It is possible that there is a pathway by which cells are able to up-regulate the tRNAs needed for the mRNAs that are being transcribed. Interestingly, the population of particular tRNA molecules aberrant in cancer cells is correlated to the codon usage of genes during proliferation. The profile of differentially expressed genes in those tissues may be due to the cell up-regulating the tRNAs necessary for efficient translation of proliferation genes [99]. It is possible that there is a pathway by which cells are able to up-regulate the particular tRNAs needed for the mRNAs that are actively transcribed [99]. Because tRNAs are a highly-folded, modified structure, they are one of the most stable RNAs in the cell with a half-life in the range of several days. Control of tRNA abundance is achieved through multiple nuclear surveillance mechanisms, such as degradation of precursor-tRNAs by the TRAMP complex, which polyadenylates the tRNAs 3' end and triggers 3' degradation by the 3' exoribonuclease of the nuclear exosome. The other method of surveillance is called rapid tRNA decay. Under modified tRNA or tRNA with deleterious mutations are degraded by this pathway. While the enzymes involved in this pathway are still being elucidated, CCAase adds the CCA to the 3' end of tRNA to selectively mark unstable tRNA, which likely allows access of Rrp44 and Xrn1 exonucleases, initiating degradation [13, 100]. Degradation of tRNA can also occur as a response to stress, and it is endonucleolytically cleaved by angiogenin in mammalian cells. The cleavage under stress can lead to tRNA degradation and the production of tRNA fragments [101].

The biogenesis of tRNA is rather different than pathways for other classes of sncRNAs, as tRNAs require substantial modifications post transcription (Figure 3). The precursor tRNA undergoes 5' cleavage via RNase P and removal of the 3' extension by tRNase Z, followed by addition of CCA sequence at the 3' terminus by CCA nucleotidyltransferase. Before the tRNA is fully mature, other posttranscriptional editing and modification of many of the nucleotides need to occur. Up to 20% of the nucleotides in tRNA are modified, and these modifications prevent degradation of pre- and mature tRNA, modulate efficiency and specificity, and alter translational speed. These modifications include C to U editing, A to I editing, dihydrouridine, pseudouridine, thiouridine, isopentenyladenosine, and methylation. Modifications in different positions alter tRNA identity. Changes to the main body structure impact folding and stability of the tRNA molecule, while modifications to the anticodon loop regulate correct codon usage and translation
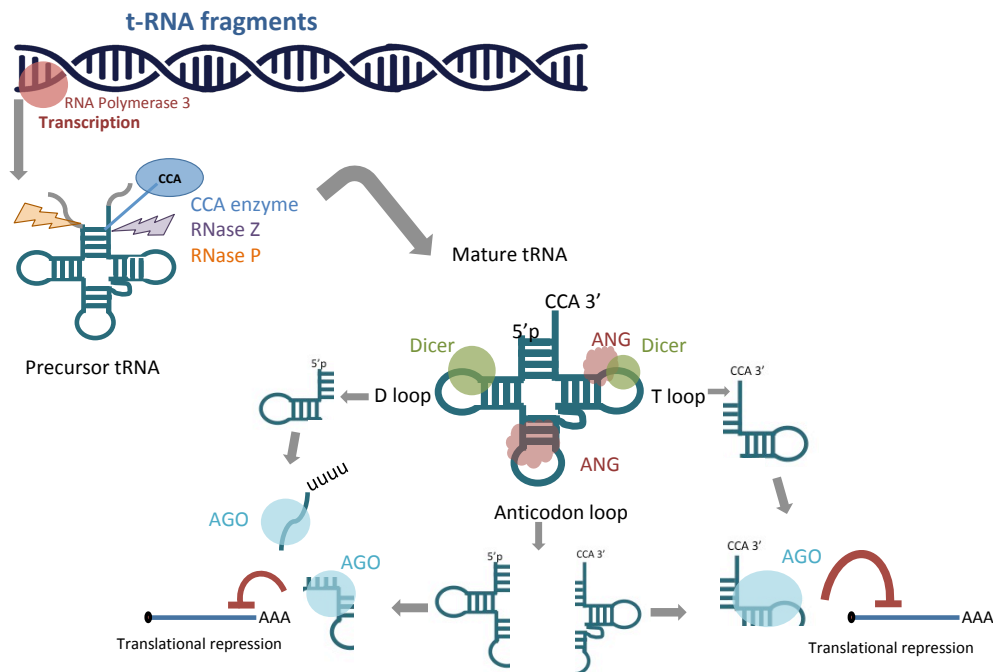
Figure 1.3. Biogenesis and action of tRFs. The details of the biogenesis pathways for tRNA fragments are discussed in the text.

[102]. Mature tRNAs must be exported by the protein Exportin from the nucleus into the cytoplasm to function in translation. Though some tRNAs are directed to the mitochondria, the molecular process regulating this targeting is not fully understood [103].

In addition to the basic housekeeping function of tRNAs for facilitating translation, these molecules also serve other functions as non-coding RNAs, which can be broadly classified into two main groups, tRNA-derived fragments (tRFs) and tRNA halves (tiRNAs), with characteristic sizes, biogenesis pathways, nucleotide compositions and cellular functions for each. tRFs are 16 to 28 nt fragments derived from mature or pre-tRNA molecules and are classified by the location of their origin (e.g., tRT-5 originates from the 5' end D loop). tiRNAs are typically 30 to 35 nt in length and originate by splitting the tRNA structure into 3' and 5' halves by cleavage within the anticodon loop. Recently, researchers determined that tRFs can participate in RNAi based on sequence complementarity. These tRFs were discovered in prostate cancer cells, in which a significant number of small RNA sequences were found to match processing from the 5' or 3' ends of tRNAs [104]. In cancer cells, it was found individual tRFs were up-regulated and controlled cell proliferation [104]. Similar to piRNAs, tRNAs may prevent transposon reactivation. Schorn et al. determined that fragments of tRNA, which included the 3' terminal CCA of mature tRNA, were abundant in mouse pre-implantation stem cells [105]. The major targets of these fragments appeared to be the two most active ERV families, which were strongly inhibited. The 22 nt-long tRNAs targeted coding-competent ERVs through post-transcriptional silencing, and the 18 nt-long tRNA interfered with transcription and the mobility of retro-transposons [105]. These tRFs and tiRNAs have been found to have major functions in numerous cancers, as well, including lung cancer, colorectal cancer, prostate cancer, breast cancer, ovarian cancer, B-cell lymphoma, etc. [106].

Small nuclear RNA

Small nuclear RNAs (snRNA) are highly conserved sncRNAs involved in splicing mRNA, performing most of the roles in the splicing reaction including catalysis. SnRNA are highly abundant, with two different classes that function in the nucleoplasm, both with a 3' stem loop. The sm-class of snRNAs are characterized by a 5' trimethylguanoside cap, binding sites for sm proteins and include U1, U2, U4, U5, U7, U11, and U12. snRNAs belonging to the lsm-class contain a monomethylphosphate cap, binding sites for lsm proteins and include U6 [107]. In their natural state, snRNAs exist as small nuclear ribonucleoprotein particles (snRNPs) with one snRNA and several proteins [1]. These snRNPs interact with a multitude of proteins to form the large complex called the spliceosome, which functions to splice introns from primary genomic transcripts [108]. Five types of snRNAs (U1, U2, U4, U5, and U6) account for about 25% of the spliceosome and are necessary for the RNA-RNA and RNA-protein interactions that constitutes its structure [1]. SnRNAs generally have a modified, hypermethylated cap structure thought to protect against decapping and degradation [13]. snRNAs are also highly 2'-O-methylated and pseudouridylated. During the assembly of the spliceosome, the snRNPs are recruited to the pre-mRNA substrate, creating the RNA-RNA duplexes. The interactions of these sites help to position the reacting groups for the first and second steps of splicing. U6 also may participate in the formation of the spliceosome active site [108]. The U1 snRNP recognizes the 5' splice site through base pair interactions and forms the commitment complex. The U2 snRNP binds to the branch site and forms a pre-splicing complex. The branch site bulges out and made available for the first chemical reaction, and the paired U4 and U6 join the complex creating complex B1 and then a series of RNA-RNA interaction rearrangements form the spliceosome. The spliceosome generates the 5' exon and the 2/3 lariat intermediate for splicing. The second chemical reaction releases the lariat intron and ligates the exon to release the mRNA product. All of the snRNAs are extensively post-transcriptionally modified, especially with pseudouridine. The modifications

on the 5' end of the U2 snRNA have been shown to be required for both snRNP biogenesis and pre-mRNA splicing [108]. The pseudouridines are almost always clustered in functionally important regions, and so have been of great interest [109]. Assembly-defective snRNAs are degraded by *Rrp6*- or *Dcp2*-dependent decapping and 5' to 3' decay [110]. The U1 species is under specific quality control, as truncated forms of this snRNA form a unique SMN complex and are diverted from the normal biogenesis pathway to processing bodies for degradation [111]. The half-life of snRNA is dependent on the type, with U1-U5 being highly stable. U6 has a shorter half-life, of about 15 hours and the two major snRNA precursors in the cytoplasm (U1 and U2) have a half-life of about 20 minutes [112].

SnRNAs also undergo specialized processing through their biogenesis pathway. The sm-class of snRNAs are transcribed by a specialized form of RNA polymerase 2, whereas the lsm class of snRNA are transcribed by polymerase 3 with specialized external promoters. In the sm-class biogenesis, RNA polymerase 2 directs the 3' end formation through a complex with Nrd1. In the lsm class, RNA polymerase 3 holds the mature 5' ends at the transcription initiation sight, and trims the 3' end of the snRNA [107]. In higher eukaryotes, the lsm class of snRNA never leaves the nucleus, while the sm-class, biogenesis requires distinct subcellular localization. sm-class snRNAs are exported by PHAX and exportin 1 to the cytosol where they undergo specific maturation steps, and are then imported back into the nucleus by import receptor importin-beta [107].

Small nucleolar RNA

Small nucleolar RNAs (snoRNAs) are named for their localization in the nucleolus, the nuclear subdomain that assembles ribosomal subunits in eukaryotic cells. However, as further functions for snoRNAs are elucidated, a range of localization that fits the function and target of individual snoRNAs has been found [107]. SnoRNAs are highly conserved and required for the

processing and modification of ribosomal RNA (rRNA). There are hundreds of types of snoRNA in vertebrate genomes, and snoRNAs are transcribed from poly-cistrons, individual genes, and introns of their host genes. Post-transcription, snoRNAs go through complex processing and maturation steps [1], and the majority of snoRNAs have specific sequences that are complementary to other RNAs [113]. SnoRNAs generally have a modified, hypermethylated cap structure thought to protect against decapping and degradation [13]. SnoRNAs guide modifying enzymes to produce site-specific associated phosphorylation, and 2'O methylation on target RNAs. Two main classes of snoRNAs have been defined: those with a C/D box structure and those with the H/ACA box structure. The C/D box class of snoRNAs include molecules 60-200 nt in length that function primarily to methylate ribosomal RNAs (rRNA). The C and D boxes are positioned at the 5' and 3' ends, respectively; these structural motifs come into close proximity in the folded molecule, which serves as a binding site for core box C/D snoRNA proteins [107]. SnoRNAs that belong to the H/ACA box class are longer at 120-250 nt and are required for pseudouridine formation in the rRNA by forming hairpins to act as guides [107]. Both classes of snoRNAs are highly conserved in eukaryotes, and both bind with protein partners to form small nucleolar ribonucleoprotein complexes, which are highly conserved as well [114]. In higher organisms, non-canonical snoRNAs that lack specific boxes, contain both boxes, are shortened, or are lengthened are expressed as well, and these snoRNAs may modulate pre-mRNA splicing or other regulatory roles [115]. Another recently discovered pathway for these non-canonical snoRNAs is processing to shorter ncRNA that act in RNAi pathways [116].

RNA Polymerase 2 is responsible for transcribing snoRNA from DNA, and co-transcriptional assembly of inactive pre-RNPs occurs to create a stable, inactive precursor snoRNA. The snoRNA undergoes functional maturation at Cajal bodies, where RNA modifications and formation of functional complexes occur. From the Cajal bodies, snoRNAs are localized to the necessary location appropriate for their specialized function [107]. SnoRNA can

also be produced from introns, so once the degradation of introns begins with de-branching

activity, biogenesis in those introns encoding snoRNAs begins. SnoRNAs are not

polyadenylated, and the box C/D snoRNAs form a closed loop, while H/ACA snoRNAs form two

stem loops linked by the H box motif [13]. SnoRNA fragments are processed in a manner

somewhat similar to miRNA processing, requiring processing from an RNase 3, the same enzyme

type as Drosha and Dicer.

In addition to their traditional housekeeping roles as guide RNAs for the post-

transcriptional modification of rRNAs, new evidence suggests that snoRNAs may be involved in

cell fate determination and oncogenesis [117, 118]. SnoRNAs undergo extensive processing, and

it appears that some snoRNAs are processed to generate stably accumulating fragments [114].

Multiple groups have reported that these snoRNA fragments associate with AGO proteins and act

in RNAi pathways similar to miRNA [118-121]. It appears that these so-called sno-miRNAs may

then target transcripts, controlling processes involved in cell behavior and carcinogenesis. As a

relatively newly identified pathway, there is still much to be elucidated on the exact function of

snoRNA fragments acting in RNAi, and what conditions regulate the pathway.

Ribosomal RNA

Ribosomes are macromolecular structures comprised of four ribosomal RNAs (rRNA)

and multiple coordinating proteins that are responsible for catalyzing protein synthesis. For every

protein coding gene produced, 10 ribosomes must be produced, on average. Thus, rRNA is

highly abundant in cells. There is growing evidence for additional functions of rRNA in the

regulation of translation efficiency, mRNA recruitment, and facilitation of ribosomal shunting

[122-124]. The investigation of mammalian rRNA functions have not been to the same level as

seen in yeast and bacteria, due to difficulties in expressing modified rRNAs and performing

functional analysis [125]. The mature 80S ribosome has two subunits, the 40S subunit (contains

the 18S rRNA) and the 60S ribosomal subunit (contains the 28s, 5.8S, and 5S rRNA). Messenger RNA passes through an mRNA channel in the ribosome, and tRNA reads the codon at the ribosome interface. The interface side of the 60S subunit contains the peptidyl transferase center (PTC) is where the peptide bond forms, and the three tRNA binding sites are located in the cavity between the two ribosomal subunits. Studies have suggested that mRNA rRNA base pairing can affect initiation of mRNA translation [126].

Ribosome assembly, similar to snoRNA, occurs co-transcriptionally on rRNA genes. The process is influenced by cell and organism growth, and requires precise coordination of rRNA and the hundreds of proteins necessary for ribosomal assembly. A dedicated set of proteins, including RNA polymerase 1, are used to transcribe rRNA, giving rRNA the ability to be regulated independently from the rest of the genome. The rRNA genes exist in tandem arrays of several hundred copies, and these genes encode the three largest RNAs of the ribosome. The formation of the 40S ribosomal subunit requires four rRNA cleavage steps, two each on the 5' end and 3' ends. [127]. The ribosomal proteins form these key secondary structures at the same time as maturation and folding of pre rRNA. The formation of the 60S subunits starts with the cleavage of nascent pre-rRNA at an A2 site, once transcription has proceeded. Once transcription is completed, a precursor particle is present, containing $27A_2$ and many of the r proteins and assembly factors for the large subunit. The association between r proteins and pre-ribosomes are strengthened as pre-rRNA is processed as the large subunit assembly proceeds in a hierarchical fashion [128]. The large subunit and small subunit undergo separate biogenesis pathways in the nucleolus and nucleus, and then both are exported and undergo final maturation steps in the cytoplasm. The rRNAs are extensively modified during the transcription and maturation, and differences in modifications may control translation and impact gene expression [129].

Epigenetic Control of Gene Expression

The term of epigenetics was first defined in the 1940s by Conrad Waddington as "the branch of biology which studies the causal interactions between genes and their products which bring the phenotype into being" [130]. Over time, this definition of has evolved to more specifically include changes in gene function that are mitotically and/or meiotically heritable and that do not entail a change in the DNA sequence. More recently, the study of epigenetics has become more inclusive, by encompassing those mechanisms that regulate gene activity via modification of DNA structure but that do not alter the genomic sequence, whether or not such changes are heritable. Epigenetic marks are crucial protectors of genomic stability, through the silencing of TEs, centromeres, telomeres and other highly repetitive, deleterious regions of the genome [131]. Some of the enzymes involved in these epigenetic modifications have been well characterized, such as those responsible for establishing the DNA methylation code or those that write or erase various histone modifications. Collectively, the full complement of specific epigenetic modifications cooperates to allow for the timely expression of only a fraction of the genome, which confers specific structure and function to the various cells of the body.

DNA methylation

The process of DNA methylation is well studied, and mechanisms controlling the establishment of the methylome are highly conserved among multicellular organisms. DNA methylation involves the addition of a methyl group (CH_3) to the fifth carbon of a cytosine residue [132]. In mammals, this epigenetic mark is mostly directed at CpG sites, which tend to be clustered in high CpG density islands. However, CpT and CpA sites are also methylated, though rarely, in oocytes, early developing embryos, and stem cells. Methylation of the DNA is highly enriched in most non-coding regions, imprinted genes, regions with high repeats, and transposon sequences. The transfer of a methyl group to a CpG site is catalyzed by DNA methyltransferases,

of which three types are well known. DNMT 1 functions in maintenance methylation, meaning

that this enzyme is responsible for copying the methylation code during the process of DNA

replication. DNMT1 preferentially methylates hemi-methylated strands of DNA following

replication, and when bound to an unmethylated CpG at the replication fork, it forms an intrinsic

auto-regulatory loop [133]. DNMT3a and 3b are very similar enzymes that primarily catalyze *de*

novo DNA methylation by adding methyl groups to CpG sites where there was no methylation

present. Thus, these isoforms are responsible for the proper establishment of the cytosine

methylation profile during development by facilitating epigenetic control of developmentally

critical processes, such as cellular differentiation, transcriptional regulation, X-inactivation,

imprinting and genome stability.

Approximately 50% of the mammalian genome contains CpG islands, and 5-

methylcytosine (5mC) constitutes approximately 1% of bases [132]. The majority of gene

promoter regions (~70%) reside within a CpG island [134], and housekeeping gene promoters are

often within a CpG island [135]. CpG methylation at these sites can block the start of

transcription, as well as interfere with recognition sites for transcription regulators, such as

CTCF. The presence of CpG islands appears to promote gene expression, as one common feature

of CpG islands is a lack of nucleosomes compared to other areas of DNA [136]. Methylation

within gene bodies also occurs and may stimulate translation as well as impact splicing [137].

However, the mechanisms behind DNA methylation as an activating phenomenon are not

currently well understood. For example, when the methylation within the gene body of repetitive

DNA elements, such as transposons, blocks transcriptional initiation of those elements while still

allowing for transcription of the host gene in which the repetitive element is located. About 45%

of the mammalian genome consists of these transposable and viral elements that must be silenced

via methylation [132]. Specific patterns of DNA methylation also influence silencing of gene

expression through association with chromatin binding proteins; these interactions may occur

through occur through trans-regulation via specific binding motifs or inhibited binding. Specific proteins have high affinity for methylated cytosines, such as MeCP2, which when bound recruit histone modification proteins, such as histone deacetylases (HDACs). Thus, DNA methylation patterns can influence the application of covalent modifications to histones, further signaling for repression of gene expression [138]. Alternatively, chromatin modifications can impact DNA methylation, as the pattern of histone modification can profoundly influence the ability for a cytosine to become methylated. DNA with the H3K4me2/3 mark is unable to be methylated, and the histone variant H2AZ has a strong correlation with unmethylated DNA [139, 140]. This trans-regulation between DNA methylation and chromatin silencing may be necessary, as there are highly repetitive regions of the genome that can become very mutagenic when unmethylated. Hypomethylation in these areas can cause global genomic instability, chromosomal abnormalities, and often leads to cancer and other diseases [141]. Because of the trans-regulation between these epigenetic marks, DNA methylation is considered a stable epigenetic state. DNA methylation can be signaled into a genomic area based on repressive H3K9me3 marks, and then recruit HDACs for very stable silencing [142].

DNA demethylation can occur in either a passive or active manner. Passive demethylation transpires during cell replication, as loss of DNMT1 during this period will cause progressive loss of DNA methylation with each cell replication that occurs. Active DNA demethylation occurs with the enzymatic reaction to process 5mC back to a cytosine. However, because the covalent bond is so strong, this process requires a series of chemical modifications as no known enzyme can accomplish this modification in a single step. Briefly, the 5mC undergoes deamination and/or oxidation reactions to produce a product that the base excision repair (BER) pathway can recognize and replace. The process occurs via Tet proteins, by converting 5mC to 5-hydroxymethylcytosine (5hmC) through oxidation, then generating a 5-dormylcytosine (5fC) and 5-carboxylxytosin (5caC). These oxidative changes prohibit DNA methylation maintenance, and

the presence of 5fC or 5caC triggers activation of BER and the restoration of an unmodified cytosine [143, 144]. The exact nature of the stability of DNA methylation marks are still a debated topic, as X-inactivation and retrotransposon silencing can be relieved through treatment with demethylating agents. Also, these marks must persist not only for the cell's lifespan, but also for the organism's lifespan [145]. However, others argue that DNA methylation is a memory signal for long-term maintenance of gene silencing and not a stable epigenetic mark, as the use of an HDAC inhibitor could reactivate gene expression in genes silenced by a hypermethylated promoter region. However, the reactivated genes were then slowly re-silenced over time, as hypermethylation was not removed but could only prevent transcription if chromatin was inactivated [146]. DNA methylation is often described as a silencing epigenetic mark, but emerging research suggests a much more complex and nuanced code for gene expression control then first realized, and that is still being uncovered.

Histone modifications

Aside from the primary structure of DNA dictated by its sequence of nucleotide base pairs, DNA molecules also have extensive secondary structures. DNA strands are wound around a core of eight histone proteins, which form a nucleosome, in a bead-on-a-string configuration. This configuration allows for tighter packing of DNA and blocks transcription of the DNA through physically impeding access to cellular transcriptional machinery. Modifications to the histone tails within the nucleosomes affects the compactness of this secondary structure, and thus influences accessibility of the chromatin[147]. Histone proteins making up the nucleosome include two dimers of H2A and H2B bound to a tetramer consisting of two H3 and two H4 molecules; a linker H1 protein is bound to the exterior of the nucleosome. The histone tail is a flexible amino acid domain that protrudes from the surface of these histone proteins, and it serves as a flexible platform for many different post-translational modifications (PTMs). The core

histone proteins are a globular domain that are small and highly basic, and are often targets of post-translational modifications at specific sites and residues. Modifications on these sites can have both direct impacts on transcription, but also can act indirectly and influence recruitment of transcription factors, chromatin modifiers, and effector proteins to activate downstream signaling and block access of remodeling complexes [147]. *Cis* and *trans* effects of covalently bonded histone tails alter the chromatin template. *Cis* changes cause a shift in the physical property of the histone tails, modulating changes in charge, structure, contact, and spacing of the DNA and histones. For example, acetylation can neutralize the charge of very basic tails, causing localized expansion of the chromatin fiber to impact transcription. Phosphorylation can generate charge by adding a negative charge to the tail, which alters the nucleosome packaging or exposes histone amino termini [148]. Linker histones will incorporate into the dyad axis of the nucleosome, which promotes packaging by shielding the negative linker histone charge. Bulky additions, such as ADP ribose or ubiquitin, can induce differential arrangements and cause de-condensation of the nucleosome [149]. *Trans* effects occur via PTMs recruiting or blocking modifying binding partners.

Methylation at the H3 and H4 histones were the first PTMs discovered, and such modifications can occur as mono-, di-, or trimethylation on lysine (K) or arginine (R) residues. Because both lysine and arginine are positively charged, this covalent modification mediates hydrogen bonding and amino aromatic interactions. These methylation marks can be either repressive or activating, depending on the site that is methylated and to what degree. For example, H3K9me and H3K27me are repressive marks, whereas H3K4me or H3K36me activating marks [150]. Polycomb proteins act as a complex to mediate mono-, di-, and trimethylation on H3K17, a hallmark of gene repression that contributes to pathways in cancer and early development [151, 152]. Histone lysine methyltransferases (KMTs) or G9a catalyzes lysine methylation, while arginine methyltransferases (PRMTs) methylate arginine residues.

Lysine demethylases (KDMs) serve to demethylate lysine residues, and are in the family of amino oxidases. The KDMs include proteins LSD and KDM1A, which use FAD, oxygen, and cofactors to target mono- and di-methylation. A second type of KDMs is hydroxylases, which have a JMJC domain as the catalytically active site and use 2-oxygluterate, iron, and cofactors [153]. Protein arginine deiminases convert unmethylated arginine amino acids to citrulline [154].

Common histone modifications also include acetylation by histone acetyltransferases (HATs), typically to lysine residues (but also serine and threonine) on histones 3 and 4. These acetyl marks are removed one of a large class enzymes known as histone deacetylase (HDACs). Histone acetylation is transcriptionally activating by weakening the charge dependent interactions, neutralizing the positive charge interactions between histones and nucleosomal DNA, linker DNA, and adjacent histones, which in turn increases the availability of DNA to transcriptional machinery [155]. Phosphorylation of serine, threonine, and tyrosine residues of histones is induced by histone kinases and removed by histone phosphatases. These phosphorylation modifications modulate DNA repair, transcription, and chromatin compaction in cell division and apoptosis [156]. ADP-ribosylation is induced by poly-ADP ribose polymerases (PARPs) and actively creates a negative charge, which may stimulate local chromatin relaxation to facilitate the DNA repair process [157]. Lysine bases can have a multitude of specific modifications, including biotinylation, formylation, propionylation, crotonylation, succinylation, O-GlcNAcylation, glutathionylation, and ubituitylation [158]. Many of these modifications are still under investigation to understand how, where and at what times they are established and how they function to modulate gene expression.

The interplay between sncRNA and epigenetic control of gene expression

Given a relative narrow definition for the study of epigenetics outlined above, the control of gene expression by sncRNA would be considered beyond that scope as these molecules are

relatively short acting, not stable and not heritable through cell replication. However, as the world of ncRNA continues to expand in its breadth of functions, it is difficult to ignore the immense impact of these molecules on gene expression. While DNA methylation and histone modifications provide more stable, long-term regulation of gene expression, miRNA and other ncRNA give cells the ability to react to their environment dynamically in fairly rapid fashion. The function of miRNAs within the cytoplasm is well-defined, although these molecules can directly cause transcriptional gene silencing or activation in the nucleus [159]. Moreover, other ncRNA types, such as lncRNAs, interact directly with epigenetic modifying proteins, including HDACs and HATs. Given the apparent importance of ncRNA in control of gene expression, perhaps a broader definition of epigenetics would be useful, one that incorporates mechanisms of control that alter the abundance of transcripts without changing the original genetic code [29].

Another complication to the more restrictive view of epigenetics that excludes contributions by sncRNAs is the apparent cross-talk and feedback that exists among these pathways, including DNA methylation, histone modifications and sncRNA. PiRNAs are an obvious example of this complex interaction, as these molecules act to recruit DNA methylation via histone marks to their target transposon genes [82]. One example of this apparent cross-talk occurs during embryo gastrulation, at which point *Oct4* must be turned off in order for the embryo to differentiate. First, *Oct4* transcription is silenced via interaction with repressor molecules, followed by transcription factor recruitment of a complex containing a methyltransferase and a deacetylase. The deacetylation of the histone allows the lysine residue to then be methylated at the H3K9 residue. Heterochromatin protein 1 (HP1) is then able to bind and facilitate heterochromatin formation, recruiting DNMT 3A and 3B [160]. However, this heterochromatin process seem to be initiated by a DICER-mediated mechanism, which recognizes double stranded RNA from the satellite sequences. The RISC complex then targets those areas, recruiting SUV39H1 [161]. Non-coding RNA has also been shown to recruit histone

methylases to imprinted loci during the X-inactivation process [162]. Additionally, studies in models of cancer have identified miRNAs that indirectly modulate the epigenome through regulation of the expression of DNMTs, HDACs, HATs, and HMTs [163]. Conversely, expression of miRNAs is also subject to control via epigenetic mechanisms. For example, knockdown of DNMT1 or DNMT3b expression in colorectal cancer cells resulted in the abnormal expression of 10% of miRNAs profiled [163]. In addition, a hallmark of cancer appears to be DNA methylation of miRNAs [163-165], and more than one-third of human miRNAs are located downstream from CpG islands [166]. Interestingly, in fetal bovine muscle tissue, 20% of miRNAs involved in cell proliferation and differentiation were methylated within the gene body, which would be predicted to have a positive impact on expression. These observations collectively point to the possible coordination of sncRNA and epigenome modifiers in the control of gene expression, with potential important roles in early development that warrant further investigation.

Oocyte Development and Fertilization

The oocyte is a highly specialized, molecularly complex product of gametogenesis that contains all the components needed to support early embryo metabolism and direct early developmental events, including successful combination of two haploid genomes into a single embryonic genome and activation of transcription. Oocytes originate from the primordial germ cells (PGCs), which migrate to the genital ridge during fetal development, then proliferate via mitosis and then differentiate into oogonia [167]. Germ cell nests are formed through asymmetric cytokinesis from these mitotic divisions. The number of mitotic divisions is species-specific, ranging from four mitotic cycles in mice, to many rounds of mitotic divisions in large mammalian species. Mitotic division then stops, and germ cells initiate meiosis to form primary oocytes, passing through the leptotene, zygotene, and pachytene stages of meiotic prophase 1 before

arresting in the diplotene stage [168]. The germ cell nest breaks down at meiotic arrest to initiate

follicle formation. The primordial follicles form before birth, and the oocyte is then surrounded

by pre-granulosa cells. These primordial follicles constitute the reservoir of germ cells available

for the female reproduction lifespan.

After puberty, once the ovary becomes responsive to gonadotropin hormone, the follicles

are activated and recruited to initiate folliculargenesis [169]. During the transition from

primordial follicles to tertiary follicles, granulosa cells undergo a change from flattened epithelial

cells to cuboidal epithelial cells in the primary follicle. Then, they become multilayer, with an

outer layer of theca cells containing a basal membrane in the tertiary follicle. The maturation of

the follicle is driven by a rise in serum follicle stimulating hormone (FSH) concentration [170].

The oocytes regulate their own maturation as well as the function of the neighboring somatic

cells, which in turn regulate oocyte transcription and promote oocyte competence. Oocyte-

derived factors include growth differentiation factor 9, which promotes the development of

cumulus cells, and bone morphogenic protein 15, which is mitogenic and stimulates granulosa

cell proliferation [171]. Proliferation of granulosa cells occurs at a high rate, as the follicle size

increases and the antral cavity eventually forms, leading to the formation of the antral follicle.

The oocyte modulates the progesterone and estradiol synthesis by the cumulus cells, induced via

FSH, and regulates the differentiation of granulosa cells to cumulus cells [172, 173]. These

cumulus cells closely surround the oocyte and provide nutrient, transcripts, and regulatory

molecules to the oocyte [174, 175]. The antrum serves as an important source of regulatory

substances derived from blood or secretions of the follicular cells. Antral fluid production is

intensified by increased vascularization and permeability of the blood vessels, which causes large

increases in follicle size. At this stage, the oocyte enters an extensive growth phase, reaching up

to 150 mm diameter in cattle and humans. As the oocyte increases in volume, complex

cytoplasmic organization occurs accompanied by large increases in RNA and protein synthesis

[176]. The number of ribosomes, mitochondria, and organelles increase as well, with an accumulation of glycogen granules, protein, lipid droplets, and multivesicular bodies, which all contribute to a high quality mature oocyte [177, 178]. The oocyte also forms the zona pellucida, a glycoprotein membrane that provides a protective coat around the oocyte.

If an oocyte is fully-grown, meiotically competent and dominant, resumption of meiosis is initiated by the preovulatory luteinizing hormone (LH) surge. The LH surge initiates both the final oocyte maturation and follicle maturation, but only for the dominant follicle that is no longer dependent on FSH levels [179]. All other follicles, which are either subordinate or growth arrested, enter atresia and degenerate. The high levels of cAMP in the oocyte maintains it in the arrested state, and a protein called maturation promoting factor (MPF) is responsible for the oocyte's resumption of meiosis. The LH surge causes a drastic drop in cAMP levels through the activation of PDE3A, as well as stimulating expansion of cumulus cells and the closure of gap junctions, which reduces diffusion of cAMP [176]. After the LH surge, the oocyte undergoes two consecutive M phase divisions, in the absence of DNA replication and S phase, then arresting at M2 until fertilization [176]. The LH surge causes ovulation of the oocyte, at which point the oocyte enters the oviduct and, if sperm is present and the oocyte is of good quality, fertilization will occur [180]. At fertilization, the sperm penetration triggers activation for the completion of the meiotic cycle [181].

The breakdown of the germinal vesicle signals an oocyte's commitment to maturation; this phase is also one of reduced transcript activity. During this period of transcriptional quiescence, changes in the transcriptome and proteome occur through interaction and intercellular transport with surrounding cells, as well as post-transcriptional modifications such as those triggered by sncRNA or RNA binding proteins. The mature oocyte is transcriptionally inactive prior to fertilization but contains reserves of transcripts and proteins needed to drive the first cleavage divisions and reprogramming events. The oocyte provides the maternal inheritance,

genetic info, as well as epigenetic info for the developing embryo before embryonic gene expression is turned on at the EGA [182]. Post-fertilization, the oocyte undergoes the second meiotic division and extrudes a second polar body, at which point both pronuclei de-condense and substantial epigenetic reprogramming begins [183]. The maternal pronucleus undergoes passive demethylation via DNA replication. DNMT1, the methyltransferase responsible for maintaining methylation in a semi-conservative fashion, is excluded from the nucleus during early embryonic division. In mice an oocyte specific version of DNMT has been found, DNMTo, which has not been identified in cattle. DNMTo accumulates in oocytes during their growth phase, and localizes to the cytoplasm in a mature oocyte; this protein is only localized to the nucleus at the 8-cell stage so as to protect imprinted genes [184]. Therefore, DNA methylation is reduced progressively over each nuclear division up to the morula stage. DNMT1 re-enters the nucleus for a single cell division, at the point when the embryonic genome is activated (varies between species), in order to re-methylate differentially methylated regions (DMRs) and prevent activation of regions, such as TEs, that need to remain silenced to protect genomic integrity [185]. While the majority of this activity occurs in the maternal pronucleus, DNMT1 also protects differentially methylated regions in the paternal pronucleus. TET3 is an enzyme that can demethylate loci and does so in the paternal pronuclei [186]. The STELLA/PGC7/DPPa3 complex binds to DNA and protects maternal imprints from demethylation via binding to H3K9me2 histone marks and blocking the activity of TET3[187, 188]. Specific histone modifications mark DNA methylation imprints that must be protected in the maternal pronuclei and the first embryonic divisions. H3K9me2 is one such mark, however it was found to be dispensable for imprinting maintenance [189]. TRIM28 is a maternal factor that regulates embryonic development, maintenance of pluripotency, imprinting, retroviral silencing and DNA damage response. TRIM28 partners with DNA binding proteins (e.g., KAP1) to induce heterochromatin activity by acting as a scaffolding protein with the histone deacetylation complex

NURD, which in turn recruits DNA methyltransferases [190]. The complex will bind to the DNA and protect these modifications from TET3-mediated demethylation. However, if a gene is aberrantly demethylated, TRIM28 will not bind and re-methylation will not occur [190].

Histone modifications of the genome undergo major changes during the process of oogenesis. HAT expression and histone acetylation increases with oocyte growth, and H3K4me3 (active mark) and H3K9me3 (silent mark) both increase with oocyte size [191]. Once the oocyte is mature, however, few changes occur pre-fertilization. Characteristics that are a component of heterochromatin, such as lysine methylation, are enriched in areas surrounding precursor nucleolar bodies, but only in the maternal pronuclei. Histone modifications are asymmetric in early development, with the maternal pronuclei having many more histone marks then the paternal pronucleus, especially lysine methylation. Comparatively, the paternal pronuclei is enriched for H4 acetylation [192]. The exact importance of these changes during oogenesis and early development is not completely clear.

In the oocyte, maternal mRNAs are stored in messenger ribonucleoprotein particles for protection until they are recruited for translation at specific time points [193]. The stored mRNAs are maintained in an untranslated state and generally have short poly (A) tails, which prevents formation of the translation initiation complex. Extending the poly (A) tail activates translation by recruiting PABPs and promoting interaction with the 5' end of the mRNA. In the oocyte, maternal mRNAs are synthesized in the nucleus and polyadenylated under the nuclear polyadenylation signal AAUAAA. Once synthesis is complete, they are transported into the cytoplasm and either become polyadenylated and translated, or deadenylated and stored for later use [194]. Because there is no active transcription, the oocyte relies heavily on post-transcriptional control of maternal transcripts. The maternal mRNA population that accumulated during maturation is very diverse, supporting a range of functions during maturation and post fertilization [195]. There are dynamic shifts in mRNA populations through early development,

following a carefully orchestrated pattern. Large groups of mRNAs are recruited, translated, or degraded at specific times. The degradation of these maternal mRNAs must occur before embryonic genome activation, as embryonic development requires precise orchestration of many molecular pathways that are controlled by a large number of different genes [196].

Within the smallest cell in the body, the sperm genome is highly specialized, as unique epigenetic marks are needed to condense the size of the DNA to fit into the compact sperm head. The majority of histone proteins are replaced with highly basic protamines in order to facilitate the packaging of the DNA into restricted space. About 1-10% of the sperm genome is still packaged into histones that can carry modifications [197]. It appears that these histone modifications are not random and may serve to prime the embryonic chromatin status and influence states of gene transcription or repression during early embryogenesis [198, 199]. Post-fertilization, the protamines are rapidly removed from the spermatic genome and replaced with histones in a DNA replication-independent fashion [191]. The histones are hypo-methylated and hyper-acetylated, which may be a function of the incorporation itself. However the paternal chromatin remains devoid of modifications associated with heterochromatin, until replication of paternal DNA at the late pronuclear stages [192]. Active demethylation of the paternal pronuclei occurs within hours of fertilization in mice and is involved with the oxidation of 5mC to 5hmC, and then to 5caC [200]. The TET hydroxylase enzyme family performs this oxidation, which explains the high expression of TET3 in the zygote, preferentially expressed in the paternal pronuclei. It is unknown if the high levels of 5hmC produced by TET3 are diluted out through cell divisions, or if another enzyme continues the chemical modifications. While TET is the dominant mechanism of demethylation in the paternal pronucleus, evidence points to the involvement of base excision repair and nucleotide excision repair, as well [191].

Maternal-to-Embryonic Transition

Following successful epigenetic reprograming of the maternal and paternal genomes,

embryonic genome activation must occur. The MET of early development is the process of

shifting genomic control from the maternal transcriptome to the activation of embryonic genome.

The shift in transcripts is quite drastic, as maternal transcripts that may have been present in the

oocyte and early embryo for weeks to months are eliminated in a matter of hours. Three major

events occur at the MET that are necessary for continued embryo development (Figure 1.4).

First, transcripts specific to the oocyte must be destroyed. Second, maternally-derived transcripts

that are expressed in both the oocyte and embryo must be replaced mRNAs derived from the

embryo's genome. And third, *de novo* production of transcripts specific for the developing

embryos must take place [201].   In this dissertation, MET will be used to refer to the entirety of

this transition, which includes the degradation of maternal transcripts as well as the process of

EGA. The term EGA will be used through this dissertation to refer to the specific step during the

MET at which the embryonic genome is first transcribed and mRNAs and miRNAs from the
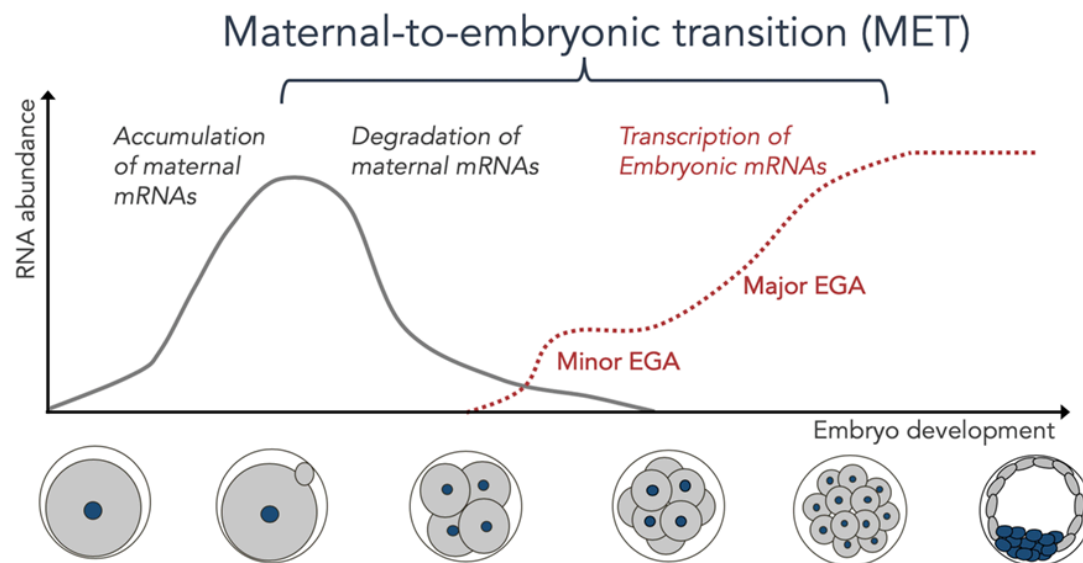
Figure 1.4. The maternal-to-embryonic transition in cattle.

embryo are activated. In cattle, the EGA occurs in two waves, the minor and the major wave, with the major wave of EGA occurring at the 8-cell stage, whereas the MET occurs over a longer period of time, approximately from fertilization to the 16-cell stage in cattle [202]. While many researchers employing the mouse model of embryonic development use the term maternal-to-zygotic transition (MZT), as the EGA coincides with the formation of the zygote in mice, this descriptor is not precise for other mammalian models for which the EGA occurs at later developmental stages, such as in cattle. (Note that Chapter 2 retains the use of MZT in place of MET, as that chapter was published prior to the completion of this dissertation.)

During the developmental window prior to EGA, maternally-derived mRNAs are regulated via stability, location, and translation. These mRNAs are capped on their 5' end and polyadenylated at the 3' end, chemical structure characteristics that control transcripts' stability and availability for translation. The 5' end cap is a 7-methylguanosine that interacts with the poly(A) tail and is required to form a loop for initiation of translation. In the arrested oocyte, mRNAs are deadenylated but dormant, and activated through polyadenylation [203]. The timing for activation of each mRNA depends on a combination of motifs in their 3' UTR, such as the nuclear polyadenylation sequence or the cytoplasmic polyadenylation element, that recruit different RNA binding proteins and determine polyadenylation and translational status. After the oocyte completes meiosis, the AU-rich element (ARE) binding protein C3H-4 recruits the CCR4-NOT complex to mRNAs that are deadenylated and contain ARE motifs. The waves of deadenylation and polyadenylation drive meiotic transitions by regulating cyclins, CDKs, APC-c, and CSFs [204]. Many of the maternal mRNAs are present in subcellular structures, such as cytoplasmic granules, which may facilitate translational repression; because these granules have no enclosing membrane, these mRNAs can be transported rapidly in response to developmental cues. [205]. Up to this point in development, all of the molecular processes of oogenesis and early development rely exclusively on maternal mRNAs. Specific mRNA become dramatically

adenylated at different stages of oocyte maturation or embryo development, correlating with their recruitment into polysomes and translational initiation [206]. Large waves of mRNA adenylation and activation occur at oocyte maturation, and again at fertilization. One protein regulator is WIPSY (aka GLD2, PAP, TUTase, or HS), which mediates the polyadenylation of the maternal transcripts and is required for oocyte maturation in mammals [207]. In *Drosophila*, this activity is performed by WISPY on maternal transcripts that are critical for early development, to prevent premature clearance [208]. Other proteins also regulate mRNA storage or action in the oocyte and early embryo, including CPEB, which is a highly conserved sequence-specific RNA binding protein that binds to a cytoplasmic polyadenylation element in the 3' region of the mRNA [209]. CPEB complexes with GLD2 in humans and PAP in *Drosophila* for polyadenylation. The CPEB protein is necessary for both nuclear and cytoplasmic polyadenylation, as it modulates mRNA location and translation [210, 211]. Polyadenylation is necessary to protect those transcripts necessary for early development from degradation.

Deadenylation in the oocyte and early embryo can cause destabilization and degradation of transcripts, although a large proportion of mRNAs remain stable even when deadenylated. Unlike somatic cells, mRNA decapping and decay are not tightly coupled to deadenylation in embryos. It is possible that embryonic activation of decapping activity is required for degradation pathways of deadenylated maternal transcripts [212]. The resistance of maternal mRNA to decay even when deadenylated may protect mRNA molecules with developmentally important functions in the MET that need to be silenced in the oocyte. In mice, P-bodies associated with mRNA degradation are only found in early stage oocytes, and these disappear as the oocyte grows. The components that normally localize to the P-body localize to subcortical aggregates containing maternal mRNA, but lack the decapping enzyme DCP1 normally associated. Similar types of specialized localization are seen in *Xenupus* and zebrafish, into a similar translationally repressed mitochondrial cloud and Balbiani body, respectively. These

changes, unique to the mature oocyte and early embryo in mammals, may reflect a function in mRNA storage and repression, but not degradation [205]. Shortening of the poly(A) tail is still the rate limiting step for degradation, and the deadenylation is followed by de-capping and degradation from the 5' end. The deadenylation complexes include the CCR4-NOT, PAN2-PAN3, and PARN complexes. Specificity is through recruitment of elements or motifs of mRNAs bound by RNA binding proteins, or miRNA binding.

At the MET, 30-40% of transcripts are completely degraded and 60% are significantly degraded [213]. The pathways for transcript degradation include those driven by the maternal or by embryonic machinery. The proportion of transcripts degraded by maternal verses embryonic modes varies across species. Tadros et al. determined that more than 1000 mRNAs were cleared after egg activation via the maternal degradation pathway in *Drosophila* and that the protein Smaug regulated two-thirds of those transcripts [214]. An additional 563 mRNAs were suggested to be subject to embryonic mode of clearance [215]. In mice, oocyte maturation caused destabilization of almost 3000 mRNA, all of which were cleared by maternal modes [216]. Immediately following fertilization almost 2300 mRNAs were degraded, consistent with a maternal mode of regulation, and at the 2-cell stage post MET, almost 500 mRNAs were cleared, consistent with an embryonic mode of regulation [217].

One major protein contributing to the maternal degradation pathway is Smaug, which is derived from the oocyte. This protein directs the degradation of a subset of maternal transcripts through recruiting a deadenylase complex made up of CCR4/POP2/NOT [218] and causes subsequent down-regulation of maternal proteins that would repress the embryonic genome [219]. In *Drosophila*, Smaug can also act as a translational repressor through binding to recognition elements in the 3'UTR and binding to the eIF4E binding protein CUP for destabilization of the transcript *Nos*. Stem loop structures termed Smaug response elements in the 3' UTR repress mRNA [220]. Smaug may control stability and translation of a large proportion of maternal

mRNA, as statistical analysis predicted that 3000 transcripts (about 50% of detectable transcripts) were under translational control, of which 339 were bound to the Smaug protein and a further two-thirds of those bound transcripts were destabilized [219]. Likely some targets of Smaug escaped detection, as many of the transcripts destabilized at this point in development contain Smaug response elements. Importantly, Smaug targets transcriptional repressors for degradation, which is a critical step to enable activation of the embryonic genome.

The embryonic pathway for clearance of maternal mRNAs is more active at the EGA, and this pathway is necessary to degrade maternally-derived Smaug. Unstable mRNAs degraded by the embryonic pathway are enriched in binding sites for Pumillo (Pum), a RNA binding protein. In support of this link, *bicoid* mRNA degradation is mediated by Pum [221]. The exact mechanism has not been identified, but Pumillo associates with the CCR4-NOT complex in other situations [222], and the CCR4-NOT complex is required both maternally and embryonically for embryonic development [222]. However, Smaug and Pum do not account for all mRNAs degraded. Thomsen et al. mined *Drosophila* deep sequencing data of unstable mRNA, and many mRNAs that did not have Pum binding sites or Smaug response sites did contain an enrichment for AREs and miRNA seeds [213]. These researchers were able to experimentally confirm the activity of miR-14 in maternal transcript degradation [213]. The potential for miRNA as an embryonic clearance mechanism of maternal transcript has promise but is complicated by species differences. MicroRNA decay pathways have been identified in both vertebrates and invertebrates. MiRNA decay-dependent transcripts were shown first in zebrafish, where one miRNA, miR-430, mediates the decay of hundreds of maternal mRNA [223]. When miRNAs are loaded into Ago proteins to target mRNA, AGO proteins may interact with the GW182 family, which in turn recruits CCR4-NOT and PAN2/3 complexes. These interactions have been demonstrated in *Drosophila* and mammalian cells, but have not been proven in the context of early development [224]. Maternal decay by both maternal and embryonic miRNAs has also

seen in *C. Elegans* [225], by the miR-430 orthologue miR-427 in *Xenopus* [226], and in *by* miR-309 in *Drosophila*, which is highly expressed at the EGA and degrades hundreds of mRNAs [227]. Interestingly, the miR-430/427/302 family of miRNAs shares a seed sequence with miR-290 in mice, which are expressed at high levels in embryos and embryonic stem cells [228, 229]. Genetic inactivation of the cluster miR-290-295 results in problems with fertility, embryonic lethality, and PGC defects [230]. However, to date, miRNAs have not been shown to be part of the embryonic mRNA degradation pathway in mammals. The contradiction with other species may be due to the heavy use of the mouse model to study mammalian early embryonic development. Discussed in more depth below, mice have an sncRNA pathway in the oocyte and early embryo that is apparently unique to rodents. PiRNA proteins appear to be divergent from what is seen in humans and cattle [72], and endo-siRNAs, which are not commonly seen in other mammalian models of early development, may be up-regulated in mice oocytes due to a rodent oocyte-specific Dicer isoform that preferentially loads siRNA over miRNA [231]. Findings in mice support a model by which endo-siRNAs are important for mRNA clearance during oocyte maturation, and perhaps at the MET, while miRNAs are crucial for mRNA regulation later in embryonic development as cell differentiation begins [232]. Because of these divergent pathways, the experimental evidence that miRNAs are not highly active or functional in mouse oocytes and early embryos [232-234] should not deter further investigation of miRNAs as a embryonic mRNA clearance mechanism in mammals

Spatial recognition is a highly regulated regulator in maternal mRNA decay and can protect developmentally important transcripts from decay. In *Drosophila*, Osk is required for translocation of mRNA to the posterior pole and prevents Smaug from binding [235]. Similarly, in zebrafish, *Nanos* is generally targeted by miR-430, but needs to be protected in the PGCs by the RNA binding protein Dead End 1 (DND1), which binds to the 3' UTR and prevents miRNA association [236]. The zebrafish maternal transcript *Tdrd7* similarly requires protection from

miRNA degradation, and in this case, the RNA binding protein DAZL causes polyadenylation and stabilization [237]. In this manner, a multitude of proteins with different functions can bind to protect transcripts from degradation. The complex system that contributes to the abundance of transcripts involving targeting by degradation proteins, miRNAs, and RNA-binding proteins makes it difficult to directly connect the abundance of an miRNA molecule to the predicted down-regulation of its target mRNA [238].

The subcortical maternal complex (SCMC) is essential for successful preimplantation development. The SCMC is assembled during oogenesis and marks cell populations during divisions. This structure also includes the proteins MATER, FLOPED, TLE6, PADI6, and Filia [239], each of which has an effect on early embryogenesis. When cell division occurs non-parallel to the apical basal axis of the polarized blastomeres, SCMC is either present or absent in the divided cells. Those cells that lack the SCMC preferentially become the inner cell mass (ICM), while those containing SCMC preferentially form the trophectoderm [240]. These observations support a model of early development by which maternally-expressed protein complexes differentially accumulate in topologically distinct blastomeres, despite the embryos' ability to adjust to the loss of these cells.

There are two phases in activation of the embryonic genome: the minor wave, during which a minority of transcripts are activated, and the major wave, which coincides with the large scale onset of transcription [196]. There are several models that may explain the initial triggering event for the EGA. The first model posits that increasing genomic material relative to the constant amount of cytoplasm triggers a release from maternal repressive transcriptional factors present in the cytoplasm [241]. The second model proposes that a maternal clock determines the timing of gene expression [242]. The maternal clock model requires that a maternal factor is increased in either quantity or activity, and once it reaches a critical level, transcription is triggered. Maternal mRNA mobilizes over time and leaves the dormant deadenylated state.

Therefore, the resulting proteins can function as transcriptional activators, removing the silencing marks on the embryonic DNA [243]. Support of this model includes the post-fertilization polyadenylated and translated transcription factors that include Nano, Pou5f1, and Sox19b. Pou5f1 can bind to specific loci, suggesting it may prime certain genes to be embryonically expressed. Knockdown of these factors causes developmental arrest, similar to what is seen with alpha-amanitin treatment, a RNA polymerase 2 inhibitor [244]. In *Drosophila*, there is evidence in support of both models, perhaps even working in a coordinated fashion to activate transcription [245]. There are also a host of maternal genes that have been shown to be required for the major wave of EGA, and chromatin structure may function early on when early expression of key genes may be required for the major wave of EGA [246].

To better understand the contributions of maternally-derived transcripts in comparison to those produced by the embryo, Graf et al. used oocytes from *Bos taurus taurus* and sperm from genetically distinct *Bos taurus indicus* for *in vitro* fertilization, allowing for the identification of embryonic transcripts by the presence of paternal single nucleotide polymorphisms [247]. It is also possible to match mRNA to an oocyte DNA library to identify transcripts that are newly present post fertilization. Researchers have used this method to identify the major and minor waves of the EGA in embryos, and to determine when leaky transcription may be occurring prior to these events. Embryonic contribution can also be measured using modified ribonucleotides, such as ^{3}H, ^{32}P or bromo uridine triphosphate, but it is difficult using this method to determine the identity of the transcribed genes [248]. RNA sequencing can identify some features of embryonic transcripts to distinguish them from maternal transcripts, as embryonic transcripts may have alternative splicing and poly adenylation patterns [249]. RNA sequencing libraries are often built using methods that rely on a poly(A) tail for mRNA isolation. However, by sequencing unspliced pre-mRNA, it may be possible to identify more SNPs and more easily identify embryonic transcripts.

The embryonic genome activation occurs following nuclear cycle 13 in *Drosophila,* at the 2-cell stage in mice, 4-to-8-cell stage in human, at the 8-cell stage in cattle, and at the mid-blastula transition (MBT) in zebrafish and *Xenopus* [212]. In bovine embryos, the major wave of the EGA occurs at the 8-to-16-cell stage. However, small amounts of transcriptional activation have been noted as early as the single cell zygotic stage. Graf et al. observed a marked increase in transcriptional activation of the embryonic genome between the 4-cell and 8-cell stage in cattle, likely due to the necessary degradation of maternal mRNA for the EGA, as many of those transcripts were down-regulated [247]. Genes activated prior to the 4-cell stage were associated with for RNA processing, translation, and transport functions, likely in preparation for genome activation and active transcription and translation. For example, *KLF4* was activated at the 4-cell stage; this gene may contribute to the major wave of EGA, as it is known to activate transcription [247]. Genes activated at the 8-cell stage were linked to functions involving transcription, nucleotide metabolic processes, protein ubiquitination, translational metabolism and RNA metabolism processes. *EIF3* is required for the initiation of protein synthesis and was first expressed at the 8-cell stage. *EIF3* also associates with the 40S subunit to facilitate formation of the pre-initiation complex, thus stimulating the process of mRNA recruitment and scanning for AUG recognition [247]. Chromatin remodeling factors, histone deacetylase and DNMT3b are also activated at this developmental stage in cattle [247].

Small Non-Coding RNAs and the Maternal-to-Embryonic Transition

The transcriptionally silent landscape of the mature oocyte and early embryo make this a unique window of development during which post-transcriptional machinery may dominate the regulatory network. As such, this period of embryo development is an interesting point at which to examine the function of miRNAs in early development. As discussed in detail above, embryos inherit numerous maternal transcripts that are key for embryo development. In many species, a

defined role for miRNA in the clearance of maternal transcripts has been well defined. However, due to differences in the sncRNA biogenesis pathway in mouse oocytes, a popular mammalian model, discrepancies and controversy has surrounded the putative role of miRNA for clearance of maternal mRNAs at the MET. During the maternal-to-embryonic transition, depletion of maternally derived mRNAs is at least partially dependent on the 3'UTR, which points to a possible of this class of sncRNA during the MET.

The debate on the role of miRNAs centers on the specialized sncRNA biogenesis pathways in rodents, which are often used as a mammalian model for embryo development. Because Dgcr8 is necessary for miRNA biosynthesis, but is not part of the siRNA biosynthesis pathway, it is possible to examine which sncRNA population carries out specific functions using a Dgcr8 knockdown. Dicer, however, is an essential processor for both siRNA and miRNA biosynthesis pathways. Therefore, studies using *Dicer* and *Drosha* genetic knockout animals provide valuable clues to the differences in function of these to biogenesis factors. Oocytes derived from *Dicer* null mice cannot mature or be correctly fertilized, whereas oocytes from Dgcr8 null mice can mature, undergo fertilization, and develop to the blastocyst stage [228, 233]. However, the *Dgcr8* knockout mouse has a lower fecundity rate, pointing to some function of miRNAs during development, even in rodents. The reduced fecundity does suggest some function of miRNA during development, even in rodents. The *DGCR8* knockout also does not affect the embryonic transcription of biogenesis factors, as it is a ZP3-cre mutation [250]. MiRNA could therefore function in the embryonic degradation pathway in the MET, as there would be no maternal miRNA in the *DGCR8* knockouts and embryonic miRNAs would not be affected. In zebrafish, miRNAs have been shown to function in the degradation of maternal transcripts [251]. MiR-430 is processed by Dicer at the EGA and targets hundreds of maternal mRNAs for degradation, linking the onset of EGA and maternal mRNA decay. No such function has been identified in mice, yet the *Dicer* knockout phenotype is more severe in mice than in

zebrafish [251]. As mice have a severe knockout phenotype, but no known function of miRNA, and zebrafish have known functions of miRNA but a less severe knockout phenotype, we cannot rule out that the miRNA functionality in mice has simply not yet been identified. Therefore, the observations from use of Dgcr8 knockout may not rule out a functionality of miRNA, but rather suggest that the loss of miRNA function does not cause problems until after the EGA, at which point cellular differentiation begins.

A study in mice assessing Dicer knockout effects showed severe defects in both chromosomal alignment and spindle organization [233, 252]. While in somatic cells, siRNA precursors would cause an interferon response, this pathway appears to be lacking in mouse oocyte [253]. Rather than function via RNAi, it is possible that endo-siRNAs serve a role in preserving the accuracy of specific developmental arrests and control of the cell cycle required in the maturing mouse oocyte and early developing embryo. If the Dicer knockdown mouse phenotype was due to siRNA serving similar roles in spindle organization, this would explain the immediate devastating consequences of the knockdown. In support of this hypothesis for the mouse model, the very low levels of siRNA in comparison to miRNA and mRNA may suggest that siRNAs may not function to degrade in maternal mRNA. The dysregulation of transcripts via aberrant miRNA expression would not necessarily cause immediate embryo death or damage, but would lead to the accumulation of increasingly aberrant transcript expression over time ultimately triggering high stress on the embryo that impairs cells' capacity to properly differentiate. These knockout study results suggest that, in mice, endogenous siRNAs are significant contributors to the development of functional oocytes. The suggested function of siRNA in oocytes, but lack of need for miRNAs, may be due to a mutant DICER only found in mouse oocytes that preferentially processes endogenous siRNA, as opposed to miRNAs [231]. However, several miRNAs have been shown to regulate genes vital in early development, even in mouse species. In fact, miRNAs have been identified that closely regulate the pluripotency

pathway such that mouse embryonic specific miRNAs can facilitate revision of differentiated cells into a pluripotent state [254]. The rodent-specific sncRNA pathway likely creates an species-specific embryonic environment in mice in which siRNAs function to degrade maternal transcripts, whereas miRNA is necessary later on as differentiation begins to occur in development [255].

MiRNA may be present concurrently with targets, but unable to degrade those transcripts until they are released from translational repression and protection [238]. The functionality of miRNA in the early embryo has been a source of controversy, as in mouse oocytes, luciferase reporters were not efficiently repressed in mature oocytes [232]. However, the miRNA-mRNA binding code can be impacted and altered by both modifications and protein binding. For example, in mammals, adenosine deaminases (ADARs) alter adenosine to either inosines or guanines in ncRNA as well as their target mRNA. Both of these actions can change binding ability of miRNA and prevent RNAi targeted degradation of targets [31]. Adenylation can also occur in miRNAs and it is possible that adenylation seen on miRNAs during early development of mice, sea urchins, and *Drosophila* could prevent miRNAs present from binding to their target mRNA, and may mark maternal miRNAs for degradation [256]. In addition, modifications such as N6 methyladenosine (M6A) of the mRNA can block A:G base-pairing and interfere the binding of miRNA to mRNA. The overall cellular M6A abundance and individual transcript M6A levels can be altered both by Dicer levels, and levels of specific miRNAs [257]. Some mRNAs express AU rich elements (AREs) in the 3'UTR, which can compete for binding to RNA binding proteins, interrupt binding of miRNAs, or act cooperatively in miRNA binding [258]. AREs interact with different proteins, such as dead end homologue 1 (DND1). When bound to AREs, DND1 physically prevents miRNA from binding to the 3'UTR of transcripts. The function of DND1 has not been characterized in mouse development, but this protein is abundant in both immature and mature pig oocytes. Direct binding of DND1 to transcripts of key

pluripotency genes has been shown in the pig model, including *OCT4*, *SOX2*, and *LIN28* [259].

DND1 may protect maternal transcripts critical for development in the period between germinal

vesicle breakdown and the EGA. Also, DND1 was down-regulated during EGA in pig embryos,

which may allow miRNAs to bind maternal mRNA targets [259]. Moreover, DND1 impacts

fertility, as it is required for embryonic development and both male and female germ line cells

require DND1 expression for viability [259, 260]. Through mediation of miRNA binding to

transcript targets, RNA binding proteins such as DND1 may be a critical mechanism for

restricting the activity of miRNA molecules during early development. The RNA-bound

proteome undergoes extensive changes at the MZT in *Drosophila*, which likely impacts the

ability of miRNAs to target transcripts [261]. It is possible that other RNA binding elements

function in similar pathways in early development, mediating interactions between sncRNA and

their targets, thus contributing to a dynamic and complex degradation code in the early embryo.

The premise that miRNAs do not have an important function in mammalian early

development has been challenged through studies in other non-rodent species, such as the pig

[262] and cattle [263, 264]. For example, using a PCR approach to profile 98 miRNAs of interest

in cattle oocytes, 8-cell and blastocyst embryos, Berg and Pfeffer determined that most miRNA

concentrations were below the threshold that could reasonably suppress transcript expression

given the apparent abundance of target transcripts [265]. However, they did determine that miR-

320 was highly expressed in embryos, and that miRNAs were generally more abundant at the 8-

cell stage compared to oocytes [265]. MiR-130b was found to control granulosa and cumulus cell

proliferation, oocyte maturation, as well as morula and blastocyst formation [266]. In the bovine

blastocyst, miR-218 may regulate *NANOG* and *CDH2*, and miR-449b may regulate *NOTCH*

[267]. In cattle, miR-212 has been demonstrated to negatively regulate the maternal effect gene

FIGLA [264], and miR-196 has been shown to negatively regulate the maternal effect gene

NOBOX [263], supporting a possible function for mammalian miRNA in the degradation of

transcripts at the MET. In addition to this work in cattle, studies using the porcine model of early embryo development have shown that *Dicer* and miRNAs are present in oocytes and embryos, and that miRNAs were dynamically regulated by developmental stage [268]. MiRNAs in porcine embryos were also found to regulate abundance of transcripts critical for development [262].

Furthermore, dynamic changes of miRNA populations have been shown in pre-implantation embryos in cattle [265, 269], and exogenous supplementation of specific miRNAs can alter the developmental potential of embryos [270]. Bovine embryos secrete miRNAs into culture media, and the profiles of miRNAs released were distinct based on cleavage time, with miR-30c and miR-10b more abundant in conditioned medium of slow-cleaving embryos [271]. Moreover, Lin et al. determined that supplementation with a miR-30c mimic resulted in increased apoptosis [271]. Furthermore, Kropp et al. showed that the secreted miRNAs that distinguished embryos that successfully developed to blastocyst stage from those that did not were also detected in culture medium of both bovine and human pre-implantation embryos [272]. The apparent lack of concordance for mouse and human models of embryo development with respect to the role of miRNAs raises concerns when considering the outsized role of rodent models in pre-clinical studies focused on early development, particularly given that human miRNAs in cumulus cells have already shown promise as a biomarker for development [273]. The apparent better concordance of cattle embryo models supports the use of bovine embryos for such studies, especially those focusing on the function of miRNA during the MET.

In mice and rats, it appears the sncRNA pathways in early development are not the same as the conserved pathway in other species such as cattle and humans, and this extends from the miRNA/siRNA dynamics to piRNA [72]. In mouse oocytes, piRNAs are relatively rare and dispensable, which came as a surprise considering the highly conserved role for piRNAs in silencing TEs and maintaining genomic stability. In mouse embryonic male germ cells, MIWI2 (PIWIL4) was shown the drive the nuclear piRNA pathway, while the Piwi paralogue MILI

(PIWIL2) initiated the loading of MIWI2, both of which specifically loaded through the secondary piRNA ping-pong cycle [274, 275]. In other models, such as cattle and *Drosophila,* piRNA is essential in spermatogenesis and oogenesis [276], whereas mouse piRNA proteins seem to only impact male fertility, not female fertility [262, 277, 278] as minimal transcripts have been detected in oocytes [279, 280]. Mice lack the PIWI protein PIWIL3, which is found in human and cattle oocytes [72]. It is possible that the specialized acquisition of the Dicer isoform has some role in maintaining PIWIL3 function, and extrapolation of mouse sncRNA pathways during development may be inappropriate across species. PiRNAs may function in early development beyond silencing of TEs. Based on piRNA populations in bovine oocytes and zygotes, researchers have hypothesized that piRNAs may also contribute to mRNA degradation during the MET [281]. Many of the piRNAs expressed in cattle embryos and zygotes are potentially derived from or targeted at specific mRNA sequences, and many of these sequences belong to transcripts destined for degradation later in embryogenesis.

In addition to a potential role in the oocyte, sperm can contain paternal RNAs that may contribute to genome reprogramming. Male gametogenesis requires the expression and function of small RNAs, and sperm appear to carry a diverse population of paternal miRNA [282]. Knockout of paternal Dicer or Drosha adversely impacted the developmental potential of those sperm. However, these miRNAs appear to have limited ability to impact maternally-derived mRNA, so their exact function in embryo development is still unclear. However, six miRNAs have recently been demonstrated to be present in sperm but absent in the oocyte. One of these, miR-34c has been shown to impact the first embryonic cleavage through interaction with *BCL2* [283]. Also, sperm miRNAs have been shown to be impacted by stress, which can cause changes in miRNA up to the morula stage [284] and may be a mechanism for intergenerational inheritance of epigenetic marks [285].

Other classes of sncRNA are vital in their supporting functions during the MET to support extreme changes in translation activity, including snRNA, snoRNA and rRNA. However, beyond these key housekeeping functions, these sncRNAs may also have more nuanced regulatory roles. For example, while tRNA has the well-known function to support translation, tRFs have been shown to modulate intergenerational inheritance. In an interesting study, researchers showed that consumption of a high fat diet altered the profile of tRFs in sperm, and this phenomenon was connected to an altered gene expression pattern and glucose intolerance in the resulting offspring [286]. This effect was attributed exclusively to the tRFs, not miRNA or lncRNA. SnRNA regulate the splicing of mRNA, and specialized isoforms of each snRNA are expressed during *Drosophila* development, which is also seen in mouse and human embryos. The change in isoforms during embryogenesis is unique, and as development progresses, the single dominant isoform gradually dominates expression. The function of these changes in snRNA isoforms in early development is unknown, but in other cellular contexts, snRNA variants can regulate gene expression [287, 288]. Moreover, any depletion of snoRNAs will lead to problems with translation in early embryos [289]. While no studies have explored the potential function of newly identified snoRNA fragments in early development, this class of sncRNA warrants further study to determine their function in differentiation or pluripotency [119]. Other, very recently discovered RNA molecules, such as circRNAs, may also function in early development. These non-coding RNAs form through the 3'-5' ligation of an RNA molecule from an exon, intron, or exon-intron and regulate gene expression via their "sponge-like" interaction with miRNAs. The profile of circRNAs differs depending on embryo quality and these populations also change over the course of embryo maturation [290]. CircRNAs widely exist in embryos and are aberrantly expressed in bovine scNT embryos, but their function is still unclear [291, 292].

Agricultural Significance of scNT

When working with agricultural clients, live offspring is the most important outcome. In order to get healthy livestock offspring, the practical application of assisted reproductive technologies (ART) requires integration of laboratory techniques and veterinary management. In breeding programs, genetic improvements are generated by selecting animals based on the animal's breeding value for the relevant valuable traits. Dissemination of the superior genetic material to the commercial population is then necessary, and genetic improvement is driven by a small fraction of the total population.

Due to the low efficiency of scNT, commercial applications of this technology in agriculture are presently limited to the production of animals of high genetic merit or the most elite show cattle [293]. Conventional breeding for these traits is difficult because they are typically scored postmortem. Even if such traits are assessed in living animals, breeding is often not an option as many superior beef cattle are sterile steers. Cloning animals from cells obtained from superior slaughtered animals has the potential to capture superior carcass trait characteristics. One of the benefits of ART is to increase the presence of desirable characteristics in production herds (e.g., improved carcass quality, increased feed efficiency, reduced waste, improved disease resistance). A recent large-scale longitudinal study that encompassed nearly 2,000 assisted reproductive procedures and production of more than 9,200 transferable embryos convincingly showed that the reproductive performance of clones is comparable to their genetic donors [294]. Therefore, cloning technology combined with superovulation, artificial insemination and embryo collection provides a valuable tool for faster dissemination of superior maternal genetics. Improvement in ART efficiency combined with genomics technologies would make a significant impact on the rate and dissemination of genetic improvement.

Medical Significance of scNT

The methods of scNT have created opportunities for significant improvements in human medicine as well. Livestock cloning can create production of animal organs and tissues that could be appropriate for human transplantation, including pig organs that can be genetically modified for the human immune system [295]. ScNT also provides the opportunity for the addition of genetic material or production of a transgenic animal for the study of human diseases in appropriate animal models [296]. It is possible some of the knowledge gained from livestock scNT, including a better understanding of epigenetic reprogramming of somatic cells to pluripotency, could improve techniques for creating induced pluripotent stem cells (iPSCs). Induced PSCs could be created from somatic cells taken from a human patient and then used for stem cell therapy in that same individual. Through this method, many of the ethical concerns of stem cell therapy could be bypassed, and treatments that match to the individual's immune system and genetics could be utilized, if current issues with mutations could be addressed [297]. The process of cloning in livestock also has direct applicability to the basic understanding of the molecular process of epigenetic reprogramming that must occur in natural development [296, 298]. The research community lacks a complete understanding of gene expression and regulation in early development and pluripotency involved with reprogramming and development, and cattle and other agricultural species may serve as a superior model to mice in this realm [299].

Methodological and Molecular Aspects of scNT

Oocytes are required for scNT, and MII stage oocytes are the normal cytoplast recipient. Oocytes obtained from a abattoir have benefits over *in vivo* oocytes derived from super-ovulatory follicles because abattoir-derived oocytes have not been subject to hormone treatment that may alter the oocyte epigenome and granulosa cell transcriptome [300, 301]. Also, the use of *in vitro*-matured oocytes provides a higher degree of control and consistency in the maturation process.

The production of animals via scNT is completely reliant on many other technologies such as oocyte recovery, oocyte maturation, enucleation, cell fusion, embryo culture, embryo vitrification and storage, and embryo transfer. Oocyte quality can be defined as the competence to yield a blastocyst within an *in vitro* production system [302], which dictates that developmental competence is not only dependent on the intrinsic quality of an oocyte but related to laboratory-specific features, including the procedures for oocyte collection, *in vitro* maturation, and IVF.

The developmental potential of embryos is significantly impacted by conditions used for the collection, manipulation, and culture of oocytes and embryos. Optimizing the temperature and gas concentration of the incubator, pH of the culture media, media composition, and even air quality are vital for minimizing stress on the cultured embryo [303-305]. Optimal handling of oocytes requires processing rapidly to reduce detriment to developmental competence, both in recovering oocytes pre-maturation, and in polar body assessment and enucleation post maturation. The post-fertilization embryo culture environment, including medium components, mineral oil, and co-culture with other cells can also have a dramatic impact on gene expression, which has implications for how researchers culturing embryos to the blastocyst stage [306-310]. Because of the differences between laboratory procedures for embryo culturing, hormonal stimulation protocols, and oocyte collection techniques, there may be challenges in repeating findings in differences between *in vitro* and *in vivo*-produced (IVP) embryos.

The first step of oocyte maturation for use in scNT is to culture the cumulus oocyte complex (COC) in maturation medium that contains species-specific hormones to promote maturation. Most laboratories utilize TCM-199 medium supplemented with fetal bovine serum, estradiol, FSH and LH [311, 312]. It is necessary to process oocytes as quickly as possible, to minimize damage. As oocytes age, they become more prone to spontaneous activation, which can cause problems with coordinating cell cycle activities [313]. Because of this issue, the protocol employed for the research described herein calls for incubation of oocytes for exactly 22 hours in

the maturation media. Oocytes will undergo spontaneous maturation after removal from the

follicular environment [314], but the maturation medium is used to modulate cAMP levels and

stimulate meiosis with FSH and epidermal growth factor (EGF) for slow meiotic progression

[315]. After maturation, the COC is disrupted, and hyaluronidase is used to break down the

hyaluronan-based matrix surrounding the cumulus cells causing their dispersal [316]. Cumulus

cells are necessary to support oocyte competence through maturation, and thus are removed only

immediately prior to enucleation [317]. Oocytes are examined for the extruded polar body, which

indicates that the oocyte has matured to the MII stage. The DNA of the oocyte is then removed,

and the metaphase plate is removed using mechanical aspiration using a microinjection system to

minimize damage to the oocyte. The oocyte must be treated with cytochalasin B or D in order to

destabilize cortical microfilaments and disrupt the plasma membrane [318]. A variable amount of

cytoplasm can be removed with mechanical aspiration, and in larger domestic species with lipid

dense cytoplasm, this approach relies on the polar body location for removal of the metaphase

plate beneath it to minimize harmful irradiation. Successful removal of the metaphase plate is

confirmed by UV inspection of the enucleation needle to confirm DNA presence [319]. Mouse

oocytes are translucent and visualization of the meiotic spindle and pronuclei can be achieved

using plane polarized light.

Different cell types can be used as nuclear donor cells from the individual animal to be

cloned, as live offspring have been generated using nuclear donor cells from cumulus cells,

fibroblasts, neuronal, and Sertoli cells with the best results from cumulus cells [320]. Relatively

few donor cell types have been shown to not work for scNT. Moreover, scNT has been employed

for many species by many research programs, including important agricultural species, such as

sheep [321], goats [322], mules [323] and cattle [324] and to important biological or pre-clinical

model species, including as mice [325], zebrafish [326] and even early humans (not allowed to

develop past blastocyst stage) [327]. While some reports found higher efficiency when using

pluripotent cells as donor cells [328], others showed higher efficiency with more differentiated

cells [329], and generally mixed results on gene expression abnormalities were found [330].

However, reports of successful cloning using highly differentiated cells, such as lymphocytes,

indicates that successful nuclear reprogramming can occur even in these highly specialized cell

types [331]. The nuclear donor is injected into the space between the oocyte cytoplasm and the

zona pellucida, and electroporation is used to disrupt the lipid bilayer of the closely apposed

membranes, which results in fusion of the donor nuclear cell and the oocyte cytoplasm [296].

Another option is direct injection, which has the advantage of delivering a karoplast devoid of

cytoplasm, but this method requires costly specialized systems and greater micromanipulation

skills and time [332]. The donor nucleus fuses with the oocyte, which is then activated with

ionomycin t mobilize intracellular calcium stores and trigger a single calcium wave, differing

from the natural repetitive calcium wave that occurs with fertilization [333]. 4-Dimethyl-

aminopyridine or cycloheximide is used to induce mitosis promoting factor (MPF) and/or

cytostatic factor, which leads to the resumption of meiosis and entry to the first embryonic cell

cycle [334]. Extrusion of the second polar body would disrupt ploidy necessary for successful

development, and cytochalasin B can be used to prevent this through disruption of microfilaments

[255, 335].

The culture medium used for embryonic culture after activation is of vital importance as

the medium contain materials that aid in development but also alter embryo quality. The medium

composition differs markedly from species to species. For example, goat and sheep embryos are

challenging to culture *in vitro* and require culture on a ligated oviduct and then prompt transfer

into the oviduct of recipients to increase embryo viability [336, 337]. However, cattle embryos

can be successfully cultured *in vitro* to the blastocyst developmental stage, followed by non-

surgical transfer into the uterus [338]. Though necessary to successfully culture embryos, fetal

bovine serum is a complex, somewhat undefined mixture containing hormones, growth factors,

and other factors that may be detrimental to early embryonic development [339]. Though culture media are designed to promote embryo growth in vitro, it is important to recognize that these formulations do not perfectly mimic the uterine embryonic environment, which can cause stress on the embryo [340]. Therefore, careful standardization of protocols, including across research laboratories, is needed to improve the reproducibility and robustness of experiments utilizing embryo culture methods.

Another critical element in scNT is the coordination of cell cycles between the donor cell and the oocyte, so as to maintain appropriate ploidy of the genome. The oocytes used for scNT should be arrested at metaphase I, and therefore have a high amount of MPF, which causes the nucleus to break down promotes DNA replication. The donor nuclear cells should then be in the G_0 or G_1 phase of the cell cycle, awaiting DNA replication in order to conserve DNA integrity and ploidy [341]. The chromatin structure in these phases may allow access to those factors present in the oocyte cytoplasm to reprogram gene expression more efficiently, although some studies have shown a favorable mitotic window for use in scNT as well [342]. Serum starvation for 24 hours or confluent cell culture can induce the quiescent G_0 state [343].

Epigenetic Errors in Reprogramming Associated with scNT

Somatic cell nuclear transfer is a well-established method for animal cloning in livestock species, whereby the genetic material from a recipient ovum is removed and then replaced with nuclear DNA from an adult donor cell. However, compared to *in vitro* fertilization technology, with successful term pregnancy rates of approximately 50% in cattle, bovine scNT embryos are much more likely to abort during pregnancy, with term rates generally below 10% [reviewed in 344]. Regardless of the species or method used for cloning, this method is generally very inefficient, with only 1-4% of embryos surviving to term as live offspring [345]. It appears that an accumulation of death occurs at every stage, with less than half of embryos able to develop

following activation and fusion, and further losses occurring after transfer into a surrogate mother with total pre-and early post-implantation losses impacting up to 70% of pregnancies [346]. After nuclear transfer, the cellular machinery of the host egg must reprogram epigenome of the somatic DNA so that the new genome is appropriately coded for a pluripotent state with the potential for cellular differentiation. Errors associated with epigenome reprogramming of somatic DNA likely lead to inappropriate gene expression, thus placing an extraordinary demand on the early developing embryo [297]. The significant cellular stress may be responsible, in part, for the high loss rate for scNT embryos throughout early development [344], as incorrect global patterns of DNA methylation and histone modifications have been demonstrated in scNT embryos. Aberrant epigenetic patterns in cloned embryos appear to be similar to the nuclear donor cell, indicating poor epigenome reprogramming may contribute to these low efficiencies [347]. While this technology is not used for human medicine or therapies, cloning does provide a valuable model to examine the epigenetic reprogramming that occurs in the germline cells and during early embryogenesis.

Offspring generated by scNT have been documented with a range of abnormalities, including large offspring syndrome (LOS) in particular, as well as high rates of abortion, high perinatal death, high birth weight, respiratory failure and oversized organs. Wilson et al. examined IVP, IVF, and scNT embryos and found that scNT offspring were 20% bigger, with weight variation 4- to 12- times that of other offspring, even when controlling for dam and sire effects [348]. LOS is common for offspring generated by IVF, but is most frequent for offspring generated by scNT [349]. Overall, the magnitude of overgrowth seen in LOS fetuses is correlated with the number of dysregulated imprinted genes, indicative of issues with aberrant gene expression causing LOS [350]. Placental perturbations may induce some of these developmental abnormalities in scNT, and appear to be due to failure to establish appropriate epigenetic marks in imprinted genes [351]. Placentas from scNT pregnancies harbor wide-scale aberrant gene

expression patterns that may contribute to some of the problems seen in these pregnancies, as dysregulation in maternal endometrium occurs as well [352, 353]. For example, dysregulation of genes in the VEGF pathway may be causative of vascular pathologies in scNT pregnancies [354]. Moreover, scNT pregnancies are typified by abnormal placental vascularization, higher frequency of abnormal placentomes, and a decrease in the mean number of placentomes [355, 356]. Increased placental mass can occur, with hydroallantois often occurring and contributing to scNT losses [357]. In addition, placenta growth does not slow at the end of gestation in scNT embryos, which could cause problems with glucose metabolism [346]. Abnormal estrogen production and metabolism is also evident in scNT pregnancies, which is problematic as estrone rise is vital for pregnancy progression at the beginning of the second trimester and at term for parturition onset. Abnormal estrogen production in scNT may contribute to the loss of pregnancies seen at the second trimester and the lack of parturition preparation, which can cause longer gestation length, which is often seen in scNT pregnancies [346].

A dramatic demonstration of the involvement of epigenetics in development is the dynamic reprogramming of gene expression that occurs during mammalian germ cell development and early embryonic development [see reviews 358, 359-361]. During germ cell development, the male and female gamete genomes are globally demethylated then re-methylated during ensuing development [362]. At fertilization, sperm genome is actively demethylated [363, 364]. The maternal genome is also demethylated during early development, but in a passive, replication-dependent manner. Global demethylation is subsequently followed by *de novo* methylation of the embryo's genome starting at the 8-cell stage in bovine embryos [365].

Global demethylation and re-methylation events take place following scNT, but differ from those of normal development in global DNA methylation levels. The difference in methylation levels suggests faulty reprogramming of epigenetic marks in scNT [347, 366-371]. DNA methylation is the first step of reprogramming and is vital for transcription factors critical

for establishment and maintenance of pluripotency, such as Oct4 [372]. ScNT embryos typically show abnormal patterns of DNA methylation and histone modifications, in contrast to IVF or IVP embryos. Genes that are methylated, and silenced, in differentiated cells may not be correctly turned back on during the scNT reprogramming, but require expression for embryonic development and pluripotency pathways [373-375]. *Oct4* was seen to be only gradually demethylated in reprogramming scNT embryos, and ineffective demethylation caused developmental retardation [376]. *Oct4* expression is tightly correlated with DNA methylation, and this gene is often aberrantly methylated in scNT embryos [377, 378].

Other pluripotency genes show similar aberrant methylation patterns in scNT embryos as well, including *NANOG* [379]. A loss of the differential methylated regions (DMRs) has been found in scNT embryos [380-384], as well as hypermethylation of the entire genome [380]. Specific methylation marks ranged from hyper- to hypomethylated at different locations in scNT embryos compared to IVF embryos, and there are examples of pluripotency genes and DMRs that were instead hypomethylated in scNT embryos [380]. Persistent cellular memory from the nuclear donor cell is likely due to the incomplete DNA methylation seen in scNT embryos. One likely cause of this incomplete reprogramming is the lack of complete removal of DNA methylation marks in scNT embryos. In IVF embryos, the DNA methylation levels were close to 0% at the 2-cell stage for the male pronuclei, and at the morula stage for the female pronuclei, before methylation starts to rise again. However, in scNT embryos, demethylation reached only about 50% before methylation increased again, and the start of *de novo* methylation occurred much earlier in the 4-cell scNT embryo, as opposed to the blastocyst stage seen in IVF embryos [379, 385].

Aberrant DNA demethylation in cloned embryos appears to be related to repressive complexes that take too long for the cell to disassemble, creating persistent chromatin states resistant to demethylation in cloned embryos. In support of this hypothesis, cloned embryos

exhibited H3K9 hypermethylation, whereas H3K9 methylation is reprogrammed in parallel with DNA methylation in normal embryos [347, 386]. Other chromatin irregularities have been found in scNT embryos as well, including abnormally high H3K4me levels [387]. Transcription factors are required to be present in the cytoplasm of the oocyte and to be able to access the DNA, which means that chromatin must be open and amenable to change. When histone modifications of the donor cell are inappropriately maintained, they act as a major barrier for efficient reprogramming and prevent correct epigenetic reprogramming, which has been shown to occur in scNT embryos [388]. Interestingly, Kang et al. showed that the transcriptome of scNT blastocyst embryos was a hybrid of somatic and embryonic transcripts [389]. Higher levels of DNMT1 were present in scNT embryos, which may stem from active expression in the nuclear donor cell that is partially maintained, and may also contribute to incomplete demethylation seen [385]. The somatic isoform of DNMT1 was also higher in scNT embryos compared to the normal expression of DNMT1o seen in early embryos during reprogramming [390]. Overall, aberrant DNA methylation was unique to the gene, whereas hypermethylation was noted genome-wide, and these abnormal levels may be due to abnormal DNMT1 activity or loss of normal demethylation pathways. The result is a persistence of DNA methylation profile from the donor nuclear cell [389], which can cause inappropriate gene expression and loss of embryo developmental competence.

These aberrant epigenome states, unsurprisingly, lead to marked aberrant gene expression. Overall, transcriptome analyses of scNT embryos have shown a multitude of aberrant pathways as compared to IVF embryos. In pigs, dozens of dysregulated transcription factors were expressed in cloned embryos, as well as genes involved with histone lysine methylation [391]. In buffalo, transcripts related to developmental regulation, epigenetic modifications, and pluripotency were aberrantly expressed in scNT blastocyst embryos [392]. In cattle, transcriptome features that distinguished blastocysts or elongated conceptus generated by scNT

compared from IVF or IVP controls, respectively, included pluripotency genes, TE development genes, developmental regulators, metabolism and epigenetic modifiers [393, 394]. Histone modifiers were found to be vitally important for development in scNT as well, and dysregulated [395]. In mice, after the EGA, 20% of genes with abundant expression in the donor cell maintained a high expression level inappropriate for embryos, and 15% of genes silenced in the donor cell and maintained a low expression level inappropriate for embryos, indicative of incomplete programming [297]. These large-scale changes in transcriptome represent cellular stresses to the embryo, and when many of these genes are inaccurately expressed in the early embryo, embryonic death occurs.

Differences in miRNA profiles between scNT embryos and IVF embryos have also been identified. For example, Hossain et al. revealed that 278 of the 377 miRNAs examined were down-regulated in bovine scNT embryos as compared to bovine IVF embryos at the blastocyst stage [396]. It also appears that incomplete reprogramming of the donor cell's genome occurred in those scNT bovine embryos, as some miRNAs maintained expression patterns similar to the donor cells. Moreover, differences in scNT placentas as compared to IVP placentas and placentas from artificial insemination pregnancies in miRNAs were also identified, as 62% of miRNAs were differentially expressed in scNT placentas compared to IVP placentas with the majority of those down-regulated [396]. The aberrations in miRNA profiles for placental tissues for scNT pregnancies were likely associated with dysregulation of gene expression that may contribute to pregnancy loss. In another study, researchers examined miRNA profiles in elongated bovine embryos at day 17 and noted that not all miRNAs measured by microarray appeared appropriately reprogrammed in the scNT embryos, as their expression was more similar to the somatic donor cell [397]. Of the 50 miRNAs subject to reprogramming, ten were aberrantly reprogrammed and another nine were not reprogrammed [397]. To date, these studies are the only reports to assess expression of miRNA in bovine scNT. Importantly, these studies may not have captured some

miRNAs of developmental importance, as they did not utilize RNA sequencing or explore the dynamics of miRNA expression during the activation of the embryonic genome in cloned embryos.

Improvements in scNT rates have been attempted in a multitude of ways, often by targeting improvements in epigenetic reprogramming. Caffeine was used to increase MPF and MAPK activity in sheep oocytes in order to release chromatin-bound factors and allow greater access for reprogramming the donor chromatin. Researchers have also attempted to change some of the donor cells' somatic epigenome through use of chemicals that impact DNA methylation or histone modification, such as 5-aza-2 deoxycytidine (5-aza-dC) and trichostatin A (TSA). Overall, many of these treatment options give different results when used in combination, or when used on donor cells as opposed to use on scNT embryos; also, different results were obtained by different research groups [398]. The fact that these chemicals have the ability, even in limited situations, to improve scNT outcomes demonstrates that these epigenetic errors in reprogramming may be the cause of low efficiency.

Overall, the weakness in these strategies is that they act in a genome-wide, non-specific manner. Alternatively, targeting specific miRNAs identified as aberrant in scNT embryos may be a more focused approach to correcting transcriptional abnormalities in an effort to improve development rates. MiRNA mimics or miRNA sponges are easily synthesized and can be applied to culture methods by microinjection directly into the embryo or may be absorbed into the embryo when added to culture media. In fact, a several studies have already shown that modulation of specific miRNAs can impact rates of development in embryos [270, 399], including the improvement of scNT embryo development [400-405].

Bovine Model for Embryo Development

The bovine scNT embryo is an excellent model for the study of epigenetic reprogramming and gene regulation in early development. The results of experiments designed to better understand critical aspects of epigenetic remodeling, including a potential role of miRNA, would prove extremely useful when applied to iPSC technologies. In addition, early embryogenesis in mouse differs from other mammals in several key aspects. Because of the rodent-specific pathways for biogenesis of sncRNAs in early development, as outlined above, moue models are not ideal for study of early embryo development with the intention to learn about processes that inform human early embryogenesis. Furthermore, studies have cast doubt on the suitability of the mouse embryo as an optimal model for all mammalian fertilization and development [reviewed in 406]. Cattle are also one of the most important agricultural species around the world, and significant resources have been dedicated to improving ART techniques and breeding efficiency in this species. Therefore, the bovine model is not only the more appropriate model (over rodents) to examine sncRNA in early embryogenesis and possible conserved functions in the MET, but also to improve our the understanding of the cause of low efficiency rates for scNT. Such new knowledge could lead to improvements in scNT an important agricultural species.

Project Objectives and Overarching Hypothesis

Cloned embryos harbor substantial problems in genome reprogramming, including extensive aberrant methylation and histone marks for key genes necessary for development. While investigators have studied DNA methylation and histone modifications as mechanisms for aberrant epigenome programming, little focus has been given to the role of sncRNA in this process and no researchers have explored the dynamic expression of sncRNAs during the period of embryonic genome activation in cattle. Thus, the goal of this project was to determine the

dynamic expression of sncRNAs and mRNAs in scNT and IVF bovine embryos during early development through the MET. The overarching hypothesis for this work was that aberrant genome reprogramming of the donor cell genome would lead to dysregulation of sncRNAs, particularly for miRNAs, that would in turn drive abnormal gene expression during this critical window of embryonic development. New knowledge gained from these studies will greatly increase our understanding of mechanisms controlling epigenetic reprogramming of somatic DNA and will significantly contribute to our long-term goal of reliably reprogramming the nuclei of multi-potent and differentiated cells. Chapter 2 describes initial work to profile the array of miRNAs present in bovine oocytes, 8-cell embryos and blastocyst-staged embryos generated by IVF using both unbiased (RNAseq) and targeted (PCR) methods.   Chapter 3 details comparison of sncRNA profiles – focusing on miRNA, tRFs and piRNAs – in cattle embryos produced by scNT and IVF over the course of early embryo development, from the 2-cell to blastocyst stage, that encompasses the MET in cattle. Chapter 4 presents a parallel set of data describing the mRNA transcriptome in cloned and IVF embryos, also during the MET, and maps those transcriptional changes to sncRNA profiles obtained for the same samples. Last, Chapter 5 summarizes our major findings and presents a perspective on opportunities and challenges for the field of reproductive cloning and the study of molecular aspects of early development.

References

1.   Krebs J, Goldstein ES, Kilpatrick ST (ed.). Lewin's Genes X. Sadbury Massachusetts: Jones and Bartlett publishers; 2011;p. 549-578

2.   Preiss T. The End in Sight: Poly(A), translation and mRNA stability in eukaryotes. Austin TX: Landes Bioscience; 2013.

3.   de Klerk E, t Hoen PA. Alternative mRNA transcription, processing, and translation: insights from RNA sequencing. Trends Genet. 2015;31(3):128-139.

4.   Yang E, van Nimwegen E, Zavolan M, Rajewsky N, Schroeder M, Magnasco M, Darnell JE, Jr.. Decay rates of human mRNAs: correlation with functional characteristics and sequence attributes. Genome Res. 2003;13(8):1863-1872.

5.   Gebauer F, Preiss T, Hentze MW. From cis-regulatory elements to complex RNPs and back. Cold Spring Harb Perspect Biol. 2012;4(7):a012245.

6.   Harcourt EM, Kietrys AM, Kool ET. Chemical and structural effects of base modifications in messenger RNA. Nature. 2017;541(7637):339-346.

7.   Dreyfuss G, Kim VN, Kataoka N. Messenger-RNA-binding proteins and the messages they carry. Nat Rev Mol Cell Biol. 2002;3(3):195-205.

8.   Moore MJ. From birth to death: the complex lives of eukaryotic mRNAs. Science. 2005;309(5740):1514-1518.

9.   Spitz F, Furlong EE. Transcription factors: from enhancer binding to developmental control. Nat Rev Genet. 2012;13(9):613-626.

10.   Riechmann JL. Transcriptional regulation: a genomic overview. Arabidopsis Book. 2002;1:e0085.

11.   Brawerman G. mRNA degradation in eukaryotic cells: an overview. In: Control of Messenger RNA stability. Edited by Brawerman JGBaG. Elsevier; 1993: 149-159.

12.   Stevens A. Eukaryotic nucleases and mRNA turnover. In: Control of mRNA stability. Edited by Brawerman JGBaG; 1991: 449-471.

13.   Houseley J, Tollervey D. The many pathways of RNA degradation. Cell. 2009;136(4):763-776.

14.   Tourriere H, Chebli K, Tazi J. mRNA degradation machines in eukaryotic cells. Biochimie. 2002;84(8):821-837.

15.   Wu X, Brewer G. The regulation of mRNA stability in mammalian cells: 2.0. Gene. 2012;500(1):10-21.

16.   Parker R, Sheth U. P bodies and the control of mRNA translation and degradation. Mol Cell. 2007;25(5):635-646.

17.   Decker CJ, Parker R. P-bodies and stress granules: possible roles in the control of translation and mRNA degradation. Cold Spring Harb Perspect Biol. 2012;4(9):a012286.

18.   Ramchandran R, Chaluvally-Raghavan P. miRNA-mediated RNA activation in mammalian cells. Adv Exp Med Biol. 2017;983:81-89.

19.   Dalmay T. Mechanism of miRNA-mediated repression of mRNA translation. Essays Biochem. 2013;54:29-38.

20.   Carthew RW, Sontheimer EJ. Origins and mechanisms of miRNAs and siRNAs. Cell. 2009;136(4):642-655.

21.   Valencia-Sanchez MA, Liu J, Hannon GJ, Parker R. Control of translation and mRNA degradation by miRNAs and siRNAs. Genes Dev. 2006;20(5):515-524.

22.   Fabian MR, Sonenberg N, Filipowicz W. Regulation of mRNA translation and stability by microRNAs. Annu Rev Biochem. 2010;79:351-379.

23.   Humphreys DT, Westman BJ, Martin DI, Preiss T. MicroRNAs control translation initiation by inhibiting eukaryotic initiation factor 4E/cap and poly(A) tail function. Proc Natl Acad Sci USA. 2005;102(47):16961-16966.

24.   Stefani G, Slack FJ. Small non-coding RNAs in animal development. Nat Rev Mol Cell Biol. 2008;9(3):219-230.

25.   Hu W, Coller J. What comes first: translational repression or mRNA degradation? The deepening mystery of microRNA function. Cell Res. 2012;22(9):1322-1324.

26.   Wu L, Fan J, Belasco JG. MicroRNAs direct rapid deadenylation of mRNA. Proc Natl Acad Sci USA. 2006;103(11):4034-4039.

27.   Eichhorn SW, Guo H, McGeary SE, Rodriguez-Mias RA, Shin C, Baek D, Hsu SH, Ghoshal K, Villen J, Bartel DP. mRNA destabilization is the dominant effect of mammalian microRNAs by the time substantial repression ensues. Mol Cell. 2014;56(1):104-115.

28.   Cottrell KA, Szczesny P, Djuranovic S. Translation efficiency is a determinant of the magnitude of miRNA-mediated repression. Sci Rep. 2017;7(1):14884.

29.   Catalanotto C, Cogoni C, Zardo G. MicroRNA in control of gene expression: an overview of nuclear functions. Int J Mol Sci. 2016;17(10).

30.   Clarke SL, Davis MR, Sunkar R. Biogenesis of mammalian miRNA. In: MicroRNAs as Tools in Biopharmaceutical Production. Springer Science; 2012.

31.   Kim VN, Han J, Siomi MC. Biogenesis of small RNAs in animals. Nat Rev Mol Cell Biol. 2009;10(2):126-139.

32.   Winter J, Jung S, Keller S, Gregory RI, Diederichs S. Many roads to maturity: microRNA biogenesis pathways and their regulation. Nat Cell Biol. 2009;11(3):228-234.

33.   Maurin T, Cazalla D, Yang S, Jr., Bortolamiol-Becet D, Lai EC. RNase III-independent microRNA biogenesis in mammalian cells. RNA. 2012;18(12):2166-2173.

34.   Ruby JG, Jan CH, Bartel DP. Intronic microRNA precursors that bypass Drosha processing. Nature. 2007;448(7149):83-86.

35.   Babiarz JE, Ruby JG, Wang Y, Bartel DP, Blelloch R. Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. Genes Dev. 2008;22(20):2773-2785.

36.   Xie M, Li M, Vilborg A, Lee N, Shu MD, Yartseva V, Sestan N, Steitz JA. Mammalian 5'-capped microRNA precursors that generate a single microRNA. Cell. 2013;155(7):1568-1580.

37. Yang JS, Maurin T, Robine N, Rasmussen KD, Jeffrey KL, Chandwani R, Papapetrou EP, Sadelain M, O'Carroll D, Lai EC. Conserved vertebrate mir-451 provides a platform for Dicer-independent, Ago2-mediated microRNA biogenesis. Proc Natl Acad Sci USA. 2010;107(34):15163-15168.

38. Park CY, Choi YS, McManus MT. Analysis of microRNA knockouts in mice. Hum Mol Genet. 2010;19(R2):R169-175.

39. Kim YK, Kim B, Kim VN. Re-evaluation of the roles of DROSHA, Export in 5, and DICER in microRNA biogenesis. Proc Natl Acad Sci USA. 2016;113(13):E1881-1889.

40. Liu J, Valencia-Sanchez MA, Hannon GJ, Parker R. MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies. Nat Cell Biol. 2005;7(7):719-723.

41. Chan SP, Slack FJ. microRNA-mediated silencing inside P-bodies. RNA Biol. 2006;3(3):97-100.

42. Eulalio A, Behm-Ansmant I, Izaurralde E. P bodies: at the crossroads of post-transcriptional pathways. Nat Rev Mol Cell Biol. 2007;8(1):9-22.

43. Eulalio A, Behm-Ansmant I, Schweizer D, Izaurralde E. P-body formation is a consequence, not the cause, of RNA-mediated gene silencing. Mol Cell Biol. 2007;27(11):3970-3981.

44. O'Brien J, Hayder H, Zayed Y, Peng C. Overview of microRNA biogenesis, mechanisms of actions, and circulation. Front Endocrinol (Lausanne). 2018;9:402.

45. Filipow S, Laczmanski L. Blood circulating miRNAs as cancer biomarkers for diagnosis and surgical treatment response. Front Genet. 2019;10:169.

46. Kosaka N, Iguchi H, Yoshioka Y, Takeshita F, Matsuki Y, Ochiya T. Secretory mechanisms and intercellular transfer of microRNAs in living cells. J Biol Chem. 2010;285(23):17442-17452.

47. Guduric-Fuchs J, O'Connor A, Camp B, O'Neill CL, Medina RJ, Simpson DA. Selective extracellular vesicle-mediated export of an overlapping set of microRNAs from multiple cell types. BMC Genomics. 2012;13:357.

48. Zhang J, Li S, Li L, Li M, Guo C, Yao J, Mi S. Exosome and exosomal microRNA: trafficking, sorting, and function. Genomics Proteomics Bioinformatics. 2015;13(1):17-24.

49. Zhang L, Hou D, Chen X, Li D, Zhu L, Zhang Y, Li J, Bian Z, Liang X, Cai X. Exogenous plant MIR168a specifically targets mammalian LDLRAP1: evidence of cross-kingdom regulation by microRNA. Cell Res. 2012;22(1):107-126.

50. Witwer KW, Hirschi KD. Transfer and functional consequences of dietary microRNAs in vertebrates: concepts in search of corroboration: negative results challenge the hypothesis

that dietary xenomiRs cross the gut and regulate genes in ingesting vertebrates, but important questions persist. Bioessays. 2014;36(4):394-406.

51. Li J, Zhang Y, Li D, Liu Y, Chu D, Jiang X, Hou D, Zen K, Zhang CY. Small non-coding RNAs transfer through mammalian placenta and directly regulate fetal gene expression. Protein Cell. 2015;6(6):391-396.

52. Zhang Z, Qin YW, Brewer G, Jing Q. MicroRNA degradation and turnover: regulating the regulators. Wiley Interdiscip Rev RNA. 2012;3(4):593-600.

53. Bail S, Swerdel M, Liu H, Jiao X, Goff LA, Hart RP, Kiledjian M. Differential regulation of microRNA stability. RNA. 2010;16(5):1032-1039.

54. Zhang Z, Zou J, Wang GK, Zhang JT, Huang S, Qin YW, Jing Q. Uracils at nucleotide position 9-11 are required for the rapid turnover of miR-29 family. Nucleic Acids Res. 2011;39(10):4387-4395.

55. Fuchs Wightman F, Giono LE, Fededa JP, de la Mata M. Target RNAs strike back on microRNAs. Front Genet. 2018;9:435.

56. Asada K, Canestrari E, Fu X, Li Z, Makowski E, Wu YC, Mito JK, Kirsch DG, Baraban J, Paroo Z. Rescuing dicer defects via inhibition of an anti-dicing nuclease. Cell Rep. 2014;9(4):1471-1481.

57. Ha M, Kim VN. Regulation of microRNA biogenesis. Nat Rev Mol Cell Biol. 2014;15(8):509-524.

58. Wang L, Nam Y, Lee AK, Yu C, Roth K, Chen C, Ransey EM, Sliz P. LIN28 zinc knuckle domain is required and sufficient to induce let-7 oligouridylation. Cell Rep. 2017;18(11):2664-2675.

59. Nowak JS, Hobor F, Downie Ruiz Velasco A, Choudhury NR, Heikel G, Kerr A, Ramos A, Michlewski G. Lin28a uses distinct mechanisms of binding to RNA and affects miRNA levels positively and negatively. RNA. 2017;23(3):317-332.

60. Kim YK, Heo I, Kim VN. Modifications of small RNAs and their associated proteins. Cell. 2010;143(5):703-709.

61. Neilsen CT, Goodall GJ, Bracken CP. IsomiRs--the overlooked repertoire in the dynamic microRNAome. Trends Genet. 2012;28(11):544-549.

62. Krol J, Loedige I, Filipowicz W. The widespread regulation of microRNA biogenesis, function and decay. Nat Rev Genet. 2010;11(9):597-610.

63. Haas G, Cetin S, Messmer M, Chane-Woon-Ming B, Terenzi O, Chicher J, Kuhn L, Hammann P, Pfeffer S. Identification of factors involved in target RNA-directed microRNA degradation. Nucleic Acids Res. 2016;44(6):2873-2887.

64. D'Ambrogio A, Gu W, Udagawa T, Mello CC, Richter JD. Specific miRNA stabilization by Gld2-catalyzed monoadenylation. Cell Rep. 2012;2(6):1537-1545.

65. Song J, Song J, Mo B, Chen X. Uridylation and adenylation of RNAs. Sci China Life Sci. 2015;58(11):1057-1066.

66. Heo I, Ha M, Lim J, Yoon MJ, Park JE, Kwon SC, Chang H, Kim VN. Mono-uridylation of pre-microRNA as a key step in the biogenesis of group II let-7 microRNAs. Cell. 2012;151(3):521-532.

67. Lee M, Choi Y, Kim K, Jin H, Lim J, Nguyen TA, Yang J, Jeong M, Giraldez AJ, Yang H et al. Adenylation of maternally inherited microRNAs by Wispy. Mol Cell. 2014;56(5):696-707.

68. Okamura K, Lai EC. Endogenous small interfering RNAs in animals. Nat Rev Mol Cell Biol. 2008;9(9):673-678.

69. Ghildiyal M, Zamore PD. Small silencing RNAs: an expanding universe. Nat Rev Genet. 2009;10(2):94-108.

70. Flemr M, Malik R, Franke V, Nejepinska J, Sedlacek R, Vlahovicek K, Svoboda P. A retrotransposon-driven dicer isoform directs endogenous small interfering RNA production in mouse oocytes. Cell. 2013;155(4):807-816.

71. Ozata DM, Gainetdinov I, Zoch A, O'Carroll D, Zamore PD. PIWI-interacting RNAs: small RNAs with big functions. Nat Rev Genet. 2019;20(2):89-108.

72. Roovers EF, Rosenkranz D, Mahdipour M, Han CT, He N, Chuva de Sousa Lopes SM, van der Westerlaken LA, Zischler H, Butter F, Roelen BA et al. Piwi proteins and piRNAs in mammalian oocytes and early embryos. Cell Rep. 2015;10(12):2069-2082.

73. Vagin VV, Sigova A, Li C, Seitz H, Gvozdev V, Zamore PD. A distinct small RNA pathway silences selfish genetic elements in the germline. Science. 2006;313(5785):320-324.

74. Gou LT, Dai P, Yang JH, Xue Y, Hu YP, Zhou Y, Kang JY, Wang X, Li H, Hua MM. Pachytene piRNAs instruct massive mRNA elimination during late spermiogenesis. Cell Res. 2014;24(6):680-700.

75. Bui LC, Evsikov AV, Khan DR, Archilla C, Peynot N, Henaut A, Le Bourhis D, Vignon X, Renard JP, Duranthon V. Retrotransposon expression as a defining event of genome reprogramming in fertilized and cloned bovine embryos. Reproduction. 2009;138(2):289-299.

76. Kabayama Y, Toh H, Katanaya A, Sakurai T, Chuma S, Kuramochi-Miyagawa S, Saga Y, Nakano T, Sasaki H. Roles of MIWI, MILI and PLD6 in small RNA regulation in mouse growing oocytes. Nucleic Acids Res. 2017;45(9):5387-5398.

77. Weick EM, Miska EA. piRNAs: from biogenesis to function. Development. 2014;141(18):3458-3471.

78. Chen D, Zhang Z, Chen B, Ji D, Hao Y, Zhou P, Wei Z, Cao Y. Altered microRNA and Piwi-interacting RNA profiles in cumulus cells from patients with diminished ovarian reserve. Biol Reprod. 2017;97(1):91-103.

79. Khurana JS, Xu J, Weng Z, Theurkauf WE. Distinct functions for the Drosophila piRNA pathway in genome maintenance and telomere protection. PLoS Genet. 2010;6(12):e1001246.

80. Rouget C, Papin C, Boureux A, Meunier AC, Franco B, Robine N, Lai EC, Pelisson A, Simonelig M. Maternal mRNA deadenylation and decay by the piRNA pathway in the early Drosophila embryo. Nature. 2010;467(7319):1128-1132.

81. Molaro A, Falciatori I, Hodges E, Aravin AA, Marran K, Rafii S, McCombie WR, Smith AD, Hannon GJ. Two waves of de novo methylation during mouse germ cell development. Genes Dev. 2014;28(14):1544-1549.

82. Aravin AA, Bourc'his D. Small RNA guides for de novo DNA methylation in mammalian germ cells. Genes Dev. 2008;22(8):970-975.

83. Yang F, Wang PJ. Multiple LINEs of retrotransposon silencing mechanisms in the mammalian germline. Semin Cell Dev Biol. 2016;59:118-125.

84. Munoz-Lopez M, Garcia-Perez JL. DNA transposons: nature and applications in genomics. Curr Genomics. 2010;11(2):115-128.

85. Elbarbary RA, Lucas BA, Maquat LE. Retrotransposons as regulators of gene expression. Science. 2016;351(6274):aac7247.

86. Nataša Lindič MB, Toni Petan, Binyamin A Knisbacher, Erez Y Levanon & Nika Lovšin Differential inhibition of LINE1 and LINE2 retrotransposition by vertebrate AID/APOBEC proteins. Retrovirology. 2013;10.

87. Richardson SR, Doucet AJ, Kopera HC, Moldovan JB, Garcia-Perez JL, Moran JV. The Influence of LINE-1 and SINE retrotransposons on mammalian genomes. Microbiol Spectr. 2015;3(2):MDNA3-0061-2014.

88. Graham T, Boissinot S. The genomic distribution of L1 elements: the role of insertion bias and natural selection. J Biomed Biotechnol. 2006;2006(1):75327.

89. Hellen EH, Brookfield JF. Alu elements in primates are preferentially lost from areas of high GC content. PeerJ. 2013;1:e78.

90. Beauregard A, Curcio MJ, Belfort M. The take and give between retrotransposable elements and their hosts. Annu Rev Genet. 2008;42:587-617.

91. Smit AF. Identification of a new, abundant superfamily of mammalian LTR-transposons. Nucleic Acids Res. 1993;21(8):1863-1872.

92. Meseure D, Alsibai KD. Part 1: The PIWI-piRNA pathway is an immune-like surveillance process that controls genome integrity by silencing transposable elements. In: Epigenetics and Chromatin. DOI: 10.5772/intechopen.79974

93. Brennecke J, Malone CD, Aravin AA, Sachidanandam R, Stark A, Hannon GJ. An epigenetic role for maternally inherited piRNAs in transposon silencing. Science. 2008;322(5906):1387-1392.

94. Zamore PD. Somatic piRNA biogenesis. EMBO J. 2010;29(19):3219-3221.

95. Malone CD, Brennecke J, Dus M, Stark A, McCombie WR, Sachidanandam R, Hannon GJ. Specialized piRNA pathways act in germline and somatic tissues of the Drosophila ovary. Cell. 2009;137(3):522-535.

96. Rak R, Dahan O, Pilpel Y. Repertoires of tRNAs: the couplers of genomics and proteomics. Annu Rev Cell Dev Biol. 2018;34:239-264.

97. Lodish H BA, Zipursky SL et al.. Section 4.4, The three roles of RNA in protein synthesis. In: Molecular Cell Biology Edited by Freeman WH, vol. 4th edition. New York; 2000.

98. Pan T. Modifications and functional genomics of human transfer RNA. Cell Res. 2018;28(4):395-404.

99. Stoecklin G, Diederichs S. tRNAs: new tricks from old dogs. EMBO J. 2014;33(18):1981-1983.

100. Megel C, Morelle G, Lalande S, Duchene AM, Small I, Marechal-Drouard L. Surveillance and cleavage of eukaryotic tRNAs. Int J Mol Sci. 2015;16(1):1873-1893.

101. Thompson DM, Parker R. Stressing out over tRNA cleavage. Cell. 2009;138(2):215-219.

102. Oberbauer V, Schaefer MR. tRNA-derived small RNAs: biogenesis, modification, function and potential impact on human disease development. Genes (Basel). 2018;9(12).

103. Nicholson, AW. Nucleic Acids and Molecular Biology, vol. 28; 2011.

104. Lee YS, Shibata Y, Malhotra A, Dutta A. A novel class of small RNAs: tRNA-derived RNA fragments (tRFs). Genes Dev. 2009;23(22):2639-2649.

105. Schorn AJ, Gutbrod MJ, LeBlanc C, Martienssen R. LTR-retrotransposon control by tRNA-derived small RNAs. Cell. 2017;170(1):61-71 e11.

106. Zhu L, Ge J, Li T, Shen Y, Guo J. tRNA-derived fragments and tRNA halves: the new players in cancers. Cancer Lett. 2019;452:31-37.

107.   Matera AG, Terns RM, Terns MP. Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs. Nat Rev Mol Cell Biol. 2007;8(3):209-220.

108.   Valadkhan S, Gunawardane LS. Role of small nuclear RNAs in eukaryotic gene expression. Essays Biochem. 2013;54:79-90.

109.   Wu G, Yu AT, Kantartzis A, Yu YT. Functions and mechanisms of spliceosomal small nuclear RNA pseudouridylation. Wiley Interdiscip Rev RNA. 2011;2(4):571-581.

110.   Shukla S, Parker R. Quality control of assembly-defective U1 snRNAs by decapping and 5'-to-3' exonucleolytic digestion. Proc Natl Acad Sci USA. 2014;111(32):E3277-3286.

111.   Ishikawa H, Nobe Y, Izumikawa K, Yoshikawa H, Miyazawa N, Terukina G, Kurokawa N, Taoka M, Yamauchi Y, Nakayama H. Identification of truncated forms of U1 snRNA reveals a novel RNA degradation pathway during snRNP biogenesis. Nucleic Acids Res. 2014;42(4):2708-2724.

112.   Sauterer RA, Feeney RJ, Zieve GW. Cytoplasmic assembly of snRNP particles from stored proteins and newly transcribed snRNA's in L929 mouse fibroblasts. Exp Cell Res. 1988;176(2):344-359.

113.   Kiss T. Small nucleolar RNAs: an abundant group of noncoding RNAs with diverse cellular functions. Cell. 2002;109(2):145-148.

114.   Scott MS, Ono M. From snoRNA to miRNA: dual function regulatory non-coding RNAs. Biochimie. 2011;93(11):1987-1992.

115.   Falaleeva M, Pages A, Matuszek Z, Hidmi S, Agranat-Tamir L, Korotkov K, Nevo Y, Eyras E, Sperling R, Stamm S. Dual function of C/D box small nucleolar RNAs in rRNA modification and alternative pre-mRNA splicing. Proc Natl Acad Sci USA. 2016;113(12):E1625-1634.

116.   Kufel J, Grzechnik P. Small nucleolar RNAs tell a different tale. Trends Genet. 2019;35(2):104-117.

117.   Williams GT, Farzaneh F. Are snoRNAs and snoRNA host genes new players in cancer? Nat Rev Cancer. 2012;12(2):84-88.

118.   Taft RJ, Glazov EA, Lassmann T, Hayashizaki Y, Carninci P, Mattick JS. Small RNAs derived from snoRNAs. RNA. 2009;15(7):1233-1240.

119.   Falaleeva M, Stamm S. Processing of snoRNAs as a new source of regulatory non-coding RNAs: snoRNA fragments form a new class of functional RNAs. Bioessays. 2013;35(1):46-54.

120.   Ender C, Krek A, Friedlander MR, Beitzinger M, Weinmann L, Chen W, Pfeffer S, Rajewsky N, Meister G. A human snoRNA with microRNA-like functions. Mol Cell. 2008;32(4):519-528.

121. Saraiya AA, Wang CC. snoRNA, a novel precursor of microRNA in Giardia lamblia. PLoS Pathog. 2008;4(11):e1000224.

122. Kondrashov N, Pusic A, Stumpf CR, Shimizu K, Hsieh AC, Ishijima J, Shiroishi T, Barna M. Ribosome-mediated specificity in Hox mRNA translation and vertebrate tissue patterning. Cell. 2011;145(3):383-397.

123. Luttermann C, Meyers G. The importance of inter- and intramolecular base pairing for translation reinitiation on a eukaryotic bicistronic mRNA. Genes Dev. 2009;23(3):331-344.

124. Mauro VP, Edelman GM. The ribosome filter redux. Cell Cycle. 2007;6(18):2246-2251.

125. Burman LG, Mauro VP. Analysis of rRNA processing and translation in mammalian cells using a synthetic 18S rRNA expression system. Nucleic Acids Res. 2012;40(16):8085-8098.

126. Bastide A, David A. Interaction of rRNA with mRNA and tRNA in translating mammalian ribosome: functional implications in health and disease. Biomolecules. 2018;8(4).

127. Moss T, Stefanovsky V, Langlois F, Gagnon-Kugler T. A new paradigm for the regulation of the mammalian ribosomal RNA genes. Biochem Soc Trans. 2006;34(Pt 6):1079-1081.

128. de la Cruz J, Karbstein K, Woolford JL, Jr.. Functions of ribosomal proteins in assembly of eukaryotic ribosomes in vivo. Annu Rev Biochem. 2015;84:93-129.

129. Sloan KE, Warda AS, Sharma S, Entian KD, Lafontaine DLJ, Bohnsack MT. Tuning the ribosome: the influence of rRNA modification on eukaryotic ribosome biogenesis and function. RNA Biol. 2017;14(9):1138-1152.

130. Waddington CH. The epigenotype. Endeavour. 1942:18-20.

131. Dupont C, Armant DR, Brenner CA. Epigenetics: definition, mechanisms and clinical perspective. Semin Reprod Med. 2009;27(5):351-357.

132. Moore LD, Le T, Fan G. DNA methylation and its basic function. Neuropsychopharmacology. 2013;38(1):23-38.

133. Lyko F. The DNA methyltransferase family: a versatile toolkit for epigenetic regulation. Nat Rev Genet. 2018;19(2):81-92.

134. Saxonov S, Berg P, Brutlag DL. A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. Proc Natl Acad Sci USA. 2006;103(5):1412-1417.

135. Gardiner-Garden M, Frommer M. CpG islands in vertebrate genomes. J Mol Biol. 1987;196(2):261-282.

136. Mohn F, Weber M, Rebhan M, Roloff TC, Richter J, Stadler MB, Bibel M, Schubeler D. Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. Mol Cell. 2008;30(6):755-766.

137. Aran D, Toperoff G, Rosenberg M, Hellman A. Replication timing-related and gene body-specific methylation of active human genes. Hum Mol Genet. 2011;20(4):670-680.

138. Razin A. CpG methylation, chromatin structure and gene silencing-a three-way connection. EMBO J. 1998;17(17):4905-4908.

139. Zilberman D, Coleman-Derr D, Ballinger T, Henikoff S. Histone H2A.Z and DNA methylation are mutually antagonistic chromatin marks. Nature. 2008;456(7218):125-129.

140. Murphy PJ, Wu SF, James CR, Wike CL, Cairns BR. Placeholder nucleosomes underlie germline-to-embryo DNA methylation reprogramming. Cell. 2018;172(5):993-1006 e1013.

141. Ehrlich M. DNA hypomethylation in cancer cells. Epigenomics. 2009;1(2):239-259.

142. Kim M, Costello J. DNA methylation: an epigenetic mark of cellular memory. Exp Mol Med. 2017;49(4):e322.

143. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science. 2009;324(5929):930-935.

144. Wu X, Zhang Y. TET-mediated active DNA demethylation: mechanism, function and beyond. Nat Rev Genet. 2017;18(9):517-534.

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

146. Raynal NJ, Si J, Taby RF, Gharibyan V, Ahmed S, Jelinek J, Estecio MR, Issa JP. DNA methylation does not stably lock gene expression but instead serves as a molecular mark for gene silencing memory. Cancer Res. 2012;72(5):1170-1181.

147. Lawrence M, Daujat S, Schneider R. Lateral thinking: how histone modifications regulate gene expression. Trends Genet. 2016;32(1):42-56.

148. Musselman CA, Lalonde ME, Cote J, Kutateladze TG. Perceiving the epigenetic landscape through histone readers. Nat Struct Mol Biol. 2012;19(12):1218-1227.

149. Bowman GD, Poirier MG. Post-translational modifications of histones that influence nucleosome dynamics. Chem Rev. 2015;115(6):2274-2295.

150. Gates LA, Foulds CE, O'Malley BW. Histone marks in the 'driver's seat': functional roles in steering the transcription cycle. Trends Biochem Sci. 2017;42(12):977-989.

151. Golbabapour S, Majid NA, Hassandarvish P, Hajrezaie M, Abdulla MA, Hadi AH. Gene silencing and Polycomb group proteins: an overview of their structure, mechanisms and phylogenetics. OMICS. 2013;17(6):283-296.

152. Laugesen A, Hojfeldt JW, Helin K. Molecular mechanisms directing PRC2 recruitment and H3K27 methylation. Mol Cell. 2019;74(1):8-18.

153. Marmorstein R, Trievel RC. Histone modifying enzymes: structures, mechanisms, and specificities. Biochim Biophys Acta. 2009;1789(1):58-68.

154. Fuhrmann J, Thompson PR. Protein arginine methylation and citrullination in epigenetic regulation. ACS Chem Biol. 2016;11(3):654-668.

155. Verdone L, Caserta M, Di Mauro E. Role of histone acetylation in the control of gene expression. Biochem Cell Biol. 2005;83(3):344-353.

156. Rossetto D, Avvakumov N, Cote J. Histone phosphorylation: a chromatin modification involved in diverse nuclear events. Epigenetics. 2012;7(10):1098-1108.

157. Alemasova EE, Lavrik OI. Poly(ADP-ribosyl)ation by PARP1: reaction mechanism and regulatory proteins. Nucleic Acids Res. 2019;47(8):3811-3827.

158. Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. Cell Res. 2011;21(3):381-395.

159. Liu H, Lei C, He Q, Pan Z, Xiao D, Tao Y. Nuclear functions of mammalian MicroRNAs in gene regulation, immunity and cancer. Mol Cancer. 2018;17(1):64.

160. Feldman N, Gerson A, Fang J, Li E, Zhang Y, Shinkai Y, Cedar H, Bergman Y. G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. Nat Cell Biol. 2006;8(2):188-194.

161. Fukagawa T, Nogami M, Yoshikawa M, Ikeno M, Okazaki T, Takami Y, Nakayama T, Oshimura M. Dicer is essential for formation of the heterochromatin structure in vertebrate cells. Nat Cell Biol. 2004;6(8):784-791.

162. Joh RI, Palmieri CM, Hill IT, Motamedi M. Regulation of histone methylation by noncoding RNAs. Biochim Biophys Acta. 2014;1839(12):1385-1394.

163. Wang F, Ma Y, Wang H, Qin H. Reciprocal regulation between microRNAs and epigenetic machinery in colorectal cancer. Oncol Lett. 2017;13(3):1048-1057.

164. Lujambio A, Ropero S, Ballestar E, Fraga MF, Cerrato C, Setien F, Casado S, Suarez-Gauthier A, Sanchez-Cespedes M, Git A. Genetic unmasking of an epigenetically silenced microRNA in human cancer cells. Cancer Res. 2007;67(4):1424-1429.

165. Lujambio A, Calin GA, Villanueva A, Ropero S, Sanchez-Cespedes M, Blanco D, Montuenga LM, Rossi S, Nicoloso MS, Faller WJ. A microRNA DNA methylation

signature for human cancer metastasis. Proc Natl Acad Sci USA. 2008;105(36):13556-13561.

166.   Zhou X, Ruan J, Wang G, Zhang W. Characterization and identification of microRNA core promoters in four model species. PLoS Comput Biol. 2007;3(3):e37.

167.   Pepling ME. From primordial germ cell to primordial follicle: mammalian female germ cell development. Genesis. 2006;44(12):622-632.

168.   Elkouby YM, Mullins MC. Coordination of cellular differentiation, polarity, mitosis and meiosis - new findings from early vertebrate oogenesis. Dev Biol. 2017;430(2):275-287.

169.   Picton HM. Activation of follicle development: the primordial follicle. Theriogenology. 2001;55(6):1193-1210.

170.   Aerts JM, Bols PE. Ovarian follicular dynamics: a review with emphasis on the bovine species. Part I: folliculogenesis and pre-antral follicle development. Reprod Domest Anim. 2010;45(1):171-179.

171.   McGinnis LK, Limback SD, Albertini DF. Signaling modalities during oogenesis in mammals. Curr Top Dev Biol. 2013;102:227-242.

172.   Diaz FJ, Wigglesworth K, Eppig JJ. Oocytes are required for the preantral granulosa cell to cumulus cell transition in mice. Dev Biol. 2007;305(1):300-311.

173.   Lucidi P, Bernabo N, Turriani M, Barboni B, Mattioli M. Cumulus cells steroidogenesis is influenced by the degree of oocyte maturation. Reprod Biol Endocrinol. 2003;1:45.

174.   Huang Z, Wells D. The human oocyte and cumulus cells relationship: new insights from the cumulus cell transcriptome. Mol Hum Reprod. 2010;16(10):715-725.

175.   Regassa A, Rings F, Hoelker M, Cinar U, Tholen E, Looft C, Schellander K, Tesfaye D. Transcriptome dynamics and molecular cross-talk between bovine oocyte and its companion cumulus cells. BMC Genomics. 2011;12:57.

176.   Sanchez F, Smitz J. Molecular control of oogenesis. Biochim Biophys Acta. 2012;1822(12):1896-1912.

177.   Dadarwal D, Adams GP, Hyttel P, Brogliatti GM, Caldwell S, Singh J. Organelle reorganization in bovine oocytes during dominant follicle growth and regression. Reprod Biol Endocrinol. 2015;13:124.

178.   Reader KL, Stanton JL, Juengel JL. The role of oocyte organelles in determining developmental competence. Biology (Basel). 2017;6(3).

179.   Aerts JM, Bols PE. Ovarian follicular dynamics. A review with emphasis on the bovine species. Part II: Antral development, exogenous influence and future prospects. Reprod Domest Anim. 2010;45(1):180-187.

180.   Verkuyl DA. Oral conception. Impregnation via the proximal gastrointestinal tract in a patient with an aplastic distal vagina. Case report. Br J Obstet Gynaecol. 1988;95(9):933-934.

181.   Rodrigues P, Limback D, McGinnis LK, Plancha CE, Albertini DF. Oogenesis: prospects and challenges for the future. J Cell Physiol. 2008;216(2):355-365.

182.   Canovas S, Ross PJ. Epigenetics in preimplantation mammalian development. Theriogenology. 2016;86(1):69-79.

183.   Fraser R, Lin CJ. Epigenetic reprogramming of the zygote in mice and men: on your marks, get set, go! Reproduction. 2016;152(6):R211-R222.

184.   Ko YG, Yun J, Park HJ, Tanaka S, Shiota K, Cho JH. Dynamic methylation pattern of the methyltransferase1o (Dnmt1o) 5'-flanking region during mouse oogenesis and spermatogenesis. Mol Reprod Dev. 2013;80(3):212-222.

185.   Cardoso MC, Leonhardt H. DNA methyltransferase is actively retained in the cytoplasm during early development. J Cell Biol. 1999;147(1):25-32.

186.   Gu TP, Guo F, Yang H, Wu HP, Xu GF, Liu W, Xie ZG, Shi L, He X, Jin SG. The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. Nature. 2011;477(7366):606-610.

187.   Nakamura T, Arai Y, Umehara H, Masuhara M, Kimura T, Taniguchi H, Sekimoto T, Ikawa M, Yoneda Y, Okabe M. PGC7/Stella protects against DNA demethylation in early embryogenesis. Nat Cell Biol. 2007;9(1):64-71.

188.   Kang J, Kalantry S, Rao A. PGC7, H3K9me2 and Tet3: regulators of DNA methylation in zygotes. Cell Res. 2013;23(1):6-9.

189.   Zhang T, Termanis A, Ozkan B, Bao XX, Culley J, de Lima Alves F, Rappsilber J, Ramsahoye B, Stancheva I. G9a/GLP complex maintains imprinted DNA methylation in embryonic stem cells. Cell Rep. 2016;15(1):77-85.

190.   Alexander KA, Wang X, Shibata M, Clark AG, Garcia-Garcia MJ. TRIM28 controls genomic imprinting through distinct mechanisms during and after early genome-wide reprogramming. Cell Rep. 2015;13(6):1194-1205.

191.   Rivera RM, Ross JW. Epigenetics in fertilization and preimplantation embryo development. Prog Biophys Mol Biol. 2013;113(3):423-432.

192.   Burton A, Torres-Padilla ME. Epigenetic reprogramming and development: a unique heterochromatin organization in the preimplantation mouse embryo. Brief Funct Genomics 2010;9(5-6):444-454.

193.   Mtango NR, Potireddy S, Latham KE. Oocyte quality and maternal control of development. Int Rev Cell Mol Biol. 2008;268:223-290.

194. Macaulay A, Scantland S, Robert C. RNA processing during early embryogenesis: managing storage, utilisation and destruction. In: RNA processing. 2011. DOI: 10.5772/20375

195. Kocabas AM, Crosby J, Ross PJ, Otu HH, Beyhan Z, Can H, Tam WL, Rosa GJ, Halgren RG, Lim B. The transcriptome of human oocytes. Proc Natl Acad Sci USA. 2006;103(38):14027-14032.

196. Tadros W, Lipshitz HD. The maternal-to-zygotic transition: a play in two acts. Development. 2009;136(18):3033-3042.

197. Yamaguchi K, Hada M, Fukuda Y, Inoue E, Makino Y, Katou Y, Shirahige K, Okada Y. Re-evaluating the localization of sperm-retained histones revealed the modification-dependent accumulation in specific genome regions. Cell Rep. 2018;23(13):3920-3932.

198. van der Heijden GW, Derijck AA, Ramos L, Giele M, van der Vlag J, de Boer P. Transmission of modified nucleosomes from the mouse male germline to the zygote and subsequent remodeling of paternal chromatin. Dev Biol. 2006;298(2):458-469.

199. Meyer RG, Ketchum CC, Meyer-Ficca ML. Heritable sperm chromatin epigenetics: a break to remember. Biol Reprod. 2017;97(6):784-797.

200. Guo F, Li X, Liang D, Li T, Zhu P, Guo H, Wu X, Wen L, Gu TP, Hu B. Active and passive demethylation of male and female pronuclear DNA in the mammalian zygote. Cell Stem Cell. 2014;15(4):447-459.

201. Schultz RM. The molecular foundations of the maternal to zygotic transition in the preimplantation embryo. Hum Reprod Update. 2002;8(4):323-331.

202. Bettegowda A, Lee KB, Smith GW. Cytoplasmic and nuclear determinants of the maternal-to-embryonic transition. Reprod Fertil Dev. 2008;20(1):45-53.

203. Vassalli JD, Stutz A. Translational control. Awakening dormant mRNAs. Curr Biol. 1995;5(5):476-479.

204. Barckmann B, Simonelig M. Control of maternal mRNA stability in germ cells and early embryos. Biochim Biophys Acta. 2013;1829(6-7):714-724.

205. Winata CL, Korzh V. The translational regulation of maternal mRNAs in time and space. FEBS Lett. 2018;592(17):3007-3023.

206. Vassalli JD, Huarte J, Belin D, Gubler P, Vassalli A, O'Connell ML, Parton LA, Rickles RJ, Strickland S. Regulated polyadenylation controls mRNA translation during meiotic maturation of mouse oocytes. Genes Dev. 1989;3(12B):2163-2171.

207. Nakanishi T, Kubota H, Ishibashi N, Kumagai S, Watanabe H, Yamashita M, Kashiwabara S, Miyado K, Baba T. Possible role of mouse poly(A) polymerase mGLD-2 during oocyte maturation. Dev Biol. 2006;289(1):115-126.

208. Cui J, Sartain CV, Pleiss JA, Wolfner MF. Cytoplasmic polyadenylation is a major mRNA regulator during oogenesis and egg activation in Drosophila. Dev Biol. 2013;383(1):121-131.

209. Mendez R, Richter JD. Translational control by CPEB: a means to the end. Nat Rev Mol Cell Biol. 2001;2(7):521-529.

210. Rouhana L, Wang L, Buter N, Kwak JE, Schiltz CA, Gonzalez T, Kelley AE, Landry CF, Wickens M. Vertebrate GLD2 poly(A) polymerases in the germline and the brain. RNA. 2005;11(7):1117-1130.

211. Benoit P, Papin C, Kwak JE, Wickens M, Simonelig M. PAP- and GLD-2-type poly(A) polymerases are required sequentially in cytoplasmic polyadenylation and oogenesis in Drosophila. Development. 2008;135(11):1969-1979.

212. Yartseva V, Giraldez AJ. The maternal-to-zygotic transition during vertebrate development: a model for reprogramming. Curr Top Dev Biol. 2015;113:191-232.

213. Thomsen S, Anders S, Janga SC, Huber W, Alonso CR. Genome-wide analysis of mRNA decay patterns during early Drosophila development. Genome Biol. 2010;11(9):R93.

214. Tadros W, Goldman AL, Babak T, Menzies F, Vardy L, Orr-Weaver T, Hughes TR, Westwood JT, Smibert CA, Lipshitz HD. SMAUG is a major regulator of maternal mRNA destabilization in Drosophila and its translation is activated by the PAN GU kinase. Dev Cell. 2007;12(1):143-155.

215. De Renzis S, Elemento O, Tavazoie S, Wieschaus EF. Unmasking activation of the zygotic genome using chromosomal deletions in the Drosophila embryo. PLoS Biol. 2007;5(5):e117.

216. Su YQ, Sugiura K, Woo Y, Wigglesworth K, Kamdar S, Affourtit J, Eppig JJ. Selective degradation of transcripts during meiotic maturation of mouse oocytes. Dev Biol. 2007;302(1):104-117.

217. Hamatani T, Carter MG, Sharov AA, Ko MS. Dynamics of global gene expression changes during mouse preimplantation development. Dev Cell. 2004;6(1):117-131.

218. Semotok JL, Cooperstock RL, Pinder BD, Vari HK, Lipshitz HD, Smibert CA. Smaug recruits the CCR4/POP2/NOT deadenylase complex to trigger maternal transcript localization in the early Drosophila embryo. Curr Biol. 2005;15(4):284-294.

219. Gotze M, Wahle E. Smaug destroys a huge treasure. Genome Biol. 2014;15(1):101.

220. Nelson MR, Leidal AM, Smibert CA. Drosophila Cup is an eIF4E-binding protein that functions in Smaug-mediated translational repression. EMBO J. 2004;23(1):150-159.

221. Gamberi C, Peterson DS, He L, Gottlieb E. An anterior function for the Drosophila posterior determinant Pumilio. Development. 2002;129(11):2699-2710.

222. Goldstrohm AC, Hook BA, Seay DJ, Wickens M. PUF proteins bind Pop2p to regulate messenger RNAs. Nat Struct Mol Biol. 2006;13(6):533-539.

223. Giraldez AJ, Mishima Y, Rihel J, Grocock RJ, Van Dongen S, Inoue K, Enright AJ, Schier AF. Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. Science. 2006;312(5770):75-79.

224. Braun JE, Huntzinger E, Izaurralde E. A molecular link between miRISCs and deadenylases provides new insight into the mechanism of gene silencing by microRNAs. Cold Spring Harb Perspect Biol. 2012;4(12): a012328.

225. Wu E, Thivierge C, Flamand M, Mathonnet G, Vashisht AA, Wohlschlegel J, Fabian MR, Sonenberg N, Duchaine TF. Pervasive and cooperative deadenylation of 3'UTRs by embryonic microRNA families. Mol Cell. 2010;40(4):558-570.

226. Lund E, Liu M, Hartley RS, Sheets MD, Dahlberg JE. Deadenylation of maternal mRNAs mediated by miR-427 in Xenopus laevis embryos. RNA. 2009;15(12):2351-2363.

227. Bushati N, Stark A, Brennecke J, Cohen SM. Temporal reciprocity of miRNAs and their targets during the maternal-to-zygotic transition in Drosophila. Curr Biol. 2008;18(7):501-506.

228. Tang F, Kaneda M, O'Carroll D, Hajkova P, Barton SC, Sun YA, Lee C, Tarakhovsky A, Lao K, Surani MA. Maternal microRNAs are essential for mouse zygotic development. Genes Dev. 2007;21(6):644-648.

229. Zeng F, Schultz RM. RNA transcript profiling during zygotic gene activation in the preimplantation mouse embryo. Dev Biol. 2005;283(1):40-57.

230. Medeiros LA, Dennis LM, Gill ME, Houbaviy H, Markoulaki S, Fu D, White AC, Kirak O, Sharp PA, Page DC. Mir-290-295 deficiency in mice results in partially penetrant embryonic lethality and germ cell defects. Proc Natl Acad Sci USA. 2011;108(34):14163-14168.

231. Much C, Auchynnikava T, Pavlinic D, Buness A, Rappsilber J, Benes V, Allshire R, O'Carroll D. Endogenous mouse dicer is an exclusively cytoplasmic protein. PLoS Genet. 2016;12(6):e1006095.

232. FL M. Maternal–zygotic transition. In: Maternal Control of Development in Vertebrates: My Mother Made Me Do It! San Rafael, CA: Morgan & Claypool Life Sciences; 2010.

233. Suh N, Baehner L, Moltzahn F, Melton C, Shenoy A, Chen J, Blelloch R. MicroRNA function is globally suppressed in mouse oocytes and early embryos. Curr Biol. 2010;20(3):271-277.

234. Svoboda P. Why mouse oocytes and early embryos ignore miRNAs? RNA Biol. 2010;7(5):559-563.

235. Zaessinger S, Busseau I, Simonelig M. Oskar allows nanos mRNA translation in Drosophila embryos by preventing its deadenylation by Smaug/CCR4. Development. 2006;133(22):4573-4583.

236. Kedde M, Strasser MJ, Boldajipour B, Oude Vrielink JA, Slanchev K, le Sage C, Nagel R, Voorhoeve PM, van Duijse J, Orom UA. RNA-binding protein Dnd1 inhibits microRNA access to target mRNA. Cell. 2007;131(7):1273-1286.

237. Takeda Y, Mishima Y, Fujiwara T, Sakamoto H, Inoue K. DAZL relieves miRNA-mediated repression of germline mRNAs by controlling poly(A) tail length in zebrafish. PLoS One. 2009;4(10):e7513.

238. Alonso CR. A complex 'mRNA degradation code' controls gene expression during animal development. Trends Genet. 2012;28(2):78-88.

239. Li L, Baibakov B, Dean J. A subcortical maternal complex essential for preimplantation mouse embryogenesis. Dev Cell. 2008;15(3):416-425.

240. Li L, Zheng P, Dean J. Maternal control of early mouse development. Development. 2010;137(6):859-870.

241. Lee MT, Bonneau AR, Giraldez AJ. Zygotic genome activation during the maternal-to-zygotic transition. Annu Rev Cell Dev Biol. 2014;30:581-613.

242. Darbo E, Herrmann C, Lecuit T, Thieffry D, van Helden J. Transcriptional and epigenetic signatures of zygotic genome activation during early Drosophila embryogenesis. BMC Genomics. 2013;14:226.

243. Jukam D, Shariati SAM, Skotheim JM. Zygotic genome activation in vertebrates. Dev Cell. 2017;42(4):316-332.

244. Langley AR, Smith JC, Stemple DL, Harvey SA. New insights into the maternal to zygotic transition. Development. 2014;141(20):3834-3841.

245. Schulz KN, Harrison MM. Mechanisms regulating zygotic genome activation. Nat Rev Genet. 2019;20(4):221-234.

246. Li L, Lu X, Dean J. The maternal to zygotic transition in mammals. Mol Aspects Med. 2013;34(5):919-938.

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

248. Kwasnieski JC, Orr-Weaver TL, Bartel DP. Early genome activation in Drosophila is extensive with an initial tendency for aborted transcripts and retained introns. Genome Res. 2019;29(7):1188-1197.

249. Harvey SA, Sealy I, Kettleborough R, Fenyes F, White R, Stemple D, Smith JC. Identification of the zebrafish maternal and paternal transcriptomes. Development. 2013;140(13):2703-2710.

250. Kim YS, Kim HR, Kim H, Yang SC, Park M, Yoon JA, Lim HJ, Hong SH, DeMayo FJ, Lydon JP. Deficiency in DGCR8-dependent canonical microRNAs causes infertility due to multiple abnormalities during uterine development in mice. Sci Rep. 2016;6:20242.

251. Suh N, Blelloch R. Small RNAs in early mammalian development: from gametes to gastrulation. Development. 2011;138(9):1653-1661.

252. Claycomb JM, Batista PJ, Pang KM, Gu W, Vasale JJ, van Wolfswinkel JC, Chaves DA, Shirayama M, Mitani S, Ketting RF. The Argonaute CSR-1 and its 22G-RNA cofactors are required for holocentric chromosome segregation. Cell. 2009;139(1):123-134.

253. Seitz H. siRNAs: the hidden face of the small RNA world. Curr Biol. 2010;20(3):R108-110.

254. Sandmaier SE, Telugu BP. MicroRNA-mediated reprogramming of somatic cells into induced pluripotent stem cells. Methods Mol Biol. 2015;1330:29-36.

255. Ohnishi Y, Totoki Y, Toyoda A, Watanabe T, Yamamoto Y, Tokunaga K, Sakaki Y, Sasaki H, Hohjoh H. Small RNA class transition from siRNA/piRNA to miRNA during pre-implantation mouse development. Nucleic Acids Res. 2010;38(15):5141-5151.

256. Libri V, Miesen P, van Rij RP, Buck AH. Regulation of microRNA biogenesis and turnover by animals and their viruses. Cell Mol Life Sci. 2013;70(19):3525-3544.

257. Fazi F, Fatica A. Interplay between N (6)-methyladenosine (m(6)A) and non-coding RNAs in cell development and cancer. Front Cell Dev Biol. 2019;7:116.

258. Plass M, Rasmussen SH, Krogh A. Highly accessible AU-rich regions in 3' untranslated regions are hotspots for binding of regulatory factors. PLoS Comput Biol. 2017;13(4):e1005460.

259. Yang CX, Wright EC, Ross JW. Expression of RNA-binding proteins DND1 and FXR1 in the porcine ovary, and during oocyte maturation and early embryo development. Mol Reprod Dev. 2012;79(8):541-552.

260. Bhattacharya C, Aggarwal S, Zhu R, Kumar M, Zhao M, Meistrich ML, Matin A. The mouse dead-end gene isoform alpha is necessary for germ cell and embryonic viability. Biochem Biophys Res Commun. 2007;355(1):194-199.

261. Sysoev VO, Fischer B, Frese CK, Gupta I, Krijgsveld J, Hentze MW, Castello A, Ephrussi A. Global changes of the RNA-bound proteome during the maternal-to-zygotic transition in Drosophila. Nat Commun. 2016;7:12128.

262. Krawczynski K, Najmula J, Bauersachs S, Kaczmarek MM. MicroRNAome of porcine conceptuses and trophoblasts: expression profile of micrornas and their potential to regulate genes crucial for establishment of pregnancy. Biol Reprod. 2015;92(1):21.

263. Tripurani SK, Lee KB, Wee G, Smith GW, Yao J. MicroRNA-196a regulates bovine newborn ovary homeobox gene (NOBOX) expression during early embryogenesis. BMC Dev Biol. 2011;11:25.

264. Tripurani SK, Wee G, Lee KB, Smith GW, Wang L, Jianboyao. MicroRNA-212 post-transcriptionally regulates oocyte-specific basic-helix-loop-helix transcription factor, factor in the germline alpha (FIGLA), during bovine early embryogenesis. PLoS One. 2013;8(9):e76114.

265. Berg DK, Pfeffer PL. MicroRNA expression in bovine preimplantation embryos. Reprod Fertil Dev. 2018;30(3):546-554.

266. Sinha PB, Tesfaye D, Rings F, Hossien M, Hoelker M, Held E, Neuhoff C, Tholen E, Schellander K, Salilew-Wondim D. MicroRNA-130b is involved in bovine granulosa and cumulus cells function, oocyte maturation and blastocyst formation. J Ovarian Res. 2017;10(1):37.

267. Goossens K, Mestdagh P, Lefever S, Van Poucke M, Van Zeveren A, Van Soom A, Vandesompele J, Peelman L. Regulatory microRNA network identification in bovine blastocyst development. Stem Cells Dev. 2013;22(13):1907-1920.

268. Stowe HM, Curry E, Calcatera SM, Krisher RL, Paczkowski M, Pratt SL. Cloning and expression of porcine Dicer and the impact of developmental stage and culture conditions on microRNA expression in porcine embryos. Gene. 2012;501(2):198-205.

269. Mondou E, Dufort I, Gohin M, Fournier E, Sirard MA. Analysis of microRNAs and their precursors in bovine early embryonic development. Mol Hum Reprod. 2012;18(9):425-434.

270. Kropp J, Khatib H. Characterization of microRNA in bovine in vitro culture media associated with embryo quality and development. J Dairy Sci. 2015;98(9):6552-6563.

271. Lin X, Beckers E, Mc Cafferty S, Gansemans Y, Joanna Szymanska K, Chaitanya Pavani K, Catani JP, Van Nieuwerburgh F, Deforce D, De Sutter P. Bovine embryo-secreted microRNA-30c is a potential non-invasive biomarker for hampered preimplantation developmental competence. Front Genet. 2019;10:315.

272. Kropp J, Salih SM, Khatib H. Expression of microRNAs in bovine and human pre-implantation embryo culture media. Front Genet. 2014;5:91.

273. B.R. McCalliea ASJC, Parksa S. Mitchella, W. Schoolcrafta, M.G.Katz-Jaffeab. Unique cumulus cell microRNA expression profiles predict IVF outcome and live birth. Fertil and Steril. 2012;98(3):S18.

274. Aravin AA, Sachidanandam R, Bourc'his D, Schaefer C, Pezic D, Toth KF, Bestor T, Hannon GJ. A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. Mol Cell. 2008;31(6):785-799.

275. Kuramochi-Miyagawa S, Watanabe T, Gotoh K, Totoki Y, Toyoda A, Ikawa M, Asada N, Kojima K, Yamaguchi Y, Ijiri TW. DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. Genes Dev. 2008;22(7):908-917.

276. Ketting RF. The many faces of RNAi. Dev Cell. 2011;20(2):148-161.

277. Carmell MA, Girard A, van de Kant HJ, Bourc'his D, Bestor TH, de Rooij DG, Hannon GJ. MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. Dev Cell. 2007;12(4):503-514.

278. Kuramochi-Miyagawa S, Kimura T, Ijiri TW, Isobe T, Asada N, Fujita Y, Ikawa M, Iwai N, Okabe M, Deng W. Mili, a mammalian member of piwi family gene, is essential for spermatogenesis. Development. 2004;131(4):839-849.

279. Tam OH, Aravin AA, Stein P, Girard A, Murchison EP, Cheloufi S, Hodges E, Anger M, Sachidanandam R, Schultz RM. Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. Nature. 2008;453(7194):534-538.

280. Watanabe T, Totoki Y, Toyoda A, Kaneda M, Kuramochi-Miyagawa S, Obata Y, Chiba H, Kohara Y, Kono T, Nakano T. Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. Nature. 2008;453(7194):539-543.

281. Russell S, Patel M, Gilchrist G, Stalker L, Gillis D, Rosenkranz D, LaMarre J. Bovine piRNA-like RNAs are associated with both transposable elements and mRNAs. Reproduction. 2017;153(3):305-318.

282. Yuan S, Schuster A, Tang C, Yu T, Ortogero N, Bao J, Zheng H, Yan W. Sperm-borne miRNAs and endo-siRNAs are important for fertilization and preimplantation embryonic development. Development. 2016;143(4):635-647.

283. Liu WM, Pang RT, Chiu PC, Wong BP, Lao K, Lee KF, Yeung WS. Sperm-borne microRNA-34c is required for the first cleavage division in mouse. Proc Natl Acad Sci USA. 2012;109(2):490-494.

284. Dickson DA, Paulus JK, Mensah V, Lem J, Saavedra-Rodriguez L, Gentry A, Pagidas K, Feig LA. Reduced levels of miRNAs 449 and 34 in sperm of mice and men exposed to early life stress. Transl Psychiatry. 2018;8(1):101.

285. Rodgers AB, Morgan CP, Leu NA, Bale TL. Transgenerational epigenetic programming via sperm microRNA recapitulates effects of paternal stress. Proc Natl Acad Sci USA. 2015;112(44):13699-13704.

286. Gapp K, Miska EA. tRNA fragments: novel players in intergenerational inheritance. Cell Res. 2016;26(4):395-396.

287. Lu Z, Matera AG. Developmental analysis of spliceosomal snRNA isoform expression. G3 (Bethesda). 2014;5(1):103-110.

288. O'Reilly D, Dienstbier M, Cowley SA, Vazquez P, Drozdz M, Taylor S, James WS, Murphy S. Differentially expressed, variant U1 snRNAs regulate gene expression in human cells. Genome Res. 2013;23(2):281-291.

289. Higa-Nakamine S, Suzuki T, Uechi T, Chakraborty A, Nakajima Y, Nakamura M, Hirano N, Suzuki T, Kenmochi N. Loss of ribosomal RNA modification causes developmental defects in zebrafish. Nucleic Acids Res. 2012;40(1):391-398.

290. Quan G, Li J. Circular RNAs: biogenesis, expression and their potential roles in reproduction. J Ovarian Res. 2018;11(1):9.

291. Liu KS, Pan F, Mao XD, Liu C, Chen YJ. Biological functions of circular RNAs and their roles in occurrence of reproduction and gynecological diseases. Am J Transl Res. 2019;11(1):1-15.

292. Su X, Gao G, Wang S, Su G, Zheng Z, Zhang J, Han L, Ling Y, Wang X, Li G. CircRNA expression profile of bovine placentas in late gestation with aberrant SCNT fetus. J Clin Lab Anal. 2019;33(6):e22918.

293. Faber DC, Ferre LB, Metzger J, Robl JM, Kasinathan P. Agro-economic impact of cattle cloning. Cloning Stem Cells. 2004;6(2):198-207.

294. Polejaeva IA, Broek DM, Walker SC, Zhou W, Walton M, Benninghoff AD, Faber DC. Longitudinal study of reproductive performance of female cattle produced by somatic cell nuclear transfer. PLoS One. 2013;8(12):e84283.

295. Vilarino M, Rashid ST, Suchy FP, McNabb BR, van der Meulen T, Fine EJ, Ahsan S, Mursaliyev N, Sebastiano V, Diab SS. CRISPR/Cas9 microinjection in oocytes disables pancreas development in sheep. Sci Rep. 2017;7(1):17472.

296. Campbell KH. A background to nuclear transfer and its applications in agriculture and human therapeutic medicine. J Anat. 2002;200(Pt 3):267-275.

297. Matoba S, Zhang Y. Somatic cell nuclear transfer reprogramming: mechanisms and applications. Cell Stem Cell. 2018;23(4):471-485.

298. Buganim Y, Faddah DA, Jaenisch R. Mechanisms and models of somatic cell reprogramming. Nat Rev Genet. 2013;14(6):427-439.

299. Wasson JA, Ruppersburg CC, Katz DJ. Restoring totipotency through epigenetic reprogramming. Brief Funct Genomics. 2013;12(2):118-128.

300. Dias FCF, Khan MIR, Sirard MA, Adams GP, Singh J. Transcriptome analysis of granulosa cells after conventional vs long FSH-induced superstimulation in cattle. BMC Genomics. 2018;19(1):258.

301. Gardner DK, Lane M. Ex vivo early embryo development and effects on gene expression and imprinting. Reprod Fertil Dev. 2005;17(3):361-370.

302. Coticchio G, Sereni E, Serrao L, Mazzone S, Iadarola I, Borini A. What criteria for the definition of oocyte quality? Ann N Y Acad Sci. 2004;1034:132-144.

303. Khoudja RY, Xu Y, Li T, Zhou C. Better IVF outcomes following improvements in laboratory air quality. J Assist Reprod Genet. 2013;30(1):69-76.

304. Swain JE, Carrell D, Cobo A, Meseguer M, Rubio C, Smith GD. Optimizing the culture environment and embryo manipulation to help maintain embryo developmental potential. Fertil Steril. 2016;105(3):571-587.

305. Kasterstein E, Strassburger D, Komarovsky D, Bern O, Komsky A, Raziel A, Friedler S, Ron-El R. The effect of two distinct levels of oxygen concentration on embryo development in a sibling oocyte study. J Assist Reprod Genet. 2013;30(8):1073-1079.

306. Gardner DK. Dissection of culture media for embryos: the most important and less important components and characteristics. Reprod Fertil Dev. 2008;20(1):9-18.

307. Gad A, Schellander K, Hoelker M, Tesfaye D. Transcriptome profile of early mammalian embryos in response to culture environment. Anim Reprod Sci. 2012;134(1-2):76-83.

308. Fauque P, Mondon F, Letourneur F, Ripoche MA, Journot L, Barbaux S, Dandolo L, Patrat C, Wolf JP, Jouannet P. In vitro fertilization and embryo culture strongly impact the placental transcriptome in the mouse model. PLoS One. 2010;5(2):e9218.

309. Geber S, Sales L, Sampaio MA. Laboratory techniques for human embryos. Reprod Biomed Online. 2002;5(2):211-218.

310. Arias ME, Ross PJ, Felmer RN. Culture medium composition affects the gene expression pattern and in vitro development potential of bovine somatic cell nuclear transfer (SCNT) embryos. Biol Res. 2013;46(4):452-462.

311. Moor RM, Kruip, Th.A.M., Green, D. Intraovarian control of folliculogenesis: limits to superovulation? Theriogenology. 1984;21(1):103-116.

312. Younis AI, Brackett BG, Fayrer-Hosken RA. Influence of serum and hormones on bovine oocyte maturation and fertilization in vitro. Gamete Res. 1989;23(2):189-201.

313. Xu Z, Abbott A, Kopf GS, Schultz RM, Ducibella T. Spontaneous activation of ovulated mouse eggs: time-dependent effects on M-phase exit, cortical granule exocytosis, maternal messenger ribonucleic acid recruitment, and inositol 1,4,5-trisphosphate sensitivity. Biol Reprod. 1997;57(4):743-750.

314. Sirard MA. Follicle environment and quality of in vitro matured oocytes. J Assist Reprod Genet. 2011;28(6):483-488.

315. Farin CE, Rodriguez KF, Alexander JE, Hockney JE, Herrick JR, Kennedy-Stoskopf S. The role of transcription in EGF- and FSH-mediated oocyte maturation in vitro. Anim Reprod Sci. 2007;98(1-2):97-112.

316. Camaioni A, Hascall VC, Yanagishita M, Salustri A. Effects of exogenous hyaluronic acid and serum on matrix organization and stability in the mouse cumulus cell-oocyte complex. J Biol Chem. 1993;268(27):20473-20481.

317. Zhang L, Jiang S, Wozniak PJ, Yang X, Godke RA. Cumulus cell function during bovine oocyte maturation, fertilization, and embryo development in vitro. Mol Reprod Dev. 1995;40(3):338-344.

318. Sugimura S, Kawahara M, Wakai T, Yamanaka K, Sasada H, Sato E. Effect of cytochalasins B and D on the developmental competence of somatic cell nuclear transfer embryos in miniature pigs. Zygote. 2008;16(2):153-159.

319. Iwamoto D, Yamagata K, Kishi M, Hayashi-Takanaka Y, Kimura H, Wakayama T, Saeki K. Early development of cloned bovine embryos produced from oocytes enucleated by fluorescence metaphase II imaging using a conventional halogen-lamp microscope. Cell Reprogram. 2015;17(2):106-114.

320. Tian XC, Kubota C, Enright B, Yang X. Cloning animals by somatic cell nuclear transfer-biological factors. Reprod Biol Endocrinol. 2003;1:98.

321. Wilmut I, Schnieke AE, McWhir J, Kind AJ, Campbell KH. Viable offspring derived from fetal and adult mammalian cells. Nature. 1997;385(6619):810-813.

322. Baguisi A, Behboodi E, Melican DT, Pollock JS, Destrempes MM, Cammuso C, Williams JL, Nims SD, Porter CA, Midura P. Production of goats by somatic cell nuclear transfer. Nat Biotechnol. 1999;17(5):456-461.

323. Woods GL, White KL, Vanderwall DK, Li GP, Aston KI, Bunch TD, Meerdo LN, Pate BJ. A mule cloned from fetal cells by nuclear transfer. Science. 2003;301(5636):1063.

324. Cibelli JB, Stice SL, Golueke PJ, Kane JJ, Jerry J, Blackwell C, Ponce de Leon FA, Robl JM. Cloned transgenic calves produced from nonquiescent fetal fibroblasts. Science. 1998;280(5367):1256-1258.

325. Wakayama T, Perry AC, Zuccotti M, Johnson KR, Yanagimachi R. Full-term development of mice from enucleated oocytes injected with cumulus cell nuclei. Nature. 1998;394(6691):369-374.

326. Lee KY, Huang H, Ju B, Yang Z, Lin S. Cloned zebrafish by nuclear transfer from long-term-cultured cells. Nat Biotechnol. 2002;20(8):795-799.

327. Hwang WS, Ryu YJ, Park JH, Park ES, Lee EG, Koo JM, Jeon HY, Lee BC, Kang SK, Kim SJ. Evidence of a pluripotent human embryonic stem cell line derived from a cloned blastocyst. Science. 2004;303(5664):1669-1674.

328. Blelloch R, Wang Z, Meissner A, Pollard S, Smith A, Jaenisch R. Reprogramming efficiency following somatic cell nuclear transfer is influenced by the differentiation and methylation state of the donor nucleus. Stem Cells. 2006;24(9):2007-2013.

329. Sung LY, Gao S, Shen H, Yu H, Song Y, Smith SL, Chang CC, Inoue K, Kuo L, Lian J. Differentiated cells are more efficient than adult stem cells for cloning by somatic cell nuclear transfer. Nat Genet. 2006;38(11):1323-1328.

330. Inoue K, Kohda T, Lee J, Ogonuki N, Mochida K, Noguchi Y, Tanemura K, Kaneko-Ishino T, Ishino F, Ogura A. Faithful expression of imprinted genes in cloned mice. Science. 2002;295(5553):297.

331. Dutta R, Malakar D, Khate K, Sahu S, Akshey Y, Mukesh M. A comparative study on efficiency of adult fibroblast, putative embryonic stem cell and lymphocyte as donor cells for production of handmade cloned embryos in goat and characterization of putative ntES cells obtained from these embryos. Theriogenology. 2011;76(5):851-863.

332. Kishigami S, Wakayama S, Thuan NV, Ohta H, Mizutani E, Hikichi T, Bui HT, Balbach S, Ogura A, Boiani M. Production of cloned mice by somatic cell nuclear transfer. Nat Protoc. 2006;1(1):125-138.

333. Kim HS, Lee JY, Jeong EJ, Yang CJ, Hyun SH, Shin T, Hwang WS. Effects of repetitive ionomycin treatment on in vitro development of bovine somatic cell nuclear transfer embryos. J Reprod Dev. 2012;58(1):132-139.

334. Bhak JS, Lee SL, Ock SA, Mohana Kumar B, Choe SY, Rho GJ. Developmental rate and ploidy of embryos produced by nuclear transfer with different activation treatments in cattle. Anim Reprod Sci. 2006;92(1-2):37-49.

335. Beebe LF, McIlfatrick SJ, Nottle MB. Cytochalasin B and trichostatin a treatment postactivation improves in vitro development of porcine somatic cell nuclear transfer embryos. Cloning Stem Cells. 2009;11(4):477-482.

336. Fukui Y, Glew AM, Gandolfi F, Moor RM. In vitro culture of sheep oocytes matured and fertilized in vitro. Theriogenology. 1988;29(4):883-891.

337. Prichard JF, Thibodeaux JK, Pool SH, Blakewood EG, Menezo Y, Godke RA. In-vitro co-culture of early stage caprine embryos with oviduct and uterine epithelial cells. Hum Reprod. 1992;7(4):553-557.

338. Petters RM. Embryo development in vitro to the blastocyst stage in cattle, pigs and sheep. Anim Reprod Sci. 1992;28(1-4):415-421.

339. Wang S, Liu Y, Holyoak GR, Bunch TD. The effects of bovine serum albumin and fetal bovine serum on the development of pre- and postcleavage-stage bovine embryos cultured in modified CR2 and M199 media. Anim Reprod Sci. 1997;48(1):37-45.

340. Cagnone G, Sirard MA. The embryonic stress response to in vitro culture: insight from genomic analysis. Reproduction. 2016;152(6):R247-R261.

341. Campbell KH, Loi P, Otaegui PJ, Wilmut I. Cell cycle co-ordination in embryo cloning by nuclear transfer. Rev Reprod. 1996;1(1):40-46.

342. Ma Y, Kanakousaki K, Buttitta L. How the cell cycle impacts chromatin architecture and influences cell fate. Front Genet. 2015;6:19.

343. Kubota C, Yamakuchi H, Todoroki J, Mizoshita K, Tabara N, Barber M, Yang X. Six cloned calves produced from adult fibroblast cells after long-term culture. Proc Natl Acad Sci USA. 2000;97(3):990-995.

344. Wilmut I, Beaujean N, de Sousa PA, Dinnyes A, King TJ, Paterson LA, Wells DN, Young LE. Somatic cell nuclear transfer. Nature. 2002;419(6907):583-586.

345. Solter D. Mammalian cloning: advances and limitations. Nat Rev Genet. 2000;1(3):199-207.

346. Chavatte-Palmer P, Camous S, Jammes H, Le Cleac'h N, Guillomot M, Lee RS. Review: Placental perturbations induce the developmental abnormalities often observed in bovine somatic cell nuclear transfer. Placenta. 2012;33 Suppl:S99-S104.

347. Santos F, Zakhartchenko V, Stojkovic M, Peters A, Jenuwein T, Wolf E, Reik W, Dean W. Epigenetic marking correlates with developmental potential in cloned bovine preimplantation embryos. Curr Biol. 2003;13(13):1116-1121.

348. J.M Wilson JDW, K.R Bondioli, C.R Looney, M.E Westhusin, D.F McCalla. Comparison of birth weight and growth characteristics of bovine calves produced by nuclear transfer (cloning), embryo transfer and natural mating. Anim Reprod Sci. 1995;38(1-2):73-83.

349. Farin PW, Piedrahita JA, Farin CE. Errors in development of fetuses and placentas from in vitro-produced bovine embryos. Theriogenology. 2006;65(1):178-191.

350. Chen Z, Hagen DE, Elsik CG, Ji T, Morris CJ, Moon LE, Rivera RM. Characterization of global loss of imprinting in fetal overgrowth syndrome induced by assisted reproduction. Proc Natl Acad Sci USA. 2015;112(15):4618-4623.

351. Su J,Wang Y, Liu Q, Yang B, Wu Y, Luo Y, Hu G, Zhang Y. Aberrant mRNA expression and DNA methylation levels of imprinted genes in cloned transgenic calves that died of large offspring syndrome. Livest Sci. 2011;141(1):24-35.

352. Salilew-Wondim D, Tesfaye D, Hossain M, Held E, Rings F, Tholen E, Looft C, Cinar U, Schellander K, Hoelker M. Aberrant placenta gene expression pattern in bovine pregnancies established after transfer of cloned or in vitro produced embryos. Physiol Genomics. 2013;45(1):28-46.

353. Biase FH, Rabel C, Guillomot M, Hue I, Andropolis K, Olmstead CA, Oliveira R, Wallace R, Le Bourhis D, Richard C. Massive dysregulation of genes involved in cell signaling and placental development in cloned cattle conceptus and maternal endometrium. Proc Natl Acad Sci USA. 2016;113(51):14492-14501.

354. Arnold DR, Fortier AL, Lefebvre R, Miglino MA, Pfarrer C, Smith LC. Placental insufficiencies in cloned animals - a workshop report. Placenta. 2008;29 Suppl A:S108-110.

355. Chavatte-Palmer P, de Sousa N, Laigre P, Camous S, Ponter AA, Beckers JF, Heyman Y. Ultrasound fetal measurements and pregnancy associated glycoprotein secretion in early pregnancy in cattle recipients carrying somatic clones. Theriogenology. 2006;66(4):829-840.

356. Lee RS, Peterson AJ, Donnison MJ, Ravelich S, Ledgard AM, Li N, Oliver JE, Miller AL, Tucker FC, Breier B. Cloned cattle fetuses with the same nuclear genetics are more variable than contemporary half-siblings resulting from artificial insemination and exhibit fetal and placental growth deregulation even in the first trimester. Biol Reprod. 2004;70(1):1-11.

357. Hill JR, Roussel AJ, Cibelli JB, Edwards JF, Hooper NL, Miller MW, Thompson JA, Looney CR, Westhusin ME, Robl JM. Clinical and pathologic features of cloned transgenic calves and fetuses (13 case studies). Theriogenology. 1999;51(8):1451-1465.

358. Palini S, De Stefani S, Scala V, Dusi L, Bulletti C. Epigenetic regulatory mechanisms during preimplantation embryo development. Ann N Y Acad Sci. 2011;1221:54-60.

359. Cantone I, Fisher AG. Epigenetic programming and reprogramming during development. Nat Struct Mol Biol. 2013;20(3):282-289.

360. Shi L, Wu J. Epigenetic regulation in mammalian preimplantation embryo development. Reprod Biol Endocrinol. 2009;7:59.

361. Kiefer JC. Epigenetics in development. Developmental dynamics : an official publication of the American Association of Anatomists. 2007;236(4):1144-1156.

362. Obata Y, Kaneko-Ishino T, Koide T, Takai Y, Ueda T, Domeki I, Shiroishi T, Ishino F, Kono T. Disruption of primary imprinting during oocyte growth leads to the modified expression of imprinted genes during embryogenesis. Development. 1998;125(8):1553-1560.

363. Weiss A, Keshet I, Razin A, Cedar H. DNA demethylation in vitro: involvement of RNA. Cell. 1996;86(5):709-718.

364. Yamaguchi S, Hong K, Liu R, Shen L, Inoue A, Diep D, Zhang K, Zhang Y. Tet1 controls meiosis by regulating meiotic gene expression. Nature. 2012;492(7429):443-447.

365. Reik W, Dean W, Walter J. Epigenetic reprogramming in mammalian development. Science. 2001;293(5532):1089-1093.

366. Cezar GG, Bartolomei MS, Forsberg EJ, First NL, Bishop MD, Eilertsen KJ. Genome-wide epigenetic alterations in cloned bovine fetuses. Biol Reprod. 2003;68(3):1009-1014.

367. Enright BP, Kubota C, Yang X, Tian XC. Epigenetic characteristics and development of embryos cloned from donor cells treated by trichostatin A or 5-aza-2'-deoxycytidine. Biol Reprod. 2003;69(3):896-901.

368. Han YM, Kang YK, Koo DB, Lee KK. Nuclear reprogramming of cloned embryos produced in vitro. Theriogenology. 2003;59(1):33-44.

369. Kang YK, Park JS, Koo DB, Choi YH, Kim SU, Lee KK, Han YM. Limited demethylation leaves mosaic-type methylation states in cloned bovine pre-implantation embryos. EMBO J. 2002;21(5):1092-1100.

370. Mann MR, Chung YG, Nolen LD, Verona RI, Latham KE, Bartolomei MS. Disruption of imprinted gene methylation and expression in cloned preimplantation stage mouse embryos. Biol Reprod. 2003;69(3):902-914.

371. Shi W, Haaf T. Aberrant methylation patterns at the two-cell stage as an indicator of early developmental failure. Mol Reprod Dev. 2002;63(3):5.

372. Simonsson S, Gurdon J. DNA demethylation is necessary for the epigenetic reprogramming of somatic cell nuclei. Nat Cell Biol. 2004;6(10):984-990.

373. Pesce M, Scholer HR. Oct-4: gatekeeper in the beginnings of mammalian development. Stem Cells. 2001;19(4):271-278.

374. Niwa H, Miyazaki J, Smith AG. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nat Genet. 2000;24(4):372-376.

375. Nichols J, Zevnik B, Anastassiadis K, Niwa H, Klewe-Nebenius D, Chambers I, Scholer H, Smith A. Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. Cell. 1998;95(3):379-391.

376. Yamazaki Y, Fujita TC, Low EW, Alarcon VB, Yanagimachi R, Marikawa Y. Gradual DNA demethylation of the Oct4 promoter in cloned mouse embryos. Mol Reprod Dev. 2006;73(2):180-188.

377. Bortvin A, Eggan K, Skaletsky H, Akutsu H, Berry DL, Yanagimachi R, Page DC, Jaenisch R. Incomplete reactivation of Oct4-related genes in mouse embryos cloned from somatic nuclei. Development. 2003;130(8):1673-1680.

378. Boiani M, Eckardt S, Scholer HR, McLaughlin KJ. Oct4 distribution and level in mouse clones: consequences for pluripotency. Genes Dev. 2002;16(10):1209-1219.

379. Bourc'his D, Le Bourhis D, Patin D, Niveleau A, Comizzoli P, Renard JP, Viegas-Pequignot E. Delayed and incomplete reprogramming of chromosome methylation patterns in bovine cloned embryos. Curr Biol. 2001;11(19):1542-1546.

380. Zhang S, Chen X, Wang F, An X, Tang B, Zhang X, Sun L, Li Z. Aberrant DNA methylation reprogramming in bovine SCNT preimplantation embryos. Sci Rep. 2016;6:30345.

381. Suzuki J, Jr., Therrien J, Filion F, Lefebvre R, Goff AK, Smith LC. In vitro culture and somatic cell nuclear transfer affect imprinting of SNRPN gene in pre- and post-implantation stages of development in cattle. BMC Dev Biol. 2009;9:9.

382. Chang G, Liu S, Wang F, Zhang Y, Kou Z, Chen D, Gao S. Differential methylation status of imprinted genes in nuclear transfer derived ES (NT-ES) cells. Genomics. 2009;93(2):112-119.

383. Hirose M, Hada M, Kamimura S, Matoba S, Honda A, Motomura K, Ogonuki N, Shawki HH, Inoue K, Takahashi S. Aberrant imprinting in mouse trophoblast stem cells established from somatic cell nuclear transfer-derived embryos. Epigenetics. 2018;13(7):693-703.

384. Shen CJ, Lin CC, Shen PC, Cheng WT, Chen HL, Chang TC, Liu SS, Chen CM. Imprinted genes and satellite loci are differentially methylated in bovine somatic cell nuclear transfer clones. Cell Reprogram. 2013;15(5):413-424.

385. Dean W, Santos F, Stojkovic M, Zakhartchenko V, Walter J, Wolf E, Reik W. Conservation of methylation reprogramming in mammalian development: aberrant reprogramming in cloned embryos. Proc Natl Acad Sci USA. 2001;98(24):13734-13738.

386. Liu X, Wang Y, Gao Y, Su J, Zhang J, Xing X, Zhou C, Yao K, An Q, Zhang Y. H3K9 demethylase KDM4E is an epigenetic regulator for bovine embryonic development and a defective factor for nuclear reprogramming. Development. 2018;145(4).

387. Hormanseder E, Simeone A, Allen GE, Bradshaw CR, Figlmuller M, Gurdon J, Jullien J. H3K4 methylation-dependent memory of somatic cell identity inhibits reprogramming and development of nuclear transfer embryos. Cell Stem Cell. 2017;21(1):135-143 e136.

388. Matoba S, Liu Y, Lu F, Iwabuchi KA, Shen L, Inoue A, Zhang Y. Embryonic development following somatic cell nuclear transfer impeded by persisting histone methylation. Cell. 2014;159(4):884-895.

389. Kang YK, Koo DB, Park JS, Choi YH, Chung AS, Lee KK, Han YM. Aberrant methylation of donor genome in cloned bovine embryos. Nat Genet. 2001;28(2):173-177.

390. Song X, Liu Z, He H, Wang J, Li H, Li J, Li F, Jiang Z, Huan Y. Dnmt1s in donor cells is a barrier to SCNT-mediated DNA methylation reprogramming in pigs. Oncotarget. 2017;8(21):34980-34991.

391. He X, Tan C, Li Z, Zhao C, Shi J, Zhou R, Wang X, Jiang G, Cai G, Liu D. Characterization and comparative analyses of transcriptomes of cloned and in vivo fertilized porcine pre-implantation embryos. Biol Open. 2019;8(4).

392. Sood TJ, Lagah SV, Mukesh M, Singla SK, Chauhan MS, Manik RS, Palta P. RNA sequencing and transcriptome analysis of buffalo (Bubalus bubalis) blastocysts produced by somatic cell nuclear transfer and in vitro fertilization. Mol Reprod Dev. 2019.

393. Min B, Cho S, Park JS, Lee YG, Kim N, Kang YK. Transcriptomic features of bovine blastocysts derived by somatic cell nuclear transfer. G3 (Bethesda). 2015;5(12):2527-2538.

394. Betsha S, Hoelker M, Salilew-Wondim D, Held E, Rings F, Grosse-Brinkhause C, Cinar MU, Havlicek V, Besenfelder U, Tholen E. Transcriptome profile of bovine elongated conceptus obtained from SCNT and IVP pregnancies. Mol Reprod Dev. 2013;80(4):315-333.

395. Liu W, Liu X, Wang C, Gao Y, Gao R, Kou X, Zhao Y, Li J, Wu Y, Xiu W. Identification of key factors conquering developmental arrest of somatic cell cloned embryos by combining embryo biopsy and single-cell sequencing. Cell Discov. 2016;2:16010.

396. Hossain MM, Tesfaye D, Salilew-Wondim D, Held E, Proll MJ, Rings F, Kirfel G, Looft C, Tholen E, Uddin J. Massive deregulation of miRNAs from nuclear reprogramming errors during trophoblast differentiation for placentogenesis in cloned pregnancy. BMC Genomics. 2014;15:43.

397. Castro FO, Sharbati S, Rodriguez-Alvarez LL, Cox JF, Hultschig C, Einspanier R. MicroRNA expression profiling of elongated cloned and in vitro-fertilized bovine embryos. Theriogenology. 2010;73(1):71-85.

398. Akagi S, Geshi M, Nagai T. Recent progress in bovine somatic cell nuclear transfer. Anim Sci J. 2013;84(3):191-199.

399. Zhang J, Wang Y, Liu X, Jiang S, Zhao C, Shen R, Guo X, Ling X, Liu C. Expression and potential role of microRNA-29b in mouse early embryo development. Cell Physiol Biochem. 2015;35(3):1178-1187.

400. Mengyun Wang YG, Pengxiang Qu, Suzhu Qing, Fang Qiao, Yong Zhang, Jesse Mager & Yongsheng Wang Sperm-borne miR-449b influences cleavage, epigenetic reprogramming and apoptosis of SCNT embryos in bovine. Sci Rep. 2017;7.

401. Chang HY, Xie RX, Zhang L, Fu LZ, Zhang CT, Chen HH, Wang ZQ, Zhang Y, Quan FS. Overexpression of miR-101-2 in donor cells improves the early development of Holstein cow somatic cell nuclear transfer embryos. J Dairy Sci. 2019;102(5):4662-4673.

402. Rashmi, Sah S, Shyam S, Singh MK, Palta P. Treatment of buffalo (Bubalus bubalis) SCNT embryos with microRNA-21 mimic improves their quality and alters gene expression but does not affect their developmental competence. Theriogenology. 2019;126:8-16.

403.	Wang P, Li X, Cao L, Huang S, Li H, Zhang Y, Yang T, Jiang J, Shi D. MicroRNA-148a overexpression improves the early development of porcine somatic cell nuclear transfer embryos. PLoS One. 2017;12(6):e0180535.

404.	Li W, Xiong Y, Wang F, Liu X, Gao Y, Wang Y, Zhang Y, Jin Y. MicroRNA-145 inhibitor significantly improves the development of bovine somatic cell nuclear transfer embryos in vitro. Cell Reprogram. 2016;18(4):230-236.

405.	Liang S, Nie ZW, Guo J, Niu YJ, Shin KT, Ock SA, Cui XS. Overexpression of microRNA-29b decreases expression of DNA methyltransferases and improves quality of the blastocysts derived from somatic cell nuclear transfer in cattle. Microsc Microanal. 2018;24(1):29-37.

406.	Menezo YJ, Herubel F. Mouse and bovine models for human IVF. Reprod Biomed Online. 2002;4(2):170-175.

CHAPTER 2

THE MATERNAL-TO-ZYGOTIC TRANSITION IN BOVINE IN

VITRO-FERTILIZED EMBRYOS IS ASSOCIATED WITH MARKED

CHANGES IN SMALL NON-CODING RNAS [1]

Abstract

In mammals, small non-coding RNAs (sncRNAs) have been reported to be important during early embryo development. However, a comprehensive assessment of the inventory of sncRNAs during the maternal-to-zygotic transition (MZT) has not been performed in an animal model that better represents the sncRNAs biogenesis pathway in human oocytes and embryos. The objective of this study was to examine dynamic changes in expression of sncRNAs during the MZT in bovine embryos produced by *in vitro* fertilization (IVF), which occurs at the 8-cell stage. An unbiased, discovery-based approach was employed using small RNAseq to profile sncRNAs in bovine oocytes, 8-cell stage embryos and blastocyst stage embryos followed by network and ontology analyses to explore the functional relevance of differentially expressed microRNAs (miRNAs). The relative abundance of miRNAs was markedly higher in 8-cell stage embryos compared to oocytes or blastocyst stage embryos. This shift in miRNA population was largely associated with up-regulation of miRNAs predicted to target genes involved in the biological processes of cell development, cell division, Wnt signaling, and pluripotency, among others. Distinct populations of piwi-interacting-like RNAs (pilRNAs) were identified in bovine oocytes and blastocyst stage embryos, though pilRNAs were nearly absent in 8-cell stage embryos. Also, small nucleolar RNAs (snoRNAs) were highly expressed in 8-cell stage embryos.

Overall, these data reveal a strong dynamic shift in the relative abundance of sncRNAs associated with the MZT in bovine oocytes and embryos, suggesting that these molecules may play important roles in the shift from maternal to zygotic control of gene expression.

Introduction

In vitro embryo production has played an increasingly important role in both medicine and agriculture by improving methods for treating human infertility and by selecting for superior genetic traits in key agricultural species, such as cattle. Determinants of embryo developmental competence are rooted in the oocyte [1]. In addition to the maternal genome, the mature oocyte provides massive stores of RNA and protein reserves necessary for the first cleavage divisions as well as embryonic programming events [1]. After fertilization, the epigenetic marks specific to the differentiated sperm and oocyte must be erased and re-established as embryonic epigenetic marks for successful development to occur. This process is monumentally complex as it requires large scale, highly coordinated changes in epigenome programming during a very narrow developmental window. Coincident with this large-scale reprogramming of the epigenome is the maternal-to-zygotic transition (MZT), a highly orchestrated process during which control of embryonic development shifts from the maternally deposited RNA and proteins to those of the zygotic genome [2]. This process requires degradation of the maternal mRNA transcripts derived from the oocyte and activation of the zygotic genome via extensive epigenome programming. The dramatic turnover of the transcriptome is integral in order to prepare the embryo for cell differentiation and further development. Maternal transcripts are removed by both maternal and zygotic derived degradation pathways, and although there are many different maternal pathways for degradation, zygotic degradation activity plays a more prevalent role [3]. Graf and colleagues pinpointed the 8-cell stage as the major point for the MZT in cattle, with the earliest activation of genes related to RNA processing occurring at the 4-cell embryonic stage, initiating large-scale

embryonic genome activation (EGA) [4]. In this period of quiescence, there is high potential for sncRNA to play a role in the degradation of maternal transcripts, as sncRNA have been found to act in post-transcriptional regulation of many genes.

Once the MZT has successfully occurred, the newly activated embryonic transcriptome guides cells through the first of many cell fate decisions at the blastocyst developmental stage. The first cell lineage specifications are the differentiation of early embryonic cells into the inner cell mass (ICM), destined to become embryonic tissue, and the trophectoderm (TE) that will differentiate into extra embryonic tissues. The cells of ICM must maintain a state of pluripotency for differentiation into many distinct embryonic tissues, whereas TE cells must block pluripotency signals in order to acquire their specialized functions that support embryo implantation. In both bovine and human ICM cells, the transcription factors SOX2 and NANOG direct the change in expression of many other genes to continue the molecular maintenance of pluripotency [5]. Changes in gene expression patterns must be established correctly through the epigenetic programming that takes place during the MZT, as any errors in this process could cause embryonic loss due to failure of pluripotency and cellular lineage specification pathways.

Non-coding RNAs (ncRNAs) were once considered to be non-functional transcription products of junk DNA. While ncRNAs are not translated into proteins, they are now known to function in RNA interference (RNAi), acting as powerful post-transcriptional regulators of gene expression. RNAi involves blocking the translation or reducing the stability of a messenger RNA (mRNA) through complimentary binding, and this process may control as much as 60% of the protein coding genes expressed in mammals [6]. In mammals, three classes of sncRNA have been identified as playing substantial roles in early development in mammals, including microRNAs (miRNAs), PIWI-interacting RNAs (piRNAs), and endogenous small interfering RNAs (siRNAs). All three of these sncRNA classes function through RNAi via degradation of either mRNA or transposable element (TE) targets, although their biogenesis and targeting

mechanisms differ [7]. Other important sncRNAs include ribosomal (rRNA), transfer (tRNA), small nucleolar (snoRNA) and small nuclear (snRNA) RNAs. Ribosomal RNA (rRNA) is necessary for protein synthesis and are of particular importance during cell growth to maintain ribosomes in subsequent generations [8]. While the amount of rRNA present differs in oocytes and embryos, the population of rRNA variants is largely similar, suggesting that these rRNA variants do not have a particular specialized function in embryos [9]. Transfer RNAs play a critical role in the translation of mRNA sequences to protein, and recently fragments of tRNAs have been implicated in RNAi through regulation of transcripts and possibly TE [10, 11]. SnoRNAs control ribosome biogenesis by guiding the modification or processing of pre-rRNA. Specifically, C/D box snoRNAs are associated primarily with methylation modifications and H/ACA box snoRNAs are associated with pseudouridylation. Similar to tRNA, snoRNAs are also processed into smaller fragments which may participate in RNAi [12]. Finally, snRNAs have well defined functions in mRNA splicing and 3'-end formation [13].

MicroRNAs are short, single-stranded RNAs approximately 22 nucleotides (nt) in length that are a highly conserved subset of ncRNA [7]. RNA polymerase II transcribes the majority of primary-miRNA transcripts (pri-miRNA) from the genome, which are then cleaved by Drosha and its mammalian cofactor DGCR8 to generate a 70 nt stem-loop precursor miRNA (pre-miRNA). Exportin 5 then exports the pre-miRNA from the nucleus to the cytoplasm for processing, where the Dicer complex cleaves the stem-loop releasing double stranded, non-hairpin miRNA. The miRNA duplex is then loaded onto an AGO protein, which unwinds the duplex and loads the guide strand of mature miRNA onto the RNA-induced silencing complex (RISC), which directs the miRNA to its target mRNA [7]. While this is the normal biogenesis pathway, a minority of miRNAs can also be processed independently of both Drosha and Dicer using alternative biogenesis pathways [14]. MicroRNAs bind their target RNAs via imperfect complementary base pairing to target sites in the 3' untranslated regions (UTR), and often one

mRNA is targeted by more than one miRNA for maximal repression. In order to destabilize targets, miRNAs promote the decapping and deadenylation of their target mRNA, localizing the target to the site of degradation in cytoplasmic processing bodies [15]. The majority of miRNAs investigated experimentally inhibit gene expression through binding interactions that block translation and decrease stability of the target transcript.

Several studies have determined that miRNA populations are dynamically regulated through oocyte maturation and development in multiple species, including zebrafish, mice and pigs [16]. For example, Giradlez and colleagues reported that miR-430 facilitated the degradation of hundreds of maternal transcripts in zebrafish embryos, an observation that pointed to the importance of this miRNA during the MZT [17]. No equivalent function of miR-430 is known in mammalian development, however several groups have shown that miR-302 (a mammalian orthologue of miR-430) enhances the reprogramming of fibroblasts to generate induced pluripotent stem cells (iPSC) in both humans and mice [18]. MicroRNAs have also been shown to regulate both pluripotency and differentiation of embryonic stem cells by indirectly targeting the expression of key pluripotency factors [19]. In cattle, miRNAs are differentially expressed over the course of oocyte maturation [20, 21], in blastocyst stage embryos [22], and in blastocysts as they mature to the hatching stage [23]. The functions of some miRNA have been investigated in cattle. For example, Tripurani and colleagues [24] reported that bovine miR-196a was essential for degrading the maternal effect gene *NOBOX* in the early MZT. To our knowledge, the dynamics of miRNA expression throughout the MZT has only been studied using targeted approaches, such as qPCR [25].

Aside from miRNA, other classes of sncRNA may also be important for the MZT, as their abundance changes over the course of parental genome silencing and embryonic genome activation. In mammalian gametes, siRNA and piRNA are also abundant, but due to their specialized function, different functions of these sncRNA populations have been less thoroughly

investigated [26]. Endogenous siRNAs are structurally similar to miRNAs and are processed from long double stranded RNA duplexes by Dicer. However, endogenous siRNAs are less conserved among species, and their pre-cursors are not processed by Drosha/DGCR8 [27]. The siRNAs exert a repressive effect on mRNA, similar to that of miRNA, through the RISC complex and AGO proteins. Small interfering RNAs bind with perfectly complementary to target sequences, which then induces endonucleolytic degradation [28]

Piwi-interacting RNAs are single stranded RNA molecules associated with the PIWI subfamily of proteins. Piwi-interacting RNAs bind RNA targets through complementarity and repress targets via slicing or chromatin mark recruitment. These piRNAs repress transposable elements (TEs), both transcriptionally and post-transcriptionally during genomic remodeling periods [29]. Piwi-interacting RNAs were originally discovered to function in gametogenesis and have been shown to silence TEs during reprogramming events throughout primordial germline cell development [30], however the role of piRNA silencing of TEs during embryogenesis is unknown. Bui and colleagues observed transcriptional activation of TEs during the MZT in bovine somatic cell nuclear transfer embryos, suggesting that TEs may become expressed during this phase as a consequence of somatic cell nuclear reprogramming [31]. Relatively high expression of TEs during the MZT has been found in other species as well, including humans [2]. Elevated expression of piRNAs prior to the major wave of the MZT may be necessary to prevent TE activation and maintain genomic integrity as epigenetic programming takes place. During pachytene spermatogenesis, piRNAs target not only TEs, but also target mRNA for degradation, suggesting that piRNAs may have another mechanism to suppress gene expression [32]. Previously, Roovers and colleagues detected piRNAs and high expression of PIWIL3 in bovine oocytes and 1-2 cell embryos, as well as piRNAs and PIWIL1 in bovine ovaries [29]. As members of the argonaute family, PIWIL proteins build complexes with piRNAs to regulate gene expression making them essential for recognition and degradation of piRNA targets. In addition,

Russell and colleagues examined the expression of piRNAs in cattle oocytes and zygotes and determined that expression of some piRNAs appeared to correlate with genes destined for degradation in the embryo [33]. However, mouse oocytes do not express PIWIL3 and contain relatively low amounts of piRNA. While mouse oocytes do express other PIWI proteins, including PIWIL1, PIWIL2 and PIWIL4, functional knock out studies suggest that their expression is not required to maintain female fertility [29].

These observations suggest that the function of the piRNA-PIWI pathway is not conserved in mammals, and bovine models may better represent these processes for human development. While it is unclear why the mouse is different from other mammals with respect to the piRNA-PIWI pathway, it is possible that high production of siRNAs in mouse oocytes may compensate for the lack of piRNA activity [29]. Although the expression of piRNAs has been determined in bovine oocytes and presumptive zygotes [33], the dynamics of piRNA expression during the MZT have not yet been explored. In addition to the similarities in the sncRNA biogenesis pathways during early development, other similarities exist between bovine and human embryos. For example, loss of the pluripotency factor *Oct4* in mice did not affect acquisition of pluripotency for cells destined to become the epiblast [34]. Conversely, absence of *OCT4* in both cattle and human embryos inhibited the emergence of these pluripotent epiblast cells, whereas cells that would become the primitive endoderm continued their development [34]. Additionally, both bovine and human pre-implantation embryos expressed similar miRNAs into culture media, including miR-25, miR-302c, miR-196a2, and miR-181a, which may serve as biological markers for selection of high quality embryos [35]. Because of these key similarities in sncRNA biogenesis and/or activity pathways as well as pluripotency pathways in early embryogenesis, the use of bovine embryos for the study of sncRNA in early development is appropriate.

As outlined above, others have investigated the contribution of specific classes of

sncRNA acting through RNAi to modulate gamete maturation, embryogenesis or acquisition of

pluripotency. Also, recently Yang and colleagues described RNAseq of sncRNAs associated

with the MZT in mouse [36], although the profile of sncRNAs in this mammalian model is likely

influenced by distinct biogenesis pathways for siRNA and miRNA. However, a comprehensive

assessment of the inventory of sncRNAs during the MZT has not been performed in an animal

model that may better represent the sncRNAs biogenesis pathway in human oocytes and embryos.

Thus, the objective of this study was to examine dynamic changes in expression of sncRNAs

during the maternal-to-zygotic transition in bovine embryos produced by *in vitro* fertilization

(IVF). Based on the evidence discussed above, we hypothesized that miRNAs would be more

abundant in 8-cell embryos at the start of the MZT and that abundant miRNAs would target

maternal mRNAs with important developmental functions associated with epigenetic

programming and development. An unbiased, discovery-based approach was employed by using

small RNAseq to profile all sncRNA classes in bovine oocytes, 8-cell stage embryos and

blastocyst stage embryos followed by network and otology analyses to explore the functional

relevance of differentially expressed miRNAs.

Methods

Oocyte collection and in vitro maturation

Bovine ovaries were collected from a local abattoir (JBS, Hyrum, UT) and transported in

a cooler containing 0.9% saline solution. The cumulus-oocyte complexes (COCs) were aspirated

from 3-8 mm follicles by using an 18-gauge needle and vacuum system. Only compact COCs

with homogenous ooplasm and intact layers of cumulus cells were used. Following aspiration,

COCs were cultured at 39 °C with 5% CO_2 for 22 to 24 hr. The oocytes were cultured in TCM199

maturation medium with Earle's salts, L-glutamine, and sodium bicarbonate (Hyclone, Logan,

Utah) supplemented with 10% fetal bovine serum (FBS), 0.05 mg/ml bovine follicle stimulating

hormone (Sioux Biochemicals, Sioux city, IA), 5 mg/ml bovine luteinizing hormone (Sioux

Biochemicals), 100 U/ml penicillin, and 100 mg/ml streptomycin.

In vitro fertilization and embryo culture

Following 22 to 24 hr of culture, MII oocytes were fertilized using the laboratory's

standard *in vitro* fertilization (IVF) protocol [37]. Briefly, one straw of cryopreserved bovine

semen obtained from a Holstein bull (Hoffman AI, Logan, UT) was removed from the liquid

nitrogen tank and placed into a 35 °C water bath to thaw. Live sperm were isolated by

centrifugation through a 45%/90% percoll gradient, suspended (final concentration $1x10^6$/ml) in

Tyrode's albumin lactate pyruvate containing 10 μg/mL heparin and used to fertilize the mature

oocytes. Twenty to 22 hours post-IVF, cumulus cells were removed by vortexing, and fertilized

zygotes were washed in phosphate buffered saline with 0.32 mM sodium pyruvate, 5.55 mM

glucose, and 3 mg/ml bovine serum albumin (PB1+). After washing, zygotes were cultured in

Charles Rosenkrans 2 medium (CR2) at 39 °C with 5% CO_2. Half of the CR2 media was

removed and replaced with fresh, equilibrated media every other day starting the day after *in vitro*

culture. Historically, production of IVF embryos via these methods followed by blastocyst stage

embryo transfer to cattle recipients has generated successful pregnancies at an approximate rate

of 50% by our research group (unpublished observations).

Tissue collection

Three pools of 40 mature oocytes were collected at 22 hr after maturation. To remove

cumulus cells, the mature COCs were treated with 10 mM hyaluronidase for 10 min and then

vortexed for 2 min. After visual inspection for complete removal of cumulous cells and the

presence of a polar body, denuded mature oocytes were washed through four droplets of PB1+.

Oocytes were then snap frozen in cryotubes containing RNA/DNA shield (Zymo, Irvine,

California), and stored at -80 °C until RNA isolation. Three pools of 40 8-cell staged embryos and three pools of 40 blastocyst staged embryos were collected at 3 and 8 days post IVF, washed in PBS, and stored similarly. Example microscopy images of collected tissues are provided in Supplementary Figure 2.S1.

Small RNA isolation and sequencing

RNA was isolated based on size using the RNA Clean & Concentrator™ 5 kit (Zymo, Irvine, CA) from three pools each of oocytes, 8-cell stage embryos or blastocyst stage embryos according to the manufacturer's protocol for purification of small RNA (<200nt) and large RNA (>200 nt) as separate fractions. Small RNA sequencing was performed on the Ion Proton™ Sequencer using the Ion Total RNA-seq kit v2 (Thermo Fisher Scientific, Waltham MA) according to manufacturer's procedure for small RNA library preparation with no deviations from the specified protocol. By using an RNA isolation protocol that yielded a specific fraction of small RNA, an enrichment step was not needed. Sample volumes were reduced to 3 μl by vacuum centrifugation, and the entire sample was used to prepare the small RNA library for sequencing. The cDNA sample was then processed on the Agilent Bioanalyzer (Agilent, Santa Clara CA) to ensure the presence of small RNA bound to a barcode (86 to 106 nt).

MicroRNA data processing and analyses

Sequence data were processed to remove low quality reads and any artificial reads introduced during library preparation (Perl script Trim Galore, Python script SortMeRNA) [38, 39]. Next, miRDeep2 (v 2.0.0.8) was used to align the small RNA reads to the bovine miRNA database with all options at program defaults [40]. The output from this program was the number of reads per bovine miRNA identifier (miR). Following sequence alignment to miRDeep2, read counts for each mapped miRNA were exported to Excel (Supplementary File 2.1; Appendix C;

Appendix C) and then filtered to exclude miRNAs with total sequence reads <100 across all

samples. Read counts were then transformed to reads-per-million (RPM) values by dividing the

number of mapped sequence reads to a miRNA ID by the total reads for the sample and

multiplying by 10^6.

Quantitative RT-PCR

RNA was isolated based on size as described above from pools of 20 oocytes, 8-cell stage

embryos or blastocyst stage embryos for analysis by qRT-PCR (n = 5 to 6 per tissue type).

Reverse transcription was performed on the isolated mRNA using the GoScript Reverse

Transcriptase kit (Promega, Madison WI). The Fluidigm Biomark HD instrument was used for

quantitative, nano-scale PCR using a 48.48 Dynamic Array chip (Fluidigm Corporation, South

San Francisco, CA). Target mRNAs were selected for the mature miRNA sequence (miR) IDs

with most pronounced changes in abundance (two-fold difference), for both over- and under-

expressed miRNAs. Selection of target mRNAs was performed using miRmap, which predicts

mRNA targets of miRNAs based on computed repression strength using thermodynamic,

evolutionary, probabilistic and sequence based features [41]. Using the lists of predicted miRNA

targets for each miRNA, 30 genes previously shown to be expressed in early developing cattle

embryos [42] were selected for analysis by qRT-PCR (see Supplementary Table 2.1). Custom

primers were designed to span exons using Integrated DNA Technologies (Coralville, IA)

RealTime qPCR tool and the NCBI reference genome for *Bos taurus* (Supplementary Table 2.2).

Fluidigm's recommended protocol (ADP14) was followed for performing quantitative PCR gene

expression analysis using the Biomark HD system. The protocol included a pre-amplification

step (activation at 95°C for 10 min followed by 14 cycles of 95°C for 15 sec and then 58°C for 4

min) followed by standard PCR (polymerase activation/denaturation 95°C for 10 min followed by

30 cycles of 95°C for 10 sec and 60°C for 1 min). Data were analyzed using the standard ΔΔCt method with housekeeping genes *GAPDH* and *SDHA* for normalization.

Network analyses of predicted gene targets of differentially expressed miRNAs

Gene targets for miRNAs identified as over- or under-expressed in 8-cell stage embryos as compared to oocytes or blastocyst stage embryos were predicted using TargetScan (release 7.1) [43] with the *Bos taurus* miRNA database (miRBase release 21). TargetScan predicts biological targets of miRNAs by searching for the presence of conserved 8-mer, 7-mer, and 6-mer sites that match the seed region of each miRNA. TargetScan provides an additional advantage by curating individual miRNAs by family (identical seed ±m8 sequence). Only transcripts with conserved sites were included in the predicted gene target list, which was further filtered for total context++ score <-0.35 (efficacy of targeting miRNA site) (Supplementary Table 2.3). The final gene lists for network analyses included 1079 predicted mRNA targets of over-expressed miRNAs and 1670 predicted targets of under-expressed miRNAs in bovine 8-cell stage embryos (Supplementary File 2.2; Appendix D). Note that the output of TargetScan is the human orthologue of the predicted target gene; the representative transcript shown is the transcript with the UTR profile that is most prevalent (highest number of supporting 3P-seq tags). However, TargetScan uses the cattle species sequence for prediction, not the human sequence for prediction. In Supplementary File 2.2; Appendix D, the representative miRNA listed is the miRNA with the lowest cumulative weighted context++ score within that miRNA family; although only one miRNA is shown as matched to the target transcript, all other members of the miRNA family were also predicted to target the same site for that transcript. Supplementary File 2.2; Appendix D also includes a table defining bovine miRNA families by level of conservation with other species.

Predicted mRNA targets were subject to gene set enrichment analysis using Metascape [44]. For these two gene lists (targets of over- or under-expressed miRNAs), pathway and process enrichment meta-analyses were carried out using the following ontology sources: GO Biological Processes, KEGG Pathways and Reactome Gene Sets. All genes in the *Bos taurus* genome were used as the enrichment background. Terms with *p*-value <0.05, minimum count of three associated genes and enrichment factor >1.5 (enrichment factor is the ratio between number of genes associated with the term observed in the query set and the number expected by chance) were collected and grouped into clusters based on their membership similarities. More specifically, *p*-values were calculated based on accumulative hypergeometric distribution, and FDR *q*-values were calculated using the Benjamini-Hochberg procedure to account for multiple testing. Kappa scores were used as the similarity metric when performing hierarchical clustering on the enriched terms, and sub-trees with similarity >0.3 were considered a cluster. The most statistically significant term within a cluster was chosen to label the cluster.

To further capture the relationship among terms, a subset of enriched terms was selected and rendered as a network plot, where terms with similarity >0.3 were connected by edges. Terms with the highest *q*-values for each cluster were used to build the network, with constraints of no more than 15 terms per cluster and fewer than 250 terms total. Networks were visualized in Cytoscape [45], and separate network plots were generated for enrichment analyses using GO biological process terms. The term with the most significant *q*-value within each cluster was used to name the cluster, and networks were colored according to cluster identity, *q*-value or relative contribution of terms to the node from each gene list.

Results of Metascape ontology analyses were verified by performing enrichment analysis using the DAVID functional annotation tool (v. 6.8) [46] with the complete predicted mRNA target list (for both over- and under-expressed miRNAs) and the *Bos taurus* genome as the background data set.

Network analyses of differentially expressed miRNAs and expressed genes in bovine oocytes

A network depicting connections between differentially expressed miRNAs in bovine 8-cell stage embryos as compared to oocytes and their predicted mRNA targets was built using Cytoscape. We expected this network to reveal broad scale patterns in mRNA expression that corresponded to changes in abundance of the target mRNAs. Many of the mRNA targets were predicted to be regulated by several (or many) miRNAs. Thus, this network also allowed for visualization of the interconnectedness, or conversely independence, of all the differentially expressed miRNAs and their mRNA targets. Central network nodes were defined as miRNAs over- or under-expressed in 8-cell stage embryos (as compared to oocytes) and were linked by edges to predicted mRNA targets (TargetScan total context score <-0.4 to reduce network size). Nodes for predicted mRNA targets were colored using data available from the Gene Expression Omnibus dataset GSE52415 [42], which includes RNAseq transcriptome data for cattle oocytes and early developing embryos. These data were expressed either as the variance-stabilizing transformed read abundance or the calculated $\log_2$ ratio of expression in 8-cell stage embryo with respect to MII oocytes. Because the main effect of miRNA is to target mRNA transcripts for degradation, nodes representing genes not represented in this dataset were retained in the network and represented as gray.

Annotation and expression of other sncRNAs

Annotation and expression analysis of other sncRNAs was performed as described previously [33]. Briefly, data processing and bioinformatics analysis for other sncRNA were performed primarily with command line tools available at http://www.smallrnagroup.uni-mainz.de/software.html. First data were filtered for sequence length between 17 and 32 nt, which includes the canonical size range for mature small RNAs (miRNA, siRNA, piRNA) (Supplementary File 3). A separate analysis was performed for sequences between 33-93 nt,

which encompassed longer reads characteristic of snoRNA, tRNA and mitochondrial rRNA. This split analysis approach avoided possible skewing of the data when comparing between these diverse sncRNA classes. Sequence annotation was performed with Unitas v1.4.2, which uses the latest available public small RNA databases to annotate input sequences [47]. Non-annotated sequences were then mapped to the bovine genome (BosTau8) with sRNAmapper [48] to determine which non-annotated sequences were likely of biological, not technical, origin. The mapped sequence reads were redistributed across the genome based on number of mapping locations using reallocate with the parameters '5000 1000 b 0' (reallocate.pl). Reallocated map files were then analyzed with ProTRAC v2.3.1 [49], which identifies probable piRNAs through their genomic clustering based on size, sequence, and cluster characteristics (all options at default settings, except for increasing allowable piRNA size range to 32 nt). Predicted piRNA clusters were compared among samples by using the Galaxy "merge" tool (usegalaxy.org) to first generate lists of predicted clusters by sample type followed by the "join" tool to identify overlapping genomic intervals among samples. Finally, repeat masker was used to identify potential piRNAs overlapping with known transposable elements in both the sense and antisense directions.

Data analysis

For miRNAs and snoRNAs, normalized read count data (RPM values) were $\log_{10} (y\pm1)$ transformed and then analyzed by one-way ANOVA followed by post-hoc tests for all pair-wise comparisons using TM4 MeV [50]. Because the post-hoc tests must be performed independent of the ANOVA in MeV, the resulting *p*-values are not corrected for multiple testing across sample types. Thus, all pairwise *p*-values were processed using GraphPad Prism (v.7) as a single family to correct for multiple comparisons across all samples and miRNA or snoRNA species by controlling the false discovery rate (FDR) (two-stage step-up method of Benjamini, Krieger and

Yekutieli). For further analysis of miRNA or snoRNA abundance profiles, data were expressed as the $\log_2$ ratio of normalized read counts for each individual sample to the average normalized read count for oocytes (Supplementary Files 2.1 and 2.3). A significant difference in miRNA or snoRNA abundance was inferred when FDR $q<0.05$ and $\log_2 |R|\geq1$ (a two-fold difference). ClustVis [51] was used to perform unsupervised, bi-directional hierarchical cluster analyses using both RPM abundance and $\log_2$ ratio data, including either all mapped IDs or those determined to be differentially expressed.

Results

Classification of sncRNA sequences

RNA sequencing of sncRNA obtained from bovine oocytes, 8-cell embryos, and blastocyst stage embryos generated a total of about 11 million sequence reads, after filtering low quality reads and artifacts, of which approximately 25% aligned to known bovine sncRNAs (Supplementary Table 2.4). The number of reads obtained per sample was markedly higher for 8-cell stage embryos (approximately 2.8 million reads/sample) than oocytes (about 0.66 million reads/sample) or blastocyst stage embryos (about 0.45 million reads/sample) (Figure 2.1A). RNA input was not normalized by sample type when generating the sample libraries for sequencing due to the exceedingly small amount of small RNA obtained. However, had the number of cells per embryo been the driving factor in yield of reads per sample type, then a different pattern would have been observed with blastocyst stage embryos having the greatest number of sequence reads.

As shown in Figure 2.1B, sequence reads of 17-32 nt in length, corresponding to endogenous siRNA and miRNA classes of sncRNA, were most abundant across sample types. Also, peaks for read lengths corresponding to snoRNAs (70 to 90 nt) and tRNAs (76-90 nt) were also evident for all sample types, with other minor peaks at various lengths likely corresponding to rRNA, snRNA or other miscellaneous RNA fragments. As a fraction of total reads per sample,

markedly more sequences were annotated for 8-cell stage embryos (36.5% on average) than

oocytes (6.1%) or blastocyst stage embryos (8.5%) (Figure 2.1C). A similar pattern was also

evident for sequences ≥33 nt, with 14.1%, 37.8% and 18.0% of reads annotated to non-coding

RNAs in oocytes, 8-cell stage embryos and blastocyst stage embryos, respectively (Figure 2.1D).

When considered as a fraction of total reads for each sample, the sncRNA population was clearly

distinct for each tissue (Figure 2.1C and D, Table 2.1). The relative abundance of miRNAs was

significantly greater in 8-cell stage embryos (14%) compared to just 2.1% or 2.7% for oocytes or

blastocyst stage embryos, respectively (Table 2.1, Supplementary Figure 2.S2A). When

calculated as a fraction of the total annotated reads, the proportion of miRNAs was similar, at

31%, 38% and 33% on average for oocytes, 8-cell stage embryo and blastocyst stage embryo.

However, one should note that the proportion of annotated reads varied markedly by sample type.

Also, the majority of the unannotated reads for oocyte, for example, were classified as putative

piwi-like RNAs (details below). Other classes of sncRNA were also proportionally more

abundant, including genomic tRNA and snoRNA. Distinct sncRNA profiles were also evident

for the ≥33 nt sncRNA fraction, with the most notable difference being the marked, significant

increase in relative abundance of snoRNA in 8-cell stage embryos (28.4%) compared to oocytes

(6.2%) or blastocyst stage embryos (7.7%) (Table 2.1, Supplementary Figure 2.S2B).

miRNA expression in oocytes, 8-cell stage
embryos and blastocyst stage embryos

MicroRNA annotated with miRDeep2 were aggregated by bovine miRNA family, which

groups miRNA species by the seed sequence. After filtering out miRNAs with very low

abundance reads across samples (sum of <100 reads across all 9 samples), 151 miRNA species

representing 102 bovine miRNA families were detected across all sample types (Supplementary

File 2.1; Appendix C, Supplementary Figure 2.S3). The Venn diagram in Figure 2.2A depicts

miRNA species detected by sample type. While a core set of 20 miRNAs were detected by RNAseq in all sample types, a large set of 45 miRNAs were identified only in 8-cell stage embryos indicating a unique miRNA expression profile for that tissue. A subset of 35 additional miRNAs were identified as common to both 8-cell stage embryos and oocytes, whereas all the miRNAs expressed in blastocyst stage embryos were also present in oocytes or 8-cell stage embryos.

The sequence data were further examined to identify specific miRNA families differentially expressed in any one sample type. Of the 102 families detected by RNAseq, 37 were differentially expressed ($|\log2 R|>1$ and FDR q-value <0.01) in 8-cell stage embryos as compared to oocytes, including 11 miRNA families under-expressed and 26 miRNAs families over-expressed (Table 2.2, Supplementary Figure 2.S4). No significant differences in miRNA expression were noted when comparing blastocyst stage embryos to oocytes, and only three significantly different miRNAs were observed when comparing 8-cell stage embryos to blastocyst stages, likely due to apparent high variation in expression profiles among the blastocyst stage samples.

Principal components analysis of miRNA expression data revealed clear grouping of data sets by sample type, with very tight clusters evident for oocyte and 8-cell stage embryo samples, indicating high consistency in the overall miRNA abundance profile for those samples (Figure 2.2B). In contrast, the miRNA expression profile for blastocyst stage embryos was more variable, with one sample notably separated from other samples in the same developmental stage. These miRNA profiles are illustrated more clearly via unsupervised, bi-directional hierarchical clustering of miRNA species and samples using miRNA RPM abundance data (Figure 2.2C) or relative expression with respect to oocytes (Figure 2.2D). For each clustering map, the samples cluster according to sample type, with the 8-cell stage embryo group clearly distinct from oocytes and blastocyst stage embryos.

Ontology and pathway analyses of predicted mRNA targets for
differentially expressed miRNAs in 8-cell stage embryos

Results of biological process gene ontology, KEGG pathway, and Reactome pathway

analyses of TargetScan predicted gene targets of differentially expressed miRNAs in 8-cell stage

embryos are provided in Supplementary File 2.2; Appendix D and summarized in Figure 2.3A.

Note that the heatmaps represent the top 20 terms (clustered by kappa score for similarity) for the

overall data set. Although the input data were separated into subsets corresponding to mRNA

targets of over- or under-expressed miRNAs, the top 20 biological process terms are largely

similar for both subsets including terms related to "developmental processes", "Wnt signaling",

"embryonic morphogenesis", "covalent chromatin modification", and others. Of note, "cell

division" and "developmental growth" are among the top 20 most significant terms for gene

targets of under-expressed miRNAs, but these terms are not among the 20 most significant for

targets of over-expressed miRNAs (though the terms are indeed significantly enriched.) Results

of this ontology analysis are also represented as a Cytoscape network in Figure 2.3B, with each

cluster corresponding to a term shown in Figure 2.3A colored by the analysis FDR q value.

Clusters related to developmental biological processes (embryo morphogenesis and various

developmental pathways) form a central network further connected to other critical processes

related to embryo development, such as "regulation of cell growth", "Wnt signaling", and

"regulation of kinase activity". Other enriched terms are disconnected from the network, such as

"cell division", "covalent chromatin modification" and "endomembrane system organization".

Supplementary Figure 2.S5 depicts this network colored by biological process cluster or by count

of predicted mRNAs associated with each node. Most nodes are represented by terms associated

with mRNA targets of both over- and under expressed mRNAs in 8-cell stage embryos, with a

proportion that approximately represents the number of terms in each query list (1078 or 40% for

over-expressed miRNAs and 1670 or 60% for under-expressed miRNAs). However, some

biological processes were preferentially associated with mRNA targets of under-expressed miRNAs, such as terms "negative regulation of cell morphogenesis involved in differentiation", "negative regulation of developmental growth", "cell division", "positive regulation of Wnt signaling pathway", "negative regulation of cell projection organization" and all nodes represented in the cluster named "endomembrane organization". In contrast, processes associated preferentially with mRNA targets of over-expressed miRNAs were fewer, including "positive regulation of cartilage development", "positive regulation of chondrocyte development" and "embryonic skeletal system development".

Of the enriched KEGG pathways, pathways in cancer was the most significant for targets of under-expressed miRNAs, whereas MAPK signaling pathway was the most enriched for mRNA targets of over-expressed miRNAs. Unique pathways for the under-expressed miR-mRNA target list included Wnt signaling and p53 signaling pathways. Interestingly, signaling pathways regulating pluripotency of stem cells (Supplementary Figure 2.S6) and pathways in cancer (Supplementary Figure 2.S7) were in the top 20 list for mRNA targets of both over- and under-expressed miRNAs. Finally, less overlap in top 20 terms for mRNA targets of over- or under-expressed miRNAs was evident for Reactome pathways, with signaling by NOTCH and signaling by BMP representing top 20 terms unique to mRNA targets of under-expressed miRNAs in 8-cell stage embryos. Common terms for both subsets include signaling by NGF, transcriptional regulation by TP53 and post-transcriptional silencing by small RNAs, among others.

Ontology analysis was also performed using the DAVID functional annotation tool with the complete predicted mRNA target list (for both over- and under-expressed miRNAs). Very similar results were obtained for GO biological processes and KEGG pathways, with significant terms including (clustered by semantic similarity) regulation of nitrogen compound metabolism,

Wnt signaling pathway, gene expression, regulation of cell communication, regulation of cell development, among others (Supplementary File 2.2; Appendix D, Supplementary Figure 2.S8).

Network analysis of differentially expressed
miRNAs and predicted mRNA targets

Although sequencing of mRNA was not possible for the samples obtained in the present study due to the nature of the protocol for isolation of sncRNA, a publicly-available RNAseq dataset is available for corresponding sample types, including bovine MII oocytes, 8-cell stage embryos and blastocyst stage embryos (GSE52415) [42]. In order to determine whether predicted mRNA targets were inversely regulated by their miRNA families, a Cytoscape network was created, where differentially-expressed miRNAs in 8-cell stage embryos are connected to their predicted mRNA targets. Figure 2.4 shows the resulting network colored according to relative expression with respect to oocytes, whereas Supplementary Figure 2.S9 depicts the same network colored by normalized abundance values. The overall number of predicted mRNA targets was limited using the TargetScan context score to reduce the complexity of the network diagram. The overarching observation of this network is that, for each miRNA-mRNA cluster, there is no clear pattern indicative of a negative correlation as would be expected given the role of miRNA to down-regulate mRNA abundance. For example, members of the miR-302 family are apparently very highly expressed and very abundant in 8-cell stage embryos with respect to oocytes, yet an overall pattern of mRNA suppression does not dominate this cluster. Similarly, miRNAs belonging to the miR-15/16/195/424-5p/497 are under-expressed relative to oocytes, yet an overall pattern of high mRNA expression is not evident.

Targeted qPCR analysis of select predicted mRNA
targets of differentially expressed miRNAs

Of the 29 mRNA targets analyzed by qPCR, 24 were differentially expressed in either 8-cell or blastocyst stage embryos compared to oocytes (Supplementary Figure 2.S10), yet the

expression profile for these genes did not entirely distinguish 8-cell stage embryos from oocytes (Figure 2.5A). Rather, more significant differences were evident for these mRNA targets in blastocyst stage embryos, which mostly segregated from other samples by unsupervised hierarchical clustering (Figure 2.5B). Notably, the observed changes in expression of predicted mRNA targets appeared negatively correlated with miRNA abundance for some of the genes examined, such as for *CDH2* and the miR-10 and miR-15/16/195/424-5p/497 families or *KLF4* and the miR-148/152 families (Figure 2.5C-G). However, for other genes examined, the expression pattern largely did not negatively correlate with their corresponding miRNAs according to the developmental stage. It was interesting to note that expression of some target mRNAs in blastocyst stages was strongly negatively correlated with miRNA expression at the 8-cell stage embryonic stage, such as observed for *MTMR3* and the miR-302 and miR-378/378c families (Figure 2.5H-I).

Expression of piwi-like RNAs in bovine oocytes,
8-cell stage embryos and blastocyst stages

In addition to miRNA analysis, the RNAseq data were mined to examine expression of other sncRNA classes, including piwi-interacting RNA (piRNA). Putative "piwi-like", or pilRNAs, were identified as sRNAs 24-32 nt in length that mapped to pilRNA-producing loci in the bovine genome. As a fraction of total reads, pilRNAs accounted for 73%, 14%, or 50% of sequence reads, on average, for oocytes, 8-cell stage embryos and blastocyst stage embryos, respectively (Table 2.1, Figure 2.6A-B). Secondary piRNA biogenesis occurs via a "ping-pong" amplification cycle, which can be detected as a 5' to 5' 10 nt overlap; this ping-pong signature was evident for sequences obtained from all sample types (Figure 2.6C). PiRNAs are also typified by a strong uracil bias at the 1' position (1'U) [47], which was also evident for all three tissues at rates of 81%, 69% and 83% for oocytes, 8-cell stage embryos and blastocyst stage embryos, respectively (Figure 2.6C).

Primary piRNAs are transcribed from dense loci located within specific clusters in the genome. Using weighted proTRAC cluster analysis, a total of 973 clusters were identified across the study samples. Oocytes were characterized by broad expression of pilRNA across the bovine genome, with highest expressing clusters located on chromosomes 3, 6, 8, 10, 13, 14,17, 23 and X (Figure 2.7A, Supplementary Figure 2.S11). In contrast, a major shift in pilRNA expression was evident in 8-cell stage embryos, with only 66 clusters identified and located randomly across chromosomes (Figure 2.7B, Supplementary Figure 2.S11). Expression of pilRNA clusters in blastocyst stages was somewhat like that of oocytes, with 473 clusters identified that were broadly expressed across the genome, though to less extent than in oocytes focused on chromosomes 3, 6, 8, 14, 17, 23 (Figure 2.7C, Supplementary Figure 2.S11). To compare cluster expression across samples, the clusters were joined on their genomic intervals. Only 22 pilRNA clusters were co-expressed in oocytes, 8-cell stage embryos and blastocyst stage embryos (Figure 2.7D). Far more clusters appeared unique to either oocytes (461 unique clusters) or blastocyst stage embryos (254 unique), whereas no clusters were uniquely expressed in 8-cell stage embryos.

Potential target transcripts of candidate pilRNAs were identified by mapping sequences to bovine TEs annotated by RepeatMasker (Figure 2.7E, Supplementary File 2.3; Appendix E). The long interspersed nuclear element (LINE) RTE-BovB and L1 and short interspersed nuclear element (SINE)-tRNA repeats were highly abundant in oocytes and blastocyst stage embryos, with pilRNAs mapping more frequently to the antisense direction. Long terminal repeat (LTR) ERVK retrotransposon and the TcMar-Tigger DNA transposon were also abundant in both oocytes and blastocyst stage embryos, with reads aligning predominantly to the antisense direction for both TEs. Alternatively, relatively few reads from 8-cell stage embryos mapped to TEs as compared to oocytes or blastocyst stage embryos, with the exception of more reads mapped to LSU-rRNA TE in the sense direction.

Expression of snoRNAs in bovine oocytes,
8-cell stage embryos and blastocyst stages

As mentioned previously, the most abundant sncRNA class for sequence reads ≥33 nt in length was snoRNA, and the relative abundance of snoRNA was markedly greater in 8-cell stage embryos compared to oocytes or blastocyst stage embryos (Figure 2.1C, Supplementary Figure 2.2SB). Moreover, principal components analysis of snoRNA expression data revealed very clear segregation of the expression profile for 8-cell stage embryos distinct from expression profiles for oocytes or blastocyst stage embryos, which overlapped (Figure 2.8A). This separation is also evident when considering expression on a per gene basis when sorting data by unsupervised, bi-directional hierarchical clustering by abundance (Figure 2.8B) or expression relative to oocytes (Figure 2.8C). Also of interest was the observation that most of the differentially expressed snoRNAs belonged to the C/D box subclass**.**

Discussion

In the developing embryo, the MZT is associated with marked changes in the control of gene expression that require coordination by multiple control mechanisms, including sncRNAs. This study is the first to employ a discovery-based, small RNA sequencing approach to determine the population of sncRNAs in bovine oocytes, 8-cell stage embryos and blastocyst stage embryos. Importantly, we identified major class changes in the population of sncRNAs associated with the 8-cell stage of embryonic development, coincident with the MZT in cattle. There were three major findings of this investigation. First, the relative abundance of most miRNAs was markedly higher in 8-cell stage embryos, the stage at which bovine embryos undergo the MZT, compared to oocytes before the MZT or blastocysts after the MZT. In contrast, the abundance of pilRNAs exhibited a mirror-image pattern, with markedly lower expression at the 8-cell stage relative to the oocyte or blastocyst stages. Finally, snoRNAs, particularly those belonging to the C/D box class that methylate rRNA, were expressed like miRNAs at much higher levels in 8-cell stage

embryos during the MZT compared to oocytes or blastocysts. Overall, these data reveal a strong dynamic shift in relative abundance of sncRNAs associated with the MZT in bovine oocytes and embryos, suggesting that these molecules may play important roles in the shift from maternal-to-zygotic control of gene expression. The patterns of these changes in particular classes of sncRNAs allow us to infer possible functions of these molecules in controlling the MZT.

Of particular note, the relative abundance of miRNAs was markedly higher in 8-cell stage embryos compared to oocytes or blastocyst stage embryos. This shift in miRNA population was largely associated with up-regulation of miRNAs predicted to target genes involved in the biological processes of cell development, cell division, Wnt signaling, pluripotency and endomembrane system organization, among others. While previous researchers identified increased expression of specific targeted miRNAs at the 8-cell stage, our results demonstrate the breadth of this up-regulation and identify specific miRNAs that are highly dynamic in relative abundance. The extreme changes in some miRNA families point to miRNAs that are likely functionally important in this transition, but that have not been previously examined in this role. Though few miRNA families were under-expressed in 8-cell stage bovine embryos, one such family was the miR-15/16/195/424-5p/497 family, which is particularly important given its high abundance with respect to other miRNAs. The miR-15/16/195/424-5p/497 family has a large number of predicted target mRNAs, which provided the potential for an enormous impact on the pathway and network analyses presented herein. The scope of these mRNA targets by this miRNA family suggest that these gene functions must be collectively suppressed in the oocyte and blastocyst-stage embryo, yet active during the MZT in cattle.

This study was also the first to examine changes in the population of potential pilRNAs during the MZT in cattle embryos. PilRNAs changed dramatically over the course of early preimplantation embryo development, as the population of potential pilRNAs was markedly reduced at the 8-cell stage as compared to oocytes, but appeared to partially rebound at the

blastocyst stage. The high level of abundance of piRNAs in blastocyst stage embryos was surprising as it contrasts findings in other species, although piRNA in bovine blastocysts has not been previously examined. Another surprising major class shift was identified in snoRNAs, notably in the C/D box subclass involved in methylation of ribosomal RNAs, which were highly expressed in 8-cell stage embryos. These snoRNAs could be of novel functional interest, as they have recently been demonstrated to participate in RNAi. However, very little is known about the role of snoRNAs in early mammalian embryogenesis. Overall, these data reveal a strong dynamic shift in relative abundance of sncRNAs associated with the developmentally critical MZT in bovine embryos, indicating that these molecules may play important roles in the shift from maternal-to-zygotic control of gene expression.

In this study, large scale changes in the relative abundance and expression profile of miRNAs were evident when comparing oocytes to 8-cell stage embryos, at which point the zygotic genome is activated in cattle. Most of the available information on expression and function of sncRNAs in embryo development has been obtained from the mouse model. However, key differences in the biogenesis and activity pathways for sncRNAs between rodents and other mammals can impact the population of small RNAs. For example, the mouse embryo expresses an oocyte-specific isoform of Dicer that preferentially binds endogenous siRNA rather than miRNA [48]. This oocyte-specific isoform efficiently produces endogenous siRNAs from long double-stranded (dsRNA) substrates and has higher cleavage activity than the other isoforms of Dicer. In contrast, in other mammalian cells and mouse somatic cells, Dicer primarily cleaves hairpin precursors, which includes the miRNA population. Thus, the population of sncRNA in mouse oocytes may be skewed in favor of production of siRNAs. Also, mouse oocytes lack PIWI3, a protein necessary for piRNA function in both human and bovine embryos [29].

To date, the data regarding the relative abundance of sncRNAs, or specific classes of sncRNAs, associated with the MZT in mammals have been inconsistent. In mice the MZT occurs

between the zygote and 2-cell developmental stages. Previously researchers determined that total miRNA content was markedly lower in 2-cell mouse embryos, post MZT, as compared to mature oocytes [52]. Recently, Yang and colleagues employed RNAseq in mouse oocytes and early embryos to examine the populations of sncRNA in mouse oocytes, zygotes, 2-cell, 4-cell, and 8-cell stage embryos [36]. In contrast to our results, they found that the relative increase in total miRNA population took place later in development at the 4- and 8-cell stages, concurrent with a marked decrease in endo-siRNA and piRNA. However, they did observe that some select miRNAs appeared to be more abundant in 2-cell mouse embryos. While these results differ from our observations, it is reasonable to expect such differences when comparing mouse to cattle pre-implantation embryos given the distinct biogenesis pathways for miRNA and endo-siRNA and activity of piRNAs. On the other hand, the biogenesis pathways for miRNAs in cattle and pig are similar, yet in porcine embryos, profiling of sncRNA revealed the lowest number of miRNA sequencing reads associated with the MZT [53]. To date, only one study has investigated expression of multiple miRNAs in cattle during the MZT in 8-cell stage embryos. Berg and Pfeffer profiled the expression of 98 mature miRNAs in bovine embryos using a stem-loop PCR assay [25]. They determined that these miRNAs were more abundant in 8-cell stage embryos compared to oocytes, and that the miRNA concentration per cell further increased from the 8-cell to blastocyst stage. However, the authors contended that most of the miRNA species were detected at amounts below the necessary threshold to functionally repress target mRNAs, with the notable exception of miR-320. In our RNAseq analysis, which measured the abundance of miRNAs in pools of bovine embryos rather than per cell, we determined that the marked increase in miRNA transcripts observed in 8-cell stage embryos was apparently not maintained through development to the blastocyst stage.

In this study, one of the most marked changes in miRNA expression observed was for the miR-302 family, which was evidently absent in oocytes and blastocyst stage embryos but very

abundant in 8-cell stage embryos. In human and mouse embryonic stem cells, miR-302 represses

multiple target genes that modulate the cell cycle, Wnt signaling, vesicular transport and

epigenetic regulation, among others [54, 55]. Deletion of miR-302a-d in mouse embryos resulted

in a fully penetrant late embryonic lethal phenotype due to neural tube defects [56], and knockout

of miR-302 in conjunction with miR-290 in mouse ESCs resulted in silencing naive pluripotency

[57]. Prediction of mRNA targets in cattle for miR-302 revealed a similar assortment of targeted

processes, including regulation of cell signaling cascades, covalent chromatin modification, and

cell fate determination. MiR-302 appears to have a key role in regulating the pluripotency of

embryonic stem cells. For example, Anokye-Danso and colleagues showed that lentiviral

transfection of miR302/367 triggered efficient reprogramming of mouse or human fibroblasts to a

pluripotent state without the need for exogenous transcription factors Oct4/Sox2/Klf4/Myc

commonly used to induce pluripotency [18]. Miyoshi and colleagues showed similar results

employing a direct transfection method with miR-200c and the miR-302 and miR-369 miRNA

families to induce pluripotency in mouse and human somatic cells [58]. In humans, the promoter

controlling expression of the miR-302 cluster is predominately active in embryonic stem cells as

opposed to differentiated cells [59, 60]. Thus, the high expression of miR-302 observed in bovine

8-cell stage embryos may be necessary to promote totipotency in early embryo development.

As with miR-302, we also observed a similar expression pattern for miR-196a, which

was very highly expressed in 8-cell stage embryos but not detected in oocytes or blastocyst stage

embryos. Tripurani and colleagues made similar observations using qPCR to measure miR-196 in

bovine oocytes and embryos, with maximal expression at 4- and 8-cell stage embryos [24]. They

also determined that miR-196a bound to the transcript for *NOBOX*, a maternal effect gene

preferentially expressed in oocytes that is required for folliculogenesis and early oocyte

development. However, *NOBOX* mRNA must undergo degradation prior to embryonic genome

activation [24]. Tripurani and colleagues determined that expression of NOBOX was greatest in

germinal vesicle oocytes and progressively decreased with oocyte maturation; similarly, NOBOX

expression was detectable in the developing zygote and progressively decreased with continued

development to blastocyst stage [61]. A similar pattern was also observed in this study, with PCR

results showing a progressive decline in *NOBOX* mRNA expression coincident with the observed

transient increase in miR-196a at the 8-cell stage. Of note, in both the Tripurani study and the

present work, expression of miR-196a did not perfectly correlate with the target *NOBOX* mRNA

expression after the 8-cell stage, as expression of both the miRNA and target declined in

blastocyst staged embryos. It is possible that other epigenetic mechanisms facilitate continued

silencing of *NOBOX* expression at these latter developmental stages and that RNAi via miR-196a

is no longer necessary.

Ontology and pathway analyses of predicted mRNA targets for miRNAs differentially

expressed in 8-cell stage embryos revealed enrichment of pathways expected to be critical for

early embryo development, including multiple pathways associated with tissue differentiation,

cell growth and embryo morphogenesis. Of interest, pathway analyses predicted miRNA

targeting of biological processes vital in early embryogenesis, including cell development, cell

division, Wnt signaling, and pluripotency. Wnt is involved in cell differentiation via pluripotency

maintenance during cell lineage specification, as well as the axis formation and cell migration

necessary later in pre-implantation embryo development [62]. The canonical Wnt pathway

regulates cell fate, whereas the non-canonical Wnt pathway regulates polarity, cell movements,

and asymmetric cell divisions. Importantly, Wnt pathway may be one of the strongest regulators

in supporting pluripotency in the early embryo, yet the timing of activation of Wnt signaling

appears critically important. Activation of Wnt signaling prior to blastocyst formation decreases

embryonic development to the blastocyst stage [63]. However, Tribulo and colleagues showed in

cattle embryos that only maternally-derived WNT signaling plays a role in pre-implantation

development [64]. Embryo-derived WNT signaling was dispensable for blastocyst formation, yet

was important for ICM proliferation [64]. Therefore, it is possible that the Wnt pathway must be inhibited until differentiation begins in the blastocyst embryo and that inhibition by miRNAs are necessary to prevent premature activation of Wnt pathways from maternally-deposited transcripts. Exactly which pathways Wnt signaling influences to impact development are still unknown. It is important to note that this is another early developmental pathway for which the observations in cattle embryos appear similar to those for the human, yet distinct from those in mouse [65].

Overall the pathway analysis suggests a general pattern of suppression of cell differentiation and promotion of pluripotency, particularly when considering those pathways that exclusively represent predicted mRNAs targets for repressed miRNAs in 8-cell stage embryos, or vice versa for induced miRNAs. However, interpretation of these pathways is complicated by a lack of one-to-one regulation of miRNAs and mRNAs. Multiple miRNAs can target a single mRNA molecule, and individual miRNAs have multiple mRNA targets. The complexity of these miRNA-mRNA interactions was exemplified by the highly intricate network depicting changes in expression of miRNAs and their predicted mRNA targets in bovine 8-cell stage embryos. Similarly, patterns of expression for miRNAs and predicted mRNA targets examined by qPCR were inconsistent. Expression for some miRNA-mRNA sets fit the expected inverse correlation (e.g., *WEE1* and miR-138 and miR-15/16/195/424-5p/497 families or *HDAC1* and miR-31). For many of the miRNA-mRNA sets, the expression pattern did not meet expectations (e.g., *DNMT1* and miR-148/152 family). Moreover, the impact of differential expression of any specific miRNA is dependent on the suite of miRNAs expressed and the functional context of the targeted gene. Dissection of the impact of any individual miRNA would require a focused, pathway-specific approach, which was not the objective of this study. That said, the ontology and network analyses reveal certain pathways for further investigation using targeted approaches, such as manipulation of expression of miR-302 or members of the miR-15/16/195/424-5p/497 family.

Furthermore, the miRNA-mRNA network does not take into account other regulatory factors that contribute to post-transcriptional gene silencing. For example, RNA binding proteins mediate binding of miRNAs to target mRNAs, allowing some mRNA targets to escape repression. One such RNA binding protein is *dead end homologue 1* (*DND1*), which binds to mRNAs containing an AU-rich element in the 3' untranslated region and blocks miRNA binding to this site. Of interest, DND1 is abundant in pig oocytes and early developing embryos, but is down regulated at the MZT; DND1 is functionally capable of binding pluripotency genes vital during early development. [16]. The role of RNA binding proteins during the MZT in modulating miRNA-mRNA binding is, as of yet, unknown. It is possible that miRNA targeting of transcripts for degradation may not be a simple one-to-one matching process, but rather a combination of degradation signals from miRNA and interactions with RNA binding proteins [66].

In the present study, few piRNA-producing loci and a low abundance of pilRNAs typified 8-cell bovine embryos, which raises the possibility that TEs targeted by these pilRNAs could be expressed during the MZT in cattle. Indeed, results of a prior study showed that transcripts derived from TEs are expressed in 8-cell bovine embryos during the MZT [31]. Also, of interest, the retrotransposon HERVK-HML-2 was expressed in 8-cell stage embryos through the blastocyst stage [67]. Alternatively, oocytes were characterized by high abundance of pilRNAs with dense piRNA clusters located on multiple chromosomes in the bovine genome, similar to the pattern previously observed [33]. Interestingly, and in contrast to other studies in mouse and pig [36, 53, 68], we observed high expression of pilRNAs in cattle embryos. Our piRNA profiling results for oocytes largely corresponds to observations made by Russell and colleagues [37], suggesting that these results are consistent across research groups and in different cattle populations. The pattern of pilRNA expression suggests that pilRNAs may be needed to block TE activation during epigenome programming events during continued development. The apparent increase in pilRNAs detected in bovine blastocyst stages in this study cannot be

explained by the higher genomic content in the tissue, as total reads per sample were consistently lower in all blastocyst stage embryo pools and the analysis was performed on read-normalized data. Another consideration is the difference in timing of the MZT, which occurs at earlier developmental stages in mouse (2-cell) and pig (4-cell) and may affect the overall population of sncRNAs accumulated in blastocyst stage embryos. Moreover, it is possible that the elevation of pilRNAs associated with the bovine blastocyst stage is unrelated to TE silencing, but rather may function to modulate mRNA expression as proposed by Russell and colleagues [33].

The discovery-based small RNA sequencing approach used in this study allowed us to investigate classes of sncRNAs not typically studied in mammalian embryos. Notably, many snoRNAs belonging to the C/D box class, those that methylate rRNAs, were expressed at much higher levels in 8-cell stage embryos during the MZT as compared to oocytes or blastocyst stages. This observed increase may reflect the function of snoRNAs in rRNA biogenesis and translation, which would be highly active during as the embryonic genome is activated. Aside from their traditional housekeeping role as guide RNAs for post-transcriptional modification of rRNAs, new evidence suggests that snoRNAs may be involved in cell fate determination and oncogenesis [69]. SnoRNAs can be processed to produce smaller sequences that appear to act as miRNAs by associating with the RISC complex to target mRNA [12, 70]. Through this processing, so called sno-miRNAs my then target transcripts that control processes involved in cell behavior and carcinogenesis. Given this potential new functional role of snoRNAs, the observed marked increase in their abundance at the 8-cell stage suggests a possible regulatory role of snoRNAs during the MZT in mammalian embryogenesis. This working hypothesis warrants further study.

The present study has some limitations with respect to sample collection and data analysis. To obtain sufficient amounts of small RNA for sequencing, pools of oocytes or embryos were used. This strategy has the advantage of reducing biological variability among samples, yet pooling can mask variability for individual oocytes or embryos. Even with pooling

of embryos, we did note fairly high variability among pooled blastocyst stage samples, which were collected based on consistency in morphology. That said, blastocyst stage embryos graded by morphology as high quality have been shown to express different transcriptomes [71], which may also impact the population of sncRNAs. Furthermore, the breed of the oocyte donor could also contribute to variability in the embryos, and the breed of oocyte donor and sperm used for IVF in this study may have been different from other studies examining gene expression. The protocol for isolating small RNA from these tissues did not allow for concurrent collection of mRNA for analysis by sequencing or other PCR methods. While that approach would have been ideal, we did examine expression of selected predicted mRNA target genes by qPCR in tissues collected by the same methods as for the small RNA sequencing analyses. Also, the miRNA-mRNA network was compiled using publicly available RNAseq data for oocyte and 8-cell bovine embryos. It is possible that the transcriptome of our blastocyst stages differs from that of the data used due to biological variability in the source material for IVF. With respect to analysis of sequencing data, the analytical approach involved normalization of sequence reads to the total reads per sample to account for differences in sample preparation and sequencing efficiency. While this approach does not allow for precise quantification of the number of sequence reads per cell or embryo, it does allow for comparison of relative contributions of specific sncRNAs classes for different sample types, such as comparing the relative amount of miRNA in oocytes versus blastocyst stage embryos. In addition, the annotation for the *Bos taurus* genome is not as thorough as for mouse or human, though it is considered excellent among domestic and agricultural species. Fewer known miRNA sequences for *Bos taurus* are available in miRBase (793 mature sequences in cattle as compared to 1,915 in mice and 2,588 in humans), and the vast majority of these miRNAs have not yet been functionally validated in the cattle model.

Overall, in every sncRNA population queried, all substantial changes in expression patterns occurred at the 8-cell stage embryos. While sncRNA plays vital roles in pluripotency

and differentiation, its function through the MZT in early mammalian development has not been elucidated. The miRNA population underwent substantial increases in expression at the 8-cell stage, suggesting a functional role for miRNA in the MZT. To our knowledge, our group is the first to examine the dynamics of sncRNA populations through the MZT in bovine embryos. Further research is needed to functionally validate roles for multiple sncRNA species through the MZT in early mammalian development.

In conclusion, this study defined key class transitions in populations of sncRNAs associated with the major MZT in bovine embryos, most notably for miRNAs, pilRNAs and snoRNAs. These regulatory molecules may prove critical for successful zygotic genome activation and embryo development in mammals. In future studies, the biological relevance and functions of pilRNAs and snoRNAs should be investigated in depth, as their roles in mammalian embryogenesis are poorly understood. Furthermore, additional work to define the function of differentially expressed miRNAs in modulating embryogenesis in mammals is needed, including a focus on the impact of these miRNAs maternally-derived transcript and activation of the embryonic genome. Functional validation could be accomplished using knock-in and knock-out approaches on embryos generated via IVF or somatic cell nuclear transfer to determine the impact of gain or loss of miRNA expression on preimplantation embryonic development.

References

1. Gilbert I, Macaulay A, Robert C. Oocyte developmental competence and embryo quality: distinction and new perspectives. Anim Repro 2015; 12:397-407.

2. Jukam D, Shariati SAM, Skotheim JM. Zygotic Genome Activation in Vertebrates. Dev Cell 2017; 42:316-332.

3. Tadros W, Lipshitz HD. The maternal-to-zygotic transition: a play in two acts. Development 2009; 136:3033-3042.

4. Graf A, Krebs S, Heininen-Brown M, Zakhartchenko V, Blum H, Wolf E. Genome activation in bovine embryos: review of the literature and new insights from RNA sequencing experiments. Anim Reprod Sci 2014; 149:46-58.

5. Hosseini SM, Dufort I, Caballero J, Moulavi F, Ghanaei HR, Sirard MA. Transcriptome profiling of bovine inner cell mass and trophectoderm derived from in vivo generated blastocysts. BMC Dev Biol 2015; 15:49.

6. Catalanotto C, Cogoni C, Zardo G. MicroRNA in Control of Gene Expression: An Overview of Nuclear Functions. Int J Mol Sci 2016; 1712.

7. Kim VN, Han J, Siomi MC. Biogenesis of small RNAs in animals. Nat Rev Mol Cell Biol 2009; 10:126-139.

8. McStay B, Grummt I. The epigenetics of rRNA genes: from molecular to chromosome biology. Annu Rev Cell Dev Biol 2008; 24:131-157.

9. Ihara M, Tseng H, Schultz RM. Expression of variant ribosomal RNA genes in mouse oocytes and preimplantation embryos. Biol Reprod 2011; 84:944-946.

10. Schorn AJ, Gutbrod MJ, LeBlanc C, Martienssen R. LTR-Retrotransposon Control by tRNA-Derived Small RNAs. Cell 2017; 170:61-71 e11.

11. Lee YS, Shibata Y, Malhotra A, Dutta A. A novel class of small RNAs: tRNA-derived RNA fragments (tRFs). Genes Dev 2009; 23:2639-2649.

12. Scott MS, Ono M. From snoRNA to miRNA: Dual function regulatory non-coding RNAs. Biochimie 2011; 93:1987-1992.

13. Matera AG, Terns RM, Terns MP. Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs. Nat Rev Mol Cell Biol 2007; 8:209-220.

14. Winter J, Jung S, Keller S, Gregory RI, Diederichs S. Many roads to maturity: microRNA biogenesis pathways and their regulation. Nat Cell Biol 2009; 11:228-234.

15. Fabian MR, Sonenberg N, Filipowicz W. Regulation of mRNA translation and stability by microRNAs. Annu Rev Biochem 2010; 79:351-379.

16. Rivera RM, Ross JW. Epigenetics in fertilization and preimplantation embryo development. Prog Biophys Mol Biol 2013; 113:423-432.

17. Giraldez AJ, Mishima Y, Rihel J, Grocock RJ, Van Dongen S, Inoue K, Enright AJ, Schier AF. Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. Science 2006; 312:75-79.

18. Anokye-Danso F, Trivedi CM, Juhr D, Gupta M, Cui Z, Tian Y, Zhang Y, Yang W, Gruber PJ, Epstein JA, Morrisey EE. Highly efficient miRNA-mediated reprogramming of mouse and human somatic cells to pluripotency. Cell Stem Cell 2011; 8:376-388.

19. Lin SL, Chang DC, Lin CH, Ying SY, Leu D, Wu DT. Regulation of somatic cell reprogramming through inducible mir-302 expression. Nucleic Acids Res 2011; 39:1054-1065.

20. Tesfaye D, Worku D, Rings F, Phatsara C, Tholen E, Schellander K, Hoelker M. Identification and expression profiling of microRNAs during bovine oocyte maturation using heterologous approach. Mol Reprod Dev 2009; 76:665-677.

21. Gilchrist GC, Tscherner A, Nalpathamkalam T, Merico D, LaMarre J. MicroRNA Expression during Bovine Oocyte Maturation and Fertilization. Int J Mol Sci 2016; 17:396.

22. Pasquariello R, Fernandez-Fuertes B, Strozzi F, Pizzi F, Mazza R, Lonergan P, Gandolfi F, Williams JL. Profiling bovine blastocyst microRNAs using deep sequencing. Reprod Fertil Dev 2017; 29:1545-1555.

23. Goossens K, Mestdagh P, Lefever S, Van Poucke M, Van Zeveren A, Van Soom A, Vandesompele J, Peelman L. Regulatory microRNA network identification in bovine blastocyst development. Stem Cells Dev 2013; 22:1907-1920.

24. Tripurani SK, Lee KB, Wee G, Smith GW, Yao J. MicroRNA-196a regulates bovine newborn ovary homeobox gene (NOBOX) expression during early embryogenesis. BMC Dev Biol 2011; 11:25.

25. Berg DK, Pfeffer PL. MicroRNA expression in bovine preimplantation embryos. Reprod Fertil Dev 2017; 30: 546-554

26. Watanabe T, Totoki Y, Toyoda A, Kaneda M, Kuramochi-Miyagawa S, Obata Y, Chiba H, Kohara Y, Kono T, Nakano T, Surani MA, Sakaki Y et al. Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. Nature 2008; 453:539-543.

27. Piatek MJ, Werner A. Endogenous siRNAs: regulators of internal affairs. Biochem Soc Trans 2014; 42:1174-1179.

28. Okamura K, Lai EC. Endogenous small interfering RNAs in animals. Nat Rev Mol Cell Biol 2008; 9:673-678.

29. Roovers EF, Rosenkranz D, Mahdipour M, Han CT, He N, Chuva de Sousa Lopes SM, van der Westerlaken LA, Zischler H, Butter F, Roelen BA, Ketting RF. Piwi proteins and piRNAs in mammalian oocytes and early embryos. Cell Rep 2015; 10:2069-2082.

30. Vagin VV, Sigova A, Li C, Seitz H, Gvozdev V, Zamore PD. A distinct small RNA pathway silences selfish genetic elements in the germline. Science 2006; 313:320-324.

31. Bui LC, Evsikov AV, Khan DR, Archilla C, Peynot N, Henaut A, Le Bourhis D, Vignon X, Renard JP, Duranthon V. Retrotransposon expression as a defining event of genome reprogramming in fertilized and cloned bovine embryos. Reproduction 2009; 138:289-299.

32. Gou LT, Dai P, Yang JH, Xue Y, Hu YP, Zhou Y, Kang JY, Wang X, Li H, Hua MM, Zhao S, Hu SD et al. Pachytene piRNAs instruct massive mRNA elimination during late spermiogenesis. Cell Res 2014; 24:680-700.

33. Russell S, Patel M, Gilchrist G, Stalker L, Gillis D, Rosenkranz D, LaMarre J. Bovine piRNA-like RNAs are associated with both transposable elements and mRNAs. Reproduction 2017; 153:305-318.

34. Simmet K, Zakhartchenko V, Philippou-Massier J, Blum H, Klymiuk N, Wolf E. OCT4/POU5F1 is required for NANOG expression in bovine blastocysts. Proc Natl Acad Sci USA 2018; 115:2770-2775.

35. Kropp J, Salih SM, Khatib H. Expression of microRNAs in bovine and human pre-implantation embryo culture media. Front Genet 2014; 5:91.

36. Yang Q, Lin J, Liu M, Li R, Tian B, Zhang X, Xu B, Liu M, Zhang X, Li Y, Shi H, Wu L. Highly sensitive sequencing reveals dynamic modifications and activities of small RNAs in mouse oocytes and early embryos. Sci Adv 2016; 2:e1501482.

37. Funston RN, Nauta WJ, Seidel GE, Jr. Culture of bovine embryos in buffalo rat liver cell-conditioned media or with leukemia inhibitory factor. J Anim Sci 1997; 75:1332-1336.

38. Krueger F. Trim Galore. In: Babraham Bioinformatics; 2016.

39. Kopylova E, Noe L, Touzet H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics 2012; 28:3211-3217.

40. Kozomara A, Griffiths-Jones S. miRBase: annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Res 2014; 42:D68-73.

41. Vejnar CE, Blum M, Zdobnov EM. miRmap web: Comprehensive microRNA target prediction online. Nucleic Acids Res 2013; 41:W165-168.

42. Graf A, Krebs S, Zakhartchenko V, Schwalb B, Blum H, Wolf E. Fine mapping of genome activation in bovine embryos by RNA sequencing. Proc Natl Acad Sci USA 2014; 111:4139-4144.

43. Agarwal V, Bell GW, Nam JW, Bartel DP. Predicting effective microRNA target sites in mammalian mRNAs. Elife 2015; 4.

44. Tripathi S, Pohl MO, Zhou Y, Rodriguez-Frandsen A, Wang G, Stein DA, Moulton HM, DeJesus P, Che J, Mulder LC, Yanguez E, Andenmatten D et al. Meta- and Orthogonal Integration of Influenza "OMICs" Data Defines a Role for UBR4 in Virus Budding. Cell Host Microbe 2015; 18:723-735.

45. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003; 13:2498-2504.

46. Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res 2009; 37:1-13.

47.   Gebert D, Hewel C, Rosenkranz D. unitas: the universal tool for annotation of small RNAs. BMC Genomics 2017; 18:644.

48.   Rosenkranz D, Han CT, Roovers EF, Zischler H, Ketting RF. Piwi proteins and piRNAs in mammalian oocytes and early embryos: From sample to sequence. Genom Data 2015; 5:309-313.

49.   Rosenkranz D, Zischler H. proTRAC--a software for probabilistic piRNA cluster detection, visualization and analysis. BMC Bioinformatics 2012; 13:5.

50.   Saeed AI, Sharov V, White J, Li J, Liang W, Bhagabati N, Braisted J, Klapa M, Currier T, Thiagarajan M, Sturn A, Snuffin M et al. TM4: a free, open-source system for microarray data management and analysis. Biotechniques 2003; 34:374-378.

51.   Metsalu T, Vilo J. ClustVis: a web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. Nucleic Acids Res 2015; 43:W566-570.

52.   Tang F, Kaneda M, O'Carroll D, Hajkova P, Barton SC, Sun YA, Lee C, Tarakhovsky A, Lao K, Surani MA. Maternal microRNAs are essential for mouse zygotic development. Genes Dev 2007; 21:644-648.

53.   Yang CX, Du ZQ, Wright EC, Rothschild MF, Prather RS, Ross JW. Small RNA profile of the cumulus-oocyte complex and early embryos in the pig. Biol Reprod 2012; 87:117.

54.   Subramanyam D, Lamouille S, Judson RL, Liu JY, Bucay N, Derynck R, Blelloch R. Multiple targets of miR-302 and miR-372 promote reprogramming of human fibroblasts to induced pluripotent stem cells. Nat Biotechnol 2011; 29:443-448.

55.   Li MA, He L. microRNAs as novel regulators of stem cell pluripotency and somatic cell reprogramming. Bioessays 2012; 34:670-680.

56.   Parchem RJ, Moore N, Fish JL, Parchem JG, Braga TT, Shenoy A, Oldham MC, Rubenstein JL, Schneider RA, Blelloch R. miR-302 Is Required for Timing of Neural Differentiation, Neural Tube Closure, and Embryonic Viability. Cell Rep 2015; 12:760-773.

57.   Gu KL, Zhang Q, Yan Y, Li TT, Duan FF, Hao J, Wang XW, Shi M, Wu DR, Guo WT, Wang Y. Pluripotency-associated miR-290/302 family of microRNAs promote the dismantling of naive pluripotency. Cell Res 2016; 26:350-366.

58.   Miyoshi N, Ishii H, Nagano H, Haraguchi N, Dewi DL, Kano Y, Nishikawa S, Tanemura M, Mimori K, Tanaka F, Saito T, Nishimura J et al. Reprogramming of mouse and human cells to pluripotency using mature microRNAs. Cell Stem Cell 2011; 8:633-638.

59.   Barroso-delJesus A, Romero-Lopez C, Lucena-Aguilar G, Melen GJ, Sanchez L, Ligero G, Berzal-Herranz A, Menendez P. Embryonic stem cell-specific miR302-367 cluster: human gene structure and functional characterization of its core promoter. Mol Cell Biol 2008; 28:6609-6619.

60. Kuo CH, Deng JH, Deng Q, Ying SY. A novel role of miR-302/367 in reprogramming. Biochem Biophys Res Commun 2012; 417:11-16.

61. Tripurani SK, Lee KB, Wang L, Wee G, Smith GW, Lee YS, Latham KE, Yao J. A novel functional role for the oocyte-specific transcription factor newborn ovary homeobox (NOBOX) during early embryonic development in cattle. Endocrinology 2011; 152:1013-1023.

62. Palini S, De Stefani S, Scala V, Dusi L, Bulletti C. Epigenetic regulatory mechanisms during preimplantation embryo development. Ann N Y Acad Sci 2011; 1221:54-60.

63. Denicol AC, Dobbs KB, McLean KM, Carambula SF, Loureiro B, Hansen PJ. Canonical WNT signaling regulates development of bovine embryos to the blastocyst stage. Sci Rep 2013; 3:1266.

64. Tribulo P, Leao B, Lehloenya KC, Mingoti GZ, Hansen PJ. Consequences of endogenous and exogenous WNT signaling for development of the preimplantation bovine embryo. Biol Reprod 2017; 96:1129-1141.

65. Tribulo P, Moss JI, Ozawa M, Jiang Z, Tian XC, Hansen PJ. WNT regulation of embryonic development likely involves pathways independent of nuclear CTNNB1. Reproduction 2017; 153:405-419.

66. Alonso CR. A complex 'mRNA degradation code' controls gene expression during animal development. Trends Genet 2012; 28:78-88.

67. Grow EJ, Flynn RA, Chavez SL, Bayless NL, Wossidlo M, Wesche DJ, Martin L, Ware CB, Blish CA, Chang HY, Pera RA, Wysocka J. Intrinsic retroviral reactivation in human preimplantation embryos and pluripotent cells. Nature 2015; 522:221-225.

68. Ohnishi Y, Totoki Y, Toyoda A, Watanabe T, Yamamoto Y, Tokunaga K, Sakaki Y, Sasaki H, Hohjoh H. Small RNA class transition from siRNA/piRNA to miRNA during pre-implantation mouse development. Nucleic Acids Res 2010; 38:5141-5151.

69. Williams GT, Farzaneh F. Are snoRNAs and snoRNA host genes new players in cancer? Nat Rev Cancer 2012; 12:84-88.

70. Falaleeva M, Stamm S. Processing of snoRNAs as a new source of regulatory non-coding RNAs: snoRNA fragments form a new class of functional RNAs. Bioessays 2013; 35:46-54.

71. Driver AM, Penagaricano F, Huang W, Ahmad KR, Hackbart KS, Wiltbank MC, Khatib H. RNA-Seq analysis uncovers transcriptomic variations between morphologically similar in vivo- and in vitro-derived bovine blastocysts. BMC Genomics 2012; 13:118.
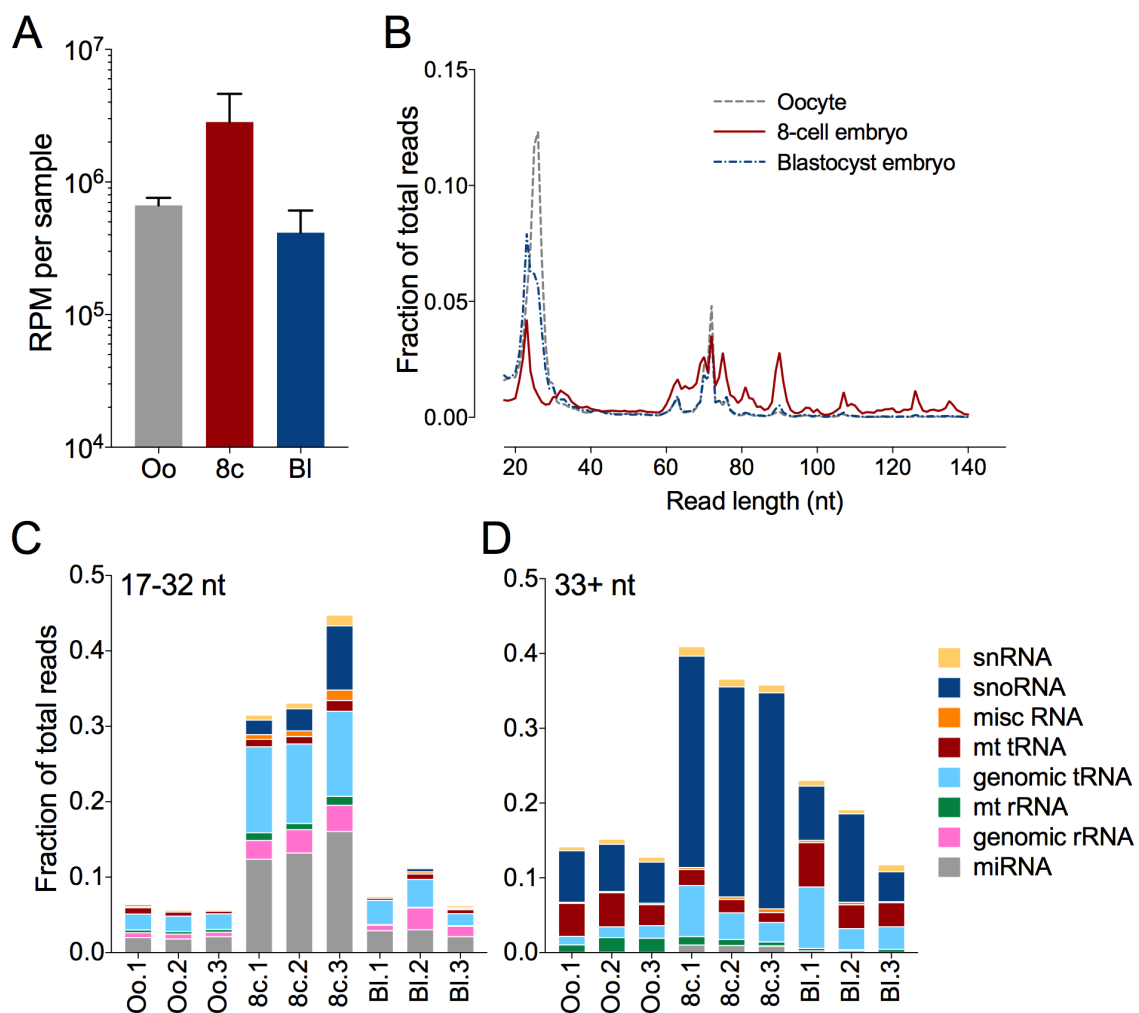
Figure 2.1. Distribution of reads length and sequence annotation. (A) Values are the mean ± SEM of the total sequence reads per sample after filtering out low quality reads and artifacts. (B) Read length distribution of sequenced samples after trimming to 17-93 nt and filtering for quality control. (C) Proportion of sncRNA annotated by class and read length 17-32 nt. (D) Proportion of sncRNA annotated reads annotated by class and read length ≥33 nt. Abbreviations: oocyte, Oo; 8-cell stage embryo, 8c; blastocyst stage embryo, Bl.
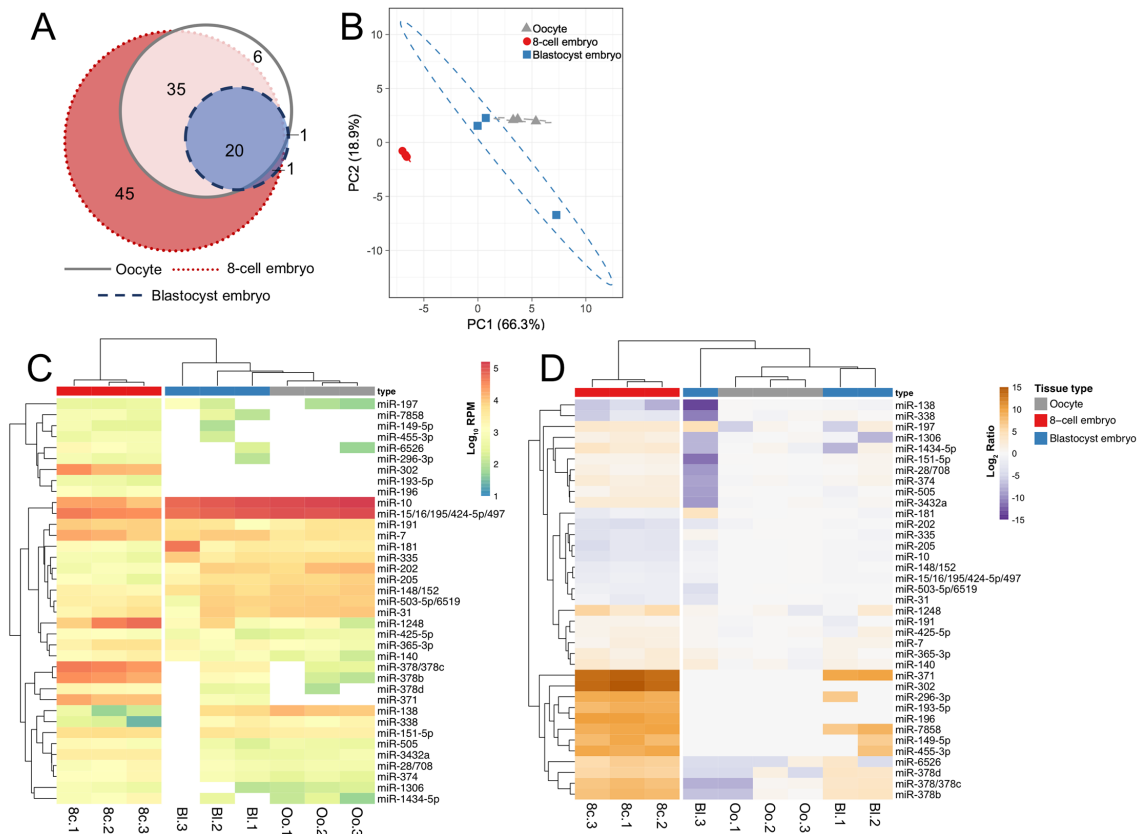
Figure 2.2. Expression of mapped miRNA in bovine oocytes, 8-cell embryos, and blastocyst stage embryos. (A) Proportional Venn diagram depicting the number miRNAs mapped to the *Bos taurus* miR database for each sample, including unique and overlapping IDs, filtered by total reads (>100 sequence reads per miR across samples) and consistency across biological replicates (>15 sequence reads for each replicate within a sample type). (B) Principal components analysis of miRNA family abundance data by sample type (oocyte, 8-cell stage embryo and blastocyst stage embryo). PC1 and PC2 are shown with 95% confidence intervals (dashed ellipses). (C-D) Unsupervised, bi-directional hierarchical cluster analyses (Euclidean distance method) of (C) abundance data ($\log_{10}$ RPM) or (D) ratio data ($\log_2$ ratio with respect to average abundance in oocytes) for miRNA families differentially expressed in any one sample ($|\log_2 R| > 1$ and FDR q-value <0.01).

Figure 2.3. (prior page) Significant ontology and pathway terms and network clustering associated with predicted mRNA targets of over- or under-expressed miRNAs in bovine 8-cell stage embryos. (A) Predicted mRNA targets (TargetScan total context++ score <-0.35) of differentially expressed miRNA families ($|\log_2 R|>1$ and FDR q-value <0.01 for 8-cell stage embryo compared to oocyte) were subject to enrichment analysis using Metascape (minimum overlap of 3, minimum enrichment 1.5 and FDR q <0.01). Only miR families annotated by TargetScan as conserved or broadly conserved were included in the Metascape analysis. Each heatmap depicts the top 20 enriched summary terms identified for Gene Ontology biological processes (top), KEGG pathways (middle) and Reactome pathways (bottom) with the left column depicting terms associated with mRNA targets of over-expressed miRs and the right column under-expressed miRs. Heatmaps are colored by $-\log_{10} q$-value; empty cells represent terms that were not among the top 20 enriched terms for the overall dataset (these terms may be significantly enriched, but not among the top 20). (B) Cytoscape network analysis of Gene Ontology biological process terms for predicted mRNA targets of differentially expressed miRNA families, including over- and under-represented miRs in 8-cell stage embryos as compared to oocytes and/or blastocyst stage embryos. Each term is represented by a circle node, for which the size is proportional to the enrichment score and the color represents the $\log_{10} q$-value. Terms with a similarity score > 0.3 are linked by an edge (the thickness of the edge represents the similarity score) using force-directed network layout and bundled edges for clarity. Each cluster is labeled by the term with the highest q-value in the cluster group. A dashed line was added to represent clusters associated with embryo development and various other development pathways. Complete results of the Metascape analyses are provided in Supplementary File 2.2; Appendix D.
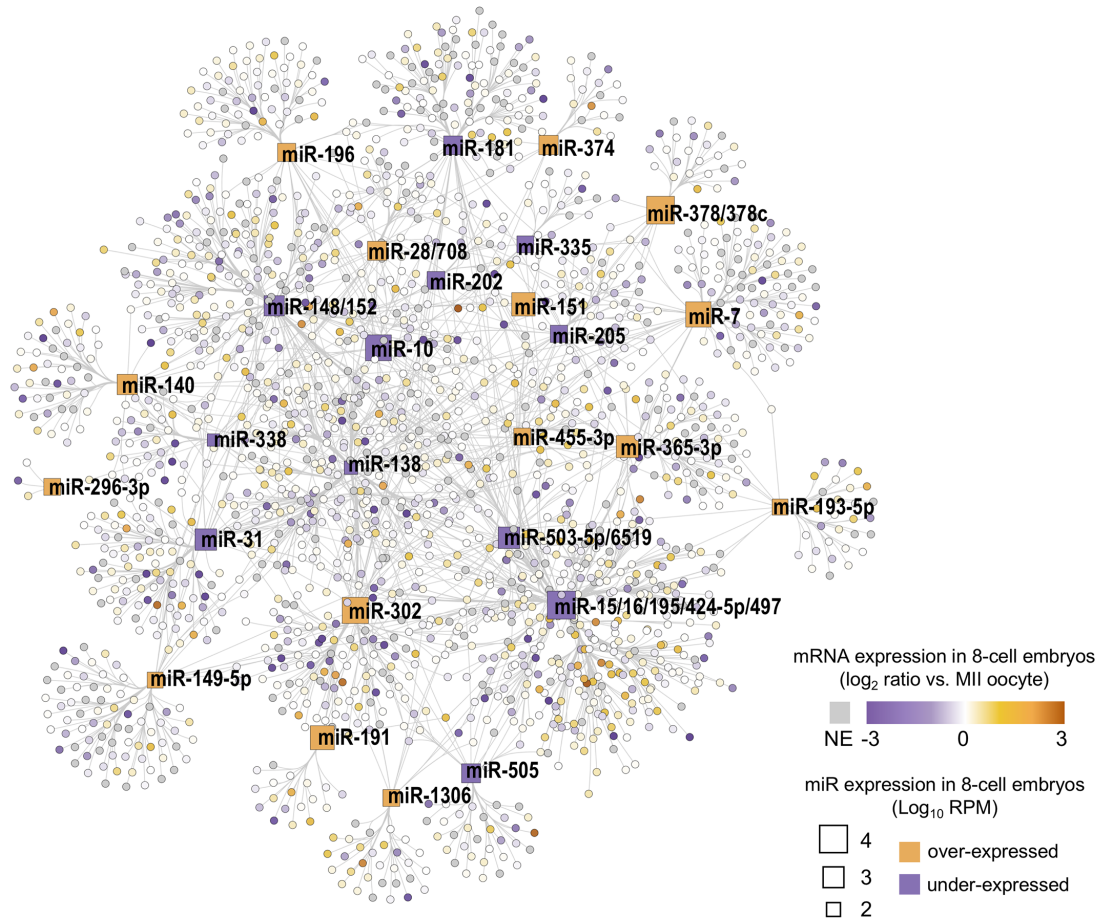
Figure 2.4. Network of differentially expressed miRNAs in 8-cell stage embryos and their predicted mRNA targets. A network of miRNA families differentially expressed in 8-cell stage embryos as compared to oocytes and their predicted mRNA targets (TargetScan total context++ score <-0.35) was created using Cytoscape with a force-directed layout. Only miR families annotated by TargetScan as conserved or broadly conserved were included in the network analysis. MiR families are represented by squares and colored according to their abundance in 8-cell stage embryos (values are log_{10} RPM), whereas target mRNAs are shown as circles and colored by their expression in 8-cell bovine IVF embryos using RNAseq data from the Gene Expression Omnibus dataset GSE52415 (calculated log_2 ratio with respect to MII oocytes). Nodes colored gray were not included in the GSE52415 data set, suggesting that they were not expressed. Edges between miR and mRNA nodes indicate that the connected mRNA is a predicted target of the miR in *Bos taurus*. Supplementary Figure 2.S9 shows the same network colored by normalized expression (log_{10} VST-normalized reads for genes and log_{10} RPM values for miRs).
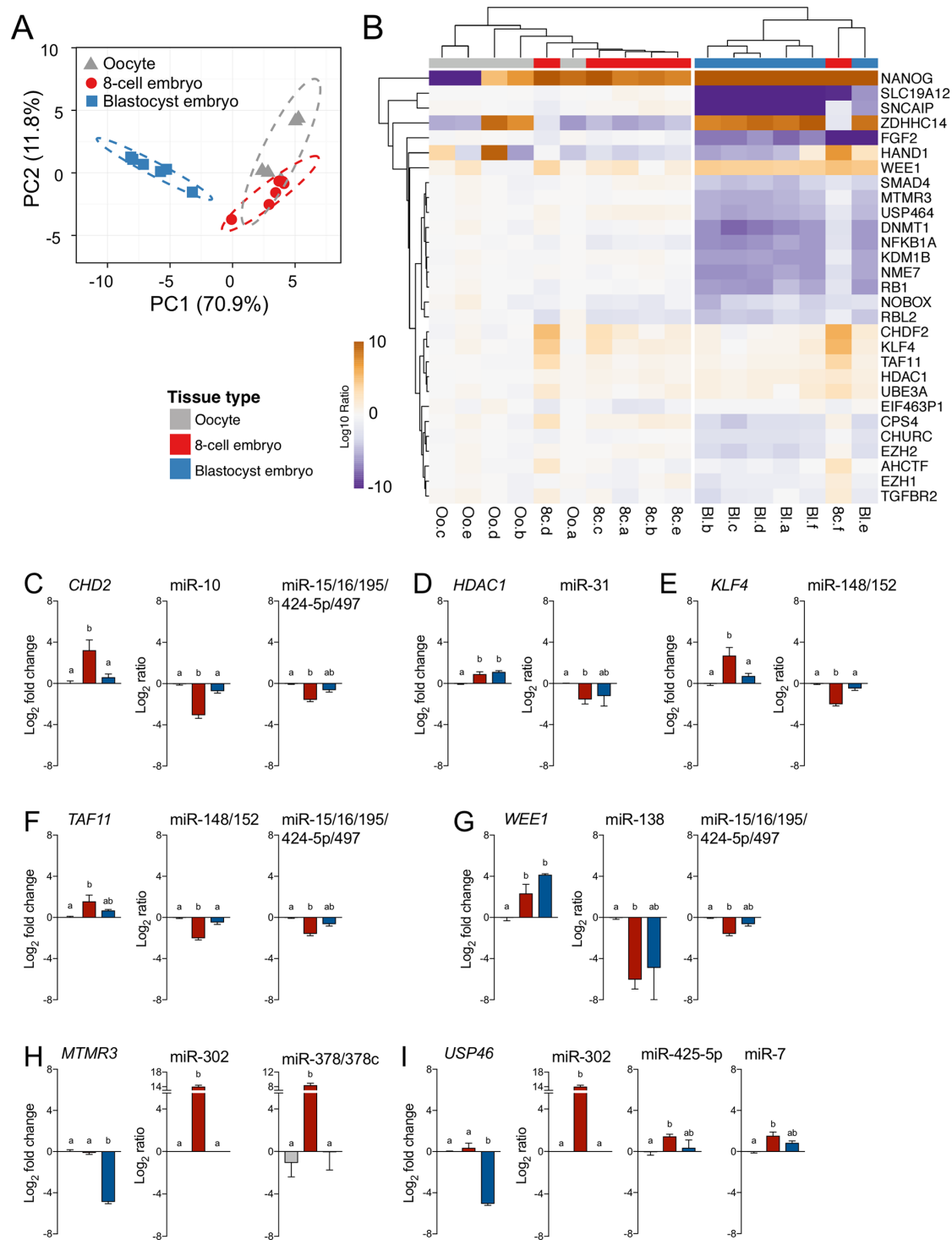
Figure 2.5. (prior page) Expression of selected mRNA targets in bovine oocytes, 8-cell stage embryos and blastocyst stage embryos. (A) Principal components analysis of mRNA abundance data by sample type (oocyte, 8-cell stage embryo and blastocyst stage embryo). PC1 and PC2 are shown with 95% confidence intervals (dashed ellipses colored by sample type). (B) Unsupervised, bi-directional hierarchical cluster analysis (Euclidean distance method) of select mRNA targets of differentially regulated miRNAs. Values shown are the $\log_2$ fold change in mRNA expression determined by qRT-PCR for each sample type normalized to expression in oocytes (i.e., oocyte/oocyte, 8-cell/oocyte and blastocyst stage/oocyte) (n=5 to 6). (C-G) Expression of selected mRNAs and their predicted targeting miRNAs that exemplify the expected inverse expression pattern by developmental stage. (H-I) Expression of selected mRNAs and their predicted targeting miRNAs that depict an expression pattern that does not match by developmental stage. Values for C-I are the mean $\log_2$ fold change vs oocytes + SEM (n=5 to 6) for mRNA expression data (qRT-PCR) or the mean $\log_2$ ratio + SEM (n=3) for miRNA expression (small RNAseq). Different letters indicate groups are significantly different from each other (FDR q<0.05). Abbreviations are Oo, oocyte; 8c, 8-cell embryo; Bl, blastocyst stage embryo.
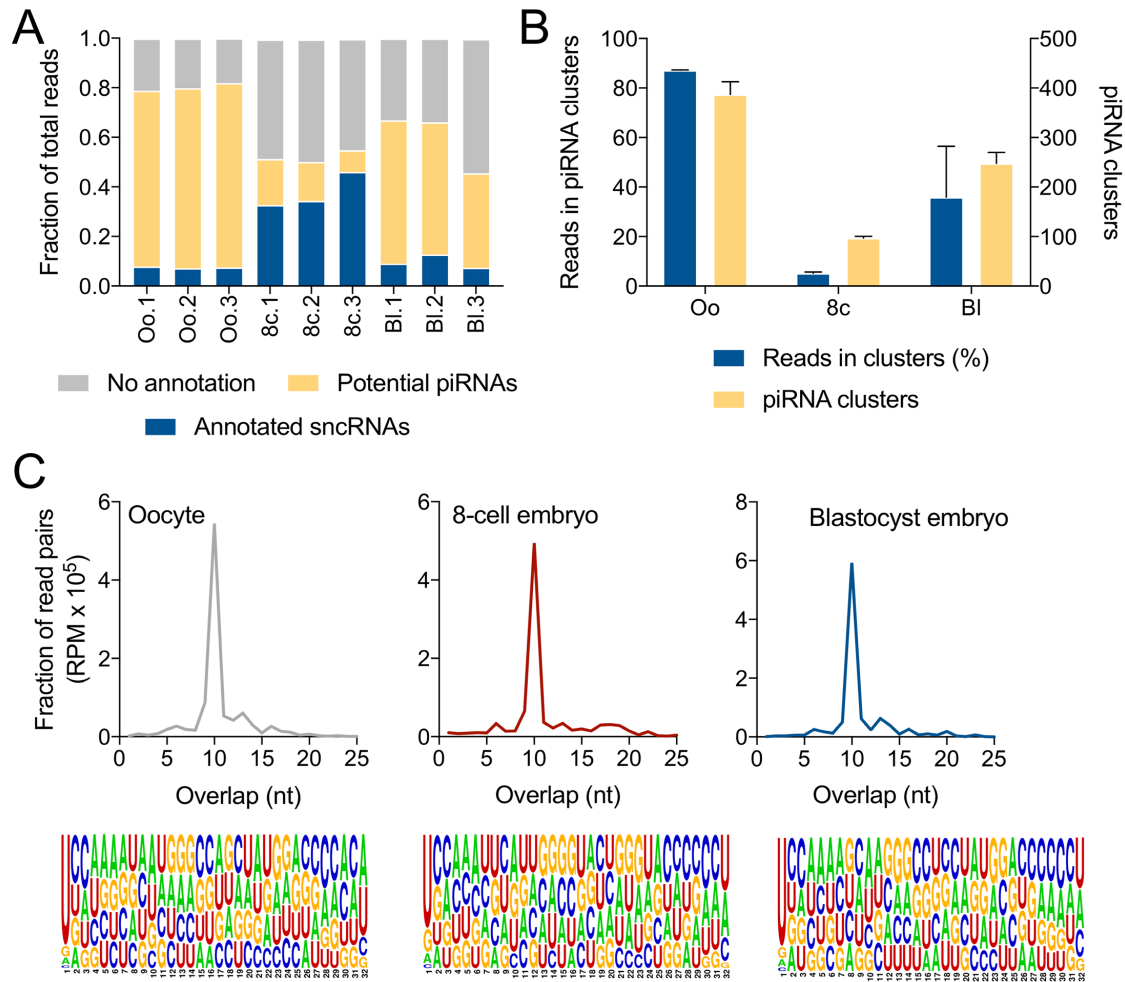
Figure 2.6. Annotation and identification of piRNA-like RNAs (pilRNAs) by ping-pong signature and 1U bias. (A) Fraction of sequence reads mapped as potential piRNAs compared to reads mapped to other sncRNA classes or not annotated by sample type. (B) Number of sequencing reads mapped to piRNA clusters (left axis) and number of piRNA clusters identified (right axis). (C) Identification of pilRNAs by ping pong signature and 1'U bias. Line graphs represent the number of pairs of reads with a 5' to 5' overlap in the datasets. The peak at 10 nt is characteristic of piRNAs produced by the ping pong cycle. The sequence logos below represent nucleotide biases at each pilRNA position for reads 32 nt in length (base position number below each logo). Abbreviations are Oo, oocyte; 8c, 8-cell stage embryo; Bl, blastocyst stage embryo.
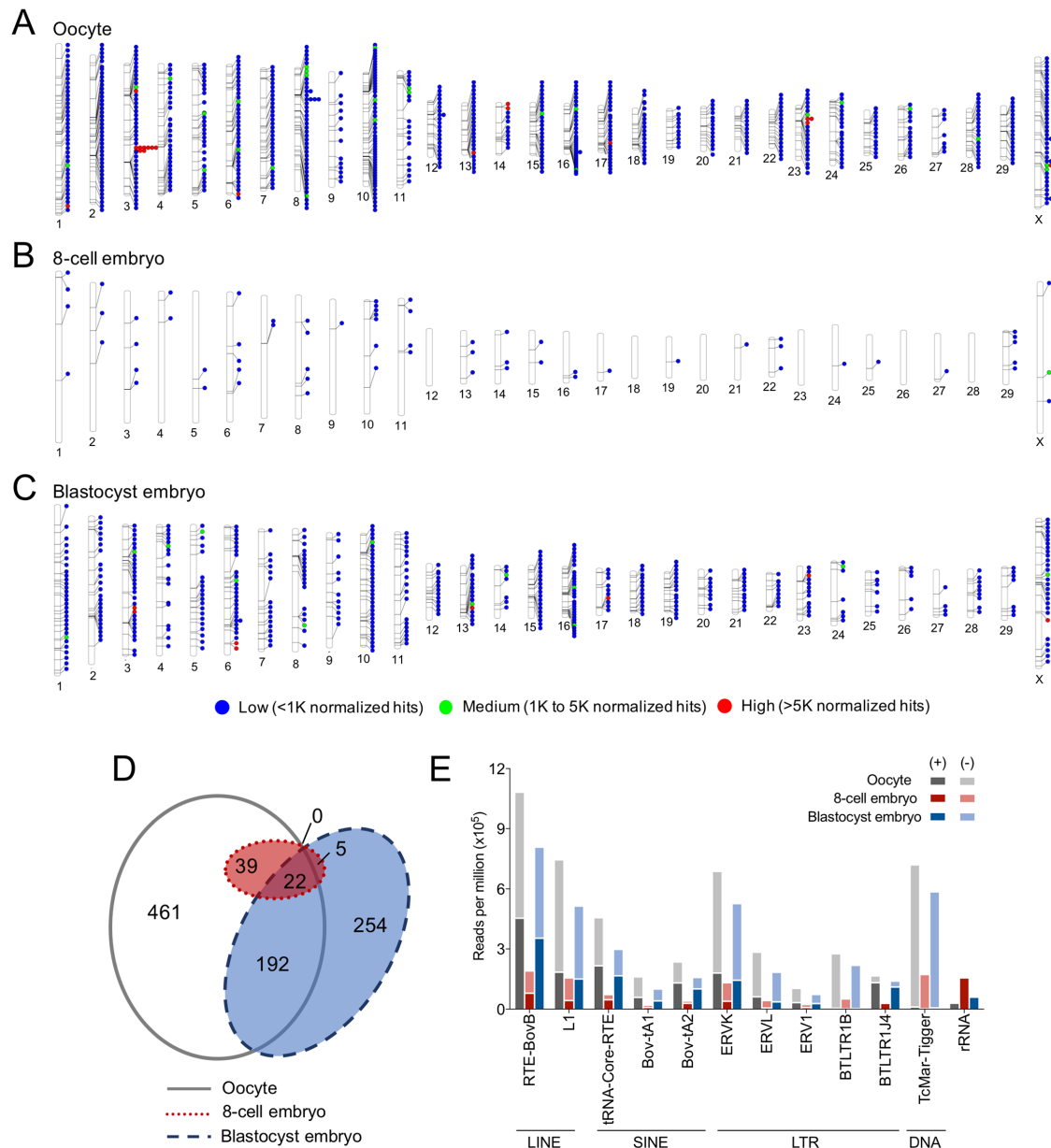
Figure 2.7. Expression of pilRNAs across chromosomes by sample type. (A-C) Chromosome diagrams for the *Bos taurus* genome were created using PhenoGram with each pilRNA cluster location marked by a circle colored according to expression level. (D) Proportional Venn diagram depicting number of predicted piRNA clusters by sample type. (E) Mapping of piRNAs to TE sites in the *Bos taurus* genome. Values shown are the reads per million for the most common TE sites for putative pilRNAs on the sense (+) or antisense (-) DNA strand. Abbreviations are: long interspersed nuclear elements, LINE; short interspersed nuclear elements, SINE; long terminal repeat, LTR; DNA transposon, DNA.
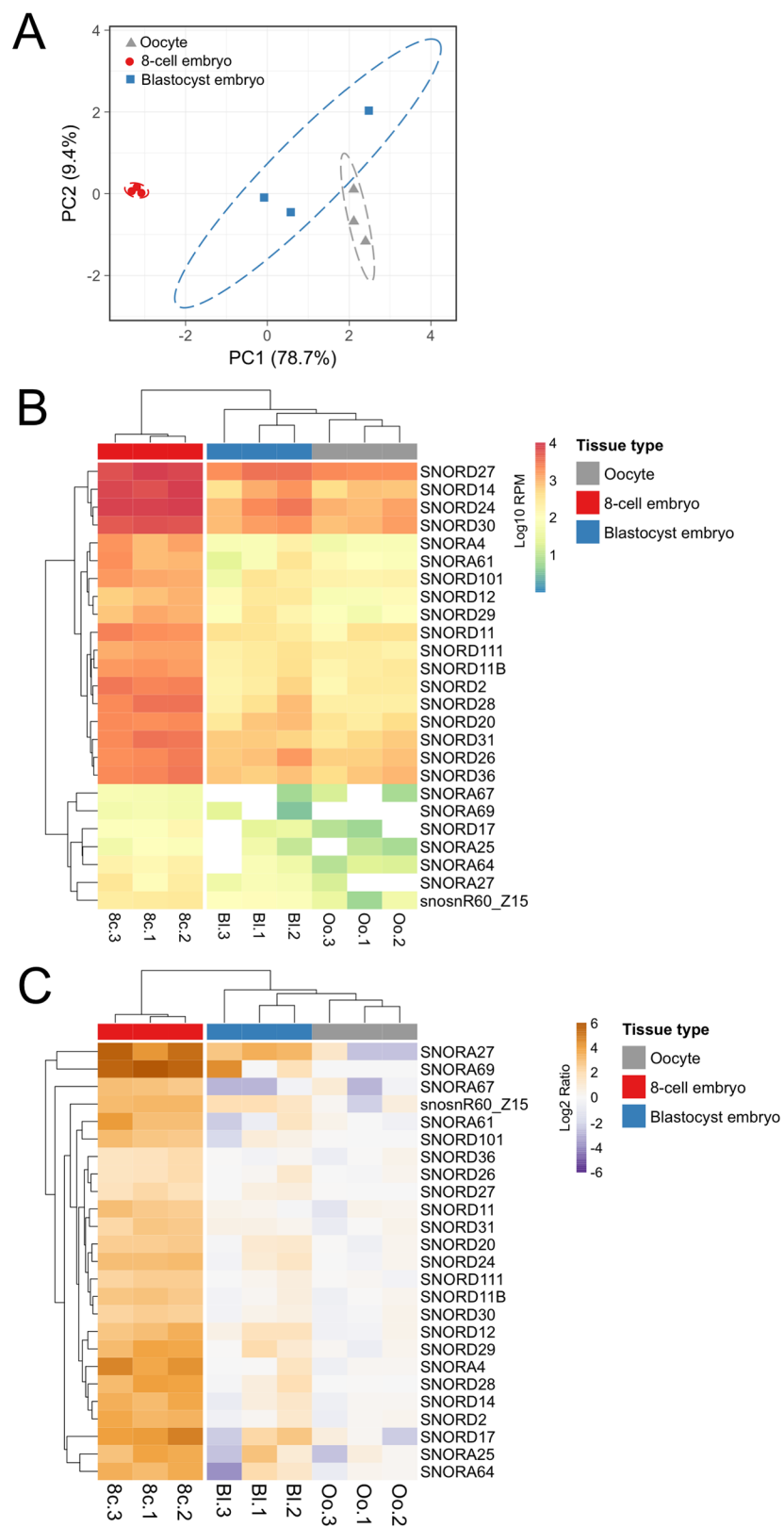
Figure 2.8. Expression of snoRNAs in bovine oocytes, 8-cell stage embryos and blastocyst stage embryos. (A) Principal components analysis of snoRNA abundance data by sample type (oocyte, 8-cell stage embryo and blastocyst stage embryo). PC1 and PC2 are shown with 95% confidence intervals (dashed ellipses colored by sample type). Abbreviations are Oo, oocyte; 8c, 8-cell embryo; Bl, blastocyst stage embryo. (B-C) Unsupervised, bi-directional hierarchical cluster analyses (Euclidean distance method) of (B) abundance data ($\log_{10}$ RPM) or (C) ratio data ($\log_2$ ratio with respect to average abundance in oocytes) for snoRNAs differentially expressed in any one sample ($|\log_2 R| > 1$ and FDR q-value < 0.01).

Table 2.1. Classification of sncRNA sequences [a]

sncRNA class	Sequence reads 17-32 nt			Sequence reads ≥33 nt		
	Oocyte	8-cell stage embryo	Blastocyst stage embryo	Oocyte	8-cell stage embryo	Blastocyst stage embryo
miRNA	2.04±0.001	13.95±0.011	2.74±0.003	0.09±0.000	1.00±0.000	0.23±0.001
genomic rRNA	0.62±0.000	3.01±0.003	1.67±0.006	0.00±0.000	0.01±0.000	0.00±0.000
mt rRNA	0.36±0.000	1.01±0.001	0.05±0.000	1.63±0.003	0.83±0.002	0.31±0.001
genomic tRNA	2.07±0.000	11.08±0.003	2.87±0.006	1.42±0.002	4.34±0.013	4.67±0.018
mt tRNA	0.58±0.001	1.14±0.001	0.50±0.001	3.93±0.006	1.75±0.002	4.11±0.009
misc RNA	0.19±0.000	0.92±0.002	0.21±0.000	0.17±0.000	0.37±0.001	0.26±0.001
snoRNA	0.12±0.000	4.46±0.021	0.21±0.001	6.24±0.004	28.41±0.002	7.71±0.023
snRNA	0.15±0.000	0.95±0.002	0.20±0.001	0.62±0.001	1.12±0.001	0.74±0.001
potential piRNA	72.68±0.010	14.40±0.029	49.77±0.060	3.10±0.008	0.52±0.001	3.36±0.009

[a] Values are the mean fraction (%) of total reads ± SEM (n=3 pools of 40 oocytes or embryos).

Table 2.2. Differentially expressed bovine miRNA families.

Cow miR family[*]	Seed sequence	Conservation[†]	Reads per million (average ± SEM)[‡]		
			Oocyte	8-cell stage embryo	Blastocyst stage embryo
miR-10	ACCCUGU	broadly conserved	5.13±0.07[a]	4.21±0.17[ab]	4.92±0.11[ab]
miR-1248	CCUUCUU	other	2.85±0.56[a]	4.44±0.44[ab]	3.33±0.52[ab]
miR-1306	CACCUCC	conserved	2.24±0.11[a]	2.77±0.14[ab]	1.35±0.60[ab]
miR-138	GCUGGUG	broadly conserved	4.03±0.09[a]	2.21±0.48[b]	2.89±1.64[ab]
miR-140	ACCACAG	broadly conserved	2.48±0.29[a]	3.31±0.16[ab]	3.00±0.22[ab]
miR-1434-5p	UACAUGA	other	2.10±0.35[a]	3.08±0.19[ab]	1.52±0.90[ab]
miR-148/152	CAGUGCA	broadly conserved	3.93±0.06[a]	3.32±0.09[ab]	3.79±0.11[ab]
miR-149-5p	CUGGCUC	conserved	0.00±0.00[a]	2.59±0.26[ab]	1.32±0.55[ab]
miR-15/16/195/424-5p/497	AGCAGCA	broadly conserved	5.02±0.05[a]	4.54±0.09[ab]	4.83±0.10[ab]
miR-151-5p	CGAGGAG	conserved	3.43±0.02[a]	3.78±0.03[ab]	2.75±1.52[ab]
miR-181	ACAUUCA	broadly conserved	3.48±0.05[a]	3.03±0.19[b]	3.84±0.75[ab]
miR-191	AACGGAA	broadly conserved	3.60±0.14[a]	3.92±0.05[ab]	3.45±0.37[ab]
miR-193-5p	GGGUCUU	broadly conserved	0.00±0.00[a]	2.60±0.18[b]	0.00±0.00[a]
miR-196	AGGUAGU	broadly conserved	0.00±0.00[a]	2.97±0.14[b]	0.00±0.00[a]
miR-197	UCACCAC	other	1.57±0.50[a]	2.57±0.07[b]	2.10±1.03[ab]
miR-202	UCCUAUG	broadly conserved	4.01±0.18[a]	2.89±0.13[ab]	3.68±0.40[ab]
miR-205	CCUUCAU	broadly conserved	3.87±0.06[a]	2.83±0.29[ab]	3.60±0.22[ab]
miR-28/708	AGGAGCU	conserved	2.79±0.06[a]	3.12±0.16[ab]	2.27±1.10[ab]
miR-296-3p	AGGGUUG	conserved	0.00±0.00[a]	2.74±0.12[ab]	1.35±0.60[ab]
miR-302	AGUGCUU	broadly conserved	0.00±0.00[a]	4.23±0.20[b]	0.00±0.00[a]
miR-31	GGCAAGA	broadly conserved	3.99±0.02[a]	3.52±0.23[ab]	3.62±0.51[ab]
miR-335	CAAGAGC	conserved	3.74±0.05[a]	2.75±0.15[b]	3.74±0.23[ab]

miR-338	CCAGCAU	broadly conserved	3.20 ± 0.17^a	2.07 ± 0.54^{ab}	2.57 ± 1.36^{ab}
miR-3432a	GCGGGAU	other	2.73 ± 0.13^a	3.49 ± 0.01^{ab}	2.16 ± 1.02^{ab}
miR-365-3p	AAUGCCC	broadly conserved	2.95 ± 0.16^a	3.61 ± 0.20^{ab}	3.21 ± 0.20^{ab}
miR-371	AGUGCCG	other	0.00 ± 0.00^a	4.11 ± 0.13^{ab}	2.25 ± 1.09^{ab}
miR-374	UAUAAUA	conserved	2.53 ± 0.17^a	3.15 ± 0.12^{ab}	2.21 ± 1.06^{ab}
miR-378/378c	CUGGACU	conserved	1.98 ± 0.85^a	4.52 ± 0.14^{ab}	2.59 ± 1.38^{ab}
miR-378b	CUUGACU	other	1.76 ± 0.67^a	4.36 ± 0.15^{ab}	2.46 ± 1.26^{ab}
miR-378d	UGGACUU	other	1.32 ± 0.55^a	3.25 ± 0.07^{ab}	2.07 ± 0.92^{ab}
miR-425-5p	UGACACG	other	2.61 ± 0.16^a	3.07 ± 0.12^{ab}	2.73 ± 0.40^{ab}
miR-455-3p	CAGUCCA	broadly conserved	0.00 ± 0.00^a	2.78 ± 0.14^{ab}	1.42 ± 0.73^{ab}
miR-503-5p/6519	AGCAGCG	conserved	3.95 ± 0.01^a	3.50 ± 0.08^{ab}	3.52 ± 0.68^{ab}
miR-505	GUCAACA	conserved	2.65 ± 0.08^a	3.13 ± 0.19^{ab}	2.02 ± 0.91^{ab}
miR-6526	CCUGUGC	other	1.25 ± 0.43^a	3.08 ± 0.19^{ab}	1.45 ± 0.77^{ab}
miR-7	GGAAGAC	broadly conserved	3.64 ± 0.08^a	4.11 ± 0.19^{ab}	3.90 ± 0.11^{ab}
miR-7858	CGCAAUU	other	0.00 ± 0.00^a	2.81 ± 0.16^{ab}	1.86 ± 0.79^{ab}

* miRNA IDs were grouped by cow miRNA family according to miRBase database (release 21).

[†]Conservation annotations are described as follows: broadly conserved, conserved across most vertebrates, usually to zebrafish; conserved, conserved across most mammals, but usually not beyond placental mammals; other, poorly conserved or other annotations in miRBase.

[‡] Values shown are the average $\log_{10}$ ($y\pm1$) transformed RPM ± SEM (n=3 pools of 40 oocytes or embryos). A value of 0 indicates that no sequences were mapped for the corresponding miR family in any of the replicate samples. Different letters indicate a significant difference according to pair-wise post-hoc Student's t-tests. One-way ANOVA was performed using log10 ($y\pm1$) transformed RPM data, followed by post-hoc pairwise tests between each tissue type. False discovery rate (FDR) q values were determined using a single family for all raw p values from all post-hoc comparisons and the two-stage step-up method of Benjamini, Krieger and Yekutieli (GraphPad Prism) with a desired Q=5%. Sample read data for all mapped miRNAs and complete statistical results are provided in Supplementary File 2.1; Appendix C.
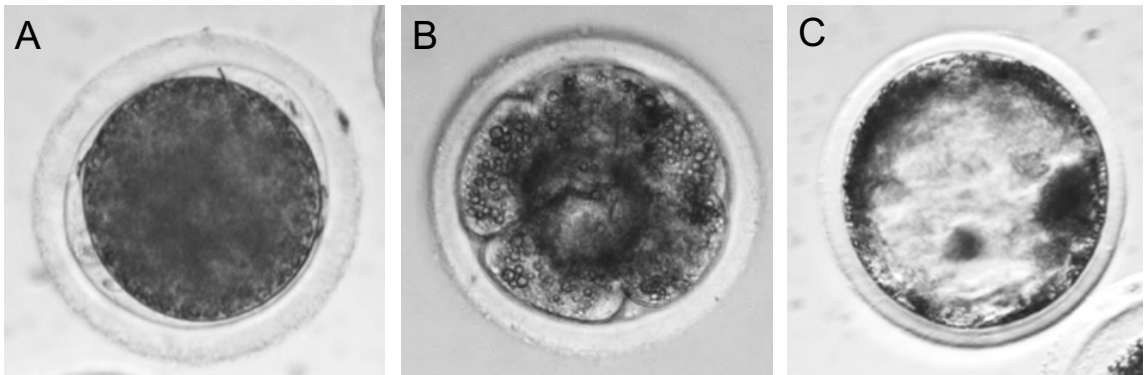
Supplementary Figure 2.S1. Representative microscopic images of bovine MII oocytes, 8-cell stage embryos, and blastocyst stage embryos from cattle. (A) Typical MII oocyte; (B) representative 8-cell stage embryo obtained about 3 days post in vitro fertilization; and (C) representative blastocyst stage embryos obtained approximately 8 days post in vitro fertilization. Images were acquired via light microscopy with 200× magnification.
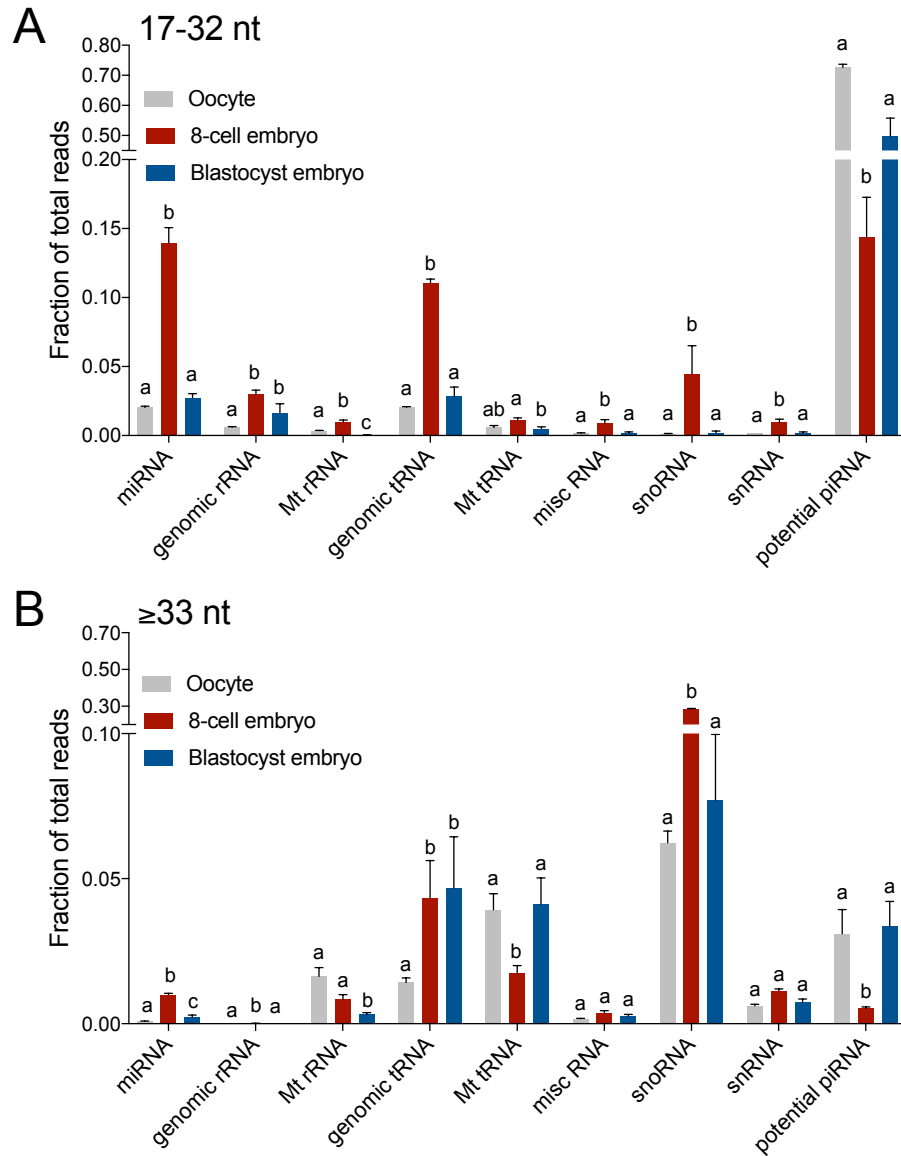
Supplementary Figure 2.S2. Fraction of total reads for sncRNAs classes 1180 for read lengths 17 to 32 nt (A) or ≥33 nt (B). Values shown are average fraction of total reads + SEM ($n = 3$ pools of oocytes or embryos). For each sncRNAs class, different letters above bars indicate that the bars are different, defined as $p < 0.05$ as determined by two-way ANOVA with factors sncRNAs class and tissue type with Tukey post hoc test for multiple comparisons. Values were $\log_{10}$ transformed to equalize variance prior to statistical analyses.
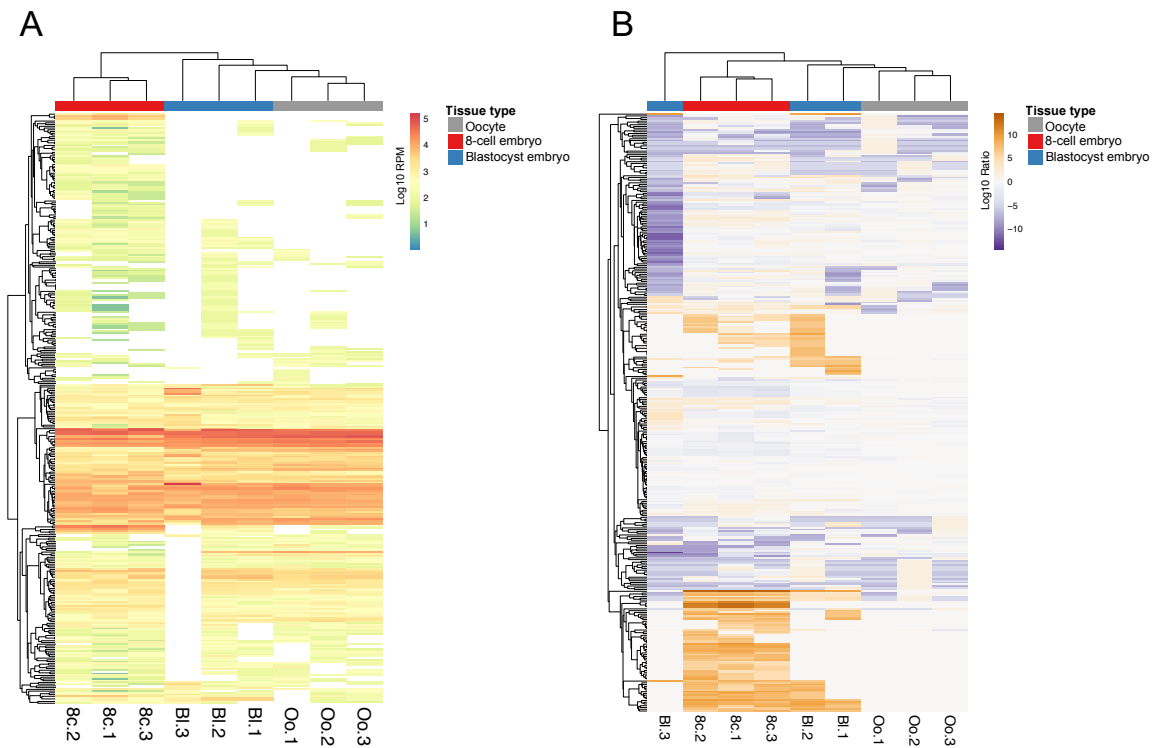
Supplementary Figure 2.S3. Expression of miRNA in bovine oocytes, 8-cell embryos, and blastocyst embryos. Unsupervised, bidirectional hierarchical cluster analyses (Euclidean distance method) of (A) abundance data (log$_{10}$ RPM) or (B) ratio data (log$_2$ ratio with respect to average abundance in oocytes) for all miRNA families detected by small RNA sequencing.
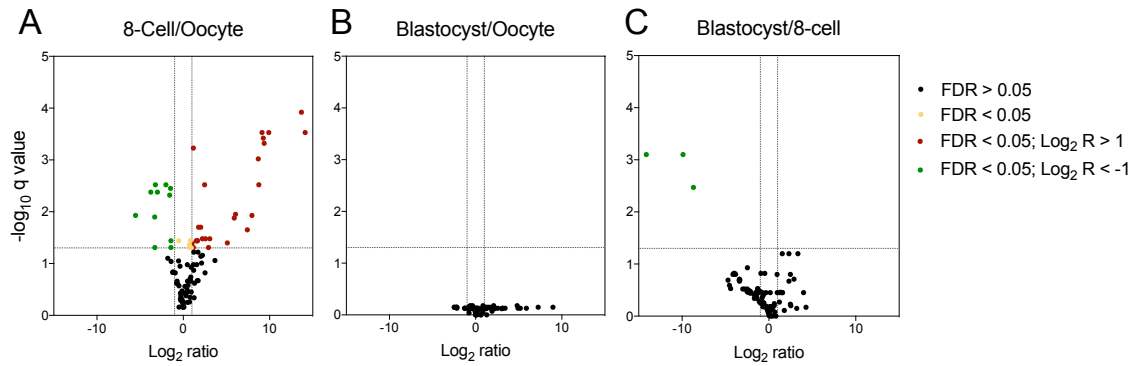
Supplementary Figure 2.S4. Volcano plot depicting changes in miRNA abundance. Values shown are the $\log_2$ ratio of expression in 8-cell embryos vs. oocytes (A), blastocysts embryos vs. oocytes (B), or blastocysts embryos vs. 8-cell embryos (C) plotted against the $-\log_{10}$ FDR q-value. A significant difference in expression was defined as $|\log_2 R| > 1$ and FDR q-value <0.01.
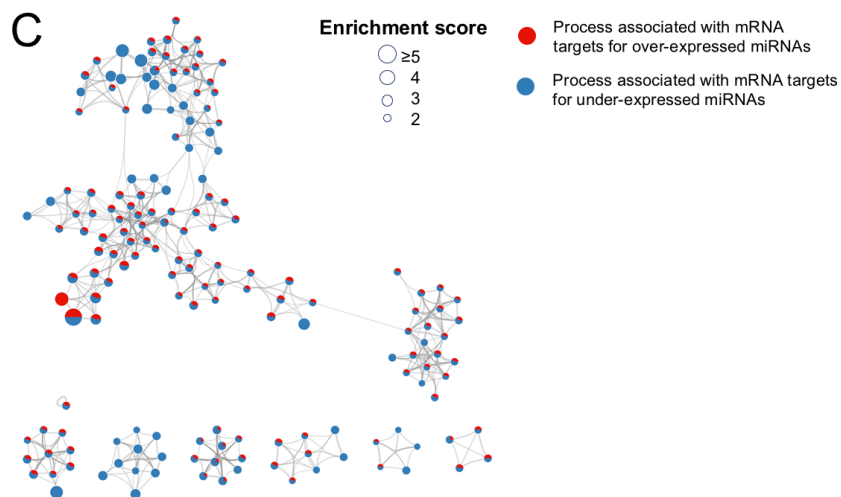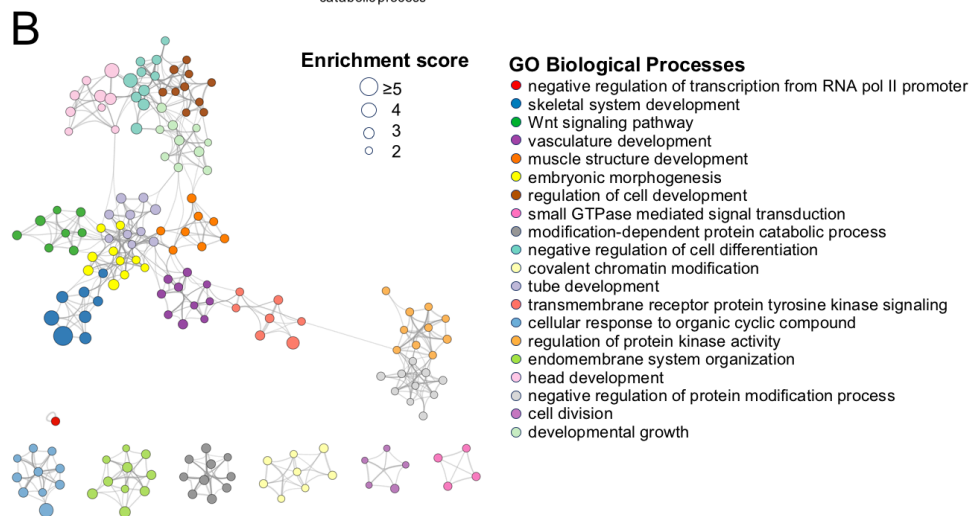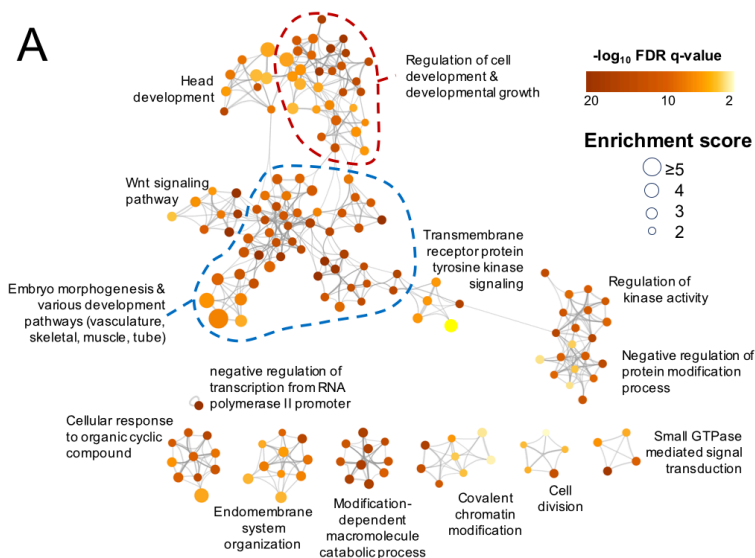
Supplementary Figure 2.S5. (prior page) Significant ontology and pathway terms and network clustering associated with predicted mRNA targets of over- or underexpressed miRNAs in bovine 8-cell embryos. Predicted mRNA targets (TargetScan total context++ score <−0.35) of differentially expressed miRNA families ($|\log_2 |R||>1$ and FDR q-value <0.01 for 8-cell embryo compared to oocyte or blastocyst embryo) were subject to enrichment analysis using Metascape (minimum overlap of 3, minimum enrichment 1.5 and FDR $q<0.01$). Only miR families annotated by TargetScan as conserved or broadly conserved were included in the Metascape analysis. Cytoscape network analysis of Gene Ontology biological process terms for predicted mRNA targets of differentially expressed miRNA families, including over- and underrepresented miRs in 8-cell embryos as compared to oocytes and/or blastocyst embryos. Each term is represented by a circle node, for which the size is proportional to the enrichment score and the color represents the $\log_{10}$ q-value. Terms with a similarity score >0.3 are linked by an edge (the thickness of the edge represents the similarity score) using force-directed network layout and bundled edges for clarity. Each cluster is labeled by the term with the highest q-value in the cluster group. A dashed line was added to represent clusters associated with embryo development and various other development pathways. Nodes within the networks are colored by the FDR q-value (A), by the associated GO biological pathway term for the cluster (B), or by the proportion of IDs in mRNA target lists for either over- or underexpressed miRs (C).
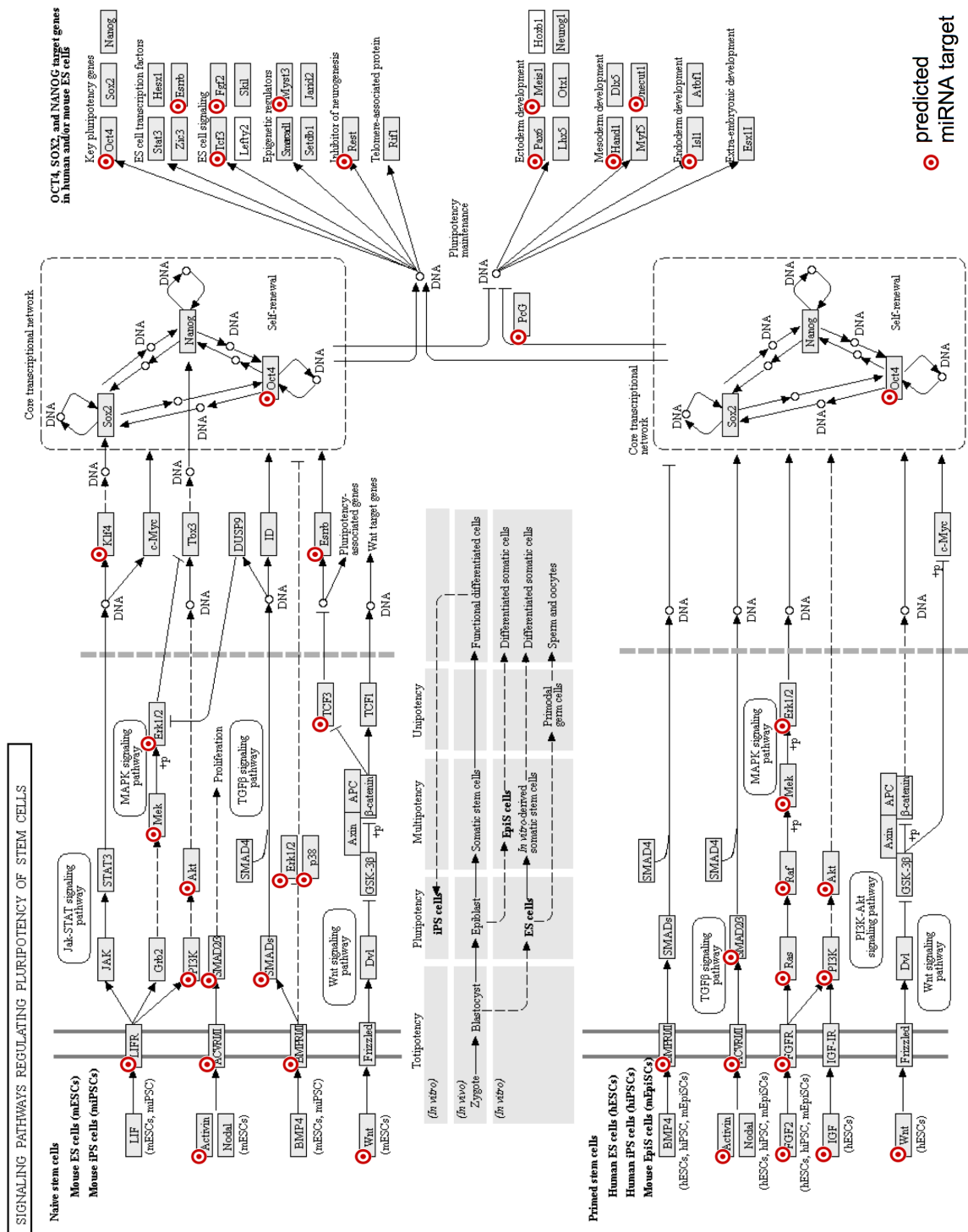
Supplementary Figure 2.S6. KEGG reference map04550 signaling pathways regulating pluripotency of stem cells for *Bos taurus*. Terms marked by a target symbol represent genes predicted as targets of over or under-represented miRNAs in 8-cell embryos with respect to oocytes
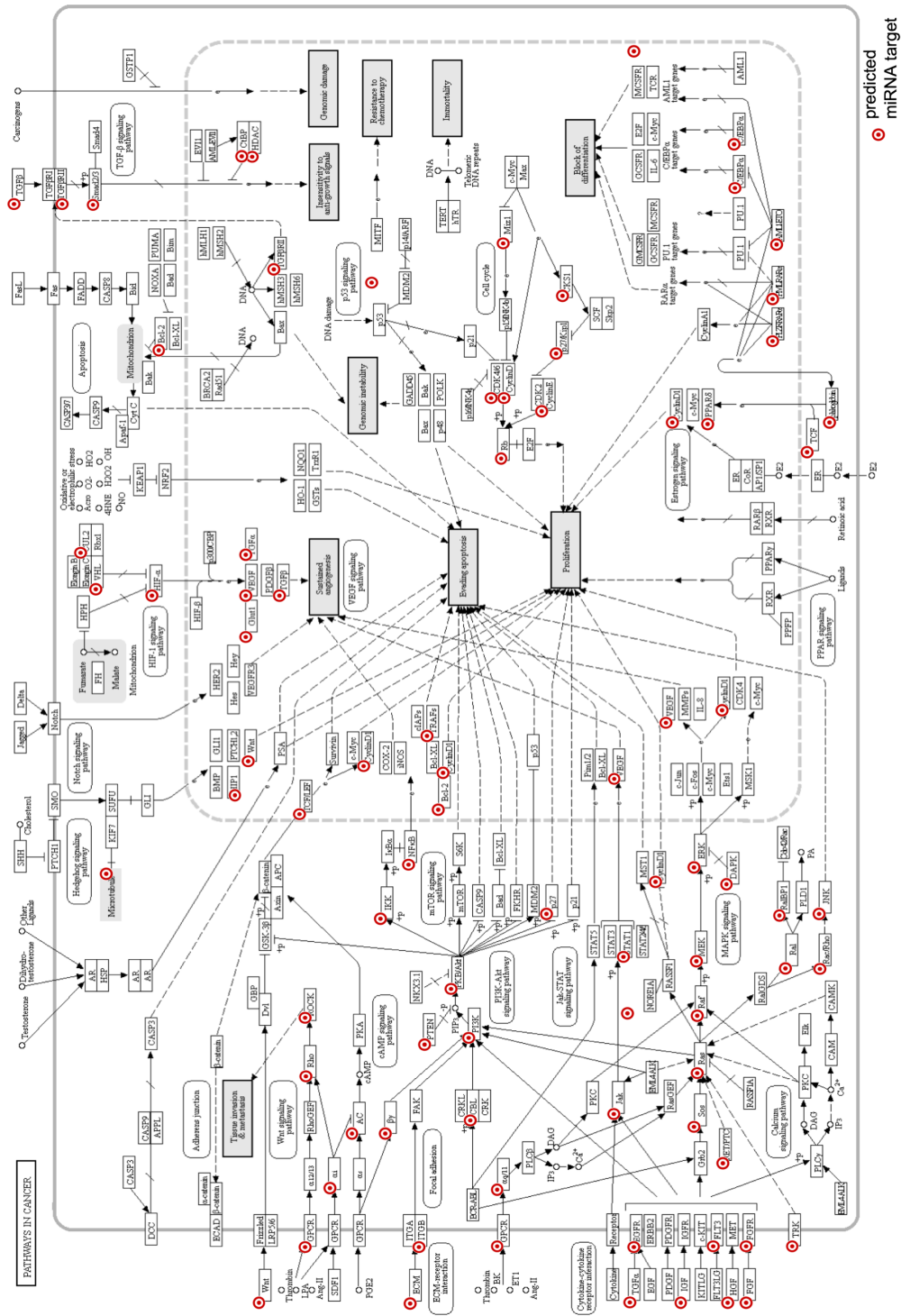
Supplementary Figure 2.S7. KEGG reference map05200 pathways in cancer for *Bos taurus*. Terms marked by a target symbol represent genes predicted as targets of over or under represented miRNAs in 8-cell embryos with respect to oocytes.
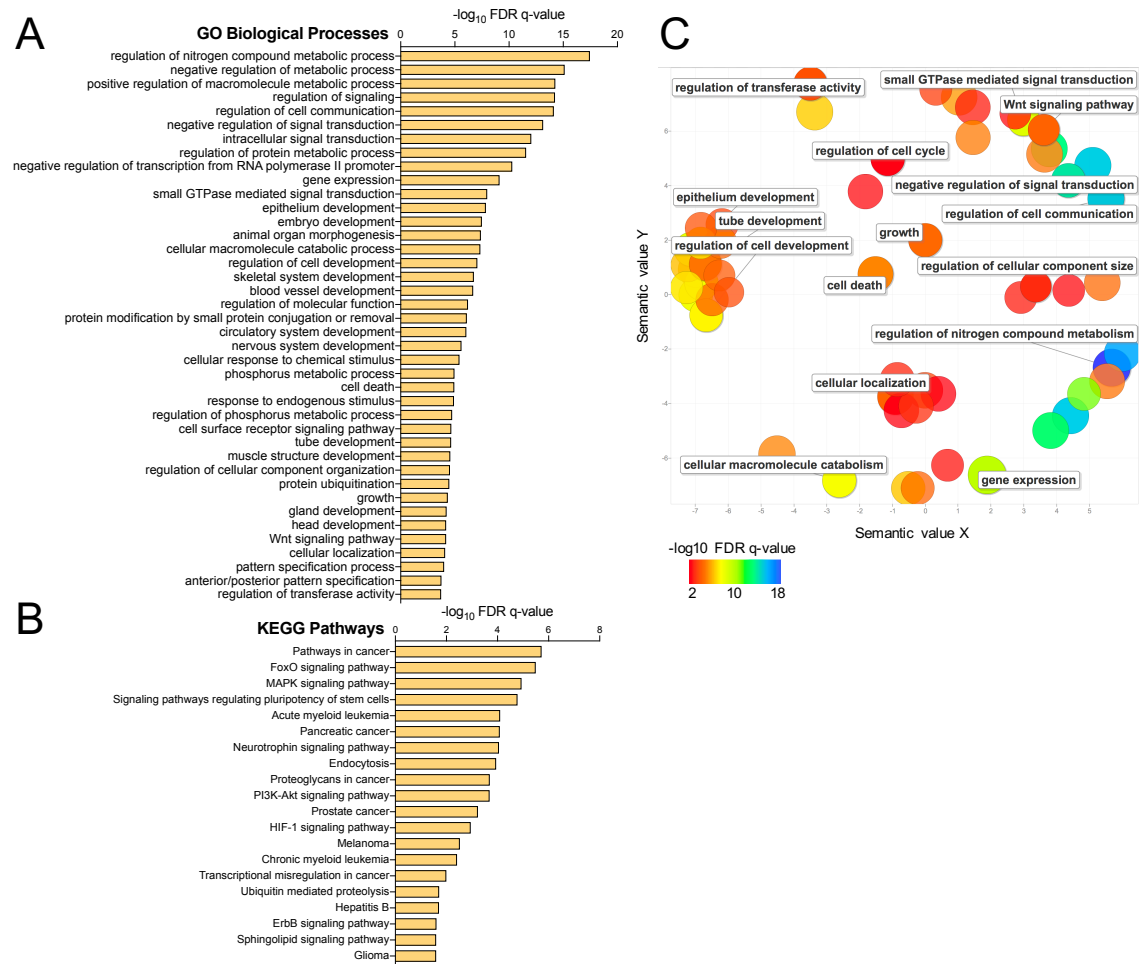
Supplementary Figure 2.S8. Predicted mRNA targets (TargetScan total context++ score <−0.35) of differentially expressed miRNA families in 8-cell embryos (|log₂ R|>1 and FDR *q*-value <0.01 compared to oocyte or blastocyst embryo) were subject to enrichment analysis using DAVID EASE against the Bos taurus genome. Each bar graph depicts the top 20 enriched summary terms identified for (A) Gene Ontology biological processes or (B) KEGG pathways. Values are the −log₁₀ enrichment FDR *q*-value. (B) REVIGO plots for significant biological process GO terms. (C) REVIGO plot (revigo.irb.hr) depicting enriched GO biological processes following reduction of the complete GO term list (all terms FDR *q*<0.05) by semantic similarity (simrel 0.7) against the *Bos taurus* database. Terms are plotted by semantic similarity scores and colored according to the−log₁₀ enrichment FDR *q*-value. Complete results are provided in Supplementary Table 2.S2.
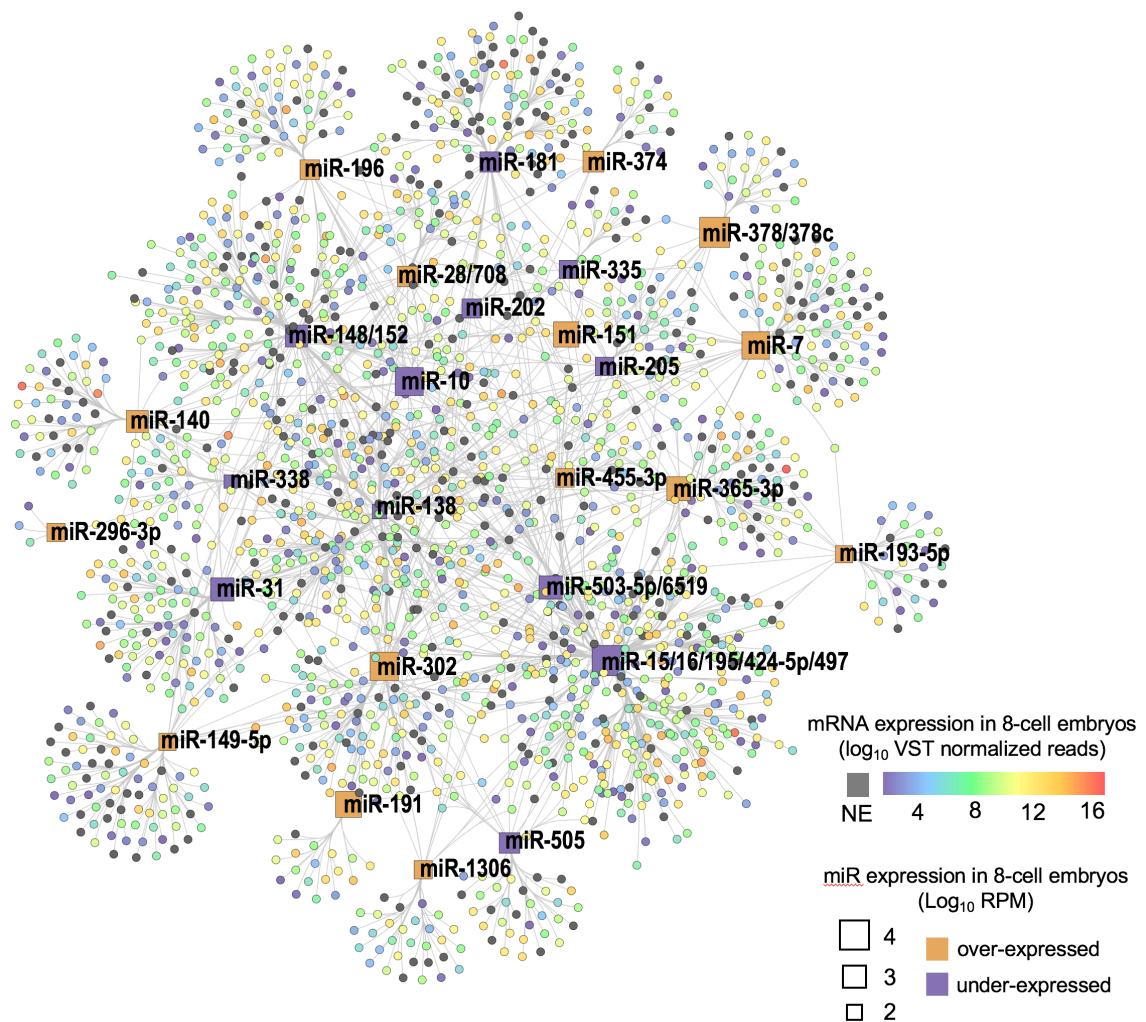
Supplementary Figure 2.S9. Network of differentially expressed miRNAs in 8-cell embryos and their predicted mRNA targets. A network of miRNA families differentially expressed in 8-cell embryos as compared to oocytes and/or blastocyst embryos and their predicted mRNA targets (TargetScan total context++ score <–0.35) was created using Cytoscape with a force-direct layout. Only miR families annotated by TargetScan as conserved or broadly conserved were included in the network analysis. MiR families are represented by squares and colored according to their abundance in 8-cell embryos (values are log_{10} RPM), whereas target mRNAs are shown as circles and colored by their expression in 8-cell bovine IVF embryos using RNAseq data from the Gene Expression Omnibus dataset GSE52415 (log_{10} VST normalized reads). Nodes colored gray were not included in the GSE52415 data set, suggesting that they were not expressed. Edges between miR and mRNA nodes indicate that the connected mRNA is a predicted target of the miR in cattle.
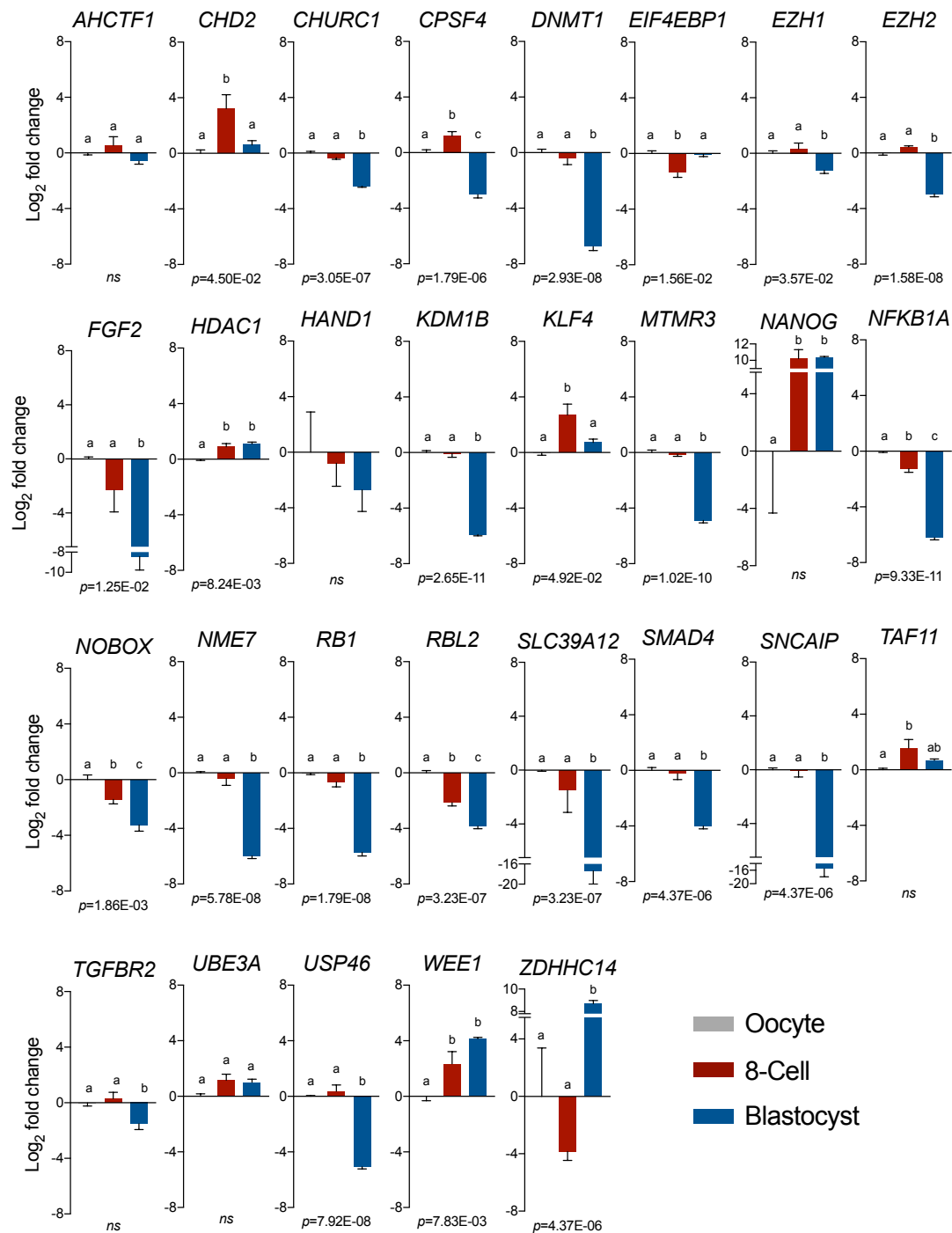
Supplementary Figure 2.S10. Expression of selected mRNA targets in bovine oocytes, 8-cell embryos, and blastocyst embryos. Values are the mean $\log_2$ fold change + SEM (n = 5 to 6). Different letters indicate groups are significantly different from each other (FDR $q<0.05$). Oo, oocyte; 8c, 8-cell embryo; Bl, blastocyst embryo.
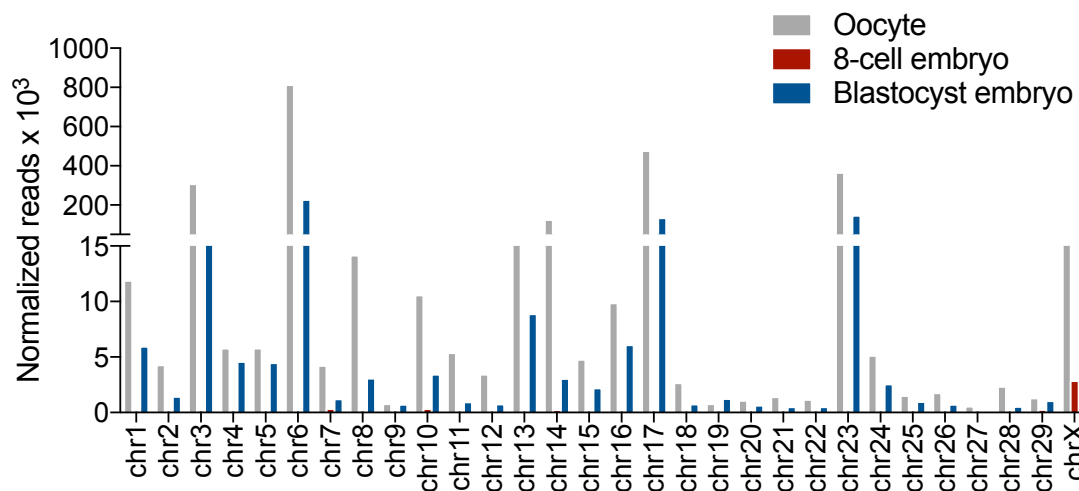
Supplementary Figure 2.S11. Potential piRNAs mapped to the *Bos taurus* genome in oocytes, 8-cell embryos, and blastocyst embryos. Values are the total normalized reads for each sample type.

CHAPTER 3

PROFILING OF SMALL NON-CODING RNA IN BOVINE SOMATIC CELL

NUCLEAR TRANSFER EMBRYOS DURING EARLY DEVELOPMENT

THROUGH THE MATERNAL-TO-EMBRYONIC TRANSITION

Abstract

Background

Somatic cell nuclear transfer (scNT) is a well-established method for animal cloning that has a transformative potential impact on production agriculture. However, scNT efficiency remains low, which represents a significant barrier to use of the technology in production and limits its potential impact. In cattle and other mammals, small non-coding RNAs (sncRNAs) have been reported to be important during early embryo development. While errors in epigenetic modifications such as DNA methylation and histone modification have been investigated thoroughly in bovine scNT embryos, a comprehensive assessment of the differences in sncRNA between scNT and *in vitro* fertilized (IVF) bovine embryos has not been performed. The objective of this study was to examine dynamic changes in expression of sncRNA during the maternal-to embryonic transition (MET) in bovine embryos produced by scNT compared to IVF.

Results

An unbiased, discovery-based approach was employed using small RNA sequencing to profile sncRNA in oocytes, bovine fibroblast cells used as scNT nuclear donors, IVF and scNT 2-cell, 8-cell, morula and blastocyst stage embryos, as well as IVF and scNT blastocyst-derived cells. The relative abundance of miRNAs was similar between IVF and scNT embryos, with the only statistically significantly different miRNAs detected between IVF and scNT morula stage embryos. These miRNAs included miR-2340, miR-345, and miR34a. Distinct populations of piwi-interacting like RNAs (pilRNAs) were identified in bovine embryos before and during the

maternal-to-embryonic transition (MET), as compared to somatic cell samples and bovine embryos post MET. Smaller distinct populations of tRNA fragments were also noted in blastocyst stage embryos, somatic cells, and earlier embryonic stages.

Conclusions

The discovery-based sncRNA sequencing analysis of preimplantation embryos revealed largely similar profiles of sncRNAs for IVF and scNT embryos at the 2-cell, 8-cell, morula and blastocyst stages of development. However, clear differences were apparent in sncRNA profiles of pre- and post-MET embryos and cultured cells with some degree of differentiation.

Introduction

Somatic cell nuclear transfer (scNT) is a well-established method for animal cloning in livestock species wherein genetic material from a recipient ovum is removed and then replaced with nuclear DNA from a donor cell. However, bovine scNT embryos are very likely to abort during pregnancy, and cloning success rates remain quite low regardless of the species or method used [1, 2]. Improvement in scNT efficiency combined with application of advanced genomics technologies could make a significant impact on the rate of genetic improvement in agricultural species.

After nuclear transfer, the cellular machinery of the host egg must reprogram the somatic DNA so that the new genome is set to a pluripotent state with the potential for cellular differentiation needed for successful development. The reprogramming of donor DNA takes place at the level of the epigenome, a layer of cellular information involving specific patterns of chromatin structure and DNA modifications that facilitate control of gene expression. Errors associated with epigenetic reprogramming of the nuclear donor genome likely lead to inappropriate gene expression in scNT embryos [3]. The cellular stress incurred by aberrant epigenome programming likely contributes to the high loss rate for scNT embryos throughout

early development and pregnancy [1], as incorrect global patterns of DNA methylation [4-7] and

inappropriate histone modifications [8-10] have been identified in scNT embryos. The epigenetic

patterns in scNT embryos appear to more closely resemble those of the nuclear donor cell,

indicating that poor epigenome reprogramming may contribute to these low efficiencies [11, 12].

One period of embryo development that would be especially detrimentally impacted by

errors in epigenetic reprogramming is the maternal-to-embryonic transition (MET). The MET is

a highly orchestrated process by which maternally deposited transcripts must be degraded in

order to allow successful embryonic genome activation (EGA). The process of degrading

massive amounts of maternal transcripts and replacing those transcripts with mRNA derived from

the embryo is required to prepare the embryo for differentiation and further development [13].

Two pathways drive transcript degradation, one using the maternal molecular machinery, and the

other utilizing the embryonic molecular machinery. One major player in the maternal

degradation pathway is the protein SMAD4, which directs the degradation of a subset of maternal

transcripts, including those that would repress the embryonic genome [14]. Maternally mediated

degradation of mRNA activates a second phase of clearance by enabling the initiation of the

embryonic mode of degradation. After the EGA, the newly activated embryonic transcriptome

guides cells through the first cell fate decisions at the blastocyst stage of development.

Small non-coding RNAs (sncRNA) are RNA molecules that are not translated into

proteins. These sncRNAs have important functions in RNA interference (RNAi), the process by

which sncRNAs block the translation or reduce the stability of protein-coding messenger RNA

(mRNA). In mammals, the microRNA (miRNA), piwi-interacting RNA (piRNA) and small

endogenous interfering RNA (siRNA) classes of sncRNA have been established as functioning in

RNAi in early development [15]. These three sncRNAs all participate in RNAi through binding

and degrading transposable elements (TEs) and/or mRNA, although biogenesis and targeting

mechanisms differ between them [16]. In addition, more recently, additional sncRNAs have been

identified that function in RNAi. Transfer RNA (tRNA) is a key player in deciphering the genetic code [17], and recently was found to be processed into fragments that act in RNAi. In cancer cell lines, individual tRNA fragments (tRFs) are up-regulated and control cell proliferation [18]. As they are further investigated, it is reasonable to expect that tRFs will be found to be involved in a myriad of cellular functions. Translation requires snoRNAs to process and modify ribosomal RNA (rRNA) [19]. Recently, snoRNA was also found to be processed to generate stably accumulating fragments that act in RNA interference [20, 21]. It appears that these so called sno-miRNAs may target transcripts controlling cell behavior and carcinogenesis [22, 23].

MiRNAs are sncRNAs that act to silence mRNA expression through binding interactions with the 3' untranslated region (UTR), which blocks translation and decreases stability of target transcripts [24]. These miRNAs are short, single stranded RNAs about 22 nt long that are highly conserved among species. The canonical pathway for miRNA biogenesis starts with RNA polymerase II-driven transcription to generate primary-miRNAs (pri-miRNAs), which are then cleaved by Drosha and its cofactor DGCR8; these miRNA processing proteins recognize cleavage sites and cleave a stem-loop and 3' 2 nt overhang on the resulting precursor miRNA (pre-miRNA). Exportin proteins transport pre-miRNA into the cytoplasm where the Dicer complex cleaves these molecules into double stranded, non-hairpin, miRNAs. AGO then unwinds and loads the miRNA duplex into the miRNA-induced silencing complex (miRISC) [25]. The miRISC directs miRNA to its target transcript and in cases of perfect complementary binding, AGO will mediate the target degradation [16]. In cases of non-complementary binding, miRNA represses translation by inhibiting translation at the initiation step [26, 27]. Also, a minority of miRNAs can be processed independently of either Drosha and/or Dicer using alternative biogenesis pathways [28]. Micro-RNAs have been shown to be powerful regulators of gene expression in many organisms and function in the majority of investigated cell functions.

Piwi-interacting RNAs (piRNAs) are single-stranded RNA molecules that are 21-34 nt long with a highly conserved function among species, which associate with the PIWI subfamily of proteins [29, 30]. Because piRNAs act as a type of immune system to protect the germ cell genomes from transposable elements (TE), they are specifically expressed in the germ cell lines, and less commonly in somatic tissue, although this is an emerging field of research. Active TEs may pose a threat to genomic stability and need to be repressed by piRNA during periods of epigenetic remodeling [31]. Piwi-interacting RNA-mediated repression occurs both transcriptionally and post-transcriptionally as they bind RNA targets through complementarity and repress targets via either endonuclease activity or chromatin mark recruitment. Piwi-interacting RNAs were originally discovered to function in gametogenesis and have since been shown be important in silencing TEs during reprogramming events throughout primordial germline cell development [32]. It is possible that piRNAs target and suppress TEs during early development, but this silencing may not be maintained during the MET, as it has been shown that TE transcripts are expressed in bovine embryos at this time [33]. Activation of a subgroup of TE transcripts is necessary at this stage, so it is likely that some down-regulation of piRNA occurs to maintain a balance between licensed and potentially detrimental TE expression [33, 34]. It is also possible that piRNAs target mRNAs for degradation. Russell and colleagues examined the expression of piRNA in cattle oocytes and zygotes and determined that piRNA populations correlated with mRNA destined for degradation in the embryo [35]. Previously, Roovers and colleagues reported that piRNAs and PIWIL3 were present at high levels in bovine oocytes and 1-2 cell embryos, which points to key functions for piRNA in cattle and is contradictory to what has been found in mice [31]. During periods of epigenetic reprogramming when methylation and other silencing marks that repress TEs are removed, piRNAs function to degrade TEs and, therefore, are key in early development.

The mature oocyte and early embryo are transcriptionally silent, setting up a unique window of development during which post-transcriptional machinery, such as sncRNA, may dominate the regulatory network. A zygotic degradation pathway mediated by miRNA was first demonstrated in zebrafish, in which a single miRNA (miR-430) mediated the decay of hundreds of maternal mRNAs [36]. Maternal mRNA decay via miRNA has also been found in *C. elegans* by both maternal and zygotic miRNA [37], in *Xenopus* by a miR-430 orthologue (miR-427) targeting mRNAs [38], and in *Drosophila* by miR-309 that is highly expressed at the EGA and degrades hundreds of mRNAs [39]. To date, mammalian miRNAs have not been shown to function in a similar manner by targeting maternal mRNA and participating in the degradation pathway, which may be due to the extensive use of the mouse mammalian model. Mice have a rodent-specific sncRNA pathway in the oocyte and early embryo dominated by miRNA and siRNA, while piRNA and associated PIWI proteins may play analogous roles in humans and cattle [31]. Endo-siRNAs may be up regulated in mouse oocytes due to a rodent oocyte-specific Dicer isoform that preferentially processes siRNA over miRNA [40]. Evidence from mouse studies supports a model in which endo-siRNAs are important for TE repression and mRNA clearance during oocyte maturation, whereas miRNAs are crucial for mRNA regulation later in embryonic development as differentiation begins [41]. However, due to the divergence in sncRNA pathways, the experimental evidence that miRNAs are not highly active or functional in mouse oocytes and early embryos [42-44] should not deter further investigation of miRNA as a embryonic clearance mechanism in mammals.

While a single miRNA that participates in the same type of maternal mRNA clearance as seen in Zebrafish and *Drosophila* has not been found in mammals, mammalian miRNAs have been identified that may function in early development. Our group previously demonstrated that in cattle, miRNA and other sncRNA are dynamically regulated with large-scale population changes through the MET in cattle *in vitro* fertilized (IVF) embryos [45]. Other groups,

quantifying only a fraction of total miRNA, also found dynamic changes in miRNA species in pre-implantation embryos in cattle including increases at the 8-cell stage [46, 47]. Specific miRNAs have been shown to function in cattle embryogenesis as well, such as miR-130b, which has been found to impact granulosa and cumulus cell proliferation and oocyte maturation, as well as morula and blastocyst formation [48]. In addition, miR-218 and miR-449b are specifically expressed in blastocyst stage cattle embryos. MiR-218 may target and regulate *NANOG* and *CDH2*, and miR-449b may target and regulate *NOTCH*, and the down-regulation of these genes would greatly impact development [49]. Additionally, miR-212 has been shown to negatively regulate the maternal effect gene *FIGLA* [50], and miR-196 has been shown to negatively regulate the maternal effect gene *NOBOX* [51], supporting a possible function for cattle miRNA in the degradation of transcripts through the MET.

Aberrant expression of miRNA may contribute to the low efficiency for production of viable scNT embryos. Incomplete reprogramming of the somatic donor cell genome has been seen in scNT embryos, as patterns of DNA methylation and histone modifications are abnormal in scNT embryos [11, 52]. Also, DNA demethylation does not appear to occur at the same levels in scNT embryos [53], and persistent nuclear donor methylation patterns appear to occur [8, 9]. As DNA methylation patterns have been shown to impact miRNA expression, it is possible that persistent somatic miRNA expression occurs as well, further impacting the aberrant transcriptome of scNT embryos [54-56]. In addition, others have observed aberrant miRNA expression in clones at more advanced stages of embryonic or fetal development [57, 58]. Differences in the expression of miRNA clusters and families in scNT placentas as compared to IVF placentas and placentas from artificial insemination pregnancies have also been identified that could significantly contribute to errors in gene expression, causing embryonic death [57]. Interestingly, one major biogenesis protein for miRNA, AGO, was down regulated in scNT placentas due to hypermethylation of the promoter region. To date, these studies are the only reports to assess

expression of miRNAs in bovine scNT. However, these reports are limited by virtue of a targeted

approach that may not have captured miRNAs of developmental importance. The importance of

miRNAs in the development of bovine embryos has been further supported by several studies that

have shown that specific miRNA supplementation can impact rates of development in embryos

[59], including the improvement of scNT embryo development rates [60-64].

A comprehensive assessment of dynamic changes in populations of sncRNAs in scNT

embryos through the MET has not been performed in any mammalian model, except for the

mouse. Because of distinct molecular differences in biogenesis of certain classes of sncRNAs in

rodents and other mammals, the mouse may not be the most useful model for these investigations.

Importantly, these processes appear to function similarly in bovine and human embryos, making

the cattle early developing embryo a useful biological model to understand dynamics of sncRNA

during early development. Importantly, to our knowledge, no one has applied a discovery based,

unbiased approach to characterize populations of sncRNA during the critical MET period in scNT

embryos, as compared to IVF embryos in any mammalian non-rodent species. Thus, the

objective of this study was to examine dynamic changes in expression of sncRNA in scNT

embryos, during early embryo development encompassing the MET period. Based on the

evidence discussed above, we hypothesized that miRNAs found in the nuclear donor cell would

be more abundant in scNT embryos as compared to IVF embryos. Also, we hypothesized that

aberrant expression patterns of miRNAs in scNT embryos would occur as compared to IVF

embryos, and the sncRNA profile of the somatic donor cells would persist in scNT embryos

potentially due to incomplete reprogramming. We predicted that the population of miRNAs

impacted would target maternal mRNAs with important functions associated with epigenetic

reprogramming and development, thus contributing to the aberrant epigenetic landscape found in

scNT embryos. RNAseq was employed to profile all sncRNA classes in bovine oocytes, 2-cell

embryos, 8-cell embryos, morula embryos, blastocyst embryos, and blastocyst-derived cells

produced using both IVF and scNT methods. Fibroblasts that had been used as donor nuclear

cells for scNT were profiled as well.

Methods

Oocyte collection and in vitro maturation

Bovine ovaries were collected from a local abattoir (JBS, Hyrum, UT) and transported in

a cooler containing 0.9% saline solution to the laboratory. The cumulus-oocyte complexes

(COCs) were then aspirated from 3-8 mm follicles by using an 18-gauge needle and vacuum

system. Only compact COCs with homogenous ooplasm and intact layers of cumulus cells were

used for scNT or IVF. Following aspiration, COCs were cultured at 39 °C with 5% CO_2 for 22 to

24 hr. The oocytes were cultured in TCM199 maturation medium with Earle's salts, L-glutamine,

and sodium bicarbonate (Hyclone, Logan, Utah) supplemented with 10% fetal bovine serum

(FBS), 0.05 mg/ml bovine follicle stimulating hormone (Sioux Biochemicals, Sioux city, IA), 5

mg/ml bovine luteinizing hormone (Sioux Biochemicals), 100 U/ml penicillin, and 100 mg/ml

streptomycin.

In vitro fertilization

Following 22 to 24 hr of maturation, MII oocytes were fertilized using the laboratory's

standard *in vitro* fertilization (IVF) protocol [65]. Briefly, one straw of cryopreserved bovine

semen obtained from a Holstein bull (Hoffman AI, Logan UT) was removed from the liquid

nitrogen tank and placed into a 35 °C water bath to thaw. Live sperm were isolated by

centrifugation through a 45%/90% percoll gradient, suspended (final concentration 1×10^6/ml) in

Tyrode's albumin lactate pyruvate containing 10 μg/ml heparin and used to fertilize the mature

oocytes. Twenty to 22 hr post-IVF, cumulus cells were removed by vortexing, and fertilized

zygotes were washed in phosphate buffered saline with 0.32 mM sodium pyruvate, 5.55 mM

glucose, and 3 mg/ml bovine serum albumin (PB1+). After washing, zygotes were cultured on a monolayer of bovine cumulus cells in 50 μl of synthetic oviductal fluid (SOFaa) with 3% FBS overlaid with mineral oil. The embryos were cultured at 39 °C in a humidified atmosphere with 5% CO_2. Half of the SOFaa medium was removed and replaced with fresh, equilibrated medium every other day starting the day after *in vitro* culture. Historically, production of IVF embryos via these methods followed by blastocyst stage embryo transfer to cattle recipients has generated successful pregnancies at an approximate rate of 50% by our research group (unpublished observations). The timing of embryo collection post-fertilization was 24-36 hr for 2-cell stage, 48-60 hr for 8-cell stage, 156-168 hr for morula stage, and 192-204 hr for blastocyst stage.

Somatic cell nuclear transfer

Primary bovine fibroblast cultures were established from ear biopsy tissues from a Brahma Spanish bull cross using well-established procedures. Frozen/thawed cells were grown to 80-100% confluence then passaged, with cells from passages 4-5 used as nuclear donors. Three days prior to a cloning session, donor fibroblast cells were thawed and propagated, and incubated in DMEM media supplemented with 15% (v/v) FBS at 39 °C with 5% CO_2. Bovine fibroblasts at 80-90% confluence were serum starved by replacing culture media with DMEM media containing 0.5% (v/v) FBS 24 hr prior to scNT. Oocytes were matured for 18-20 hr, then denuded using 100 μl of 1% (v/v) hyaluronidase, incubated for 5 min at 39˚C with 5% CO_2, followed by gentle pipetting. The removed cumulus cells were centrifuged at 500×*g* for 5 min at room temperature, and the supernatant was removed. Cumulus cells were then cultured in SOFaa under oil for use in embryo culture. The denuded oocytes were then rinsed using PB1+, and oocytes with polar bodies were selected and separated. ScNT was performed according to established protocols [66-68]. Briefly, oocytes with polar bodies were incubated in 0.6 μg/ml demecolcine for 30-40 min and a metaphase plate and a polar body were removed using a beveled

pipette. Fibroblast cells were covered in 0.25% trypsin for 1 min, and then the trypsin was removed and fibroblasts were incubated for anther 6 min at 39˚C with 5% CO_2. Once fibroblast cells were detached, they were rinsed in warm medium and centrifuged at 150xg for 6 min. The supernatant was removed, and the cellular pellet was resuspended in 100 μl Hepes SOF medium [69]. One fibroblast cell was injected into the perivitelline space of the recipient oocyte, and fused using one direct pulse of 1.2 kV/cm for 22 μs by an Electro Cell Manipulator 2001 (BTX, San Diego, CA) in 0.28 M sorbitol, 0.05% (w/v) BSA, 0.1 mM $CaCl_2$, 0.5 mM $MgCl_2$, and 0.5 mM Hepes. Following fusion, embryos were washed through Hepes SOF, and incubated in embryo culture medium for 1 hr. Activation was then performed, with successfully fused embryos cultured in 5μM ionomycin for 5 min, followed by 4 hr of incubation in activation medium composed of SOFaa medium with 10μg/ml cyclohexamide and 1 mM 6-dimethylamino-pyridine at 39˚C with 5% CO_2. Following activation, embryos were cultured on a monolayer of bovine cumulus cells in 50 µl drops of SOFaa with 3% (v/v) FBS overlaid with mineral oil. The embryos were cultured at 39 °C in a humidified atmosphere with 5% CO_2. Half of the SOFaa medium was removed and replaced with fresh, equilibrated medium every other day starting the day after *in vitro* culture. The timing of embryo collection was the same as for IVF embryos as described above.

Isolation of blastocyst-derived cells

We followed methods previously described [70] to generate putative embryonic stem cells (ESCs) from IVF and scNT embryos for comparison. Briefly, bovine expanded blastocysts produced by IVF and scNT were used for BDC isolation. Zona pellucida-free blastocysts were placed into 4-well culture dishes onto feeder layers of mitomycin C-treated mouse primary embryonic fibroblasts. Feeder layer preparation was performed as previously reported [71]. These cells were cultured in DMEM supplemented with 0.2% (w/v) LIF, 1% (w/v) bFGF, and

30% (v/v) Knock out Serum Replacement (ThermoFisher Scientific, Waltham, MA). When propagation was needed, cultured cell colonies were mechanically propagated using a small metal blade. Germ line competent ESCs have only been reported for mouse and rat [72]. However, bovine putative ESC-like cells expressing pluripotency markers have been reported by several groups [70, 73-75]. Though the objective here was to generate putative ESC-like cells, transcriptional profiling revealed that these cells did not consistently express the array of gene markers typical of pluripotent stem cells (data available with Chapter 4). Thus, we instead refer to these cells as blastocyst-derived cells (BDCs) herein.

Tissue collection

Four pools of 20 mature oocytes were collected at 22 hr after maturation. To remove cumulus cells, the mature COCs were treated with 10 mM hyaluronidase for 5 min, followed by repeatedly pipetting the mixture until cumulus cells were removed. After visual inspection for complete removal of cumulus cells and the presence of a polar body, denuded mature metaphase II (MII) oocytes were washed through four droplets of PB1+. Oocytes were then snap frozen in cryotubes containing RNA/DNA shield (Zymo, Irvine, California), and stored at -80 °C until RNA isolation. Fibroblasts prepared for scNT were collected after centrifugation and pelleting, and were snap-frozen in 100ul of RNA/DNA shield (Zymo, Irvine, California). Four pools of 20 2-cell, 8-cell, morula, and blastocyst stage embryos were collected at 2, 3, 6, and 8 days post-IVF or scNT, washed in PBS, and stored at -80 °C.

Small RNA isolation and sequencing

RNA was isolated based on size using the RNA Clean & Concentrator™ 5 kit (Zymo) from four pools each of oocytes, 2-cell stage embryos, 8-cell stage embryos, morula stage embryos, or blastocyst stage embryos as well as donor fibroblasts and BDCs according to the

manufacturer's protocol for purification of small RNA (<200 nt) and large RNA (>200 nt) as separate fractions. Small RNA sequencing was performed on the Ion Proton[TM] Sequencer using the Ion Total RNA-seq kit v2 (Thermo Fisher Scientific, Waltham MA) according to manufacturer's procedure for small RNA library preparation with no deviations from the specified protocol. By using an RNA isolation protocol that yielded a specific fraction of small RNA, an enrichment step was not needed. Sample volumes were reduced to 3 μl by vacuum centrifugation, and the entire sample was used to prepare the small RNA library for sequencing. The cDNA sample was then processed on the Agilent Bioanalyzer (Agilent, Santa Clara CA) to ensure the presence of small RNA bound to a barcode (86 to 106 nt).

MicroRNA data processing and analyses

Sequence data were processed to remove low quality reads and any artificial reads introduced during library preparation (Perl script Trim Galore, Python script SortMeRNA) [76, 77]. Annotation and expression analysis of sncRNAs was performed as described previously [35]. Briefly, data processing and bioinformatics analysis for other sncRNA were performed primarily with command line tools available at http://www.smallrnagroup.uni-mainz.de/software.html. First data were filtered for sequence length between 17 and 32 nucleotides (nt), which includes the canonical size range for mature small RNAs (miRNA, siRNA, piRNA). A separate analysis was performed for sequences between 33-93 nt, which encompassed longer reads characteristic of snoRNA, tRNA and mitochondrial rRNA. This split analysis approach avoided possible skewing of the data when comparing between these diverse sncRNA classes. Sequence annotation was performed with Unitas v1.6.1, which uses the latest available public small RNA databases to annotate input sequences [78]. Read length and annotation summaries for all samples are provided in Supplementary File 3.1; Appendix F. Non-annotated sequences were then mapped to the bovine genome (BosTau8) with sRNAmapper [79]

to determine which non-annotated sequences were likely of biological, not technical, origin. The mapped sequence reads were redistributed across the genome based on number of mapping locations using reallocate with the parameters '5000 1000 b 0' (reallocate.pl). Reallocated map files were then analyzed with ProTRAC v2.4.2 [80], which identifies probable piRNAs through their genomic clustering based on size, sequence, and cluster characteristics (all options at default settings, including repeatmasker and geneset references). Predicted piRNA clusters were compared among samples by using the Galaxy "merge" tool (usegalaxy.org) to first generate lists of predicted clusters by sample type followed by the "join" tool to identify overlapping genomic intervals among samples. Sequence logos representing nucleotide biases were generated with ggseqlogo [81]. Count tables of sncRNAs were generated from the output of Unitas. Normalization and differential expression analysis was performed with the DESeq2 R package [82]. Differentially expressed sncRNAs were called with an adjusted p-value of < 0.05. Normalized and variance stabilized data for all annotated reads are provided in Supplementary File 3.2; Appendix G. Complete results of DESeq2 differential expression analyses are provided in Supplementary File 3.3; Appendix H. BioVenn was used to create proportional Venn diagrams depicting the number of transcripts mapped to the cattle genome for each sample [83]. Unsupervised, bidirectional hierarchical cluster analyses (Euclidean distance method with average linkage) and principal component analyses were performed using ClustVis [84].

Network analysis of predicted gene targets of differentially expressed miRNAs

Gene targets for miRNAs individual species/families identified as differentially expressed between scNT and IVF embryos or by developmental stage within the embryo type were predicted using TargetScan (release 7.2) [85] with the *Bos taurus* miRNA database (miRBase release 22). TargetScan predicts biological targets of miRNAs by searching for the

presence of a conserved 8-mer, 7-mer, and 6-mer sites matching the seed region of each miRNA.

TargetScan provides an additional advantage by curating individual miRNAs by family (identical

seed ± m8 sequence). Lists of predicted target mRNAs were filtered for total context ++ score $\leq$ -

0.35 (efficacy of targeting miRNA site: Supplementary File 3.4; Appendix I), and miRNAs of

interest were defined by family, level of conservation with other species, and by seed sequence.

Note that the output of TargetScan is the human orthologue of a predicted target gene; the

representative transcript that is shown is the transcript with the UTR profile that is the most

prevalent (highest number of supporting 3P-seq tags). Lists of all predicted mRNA targets were

then subject to gene set enrichment analysis using Metascape and the Gene Ontology (GO)

biological process database with the following parameters: minimum overlap of 3 genes in a

category, minimum enrichment of 1.5, and p <0.05 [86]. Ontology analyses were performed

using DEG sets for comparison between IVF and scNT embryos at each developmental stage and

for comparisons by developmental stage within each embryo group. Resulting GO biological

process terms were then clustered by semantic similarity (kappa score >0.3) and represented as

networks in Cytoscape [87]. Clusters of related terms were named according to the term within

the cluster with the lowest p value; thus, clusters with different labels may have some similar

terms within the cluster.

Results

Classification of sncRNA sequences

RNA sequencing of sncRNA obtained from bovine oocytes, fibroblasts, 2-cell embryos,

8-cell embryos, morula stage embryos, blastocyst stage embryos, and blastocyst-derived cells

generated a total of about 115.8 million reads after filtering out low quality reads and artifacts.

Sixty percent of these reads aligned to known bovine sncRNAs, with reasonable read depth

across sample types (Supplementary File 3.1). However, we noted that the sequencing depth per

sample was somewhat variable and did not achieve the high standard of 2 million reads per sample as described by ENCODE [88], which was not surprising given the small number of starting cells that could be obtained early developing embryos. As shown in Figure 3.1A, the number of reads obtained per sample was markedly higher in the differentiated cell types, BDCs and fibroblasts. Overall, scNT embryos generally had more reads then their IVF staged counterparts, although this was not the case at the 8-cell stage, where the scNT embryos had approximately 390 thousand fewer reads. RNA input was not normalized by sample type when generating the sample libraries for sequencing due to the exceedingly small amount of small RNA obtained. However, the RNA input for somatic cells were adjusted to be similar to the low input obtained for embryos and was normalized.

As shown in Figure 3.1B, sequence reads of 21-24 nt in length, which corresponds with the miRNA class of sncRNA, were very abundant across all sample types. Also, small peaks for read lengths corresponding to snoRNAs (70-90 nt) and tRNAs (76-90 nt) were also evident for all sample types, with these peaks representing a range of sizes down to 60 nt and a larger fraction of the total reads in the blastocyst staged embryos and the BDCs. Other minor peaks at various lengths were present, likely corresponding to rRNA, snRNA, or other miscellaneous RNA fragments. Within the set of sequences 17-32 nt in length, the proportion of reads that were annotated as sncRNAs in oocyte samples was approximately 5%, whereas 2-cell and 8-cell embryos had annotation rates ranging from 4-15% and 5-17%, respectively (Figure 3.2A). Annotation of sncRNAs was greater among morulae (10-39%) and blastocyst stage embryos (14-33%) (Figure 3.2B), but the highest annotation rates were noted for more differentiated cells including fibroblasts (54-85%) and BDCs (10-71%), with notable variation among the BDCs (Figure 3.2C). For longer reads >33 nt, 5-8% of reads from oocyte samples were annotated, while annotation rates for 2-cell and 8-cell stage embryos were more variable ranging from 2-14% and 5-11%, respectively (Figure 3.2D) For mid-stage embryos, annotation rates of reads

>33 nt in length for morulae were variable ranging from 4-30% compared to rates of 12-41% for

blastocyst stage embryos (Figure 3.2E). In contrast to short sequences 17-32 nt, for sequences

>33 nt, the annotation rate for fibroblasts and BDCs were comparatively low at 4-15% and 12-

40%, respectively (Figure 3.2F). Note that these annotation percentages do not include those

identified as protein coding (see Supplementary File 3.1). Also, no consistent differences in

sequence annotation rates were noted for IVF or scNT embryos.

Figure 3.2 also shows the distribution of annotated sequence reads by sncRNA class. The

population of sncRNA in 2-cell and 8-cell embryos appears dominated by genomic tRNA, with

smaller populations of miRNA and genomic rRNA (Figure 3.2A). Later embryo stages – notably

blastocyst – had a greater fraction of sncRNA classified as miRNAs. However, this pattern was

highly variable among individual embryo pools. Interestingly, while genomic tRNA was a major

fraction of sncRNA in fibroblasts, miRNA appeared to dominate the population of sncRNA in

BDCs. Fairly strong class differences were evident for sequence reads >33 nt, with 2-cell, 8-cell

having high relative abundance of genomic tRNA and mitochondrial (mt)-RNA, whereas

snoRNA was relatively abundance in blastocysts and BDCs (Figure 3.2E-F).

Expression profiles of miRNAs in IVF and scNT embryos

After filtering out miRNAs with very low abundance reads across samples (<100 total

reads across all 48 samples), 382 miRNA species representing 71 miRNA families were detected.

Figure 3.3 shows overlap in miRNAs detected via sncRNA sequencing by developmental stage

for both IVF and scNT embryos. Over the course of early development, sizable subsets of

miRNAs were commonly expressed at proximal developmental stages (e.g., 230 miRNAs

detected in both 2-cell and 8-cell IVF embryos and 248 miRNAs detected in both morula and

blastocyst stage scNT embryos). However, at each stage, uniquely expressed miRNAs were also

identified (e.g., 390 miRNAs uniquely expressed in 2-cell scNT embryos that were not detected

in 8-cell scNT embryos). Two exceptions were noted, where all of the miRNAs detected in

oocytes were also detected in 2-cell scNT embryos, and all of the miRNAs identified in 8-cell

scNT embryos were also present in 2-cell scNT embryos. A three-way Venn analysis was used to

compare the IVF and scNT type of each embryo stage with the oocyte and fibroblast sample to

visualize similarities and differences in sample types (Figure 3.3B). Interestingly, at the 2-cell

stage, the scNT embryo expressed all miRNAs that were present in the IVF 2-cell stage embryo,

the oocyte, and the fibroblast. The 2-cell stage pattern did not hold for 8-cell embryos, as this

developmental stage – and all subsequent stages – was typified by a pattern of some miRNAs

specific for scNT or IVF embryos. A fairly large fraction of miRNAs was shared by both embryo

types but not oocytes, and another fairly large pool of miRNAs were detected in oocytes and both

types of embryos. These trends were also evident for comparisons of embryo types to fibroblast

donor cells. Also, BDCs shared a larger fraction of their identified miRNAs with fibroblasts

(288) than oocytes (124) (Figure 3.3B). Last, BDCs originated from scNT embryos had fewer

unique miRNAs as compared to those derived from IVF embryos.

The sequence data were further examined to identify specific miRNAs differentially

expressed in any one-sample type, summarized in Table 3.1. Complete results for all differential

expression analyses by DESeq2 are provided in Supplementary File 3.3; Appendix H. First,

comparisons were made at each developmental stage between IVF and scNT embryos. Of note,

in early developing 2-cell and 8-cell embryos, no differentially expressed miRNAs were

identified, whereas six miRNAs were differentially expressed at the morula stage, including cattle

miRNAs miR-34a, miR-345-5p, and miR-2340-3p (14T→A) (Supplementary Figure 3.S1). In

contrast, substantial differences in miRNA expression were noted for BDCs obtained from either

IVF or scNT embryos, including 24 miRNAs annotated for cattle (Table 3.1).

Comparisons were also made by developmental stage for each embryo type (Table 3.1).

No significant differences in miR expression were noted when comparing oocytes to 2-cell

embryos or 2-cell embryos to 8-cell embryos for either IVF or scNT embryo types. However, a major shift was evident when comparing 8-cell to morula embryos, coincident with the EGA, with 22 differentially expressed miRNAs identified in IVF (e.g., miR-450a and miR-320a) and 19 in scNT embryos (e.g., miR-302a-3p and miR-378), respectively (Table 3.1, Supplementary Figure 3.S1). An additional 19 and 33 miRNAs were differentially expressed in cattle IVF (e.g., miR196a, miR-93-5p and miR-30c) and scNT embryos (e.g., miR-145-5p and miR-451), respectively, when comparing morula to blastocyst stage embryos (Table 3.1, Supplementary Figure 3.S1). Finally, extensive changes in expression of miRNAs were evident when comparing blastocyst stage embryos to cells derived from those embryos, including 160 for IVF blastocysts vs. BDCs and 201 scNT blastocysts vs. BDCs (Table 3.1, Supplementary File 3.3; Appendix H). As expected for comparing differentiated and undifferentiated cells, sizable differences were observed when comparing donor fibroblasts to 2-cell embryos, including 191 differentially expressed miRNAs for IVF and 202 for scNT embryo types.

Principal components analysis of miRNA expression revealed clear grouping of data sets by developmental stage, though with no obvious clustering that separated IVF and scNT embryos (Figure 3.4A). Of note, fibroblasts and IVF morulae were not well clustered, indicating high overall variability for those sample types. Also, while BDCs were apparently distinct from other samples, their overall miRNA profiles were more variable as indicated by a greater spread along the two principal components. An unsupervised, bidirectional hierarchical cluster analysis (Euclidean distance method) of abundance data (VST normalized) for all miRNA samples expressed show the same pattern of clustering primarily by embryonic stage without strong segregation of IVF and scNT embryos (Figure 3.4B). Considering the miRNA expression profiles, a clear separation was noted for nearly all BDC samples, which includes a set of miRNAs expressed at markedly greater levels than all other sample types. Also evident in the hierarchical cluster analysis are the distinct miRNA profiles for some isolated samples that

clearly did not segregate with their counterparts, such as sample IDs IVF.BDC2, IVF.Mo.3 and Fb.1.

Principal components and unsupervised, bi-directional hierarchical cluster analyses were also performed for sequential comparisons of developmental stages for either IVF or scNT embryos as well as pairwise comparisons for IVF vs. scNT at each stage (Supplementary Figures 3.S2 to S4). Distinct miRNA expression profiles were not evident when comparing oocytes to 2-cell IVF embryos, nor when comparing 2-cell to 8-cell IVF embryos, whereas segregation by developmental stage was evident for all other comparisons (Supplementary Figure 3.S2A). Even so, high variability was evident across many sample types, particularly for the less abundant miRNAs. Interestingly, PCA plots revealed some different patterns with broad variability for miRNA profiles of 2-cell and morula staged IVF embryos, while profiles for oocytes, 8-cell and blastocysts were highly similar (Supplementary Figure 3.S2B). For scNT embryos, samples apparently separated earlier in development, as revealed by comparing 2-cell and 8-cell stages (Supplementary Figure 3.S3A). Again, some variability among the profiles was evident, mostly for those low-expressed miRNAs. Of interest, PCA revealed highly distinct miRNA profiles for scNT embryos when comparing 8-cell to morula stages, morula to blastocyst stages, and blastocysts to BDCs (Supplementary Figure 3.S3B). Last, comparisons at each stage revealed no clear distinctions for miRNA profiles in IVF and scNT embryos at the 2-cell or 8-cell stages, whereas clear separation was evident in morula and blastocyst embryos (Supplementary Figure 3.S4). Of note, miRNA expression profiles were highly consistent for BDCs generated from either IVF or scNT embryos with no separation evident by hierarchical clustering or PCA.

*Ontology analysis of predicted mRNA targets for differentially
expressed miRNAs in bovine IVF and scNT embryos*

Results of biological process gene ontology analysis of TargetScan predicted gene targets of differentially expressed miRNAs between separate stages are provided in the Supplementary

File 3.4; Appendix I (TargetScan predicted mRNA targets) and Supplementary File 3.5;

Appendix J (results of Metascape analysis) and summarized in Figures 3.5 and 3.6. Note that the

heat maps represent the top 20 terms (clustered by kappa score for similarity) for each data set.

For mRNA targets of the three miRNAs differentially expressed in scNT embryos at the morula

stage, the top GO biological process terms included "lateral mesoderm development", "negative

regulation of cellular component organization" and "establishment of protein localization of

organelles" as well as other terms associated with early development (Figure 3.5A). Results of

this ontology analysis were also represented as a Cytoscape network (Figure 3.5B), with each

cluster corresponding to a term shown in Figure 5A. This network shows that the pathways

associated with differentiation and development as central with other terms branch off into

smaller disconnected networks within their own terms. For mRNA targets of differentially

expressed miRNAs in BDCs obtained from scNT embryos, the top GO terms included

"regulation of protein kinase activity", "tissue morphogenesis", "embryonic morphogenesis",

"regulation of mRNA metabolic process", "regulation of growth", "response to growth factor",

and "Wnt signaling pathway", as well as other GO terms related to differentiation, growth, and

development (Figure 3.6A). The Cytoscape network for these ontology results depicts

connections between clusters of terms related to morphogenesis, development and growth (Figure

3.6B). A second, unconnected cluster included metabolic processes of mRNA and proteins, and

several other pathways were represented by isolated sub networks, such as Wnt signaling

pathway.

Ontology analyses were also performed for sets of predicted mRNA targets of miRNAs

identified as differentially expressed by developmental stage, including 8-cell vs. morula and

morula vs. blastocyst in IVF embryos (Supplementary Figure 3.S5) or scNT embryos

(Supplementary Figure 3.S6). The enriched terms for differences by type in morula stage

embryos included terms related to basic cell functions that would be important during morula

development, including "cellular migration and growth", "tissue development", "changes in

protein localization and stabilization", "cellular component organization", "organelle

organization", and "cytoskeleton organization". The differences in scNT and IVF BDC cells

included terms that more directly seemed to relate to early development, including "regulation of

growth", "Wnt signaling pathway", and terms related to morphogenesis and differentiation.

These are very similar terms to what was seen in enriched terms within type, between the stages

of morula to blastocyst. The morula to blastocyst stage embryo transition looked similar between

IVF and scNT type embryos, with enriched terms including "response to growth factor", "Wnt

signaling pathways", "negative regulation of cell differentiation", and over half enriched terms

relating to morphogenesis, differentiation, or system development. There were some differences

between the 8-cell to morula transition between IVF and scNT embryos, most notably the scNT

embryos lacked the "Wnt signaling pathway", although it is possible that since this term shows up

in the scNT morula to blastocyst stage transition, late induction of the miRNAs participating in

this pathway occurred.

Expression of tRNA fragments in bovine IVF and scNT samples

Transfer RNA fragments (tRFs) showed a similar pattern for differential expression as

miRNAs; with no significantly different tRFs expressed between IVF and scNT type embryos at

the 2-cell, 8-cell, blastocyst stage, or BDC samples (Table 3.2; Supplementary File 3.3; Appendix

H). However, at the morula stage, four tRFs were identified as significantly different when

comparing IVF and scNT embryos. Notably more tRFs were differentially expressed when

sequentially comparing developmental stages of IVF embryos, ranging from 3 altered tRFs

comparing oocytes to 2-cell embryos to 26 altered tRFs comparing morulae to blastocysts (Table

3.2). When comparing IVF embryos to more differentiated cells, more substantial differences

were observed as 192 tRFs were noted as significantly different between the blastocyst stage

embryos and BDCs and 379 tRFs were different between fibroblasts and 2-cell staged embryos. Alternatively, no differentially expressed tRFs were noted for early developing scNT embryos, only 21 tRFs were altered for 8-cell vs. morula and a large set of 139 tRFs were differentially expressed in morula vs. blastocyst scNT embryos. When comparing scNT embryos to more differentiated cells, similar numbers to those observed in IVF type were apparent, with 129 tRFs significantly different between the blastocyst staged embryos and BDCs, and 393 tRFs different between fibroblasts and 2-cell staged embryos. Interestingly, the number of predicted active tRNA genes in cattle is 439 [89]. Moreover, differential tRNA gene expression leads to changes in abundance of tRFs rather than mature tRNAs [90]. As each tRNA may generate multiple tRFs based on the cleavage location, one would expect to detect a sizable population of tRFs by sequencing [91].

Expression profiles of tRFs in IVF and scNT embryos, oocytes, and donor cell fibroblast were also examined by PCA and hierarchical clustering. As was noted for miRNAs, the profiles of tRFs appeared to group more closely by embryonic stage, with oocytes, 2-cell and 8-cell embryos clustered near each other (Figure 3.7A). Most morula samples were intermediate, between 8-cell and blastocyst embryos, though two samples appeared to cluster closely with BDCs. Of note, IVF and scNT embryos were not segregated, even within developmental stage. These observations were confirmed in the hierarchical clustering of samples, which revealed distinct tRF profiles for blastocysts, BDCs and fibroblasts separated from earlier developmental stages including oocytes, 2-cell, 8-cell and most morula embryos (Figure 3.7B).

Expression of piwi-like RNAs in bovine IVF and scNT samples

In addition to miRNA analysis, the RNAseq data were mined to examine expression of additional sncRNA classes, specifically piwi-interacting RNA (piRNA). Putative "piwi-like", or pilRNAs, were identified as sncRNAs 24-32 nt in length with a classic ping-pong signature and

1U bias (Supplementary Figure 3.S7) and that were mapped to pilRNA-producing loci in the bovine genome. As shown in Figure 3.8A, pilRNAs accounted for only 0.1%, 0.02% and 0.02% of sequence reads from fibroblasts, and IVF BDCs, or scNT BDCs – all somatic cell types. A much greater fraction of sncRNA was identified as pilRNAs in oocytes (47%), 2-cell embryos (59%), and 8-cell embryos (48%), with no clear distinctions between IVF and scNT embryos; a substantial portion of these reads were mapped to piRNA clusters (Figure 3.8B). Interestingly, post-EGA, a substantial decrease in pilRNAs was noted for morula and blastocyst embryos (4% or 2%, respectively), with a small fraction mapped to piRNA clusters and overall fewer piRNA clusters identified in these sample types. Secondary piRNA biogenesis occurs via the "ping-pong" amplification cycle, which can be detected as a 5' to 5' 10 nt overlap. The ping-pong signature was detected in all embryo sample types as shown in Supplementary Figure 3.S6, with lower peaks in the blastocyst stage embryos; no signature was detected in the somatic fibroblast or BDC samples.

The pattern of differential expression between IVF and scNT embryos occurring only at the morula stage embryo was maintained in pilRNA, with four significantly different pilRNAs identified at the morula stage embryo, 10 in BDC samples, and 0 at any other stage (Table 3.3, Supplementary File 3.3; Appendix H). When examining piRNA expression by developmental stage, many more differentially expressed piRNAs were noted, such as 49 altered piRNAs when comparing 8-cell to morula or 70 when comparing morula to blastocyst stages. Also, 24 piRNAs were differentially expressed between 8-cell and morula scNT embryos but were not different when comparing morula to blastocysts. Similar differences were seen between scNT and IVF 2-cell embryos and fibroblasts, at 4,672 and 4,586 differentially expressed piRNA respectively. These high differences are expected, as early embryos require high piRNA expression and differentiated fibroblasts should have no piRNA expression. However, there was high disparity in the differences between blastocyst stage embryos and BDC for IVF and scNT types, at 1,766

and 717 differentially expressed piRNAs respectively. These differences do suggest some aberrant piRNA reprogramming in scNT embryos, which may be too variable for statistical significance when directly comparing IVF and scNT within stage.

PCA of pilRNA expression revealed a similar pattern of sample segregation as was noted for other sncRNA classes, with clusters evident by developmental stage (Figure 3.9A). The tight clustering in the PCA plot indicates fairly consistent pilRNA expression for sample types, which was also evident in the hierarchical cluster analysis (Figure 3.9B). Expression profiles for pilRNAs were markedly different by developmental stage, with exceedingly low expression in differentiated cell types (fibroblasts and BDCs), high expression in early developmental stages (most 2-cell and some 8-cell embryos) and oocytes, and more moderate expression in some 8-cell embryos, morulae and blastocysts. No clear separation by IVF or scNT sample type was apparent in the hierarchical cluster heat map. As for other sncRNAs, two IVF morula staged embryos clustered separately from the other samples in their stage. Interestingly, we noted that morula embryos exhibited more variable expression and less tight clustering for all the sncRNA classes examined, including miRNAs, piRNAs, and tRFs.

Primary piRNAs are transcribed from the dense loci that are located in specific clusters in the genome. Supplementary Figure 3.S8 depicts the total number of pilRNAs mapped to each cattle chromosome. Using a weighted proTRAC cluster analysis, the somatic cellular samples (fibroblast cells and BDCs) were typified by very few clusters that were only expressed in a few loci in the cattle genome (Figures 3.10-11). The oocyte samples were typified by broad expression of piRNA clusters across the genome, as were both IVF and scNT 2-cell staged embryos and IVF and scNT 8-cell staged embryos (Figure 3.10). The IVF morula staged embryos were typified by few piRNA clusters, and the scNT morula staged embryos were typified with cluster expression falling between the 8-cell staged embryos and the IVF morula staged embryos with broader cluster expression then the IVF counterpart (Figure 3.11). Both IVF

and scNT blastocyst staged embryos were typified by a similar low cluster expression as compared to early embryos, but with the IVF counterpart demonstrating higher cluster expression (Figure 3.11).

Potential target transcripts of candidate pilRNAs were identified by mapping sequences to bovine TEs annotated by RepeatMasker (Supplementary File 3.6; Appendix K). Figure 3.12 depicts mapping of sequence reads to the most targeted TEs in the cattle genome. We noted relatively high numbers of sequence reads mapped to TEs in oocytes, 2- and 8-cell embryos in both IVF and scNT sample types (Figure 3.12). Also, there appeared to be a notable increase in reads mapped to 2-cells of scNT embryos as compared to oocytes. Coincident with the EGA, reads mapped to TEs markedly decreased across all TEs for both IVF and scNT embryos. Interestingly, no reads mapped to TE in BDCs obtained from IVF embryos (Figure 3.12A), which contrasted with a modest number of reads that mapped to TEs in scNT embryos, most notably for RTE-BovB, L1, ERVK and TcMar-Tigger elements (Figure 3.12B).

Discussion

This study is the first to employ a discovery-based, small RNA sequencing approach to determine the differences in sncRNA profiles for cattle embryos produced by IVF or scNT. In addition, small quantities of biological sample were successfully utilized for sncRNA sequencing, demonstrating the possibilities of utilizing these discovery-based approaches with the production of reasonable numbers of scNT embryos. Previously, we employed RNAseq in IVF embryos and reported that the MET was associated with major shifts in relative abundance of several classes of sncRNA, including miRNA, piRNA, tRFs and snoRNA. In the present study, we anticipated that sncRNA sequencing of scNT embryos would reveal marked differences in profiles of these sncRNAs as compared to those produced by IVF, particularly at the activation of the embryonic genome, and that such changes may explain the higher rate of developmental failure for scNT

embryos. By virtue of their function in RNAi, aberrant expression of miRNA or tRFs could lead to aberrant control of gene expression – such as failure to degrade maternal transcripts during the MET – and ultimately lead to poor embryonic development. However, in this study, we identified few significant differences in expression of miRNA or tRFs when comparing expression profiles of IVF and scNT embryos. PCA and hierarchical clustering analyses showed that the overall profiles of sncRNAs were not markedly different for IVF and scNT embryos for any particular developmental stage, though clear distinctions were apparent when developmental stages were compared sequentially after the EGA. Moreover, we noted that sncRNA profiles were markedly variable within sample types, particularly for sncRNAs expressed at low to moderate levels.

By virtue of their diverse array of transcript targets, miRNAs influence many cellular functions and are known to be important in the process of tissue differentiation. In this study, abundance of miRNAs was relatively low in oocytes and embryos prior to the MET (2-cell), whereas expression of miRNAs increased substantially post-EGA in morula and blastocyst embryos. Moreover, we also noted that reads of approximately 60 to 80 nt in length were increased in BDCs (and to a lesser extent in blastocysts), but not fibroblast donor cells. These read lengths correspond to snoRNAs (70 to 90 nt) and tRNAs (76 to 90 nt), which may be important for early steps in cell differentiation.

As RNAseq allows for an unbiased assessment of the entire population of miRNA, our analysis afforded the opportunity to address two questions: *What miRNAs are expressed in a sample?* and *Are those miRNAs differentially expressed when comparing specific sample types?* Of the several hundred-miRNA molecules identified, we were intrigued to note that all of those expressed in oocytes were also present in scNT 2-cell embryos. Also, all miRNAs expressed in 8-cell embryos were also expressed in early developing scNT 2-cell embryos. However, for IVF embryos, the pattern was notably different, with oocytes and 8-cell embryos having 76 or 153

unique miRNAs, respectively, which were not common to 2-cell embryos. One would expect the 2-cell scNT embryo to harbor more miRNAs, as those cells would still contain miRNAs from the donor fibroblast cell and the oocyte. However, further comparison of scNT 2-cell embryos to fibroblast cells or oocytes with stage-matched IVF embryos show that the scNT 2-cell embryos express all of the miRNAs that were identified in fibroblasts, oocytes and IVF embryos. The lack of a distinct population of miRNA molecules for IVF 2-cell embryos not expressed in fibroblasts, oocytes or scNT 2-cell embryos was surprising, as sperm sncRNA contributions would have been expected to be a unique population within the IVF 2-cell embryo [92]. However, at onset of the EGA and activation of the embryonic genome, unique sets of miRNAs were apparent for IVF 8-cell embryos, as well as their scNT counterparts, a trend that continued through the remainder of development. Also of note, post EGA, the scNT embryos did not share an outsized fraction of their miRNA pool with fibroblasts as compared to the IVF embryos. Had the donor cell miRNA population persisted through development of scNT embryos, then one might expect a markedly greater fraction of miRNAs to be shared with fibroblasts in scNT embryos than IVF embryos. The greatest difference was noted at the blastocyst stage, with 71 miRNA molecules shared between scNT blastocysts and fibroblasts as compared to 52 shared between IVF embryos and fibroblasts.

Expression of miRNAs miR-2340, miR-345-5p, and miR-34a was significantly lower in scNT morula stage cattle embryos, the only stage for which differences in miRNA expression between scNT and IVF embryos was observed. However, it is worth noting that while statistical significance was reached for these three miRNAs at the morula stage, the morula stage was the most highly variable of the sample types. Future functional validation by over-expression in IVF embryos and blocking expression in scNT embryos may provide confirmation of their importance in embryonic development. Ontology analyses for predicted target genes of these three miRNAs indicate functions to regulate cell proliferation, stabilize proteins, and regulate development of

several types of tissues. MiR-2340 has no known homologues to its seed sequence in any species as reported in MirBase, and thus may be unique to cattle. However, miR-2340 has been reported to be a miRNA present in bovine oocytes, and is differentially expressed through oocyte maturation [93]. While the function of miR-2340 in bovine embryos is not known, it is a predicted target of two circular RNA (circRNA) that were observed as abnormally expressed in bovine scNT placentas, specifically bta_circ_0006612 and bta_circ_0026700 [94]. CircRNAs sequester miRNAs like a sponge, and thus function as competitive inhibitors that suppress binding of miRNAs to their mRNA targets. Although this RNAseq analysis did not capture circRNA sequences, one might speculate that up regulation of these specific circRNAs in scNT embryos could be linked to lower expression of miR-2340.

The other two differentially expressed miRNAs, miR-34a and miR-345-5p, both have homologues in mice and humans, which allows for some inference on their possible functions in cattle based on available evidence from literature on other mammalian embryonic developmental models. MiR-34a may contribute to folliculogenesis and the ovarian cycle, as this miRNA is expressed in bovine granulosa cells and its over expression in human granulosa cells inhibits estradiol release [95]. In pig ovaries, miR-34a was found to target the *Inhibin beta B* gene, promoting granulosa cell apoptosis [96]. In bovine corpus luteum, miR-34a modulated luteal formation and function through regulation of cell proliferation and progesterone production pathways [97]. In addition to regulating oocyte development via influencing follicular development, miR-34a may also contribute more directly to early development of the embryo and acquisition of pluripotency. MiR-34a is present in bovine spermatozoa at low levels, as well as higher levels in oocytes and 2-cell cleaved embryos [98]. Because miR-34a is stably expressed in developing oocytes and through preimplantation development embryos, Tscherner and colleagues speculated that this miRNA may be functionally active during that developmental window, or that a reservoir of miR-34a may be needed in later embryonic stages. However, this group did not

measure miR-34a expression beyond embryonic cleavage [98]. In our study, miR-34a expression was generally low in MII oocytes as well as IVF and scNT 2-cell and 8-cell embryos, though a notable increase in expression was observed at the morula stage for IVF embryos compared to their stage matched scNT counterparts. For both embryo types, expression was increased in blastocysts and BDCs compared to the earlier development stages. Reports in other species point to miR-34a contributing to pluripotency and cell differentiation. For example, Choi et al. reported that miR-34a suppressed the transcription factor *GATA2* in mouse pluripotent stem cells and restricted developmental potential. Knockout of miR-34a allows generation of both embryonic and extra-embryonic lineages from pluripotent stem cells, expanding their developmental potential [99]. Furthermore, members of the miR-34 miRNA family are direct transcriptional targets of the tumor suppressor protein p53, and thus are critical mediators of p53-mediated control of genome reprogramming [100]. Suppression of reprogramming by miR-34a has been partially attributed to its targeting and suppression of key pluripotency factors Nanog, Sox2 and Myc. Thus, elevated expression of miR-34a, as was noted in bovine BDCs in this study, may facilitate early stages of differentiation by turning off these pluripotency signals. Studies of cancer cell models can provide useful insight on the function of miRNAs, as many pathways dysregulated in cancer are those also critical for early development. In thyroid cancer, miR-34a works with the long non-coding RNA (lncRNA) *XIST* to modulate cell proliferation and tumor growth, as *XIST* serves as a competing endogenous RNA (ceRNA) for miR-34a [101]. The tumor suppressor protein P53 is frequently mutated in human cancers, and regulates miRNAs including the miR-34 family at high prevalence. MiR-34a is suppressed in many cancers suggesting that this miRNA may act as a tumor suppressor for many molecular pathways under normal conditions [102]. Overall, miR-34a function in diverse pathways in both cancer pathways and pluripotency, and in-depth investigation to possible functions in early development is warranted.

While a review of the available literature does not provide evidence for miR-345 having a critical regulatory role in early development, this miRNA has been detected in some tissues of importance for development. In human embryos, miR-345 was differentially expressed in euploid embryos as compared to aneuploid embryos [103] and was also detected in a mesenchymal stem cell (or a multipotent stromal cell) line derived from human embryonic stem cells [104]. Interestingly, hsa-miR-345-5p is differentially expressed in the tongues of people affected by Beck-Wiedemann syndrome (BWS). BWS is a human disease that is notably similar to large offspring syndrome (LOS) in cattle, and LOS occurs at higher frequency in scNT-produced calves [105]. Although the function of miR-345-5p in early development is not known, studies in various cancer cell lines point to its importance in regulating cell behavior. For example, in human non-small cell lung cancer cells, a recent report showed than miR-345 suppressed cell invasion and migration by targeting and down-regulating the oncogenic transcriptional co-activator *Yap1* [106]. Also, expression of miR-345 was also suppressed in pancreatic cancer, and over-expression of miR-345 leads to increased apoptosis through targeting of *Bcl2* [107]. In colorectal cancer, miR-345 is sensitive to methylation-dependent regulation and modulates cell proliferation and invasion through targeting the anti-apoptosis protein *Bag3* [108]. ScNT embryos have been observed to maintain a higher level of genome wide methylation levels than their IVF counterparts, and *de novo* DNA methylation begins earlier in development [8, 53]. If miR-345 is regulated via DNA methylation, aberrant expression in scNT embryos may be due to aberrant methylation marks in scNT embryos. The function of miR-345-5p in the development early embryonic cells in unknown, and the difference in expression pattern between scNT and IVF morula staged embryos warrants further study to explore any potential mechanism of action.

In this study, we also explored sncRNA profiles for BDCs and identified a substantial set of 24 differentially expressed cattle miRNAs for BDCs generated from scNT embryos as compared to those obtained from IVF embryos. Interestingly, the mRNA targets of those

differentially expressed miRNAs were associated with biological processes directly related to embryo development and growth as well as protein and mRNA metabolic processes. That BDCs from IVF and scNT embryos harbored distinct differences in miRNA profiles whereas IVF and scNT blastocysts did not was an important observation. It is also likely that only those embryos with a normal profile of sncRNAs would progress to the blastocyst developmental stage, and those embryos harboring abnormal profiles would arrest prematurely. As such, the profiles of sncRNA for scNT blastocysts likely reflect those embryos that survived the developmental challenge of the MET and, thus, were more similar to their IVF counterparts.

Thus, it was interesting to note that many sncRNAs were differentially expressed in BDCs from those scNT embryos that did survive to the blastocyst stage, suggesting that subtle aberrations at the blastocyst stage may have been amplified through prolonged culture or may have appeared as differentiation progressed. Also of note, the biological processes associated with predicted mRNA targets of differentially expressed miRNAs in BDCs were more directly connected to embryonic development and cell differentiation than were the terms for mRNA targets of the three differentially expressed miRNAs identified in morula stage embryos.

Other researchers have identified aberrant patterns of miRNA expression associated with cloning, such as dysregulation of miRNAs in placental tissues from scNT pregnancies [57] or dysregulation of epigenetic reprogramming controlling miRNA expression in blastocyst scNT embryos [58]. Both of these studies used a targeted microarray approach that may not have captured all miRNAs of developmental importance. By employing an RNAseq, discovery-based approach, we expected to capture the full repertoire of miRNAs expressed in early developing embryos, and those differentially expressed in clones compared to IVF controls. Although hundreds of cattle miRNAs were sequenced in this study, very few miRNAs were identified as differentially expressed in scNT embryos compared to their stage-matched IVF counterparts and widespread differences were not apparent. Underlying the overarching hypothesis is the

presumption that errors in genome programming that ultimately control miRNA expression – particularly through the MET – would follow a pattern that could be discerned through statistical evaluation of the sequencing data. However, the alternative hypothesis is that such errors are, in fact, random in nature leading to high variability in genomic reprogramming of scNT embryos. Moreover, the process of collecting embryo samples could introduce some biologically based bias, as only competent embryos can progress through each developmental stage. Those scNT embryos harboring extensive aberrations may undergo developmental arrest or degrade, and thus, not be included in the later developmental stages. In addition, embryo samples were pooled to produce sufficient biological material, which could mask individual variability or outliers. To tackle the issue of biological variability and attempt to address the alternative hypothesis, future studies could attempt single cell or single embryo RNA sequencing to build libraries from a very small number of cells.

In addition to the comparisons between IVF and scNT embryos at each stage of development, we also examined patterns of miRNA expression over the course of embryo development for IVF and scNT embryos separately to understand the dynamics of miRNA expression in context of embryonic gene activation. First, no significant differences in miRNA expression were apparent for either IVF or scNT embryos when comparing oocytes to the 2-cell stage or the 2-cell to the 8-cell developmental stage, suggesting that the population of miRNAs was generally stable up to the onset of the EGA. In IVF embryos, the transition from 8-cell to morula stage embryos was associated with significant differences in the expression of 22 cattle miRNAs, including the same miRNAs discussed above that were differentially expressed in IVF and scNT morulae (miR-2340, miR-345-5p, and miR-34a). Other significantly different miRNAs associated with the EGA included miR-93-5p, miR-320a, miR-378, and miR-450a. Interestingly, miR-320 has been found to be *DGCR8*-independent in mouse embryonic stem cells [109], and is thought to be important for development and embryo quality in mice and humans [110]. MiR-

320 was also observed as highly expressed in cattle oocytes as well as female bovine embryos [47, 111]. Evidence from studies in cancer cells indicates that miR-320a functions to down-regulate the Wnt signaling pathway [112]. Similar activity in embryos could markedly impact early embryo development, given the importance of Wnt signaling. MiR-378c may target the transcript *Nodal* to promote survival and migration of trophoblast cells [113]. In mice, miR-450a has been shown to regulate early development through the targeting of *BUB1* to repress cell proliferation [113]. The miR-17 family includes miR-93-5p, which was shown to be expressed during early development and to regulate *STAT3*, a known embryonic stem cell regulator [114]. Thus, miR-93-5p expression may influence pluripotency.

In scNT embryos, the transition from 8-cell to the morula stage was associated with differential expression of 19 cattle miRNAs, excluding three that were regulated in IVF embryos (miR-2340, miR-345-5p, and miR-34a). Of those differentially expressed miRNAs in scNT morulae, miR-302 and miR-7 are known to be important in development. MiR-7 is expressed in the small intestine of bovine fetuses and modulates the development of the gastrointestinal tract [115]. Bick et al. determined that miR-7 was enriched in porcine blastocysts [116], while others showed that miR-7 was involved in trophoblast differentiation in mice [117]. MiR-302 is a bovine and human orthologue to the zebrafish miR-427, which was responsible for degrading hundreds of maternal transcripts during the MZT [118]. Also, miR-302 has been shown to target *Akt1*, which maintains high levels of *Oct4* and, therefore, improves self-renewal and pluripotency in humans [119]. These roles in pluripotency pathways would be vital in early embryo development, and miR-302 knock-out (in conjunction with miR-290) in mice silences naïve pluripotency [120].

In IVF embryos, the transition from morula to blastocyst was associated with changes in 19 cattle miRNAs. This set includes miR-2340 and miR-345-5p, which were apparently decreased at the blastocyst stage, suggesting that these miRNAs may have a specific function

associated with the morula developmental stage. Other notable miRNAs differentially expressed over the transition from morula to blastocyst included miR-196a and miR30c/e/f. Tripurani and colleagues observed maximal expression of miR-196a at the 4-8 cell developmental stage in cattle embryos [51]. They also noted that miR-196a bound to the *Nobox* transcript, which must be degraded prior to the MET. In this study, miR-196a expression was elevated at the blastocyst stage compared to earlier developmental stages, including 2-cell and 8-cell embryos, an observation that suggests this miRNA may have a secondary function in embryo development aside from degrading maternal effect genes. Lin et al. [121] showed that miR-30c was secreted from bovine embryos, more so from slow cleaving embryos, suggesting this miRNA may serve as a biomarker of developmental competence. Moreover, a recent review by Mao et al. [122] highlights the important regulatory functions of miR-30 family members in tissue and organ development and disease pathogenesis, pointing to the importance of this miRNA family throughout life.

In scNT embryos, the transition from morula to blastocyst developmental stage was associated with changes in 33 cattle miRNAs, including miR-34a, which was not different for this transition in IVF embryos. This set of miRNAs includes miR-106b, miR-15a, miR-145-5p and miR-451. Interestingly, miR-451 is a miRNA that can mature independent of *Dicer*, and instead is cleaved directly by *Ago2* [123], and bovine miR-451 can be measured in circulation to identify early pregnancy [124] as early as day 8 of pregnancy. In mice, miR-451 modulates implantation [125], and impacts cell fate in mouse embryonic stem cells [126]. MiR-145 influences the attachment of embryos in mice [127] and also helps regulate initiation, development, and maintenance of mouse primordial follicles [128]. MiR-145 also targets *Lif* in bovine granulosa cells in the follicle, and controls proliferation of primordial follicles through targeting *Acvr1b* [128]. *Lif* plays multiple important roles in implantation, including decidualization, blastocyst growth, and development. Thus, targeting of *Lif* by miR-145 in early development may interfere

with pregnancy [129]. MiR-145 was also found to be most highly expressed in 8-cell staged bovine embryos, with relatively lower expression in pre and post MET [128]. In contrast, in our study, miR-145-5p expression was greatest in morula IVF embryos compared to earlier stages or blastocysts, whereas expression of miR-145-5p was relatively constant from the 2-cell to morula stages in scNT embryos with a decline in blastocysts. The miR-17 family includes miR-106b, which regulates *Stat3*, a known embryonic stem cell regulator, and thus this miRNA molecule may influence pluripotency in early development [114]. By targeting transcripts of the pro-apoptotic gene *Bcl2*, miR-15 regulates apoptosis in cancer cells [130]. In mouse ESCs, miR-15 also has been shown to control the cell cycle through targeting *KLF4,* which causes down-regulation of *CYCLIN E.*

During early development, the parental epigenome must be reprogrammed for the re-establishment of cellular potency, and the required removal of epigenetic silencing marks can then make the embryo susceptible to TE activation. If in scNT embryos the process of piRNA activation is delayed or reduced, scNT embryos may become vulnerable to TE activation, reduced genomic integrity, and impaired viability [35]. In addition, the expression of more than 50% of the genome in mammals is controlled by DNA methylation or repressive chromatin marks, and releasing the TEs from these constraints could allow mobilization of these TEs [131]. However, because TEs are also required in a developmental context, complete silencing would be detrimental to the developing embryo. Researchers have shown that there is some functional requirement for TE expression during early development in mice. For example, suppression of L1 ORF1 increases the rate of embryo arrest [132], and MuERV-L knockout causes embryo arrest at the 4-cell stage [133]. Also, LINE1 expression is required for early embryogenesis, but then must be silenced by the blastocyst stage for successful development [134]. Bui et al. [33] employed a cDNA array to compare transcripts in normal IVF and scNT embryos during the MTZ and compared those profiles to donor cells. They found that LTR retrotransposons and

mitochondrial transcripts up-regulated and ribosomal proteins were down regulated [33], suggesting that specific categories of transcripts were impacted during somatic reprograming which likely affected the viability of scNT embryos.

In this study, while four pilRNAs were identified as differentially expressed in scNT morula embryos compared to their IVF counterparts, large scale differences in embryo types were not evident for this class of sncRNA. It is interesting to note, however, that this developmental stage just following the EGA in cattle was associated with the most differentially expressed sncRNAs, including miRNAs, tRFs, and pilRNAs and tRFs. Also of note, greater fractions of pilRNAs were mapped to clusters in scNT 8-cell embryos. If piRNAs are aberrantly highly expressed at the MET, it is possible that they are repressing TEs that are necessary for early development and successful EGA. We also noted variation in the ping-pong signature for IVF and scNT embryos, with an apparent increase in the number of sequence reads containing a strong ping-pong signature for scNT 2-cell embryos as compared to IVF. Indeed, the ping-pong signature for 2-cell scNT embryos was much more like that of the later 8-cell stage, for which the signature was consistent for IVF and scNT embryos. Because the ping-pong signature typifies piRNAs generated via the ping-pong pathway that are destined to target active transposon sequences, the difference in the relative abundance of sequences with this signature at the 2-cell stage suggests the possible early activation of TEs in scNT embryos. Another possibility would be differing sncRNA in scNT embryos from the nuclear donor could create a ping-pong cascade in the early embryo, but this pattern was not evident in our differential expression analyses.

As with other sncRNA classes, much more pronounced differences were noted when comparing populations of pilRNAs across developmental stages. Remarkably, pilRNA expression was more consistent within stage and much more distinct among stages than was noted for miRNA or tRFs. The observation of extreme changes in piRNA expression as a function of embryonic development aligns with the putative role for piRNA protection of the

genome during periods of remodeling. High expression of piRNAs would be required in oocytes and early embryos up to the 8-cell stage coincident with the EGA, and expression would decline as cells differentiated and virtually disappear in more differentiated cell types, such as the BDCs and fibroblasts. In agreement with this model, we observed much higher fraction of reads that were potential pilRNAs in 2-cell and 8-cell stage embryos, with significant loss of pilRNAs at the morula stage, even fewer at the blastocyst stage, and none in BDCs or fibroblasts. A very similar pattern was seen for the number of piRNA clusters, as well as the location of clusters transcribed from different bovine chromosomes, with high numbers and locations of clusters observed in pre-MET embryos, a decrease at the morula and blastocyst stages, and ultimately very few noted in BDCs and fibroblasts. Given the persistent expression of some pilRNAs through the morula stage, it is possible that these may continue to function beyond the completion of the MET to constrain activity of TEs that function later in development, such as the differentiation of trophoblast cells [135], or may be involved in other essential functions, such as renewal of germline stem cells and support of self-renewal and differentiation [136].

While the PIWI pathway has been clearly identified as a mechanism for controlling endogenous transposable elements during genomic remodeling periods, some recent research has implicated the PIWI pathway in the regulation of mRNA transcripts. PiRNAs are known to be involved in spermatogenesis, but a unique pool of pilRNAs was found to be present in oocytes and zygotes that appear to have the potential for endogenous gene expression regulation in a similar manner to miRNA RNAi pathways [35]. Russell et al. found that piRNAs that were present in the bovine zygote were sequentially complementary to mRNA destined for turnover in the early bovine embryo. Our results support the hypothesis that pilRNAs are deposited during oogenesis and remain in the early embryo to regulate TEs that arise during embryonic reprogramming. Our data suggests that the pilRNA pathway may participate in TE repression during the MET, as the pilRNA populations in 8-cell stage were associated with a very high ping-

pong signature, suggesting activation of TEs and associated activation of targeting piRNA.
Interestingly, when the magnitude of change between pilRNA and predicted target mRNA were
analyzed by Russell et al, there was low correlation between the magnitude of change and the
number of targeting pilRNAs, similar to what we observed with miRNA [35].

Our results did not completely match past bovine embryo piRNA data, although there
were patterns of similarity. For example, Russell et al found that the highest pilRNA expression
was noted for clusters on chromosomes 6, 14, 17, and 23 in immature oocytes [35]. While we
only sampled mature oocytes, higher expression for pilRNAs were noted for clusters on
chromosomes 6, 14, 17 and 23, although not to the same extent seen previously. Interestingly, we
noted that scNT morulae tended to have many more pilRNAs mapped to these clusters on
chromosomes 6, 17 and 23 compared to their IVF counterparts. While some differences in
patterns of pilRNA clusters were noted for some chromosomes at other stages, the consistency of
the pattern in scNT morula staged embryos across the bovine genome seems likely to be of
functional importance. It appears that in IVF embryos, piRNA expression before and during the
MET is high, as the genome is reprogrammed. The scNT embryos appear to maintain aberrantly
high levels of piRNA after the EGA at these loci, which could be the result of aberrant
reprogramming during the MET or could be attributed to faulty control of piRNA expression and
may act as a stress response. Based on these data, one may hypothesize that these persistently
expressed piRNAs may interfere with the normal role of TEs in early development, thus
contributing to low efficiency of scNT embryo production.

When examining the dynamic expression of sncRNA for IVF embryos over the course of
early embryo development, we noticed more variability among samples in the present study than
was observed in our prior work on IVF 8-cell and blastocysts embryos and oocytes [45]. These
differences may be explained by some methodological differences in generation and culture of
the embryos. First, although both studies utilized the same semen source for IVF, in the present

study, oocytes were collected from various cattle breeds at differing ages and originating from multiple locations at the local abattoir over the course of two years, as opposed to a short three-week period for the 2019 report. It is possible that the longer time frame required to generate the much greater number of embryos for this study introduced some variability. However, these embryos were carefully allocated to sample pools so that each pool included embryos from a minimum of 3 cloning sessions spread across the entire collection period. Another difference was the number of embryos allocated to each sample pool, 20 in this study compared to 40 in the 2019 report. Importantly, embryos for this study were cultured in SOFaa medium with oil-covered medium droplets on top of a cumulus cell feeder layer to improve development, whereas the 2019 study used Charles Rosenkrans 2 medium (CR2) with no oil and no cumulus cells. While group culture was utilized in both studies, the embryos were cultured in a smaller volume of medium in the present study. Thus, secreted factors, such as miRNA or growth factors, would have been at a much higher concentration in the cell culture media, potentially impacting variability of those embryos in a greater manner than occurred in the previous study. These secreted factors have been shown to impact development rates [137]. Moreover, cumulus cells in cattle express highly variable populations of miRNAs, which may impact embryo development, although they do not appear to affect oocyte quality [138, 139]. It is possible that the use of cumulus cells as feeder cells to support embryo development introduced an element of variability that was manifest in the sncRNA profiles of the cultured embryos. Last, we did note variability in sncRNA profiles among the fibroblast samples that was unexpected, given that all samples were from the same cell line at similar passage number (5 to 7). The aim of this study was to quantify known sncRNA rather than to identify novel sncRNAs in the bovine cattle embryo. However, it should be noted that, there appeared to be a population of reads that did not align to known sncRNAs especially in embryo samples. It is possible that these reads represent novel sncRNAs, which may be explored in future work.

In conclusion, results of this discovery-based sncRNA sequencing analysis of early developing embryos revealed largely similar profiles of sncRNAs for IVF and scNT embryos at the 2-cell, 8-cell, morula and blastocyst stage of development. Alternatively, clear distinctions were apparent when comparing sncRNA profiles by developmental stage that largely corresponded to a pre-MET phase, post-MET phase and cultured cells with some degree of differentiation. SncRNAs are known to be important regulators of pluripotency and differentiation of tissues. However, their specific functions in mediating the MET in mammalian early development remains unclear. Our team is the first to examine the dynamics of sncRNA populations through the MET in both scNT and IVF embryos. Further research will be necessary to assess the function of those miRNAs identified as differentially expressed in scNT embryos as compared to their IVF counterparts, or those miRNAs identified as differentially regulated over the course of embryo development. Furthermore, by exploring changes in the transcriptome coincident with dynamic miRNA expression, we may gain additional insight into the function of miRNAs in early embryonic development.

References

1. Wilmut I, Beaujean N, de Sousa PA, Dinnyes A, King TJ, Paterson LA, Wells DN, Young LE. Somatic cell nuclear transfer. Nature. 2002;419(6907):583-586.

2. Solter D. Mammalian cloning: advances and limitations. Nat Rev Genet. 2000;1(3):199-207.

3. Matoba S, Zhang Y. Somatic Cell Nuclear Transfer Reprogramming: Mechanisms and Applications. Cell Stem Cell. 2018;23(4):471-485.

4. Cezar GG, Bartolomei MS, Forsberg EJ, First NL, Bishop MD, Eilertsen KJ. Genome-wide epigenetic alterations in cloned bovine fetuses. Biol Reprod. 2003;68(3):1009-1014.

5. Kang YK, Park JS, Koo DB, Choi YH, Kim SU, Lee KK, Han YM. Limited demethylation leaves mosaic-type methylation states in cloned bovine pre-implantation embryos. EMBO J. 2002; 21(5):1092-1100.

6. Mann MR, Chung YG, Nolen LD, Verona RI, Latham KE, Bartolomei MS. Disruption of imprinted gene methylation and expression in cloned preimplantation stage mouse embryos. Biol Reprod. 2003;69(3):902-914.

7. Han YM, Kang YK, Koo DB, Lee KK. Nuclear reprogramming of cloned embryos produced in vitro. Theriogenology. 2003;59(1):33-44.

8. Bourc'his D, Le Bourhis D, Patin D, Niveleau A, Comizzoli P, Renard JP, Viegas-Pequignot E. Delayed and incomplete reprogramming of chromosome methylation patterns in bovine cloned embryos. Curr Biol. 2001, 11(19):1542-1546.

9. Kang YK, Koo DB, Park JS, Choi YH, Chung AS, Lee KK, Han YM. Aberrant methylation of donor genome in cloned bovine embryos. Nat Genet. 2001;28(2):173-177.

10. Matoba S, Liu Y, Lu F, Iwabuchi KA, Shen L, Inoue A, Zhang Y. Embryonic development following somatic cell nuclear transfer impeded by persisting histone methylation. Cell. 2014;159(4):884-895.

11. Santos F, Zakhartchenko V, Stojkovic M, Peters A, Jenuwein T, Wolf E, Reik W, Dean W. Epigenetic marking correlates with developmental potential in cloned bovine preimplantation embryos. Curr Biol. 2003;13(13):1116-1121.

12. Shi W, Haaf T. Aberrant methylation patterns at the two-cell stage as an indicator of early developmental failure. Mol Reprod Dev. 2002;63(3):5.

13. Schultz RM. The molecular foundations of the maternal to zygotic transition in the preimplantation embryo. Hum Reprod Update. 2002;8(4):323-331.

14. Gotze M, Wahle E. Smaug destroys a huge treasure. Genome Biol. 2014;15(1):101.

15. Ohnishi Y, Totoki Y, Toyoda A, Watanabe T, Yamamoto Y, Tokunaga K, Sakaki Y, Sasaki H, Hohjoh H. Small RNA class transition from siRNA/piRNA to miRNA during pre-implantation mouse development. Nucleic Acids Res. 2010;38(15):5141-5151.

16. Carthew RW, Sontheimer EJ. Origins and mechanisms of miRNAs and siRNAs. Cell 2009;136(4):642-655.

17. Rak R, Dahan O, Pilpel Y. Repertoires of tRNAs: The couplers of genomics and proteomics. Annu Rev Cell Dev Biol. 2018;34:239-264.

18. Lee YS, Shibata Y, Malhotra A, Dutta A. A novel class of small RNAs: tRNA-derived RNA fragments (tRFs). Genes Dev. 2009;23(22):2639-2649.

19. Kiss T. Small nucleolar RNAs: an abundant group of noncoding RNAs with diverse cellular functions. Cell. 2002;109(2):145-148.

20. Kufel J, Grzechnik P. Small nucleolar RNAs tell a different tale. Trends Genet. 2019;35(2):104-117.

21.   Scott MS, Ono M. From snoRNA to miRNA: Dual function regulatory non-coding RNAs. Biochimie. 2011;93(11):1987-1992.

22.   Taft RJ, Glazov EA, Lassmann T, Hayashizaki Y, Carninci P, Mattick JS. Small RNAs derived from snoRNAs. RNA. 2009;15(7):1233-1240.

23.   Williams GT, Farzaneh F. Are snoRNAs and snoRNA host genes new players in cancer? Nat Rev Cancer. 2012;12(2):84-88.

24.   Ramchandran R, Chaluvally-Raghavan P. miRNA-mediated RNA activation in mammalian cells. Adv Exp Med Biol. 2017;983:81-89.

25.   Clarke S, Sunkar, R. Biogenesis of Mammalian miRNA. In: MicroRNAs as tools in biopharmaceutical production. Springer Science. 2012. p. 15-27.

26.   Valencia-Sanchez MA, Liu J, Hannon GJ, Parker R. Control of translation and mRNA degradation by miRNAs and siRNAs. Genes Dev. 2006;20(5):515-524.

27.   Fabian MR, Sonenberg N, Filipowicz W. Regulation of mRNA translation and stability by microRNAs. Annu Rev Biochem. 2010;79:351-379.

28.   Winter J, Jung S, Keller S, Gregory RI, Diederichs S. Many roads to maturity: microRNA biogenesis pathways and their regulation. Nat Cell Biol. 2009;11(3):228-234.

29.   Ozata DM, Gainetdinov I, Zoch A, O'Carroll D, Zamore PD: PIWI-interacting RNAs. small RNAs with big functions. Nat Rev Genet. 2019;20(2):89-108.

30.   Parhad SS, Theurkauf WE. Rapid evolution and conserved function of the piRNA pathway. Open Biol. 2019;9(1):180181.

31.   Roovers EF, Rosenkranz D, Mahdipour M, Han CT, He N, Chuva de Sousa Lopes SM, van der Westerlaken LA, Zischler H, Butter F, Roelen BA. Piwi proteins and piRNAs in mammalian oocytes and early embryos. Cell Rep. 2015;10(12):2069-2082.

32.   Vagin VV, Sigova A, Li C, Seitz H, Gvozdev V, Zamore PD. A distinct small RNA pathway silences selfish genetic elements in the germline. Science. 2006;313(5785):320-324.

33.   Bui LC, Evsikov AV, Khan DR, Archilla C, Peynot N, Henaut A, Le Bourhis D, Vignon X, Renard JP, Duranthon V. Retrotransposon expression as a defining event of genome reprogramming in fertilized and cloned bovine embryos. Reproduction. 2009;138(2):289-299.

34.   Gerdes P, Richardson SR, Mager DL, Faulkner GJ. Transposable elements in the mammalian embryo: pioneers surviving through stealth and service. Genome Biol. 2016;17:100.

35. Russell S, Patel M, Gilchrist G, Stalker L, Gillis D, Rosenkranz D, LaMarre J. Bovine piRNA-like RNAs are associated with both transposable elements and mRNAs. Reproduction. 2017;153(3):305-318.

36. Giraldez AJ, Mishima Y, Rihel J, Grocock RJ, Van Dongen S, Inoue K, Enright AJ, Schier AF. Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. Science. 2006;312(5770):75-79.

37. Wu E, Thivierge C, Flamand M, Mathonnet G, Vashisht AA, Wohlschlegel J, Fabian MR, Sonenberg N, Duchaine TF. Pervasive and cooperative deadenylation of 3'UTRs by embryonic microRNA families. Mol Cell. 2010;40(4):558-570.

38. Lund E, Liu M, Hartley RS, Sheets MD, Dahlberg JE. Deadenylation of maternal mRNAs mediated by miR-427 in Xenopus laevis embryos. RNA. 2009;15(12):2351-2363.

39. Bushati N, Stark A, Brennecke J, Cohen SM. Temporal reciprocity of miRNAs and their targets during the maternal-to-zygotic transition in Drosophila. Curr Biol. 2008;18(7):501-506.

40. Much C, Auchynnikava T, Pavlinic D, Buness A, Rappsilber J, Benes V, Allshire R, O'Carroll D. Endogenous mouse dicer is an exclusively cytoplasmic protein. PLoS Genet. 2016;12(6):e1006095.

41. Marlow F. Maternal-zygotic transition. In: Maternal Control of Development in Vertebrates: My Mother Made Me Do It! San Rafael, CA. Morgan & Claypool Life Sciences. 2010.

42. Svoboda P. Why mouse oocytes and early embryos ignore miRNAs? RNA Biol. 2010;7(5):559-563.

43. Suh N, Baehner L, Moltzahn F, Melton C, Shenoy A, Chen J, Blelloch R. MicroRNA function is globally suppressed in mouse oocytes and early embryos. Curr Biol. 2010;20(3):271-277.

44. Ma J, Flemr M, Stein P, Berninger P, Malik R, Zavolan M, Svoboda P, Schultz RM. MicroRNA activity is suppressed in mouse oocytes. Curr Biol. 2010;20(3):265-270.

45. Cuthbert JM, Russell SJ, White KL, Benninghoff AD. The maternal-to-zygotic transition in bovine in vitro-fertilized embryos is associated with marked changes in small non-coding RNAs. Biol Reprod. 2019;100(2):331-350.

46. Mondou E, Dufort I, Gohin M, Fournier E, Sirard MA. Analysis of microRNAs and their precursors in bovine early embryonic development. Mol Hum Reprod. 2012;18(9):425-434.

47. Berg DK, Pfeffer PL. MicroRNA expression in bovine preimplantation embryos. Reprod Fertil Dev. 2018;30(3):546-554.

48. Sinha PB, Tesfaye D, Rings F, Hossien M, Hoelker M, Held E, Neuhoff C, Tholen E, Schellander K, Salilew-Wondim D. MicroRNA-130b is involved in bovine granulosa and cumulus cells function, oocyte maturation and blastocyst formation. J Ovarian Res. 2017;10(1):37.

49. Goossens K, Mestdagh P, Lefever S, Van Poucke M, Van Zeveren A, Van Soom A, Vandesompele J, Peelman L. Regulatory microRNA network identification in bovine blastocyst development. Stem Cells Dev. 2013;22(13):1907-1920.

50. Tripurani SK, Wee G, Lee KB, Smith GW, Wang L, Jianboyao. MicroRNA-212 post-transcriptionally regulates oocyte-specific basic-helix-loop-helix transcription factor, factor in the germline alpha (FIGLA), during bovine early embryogenesis. PLoS One. 2013;8(9):e76114.

51. Tripurani SK, Lee KB, Wee G, Smith GW, Yao J. MicroRNA-196a regulates bovine newborn ovary homeobox gene (NOBOX) expression during early embryogenesis. BMC Dev Biol. 2011;11:25.

52. Matoba S, Wang H, Jiang L, Lu F, Iwabuchi KA, Wu X, Inoue K, Yang L, Press W, Lee JT et al. Loss of H3K27me3 imprinting in somatic cell nuclear transfer embryos disrupts post-implantation development. Cell Stem Cell 2018, 23(3):343-354 e345.

53. Dean W, Santos F, Stojkovic M, Zakhartchenko V, Walter J, Wolf E, Reik W. Conservation of methylation reprogramming in mammalian development: aberrant reprogramming in cloned embryos. Proc Natl Acad Sci USA. 2001;98(24):13734-13738.

54. Lee YM, Chen HW, Maurya PK, Su CM, Tzeng CR. MicroRNA regulation via DNA methylation during the morula to blastocyst transition in mice. Mol Hum Reprod. 2012;18(4):184-193.

55. Cui XS, Zhang DX, Ko YG, Kim NH. Aberrant epigenetic reprogramming of imprinted microRNA-127 and Rtl1 in cloned mouse embryos. Biochem Biophys Res Commun. 2009;379(2):390-394.

56. Glaich O, Parikh S, Bell RE, Mekahel K, Donyo M, Leader Y, Shayevitch R, Sheinboim D, Yannai S, Hollander D. DNA methylation directs microRNA biogenesis in mammalian cells. Nat Commun. 2019;10(1):5657.

57. Hossain MM, Tesfaye D, Salilew-Wondim D, Held E, Proll MJ, Rings F, Kirfel G, Looft C, Tholen E, Uddin J. Massive deregulation of miRNAs from nuclear reprogramming errors during trophoblast differentiation for placentogenesis in cloned pregnancy. BMC Genomics. 2014;15:43.

58. Castro FO, Sharbati S, Rodriguez-Alvarez LL, Cox JF, Hultschig C, Einspanier R. MicroRNA expression profiling of elongated cloned and in vitro-fertilized bovine embryos. Theriogenology. 2010;73(1):71-85.

59. Kropp J, Khatib H. Characterization of microRNA in bovine in vitro culture media associated with embryo quality and development. J Dairy Sci. 2015;98(9):6552-6563.

60. Liang S, Nie ZW, Guo J, Niu YJ, Shin KT, Ock SA, Cui XS. Overexpression of microRNA-29b decreases expression of DNA methyltransferases and improves quality of the blastocysts derived from somatic cell nuclear transfer in cattle. Microsc Microanal. 2018;24(1):29-37.

61. Li W, Xiong Y, Wang F, Liu X, Gao Y, Wang Y, Zhang Y, Jin Y. MicroRNA-145 inhibitor significantly improves the development of bovine somatic cell nuclear transfer embryos in vitro. Cell Reprogram. 2016;18(4):230-236.

62. Zhang J, Qu P, Zhou C, Liu X, Ma X, Wang M, Wang Y, Su J, Liu J, Zhang Y. MicroRNA-125b is a key epigenetic regulatory factor that promotes nuclear transfer reprogramming. J Biol Chem. 2017;292(38):15916-15926.

63. Rashmi, Sah S, Shyam S, Singh MK, Palta P. Treatment of buffalo (Bubalus bubalis) SCNT embryos with microRNA-21 mimic improves their quality and alters gene expression but does not affect their developmental competence. Theriogenology. 2019;126:8-16.

64. Chang HY, Xie RX, Zhang L, Fu LZ, Zhang CT, Chen HH, Wang ZQ, Zhang Y, Quan FS. Overexpression of miR-101-2 in donor cells improves the early development of Holstein cow somatic cell nuclear transfer embryos. J Dairy Sci. 2019;102(5):4662-4673.

65. Funston RN, Nauta WJ, Seidel GE. Culture of bovine embryos in buffalo rat liver cell-conditioned media or with leukemia inhibitory factor. J Anim Sci. 1997;75(5):1332-1336.

66. Aston KI, Li GP, Hicks BA, Sessions BR, Pate BJ, Hammon D, Bunch TD, White KL. Effect of the time interval between fusion and activation on nuclear state and development in vitro and in vivo of bovine somatic cell nuclear transfer embryos. Reproduction. 2006;131(1):45-51.

67. Aston KI, Li GP, Hicks BA, Sessions BR, Pate BJ, Hammon DS, Bunch TD, White KL. The developmental competence of bovine nuclear transfer embryos derived from cow versus heifer cytoplasts. Anim Reprod Sci. 2006;95(3-4):234-243.

68. Li GP, Bunch TD, White KL, Aston KI, Meerdo LN, Pate BJ, Sessions BR. Development, chromosomal composition, and cell allocation of bovine cloned blastocyst derived from chemically assisted enucleation and cultured in conditioned media. Mol Reprod Dev. 2004;68(2):189-197.

69. Gandhi AP, Lane M, Gardner DK, Krisher RL. A single medium supports development of bovine embryos throughout maturation, fertilization and culture. Hum Reprod. 2000;15(2):395-401.

70. Polejaeva IA. White KL, Ellis LC, Reed WA. Isolation and long-term culture of mink and bovine embryonic stem-like cells. Theriogenology. 1995;43:300.

71. Polejaeva IA, Reed WA, Bunch TD, Ellis LC, White KL. Prolactin-induced termination of obligate diapause of mink (Mustela vison) blastocysts in vitro and subsequent establishment of embryonic stem-like cells. J Repro Fert. 1997;109(2):229-236.

72. Polejaeva I, Mitalipov S. Stem cell potency and the ability to contribute to chimeric organisms. Reproduction. 2013;145(3):R81-88.

73. Malaver-Ortega LF, Sumer H, Liu J, Verma PJ. The state of the art for pluripotent stem cells derivation in domestic ungulates. Theriogenology. 2012;78(8):1749-62.

74. Saito S, Strelchenko N, Niemann H. Bovine embryonic stem cell-like cell lines cultured over several passages. Roux Arch Dev Biol. 1992;201(3):134-141.

75. Cibelli JB, Stice SL, Golueke PJ, Kane JJ, Jerry J, Blackwell C, Ponce de Leon FA, Robl JM. Transgenic bovine chimeric offspring produced from somatic cell-derived stem-like cells. Nat Biotechnol. 1998;16(7):642-646.

76. Krueger F. Trim Galore. In: Babraham Bioinformatics. 2016. https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/ Accessed July 2018.

77. Kopylova E, Noe L, Touzet H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics. 2012;28(24):3211-3217.

78. Gebert D, Hewel C, Rosenkranz D. Unitas: the universal tool for annotation of small RNAs. BMC Genomics. 2017;18(1):644.

79. Rosenkranz D, Han CT, Roovers EF, Zischler H, Ketting RF. Piwi proteins and piRNAs in mammalian oocytes and early embryos: From sample to sequence. Genom Data. 2015;5:309-313.

80. Rosenkranz D, Zischler H. proTRAC--a software for probabilistic piRNA cluster detection, visualization and analysis. BMC Bioinformatics. 2012;13:5.

81. Wagih O. ggseqlogo: a versatile R package for drawing sequence logos. Bioinformatics. 2017;33(22):3645-3647.

82. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.

83. Hulsen T, de Vlieg J, Alkema W. BioVenn - a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. BMC Genomics. 2008;9:488.

84. Metsalu T, Vilo J. ClustVis: a web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. Nucleic Acids Res. 2015;43(W1):W566-570.

85. Agarwal V, Bell GW, Nam JW, Bartel DP. Predicting effective microRNA target sites in mammalian mRNAs. Elife. 2015;4.

86. Tripathi S, Pohl MO, Zhou Y, Rodriguez-Frandsen A, Wang G, Stein DA, Moulton HM, DeJesus P, Che J, Mulder LC. Meta- and orthogonal integration of influenza "OMICs" data defines a role for UBR4 in virus budding. Cell Host Microbe. 2015;18(6):723-735.

87. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13(11):2498-2504.

88. ENCODE: microRNA-seq data standards and processing pipeline. https://www.encodeproject.org/microrna/microrna-seq. Accessed April 2020.

89. Tang DT, Glazov EA, McWilliam SM, Barris WC, Dalrymple BP. Analysis of the complement and molecular evolution of tRNA genes in cow. BMC Genomics. 2009;10:188.

90. Schimmel P. The emerging complexity of the tRNA world: mammalian tRNAs beyond protein synthesis. Nat Rev Mol Cell Biol. 2018;19(1):45-58.

91. Kanai A. Disrupted tRNA genes and tRNA fragments: a perspective on tRNA gene evolution. Life (Basel). 2015;5(1):321-331.

92. Yuan S, Schuster A, Tang C, Yu T, Ortogero N, Bao J, Zheng H, Yan W. Sperm-borne miRNAs and endo-siRNAs are important for fertilization and preimplantation embryonic development. Development. 2016;143(4):635-647.

93. Gilchrist GC. Dynamics of telomeres and microRNAs during oocyte maturation and embryo development in the cow. The University of Guelph; 2015.

94. Su X, Gao G, Wang S, Su G, Zheng Z, Zhang J, Han L, Ling Y, Wang X, Li G et al. CircRNA expression profile of bovine placentas in late gestation with aberrant SCNT fetus. J Clin Lab Anal. 2019;33(6):e22918.

95. Tesfaye D, Gebremedhn S, Salilew-Wondim D, Hailay T, Hoelker M, Grosse-Brinkhaus C, Schellander K. MicroRNAs: tiny molecules with a significant role in mammalian follicular and oocyte development. Reproduction. 2018;155(3):R121-R135.

96. Tu F, Pan ZX, Yao Y, Liu HL, Liu SR, Xie Z, Li QF. miR-34a targets the inhibin beta B gene, promoting granulosa cell apoptosis in the porcine ovary. Genet Mol Res. 2014;13(2):2504-2512.

97. Maalouf SW, Smith CL, Pate JL. Changes in microRNA expression during maturation of the bovine corpus luteum: regulation of luteal cell proliferation and function by microRNA-34a. Biol Reprod. 2016;94(3):71.

98. Tscherner A, Gilchrist G, Smith N, Blondin P, Gillis D, LaMarre J. MicroRNA-34 family expression in bovine gametes and preimplantation embryos. Reprod Biol Endocrinol. 2014;12:85.

99. Choi YJ, Lin CP, Risso D, Chen S, Kim TA, Tan MH, Li JB, Wu Y, Chen C, Xuan Z et al. Deficiency of microRNA miR-34a expands cell fate potential in pluripotent stem cells. Science. 2017;355(6325).

100. Li MA, He L. MicroRNAs as novel regulators of stem cell pluripotency and somatic cell reprogramming. Bioessays. 2012;34(8):670-680.

101. Liu H, Deng H, Zhao Y, Li C, Liang Y. LncRNA XIST/miR-34a axis modulates the cell proliferation and tumor growth of thyroid cancer through MET-PI3K-AKT signaling. J Exp Clin Cancer Res. 2018;37(1):279.

102. Rokavec M, Li H, Jiang L, Hermeking H. The p53/miR-34 axis in development and disease. J Mol Cell Biol. 2014;6(3):214-230.

103. Rosenbluth EM, Shelton DN, Sparks AE, Devor E, Christenson L, Van Voorhis BJ. MicroRNA expression in the human blastocyst. Fertil Steril. 2013;99(3):855-861 e853.

104. Guo L, Zhao RC, Wu Y. The role of microRNAs in self-renewal and differentiation of mesenchymal stem cells. Exp Hematol. 2011;39(6):608-616.

105. Li Y, Hagen DE, Ji T, Bakhtiarizadeh MR, Frederic WM, Traxler EM, Kalish JM, Rivera RM. Altered microRNA expression profiles in large offspring syndrome and Beckwith-Wiedemann syndrome. Epigenetics. 2019;14(9):850-876.

106. Zhang MY, Lin J, Kui YC. MicroRNA-345 suppresses cell invasion and migration in non-small cell lung cancer by directly targeting YAP1. Eur Rev Med Pharmacol Sci. 2019;23(6):2436-2443.

107. Srivastava SK, Bhardwaj A, Arora S, Tyagi N, Singh S, Andrews J, McClellan S, Wang B, Singh AP. MicroRNA-345 induces apoptosis in pancreatic cancer cells through potentiation of caspase-dependent and -independent pathways. Br J Cancer. 2015;113(4):660-668.

108. Tang JT, Wang JL, Du W, Hong J, Zhao SL, Wang YC, Xiong H, Chen HM, Fang JY; MicroRNA 345, a methylation-sensitive microRNA is involved in cell proliferation and invasion in human colorectal cancer. Carcinogenesis. 2011;32(8):1207-1215.

109. Babiarz JE, Ruby JG, Wang Y, Bartel DP, Blelloch R. Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. Genes Dev. 2008;22(20):2773-2785.

110. Feng R, Sang Q, Zhu Y, Fu W, Liu M, Xu Y, Shi H, Xu Y, Qu R, Chai R et al. MiRNA-320 in the human follicular fluid is associated with embryo quality in vivo and affects mouse embryonic development in vitro. Sci Rep. 2015;5:8689.

111. Gross N, Kropp J, Khatib H. Sexual dimorphism of miRNAs secreted by bovine in vitro-produced embryos. Front Genet. 2017;8:39.

112. Hsieh IS, Chang KC, Tsai YT, Ke JY, Lu PJ, Lee KH, Yeh SD, Hong TM, Chen YL. MicroRNA-320 suppresses the stem cell-like characteristics of prostate cancer cells by downregulating the Wnt/beta-catenin signaling pathway. Carcinogenesis. 2013;34(3):530-538.

113. Luo L, Ye G, Nadeem L, Fu G, Yang BB, Honarparvar E, Dunk C, Lye S, Peng C. MicroRNA-378a-5p promotes trophoblast cell survival, migration and invasion by targeting Nodal. J Cell Sci. 2012;125(Pt 13):3124-3132.

114. Foshay KM, Gallicano GI. miR-17 family miRNAs are expressed during early mammalian development and regulate stem cell differentiation. Dev Biol. 2009;326(2):431-443.

115. Liang G, Malmuthuge N, McFadden TB, Bao H, Griebel PJ, Stothard P, Guan le L. Potential regulatory role of microRNAs in the development of bovine gastrointestinal tract during early life. PLoS One. 2014;9(3):e92592.

116. Bick JT, Floter VL, Robinson MD, Bauersachs S, Ulbrich SE. Small RNA-seq analysis of single porcine blastocysts revealed that maternal estradiol-17beta exposure does not affect miRNA isoform (isomiR) expression. BMC Genomics. 2018;19(1):590.

117. Viswanathan SR, Mermel CH, Lu J, Lu CW, Golub TR, Daley GQ. microRNA expression during trophectoderm specification. PLoS One. 2009;4(7):e6143.

118. Giraldez AJ. microRNAs, the cell's Nepenthe: clearing the past during the maternal-to-zygotic transition and cellular reprogramming. Curr Opin Genet Dev. 2010;20(4):369-375.

119. Li HL, Wei JF, Fan LY, Wang SH, Zhu L, Li TP, Lin G, Sun Y, Sun ZJ, Ding J et al. miR-302 regulates pluripotency, teratoma formation and differentiation in stem cells via an AKT1/OCT4-dependent manner. Cell Death Dis. 2016;7:e2078.

120. Gu KL, Zhang Q, Yan Y, Li TT, Duan FF, Hao J, Wang XW, Shi M, Wu DR, Guo WT et al. Pluripotency-associated miR-290/302 family of microRNAs promote the dismantling of naive pluripotency. Cell Res. 2016;26(3):350-366.

121. Lin X, Beckers E, Mc Cafferty S, Gansemans Y, Joanna Szymanska K, Chaitanya Pavani K, Catani JP, Van Nieuwerburgh F, Deforce D, De Sutter P et al. Bovine embryo-aecreted microRNA-30c is a potential non-invasive biomarker for hampered preimplantation developmental competence. Front Genet. 2019;10:315.

122. Mao L, Liu S, Hu L, Jia L, Wang H, Guo M, Chen C, Liu Y, Xu L. miR-30 family: a promising regulator in development and disease. Biomed Res Int. 2018;2018:9623412.

123. Yang JS, Maurin T, Robine N, Rasmussen KD, Jeffrey KL, Chandwani R, Papapetrou EP, Sadelain M, O'Carroll D, Lai EC. Conserved vertebrate mir-451 provides a platform for Dicer-independent, Ago2-mediated microRNA biogenesis. Proc Natl Acad Sci USA. 2010;107(34):15163-15168.

124. Ioannidis J, Donadeu FX. Changes in circulating microRNA levels can be identified as early as day 8 of pregnancy in cattle. PLoS One. 2017;12(4):e0174892.

125. Li Z, Jia J, Gou J, Zhao X, Yi T. MicroRNA-451 plays a role in murine embryo implantation through targeting Ankrd46, as implicated by a microarray-based analysis. Fertil Steril. 2015;103(3):834-834 e834.

126. Shokri G, Kouhkan F, Nojehdehi S, Soleimani M, Pourfathollah AA, Nikougoftar Zarif M, Tamaddon M, Obeidi N. Simultaneous regulation of miR-451 and miR-191 led to erythroid fate decision of mouse embryonic stem cell. Iran J Basic Med Sci. 2019;22(4):432-438.

127. Liang J, Wang S, Wang Z. Role of microRNAs in embryo implantation. Reprod Biol Endocrinol. 2017;15(1):90.

128. Reza A, Choi YJ, Han SG, Song H, Park C, Hong K, Kim JH. Roles of microRNAs in mammalian reproduction: from the commitment of germ cells to peri-implantation embryos. Biol Rev Camb Philos Soc. 2019;94(2):415-438.

129. Salleh N, Giribabu N. Leukemia inhibitory factor: roles in embryo implantation and in nonhormonal contraception. ScientificWorldJournal. 2014;2014:201514.

130. Cimmino A, Calin GA, Fabbri M, Iorio MV, Ferracin M, Shimizu M, Wojcik SE, Aqeilan RI, Zupo S, Dono M et al. miR-15 and miR-16 induce apoptosis by targeting BCL2. Proc Natl Acad Sci USA. 2005;102(39):13944-13949.

131. Russell SJ, LaMarre J. Transposons and the PIWI pathway: genome defense in gametes and embryos. Reproduction. 2018;156(4):R111-R124.

132. Beraldi R, Pittoggi C, Sciamanna I, Mattei E, Spadafora C. Expression of LINE-1 retroposons is essential for murine preimplantation development. Mol Reprod Dev. 2006;73(3):279-287.

133. Kigami D, Minami N, Takayama H, Imai H. MuERV-L is one of the earliest transcribed genes in mouse one-cell embryos. Biol Reprod. 2003;68(2):651-654.

134. Ge SX. Exploratory bioinformatics investigation reveals importance of "junk" DNA in early embryo development. BMC Genomics. 2017;18(1):200.

135. Lin L, Xu B, Rote NS. Expression of endogenous retrovirus ERV-3 induces differentiation in BeWo, a choriocarcinoma model of human placental trophoblast. Placenta. 1999;20(1):109-118.

136. Juliano C, Wang J, Lin H. Uniting germline and stem cells: the function of Piwi proteins and the piRNA pathway in diverse organisms. Annu Rev Genet. 2011;45:447-469.

137. Ebner T, Shebl O, Moser M, Mayer RB, Arzt W, Tews G. Group culture of human zygotes is superior to individual culture in terms of blastulation, implantation and life birth. Reprod Biomed Online. 2010;21(6):762-768.

138. Uhde K, van Tol HTA, Stout TAE, Roelen BAJ. MicroRNA expression in bovine cumulus cells in relation to oocyte quality. Noncoding RNA. 2017;3(1). doi: 10.3390/ncrna3010012.

139. Assou S, Al-edani T, Haouzi D, Philippe N, Lecellier CH, Piquemal D, Commes T, Ait-Ahmed O, Dechaud H, Hamamah S. MicroRNAs: new candidates for the regulation of the human cumulus-oocyte complex. Hum Reprod. 2013;28(11):3038-3049.

140. PhenoGram Plot. http://visualization.ritchielab.org/phenograms/plot. Accessed August 2019
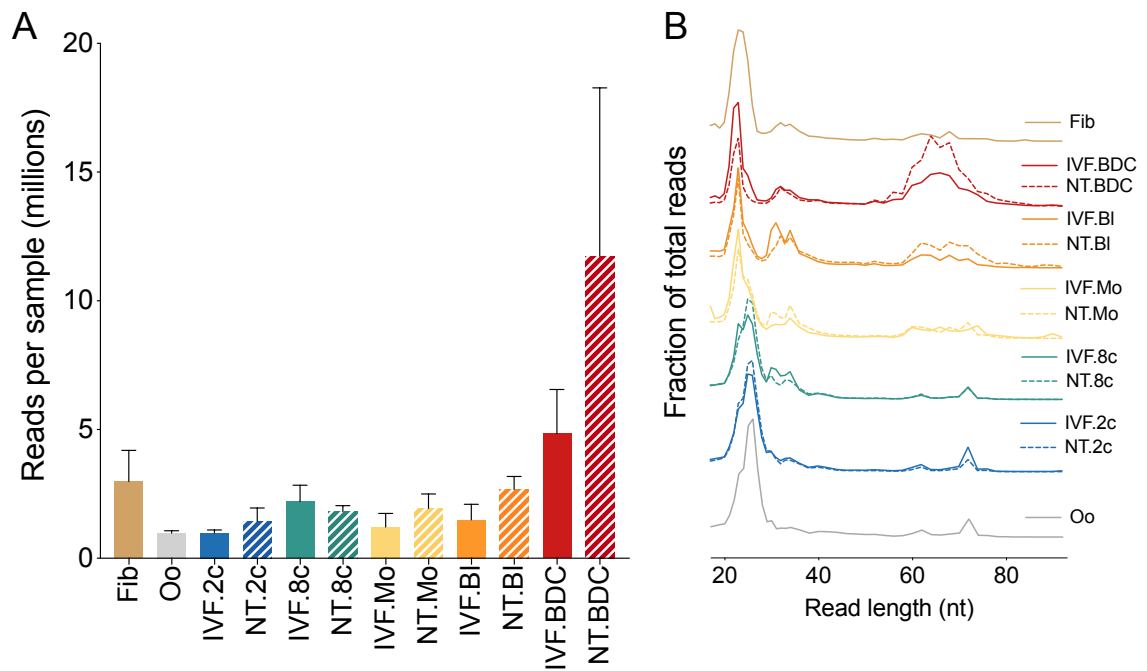
Figure 3.1. Distribution of reads. (A) Total reads per sample. (B) Read length distribution of sequenced samples after trimming to 17–93 nt and filtering for quality control. Abbreviations are: Oo, oocyte; 2c, 2-cell; 8c, 8-cell; Mo, morula; Bl, blastocyst; BDC, blastocyst-derived cells; Fb, fibroblast; IVF, in vitro-fertilized; NT, somatic cell nuclear transfer; miRNA, microRNA; ID, identification number.
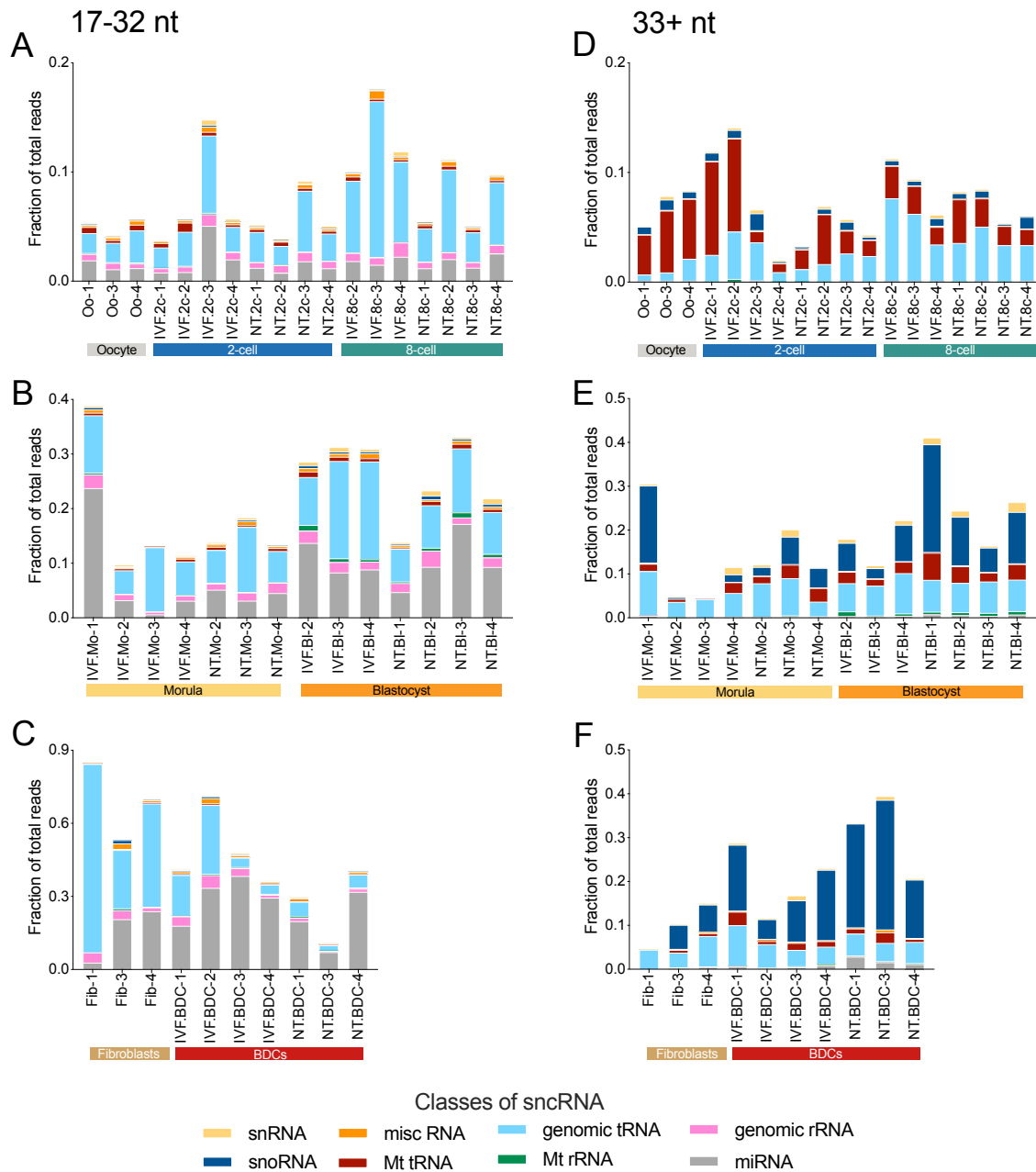
Figure 3.2. Distribution of sequence annotations. (A-C) Proportion of sncRNA annotated reads by class and read length 17–32 nt. (D-F) Proportion of sncRNA annotated reads by class and read length ≥33 nt. Abbreviations are: Oo, oocyte; 2c, 2-cell; 8c, 8-cell; Mo, morula; Bl, blastocyst; BDC, blastocyst-derived cells; Fb, fibroblast; IVF, in vitro-fertilized; NT, somatic cell nuclear transfer; miRNA, microRNA; ID, identification number.
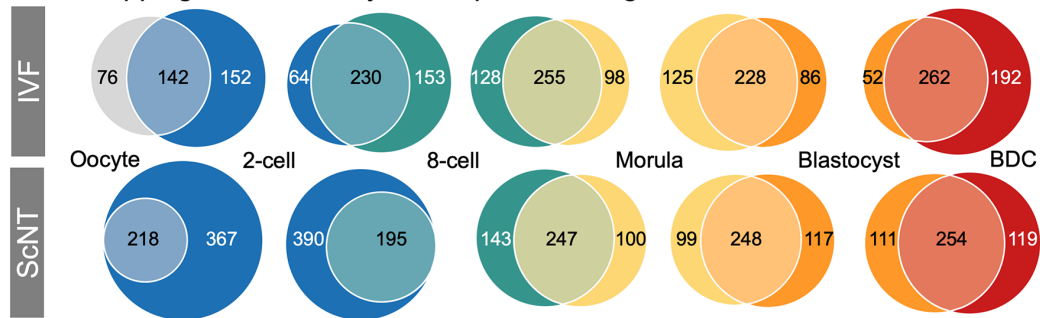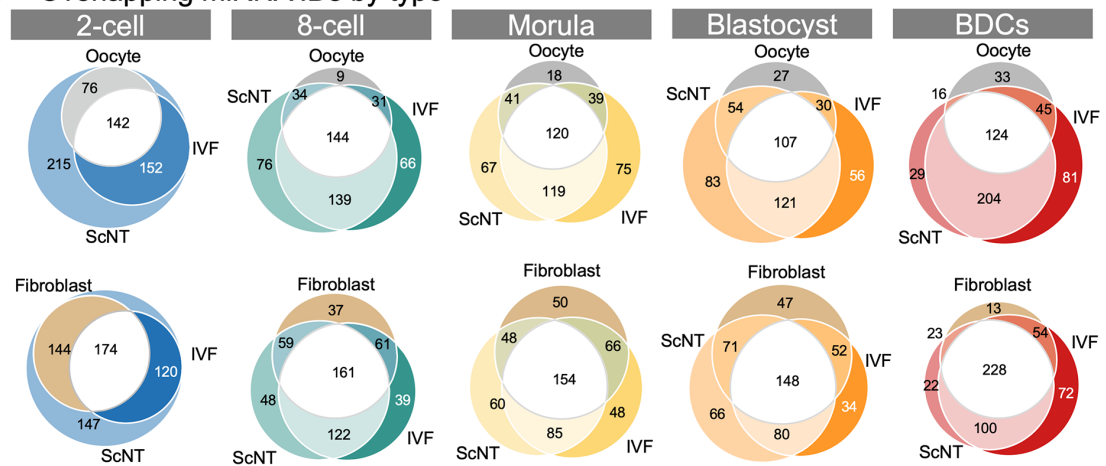
Figure 3.3. Overlap in detection of all mapped miRNAs in bovine oocytes and donor fibroblast cells as well as early developing embryos and blastocyst-derived cells generated via IVF or scNT. Proportional Venn diagrams depicting miRNAs mapped to the *Bos taurus* miR database, including unique and overlapping IDs. (A) miRNAs identified by sequential comparison of developmental stages for IVF or scNT embryos. (B) miRNAs identified in IVF, scNT and either oocytes or donor fibroblast cells at each developmental stage. Colors represent the following stages and embryo types: grey, oocyte; tan, fibroblast; blue, 2-cell (light, scNT; dark, IVF); green, 8-cell (light, scNT; dark, IVF); yellow, morula (light, scNT; dark, IVF); orange, blastocyst (light, scNT; dark, IVF); red, BDC (light, scNT; dark, IVF).

Figure 3.4. Patterns of miRNA expression in bovine oocytes and donor fibroblast cells as well as early developing embryos and blastocyst-derived cells generated via IVF or scNT. (A) Principal components analysis of bovine miRNAs using the standard singular value decomposition method with imputation. (B) Unsupervised, bi-directional hierarchical cluster analysis (Euclidean distance method) of variance stabilized normalized read counts for all identified miRNAs. Data were categorized by sample type (Fb, fibroblast; Oo, oocyte; IVF, *in vitro*-fertilized embryo; scNT, scNT embryo) and stage (2c, 2-cell; 8c, 8-cell; Mo, morula; Bl, blastocyst; BDC, blastocyst-derived cells).

A

-log *p* value

GO:0048368: lateral mesoderm development
GO:0000226: microtubule cytoskeleton organization
GO:0072594: establishment of protein localization to organelle
GO:0051129: negative regulation of cellular component organization
GO:0045844: positive regulation of striated muscle tissue development
GO:0021871: forebrain regionalization
GO:0030194: positive regulation of blood coagulation
GO:0070201: regulation of establishment of protein localization
GO:0000910: cytokinesis
GO:0061005: cell differentiation involved in kidney development
GO:0050821: protein stabilization
GO:0090050: positive regulation of cell migration involved in sprouting angiogenesis
GO:0071636: positive regulation of transforming growth factor beta production
GO:1990182: exosomal secretion
GO:1900182: positive regulation of protein localization to nucleus
GO:0007005: mitochondrion organization
GO:0043624: cellular protein complex disassembly
GO:0044089: positive regulation of cellular component biogenesis
GO:0030308: negative regulation of cell growth
GO:0033059: cellular pigmentation

B

Figure 3.5. (prior page) Significant ontology and pathway terms and network clustering associated with predicted mRNA targets of over or under expressed miRNAs in bovine scNT morula stage embryos as compared to IVF morula stage embryos. (A) Predicted mRNA targets (TargetScan total context ++ score <-0.35) of differentially expressed miRNAs ($|\log_2 R|>1$ and p-value <0.05 for scNT morula stage embryos compared to IVF morula stage embryos) were subject to enrichment analysis using Metascape (Minimum overlap of 3, minimum enrichment 1.5 and $p<0.01$). Only miRNAs whose families annotated by TargetScan as conserved or broadly conserved were included in the Metascape analysis. Each heat-map depicts the top 20 enriched summary terms identified for Gene Ontology biological processes, with the length of the bar determined by the $-\log_{10} p$-value. (B) Cytoscape network analysis of Gene ontology biological process terms for predicted mRNA targets of differentially expressed miRNAs, including over- and under- expressed miRs in the scNT morula staged embryos as compared to IVF morula staged embryos. Each term is represented by a circle node, for which the size is proportional to the enrichment score and the color represents the enrichment pathway term. Terms with a similarity score >0.3 are linked by an edge (the thickness of the edge represents the similarity score) using force-directed network layout and bundled edges for clarity. Each cluster is labeled by the term with the highest p-value in the cluster group. Complete results of the Metascape analysis are provided in Supplementary File 3.5; Appendix J.
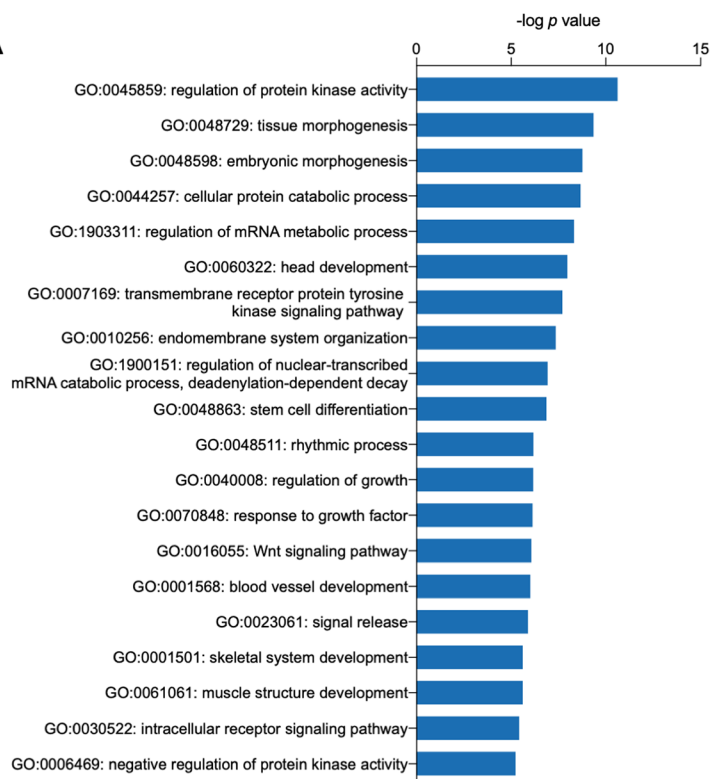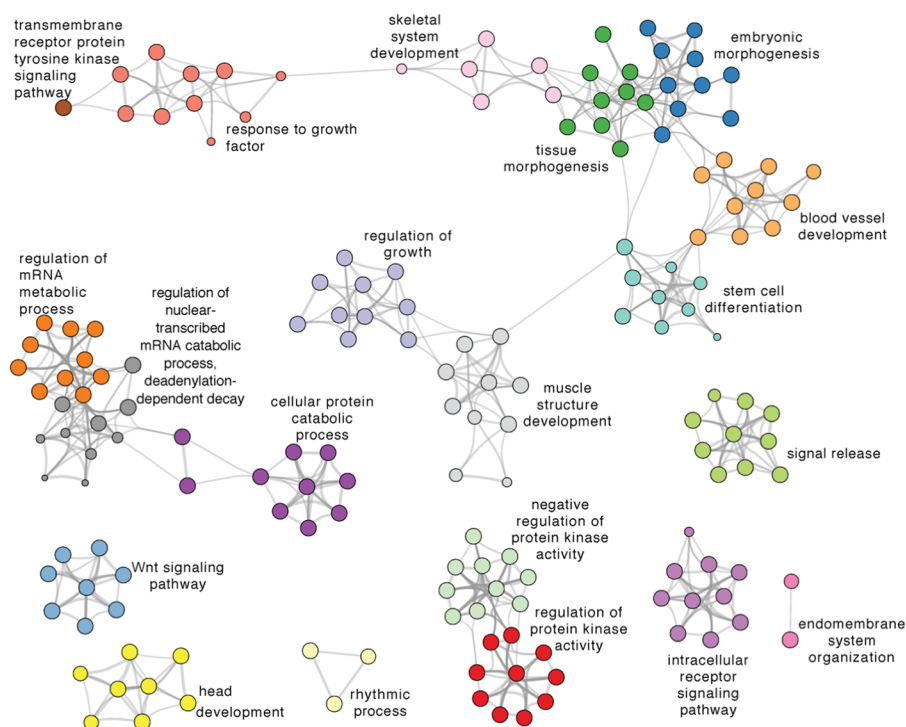
Figure 3.6. (prior page) Significant ontology and pathway terms and network clustering associated with predicted mRNA targets of over or under expressed miRNAs in bovine scNT BDCs as compared to IVF BDCs. (A) Predicted mRNA targets (TargetScan total context ++ score <-0.35) of differentially expressed miRNAs ($|\log_2 R|>1$ and *p*-value <0.05 for scNT BDCs compared to IVF BDCs) were subject to enrichment analysis using Metascape (Minimum overlap of 3, minimum enrichment 1.5 and *p*<0.01). Only miRNAs whose families annotated by TargetScan as conserved or broadly conserved were included in the Metascape analysis. Each heat map depicts the top 20 enriched summary terms identified for Gene Ontology biological processes, with the length of the bar determined by the $-\log_{10}$ *p*-value. (B) Cytoscape network analysis of Gene ontology biological process terms for predicted mRNA targets of differentially expressed miRNAs, including over- and under- expressed miRs in the scNT BDCs as compared to IVF BDCs. Each term is represented by a circle node, for which the size is proportional to the enrichment score and the color represents the enrichment pathway term. Terms with a similarity score >0.3 are linked by an edge (the thickness of the edge represents the similarity score) using force-directed network layout and bundled edges for clarity. Each cluster is labeled by the term with the highest *p*-value in the cluster group. Complete results of the Metascape analysis are provided in Supplementary File 3.5; Appendix J.
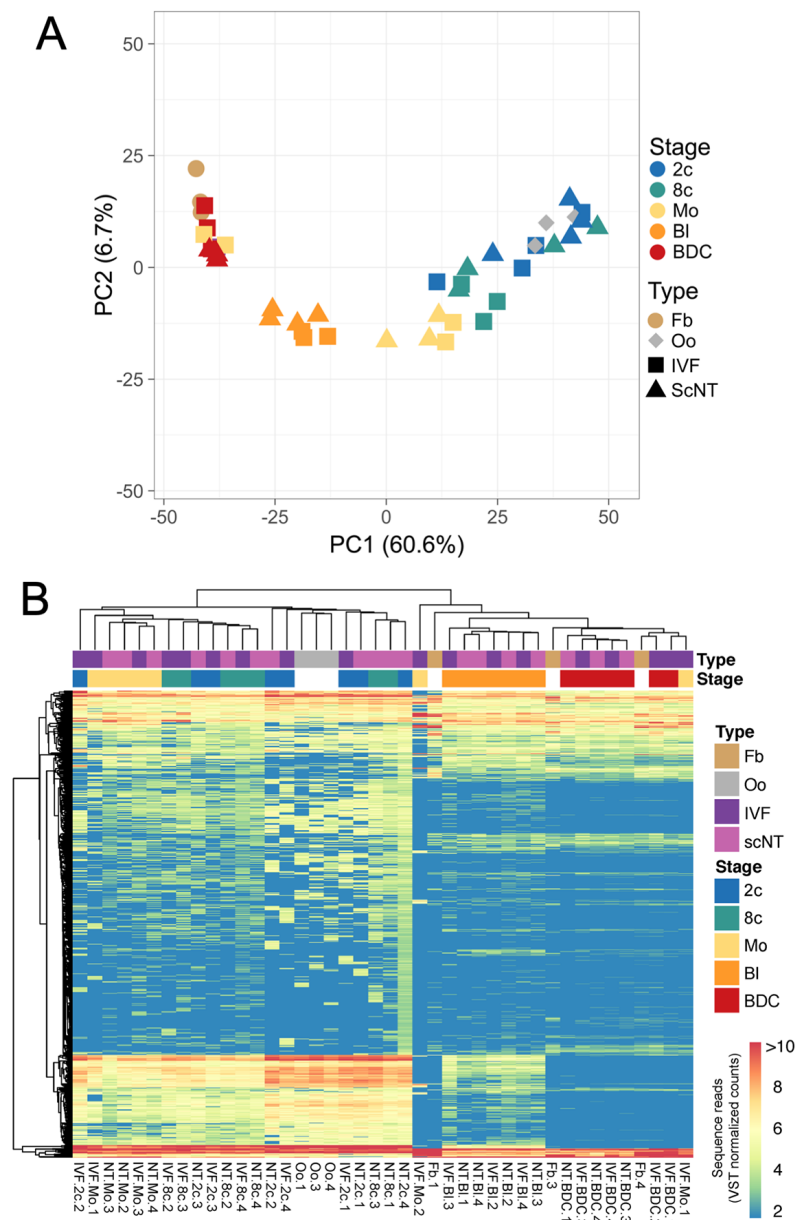
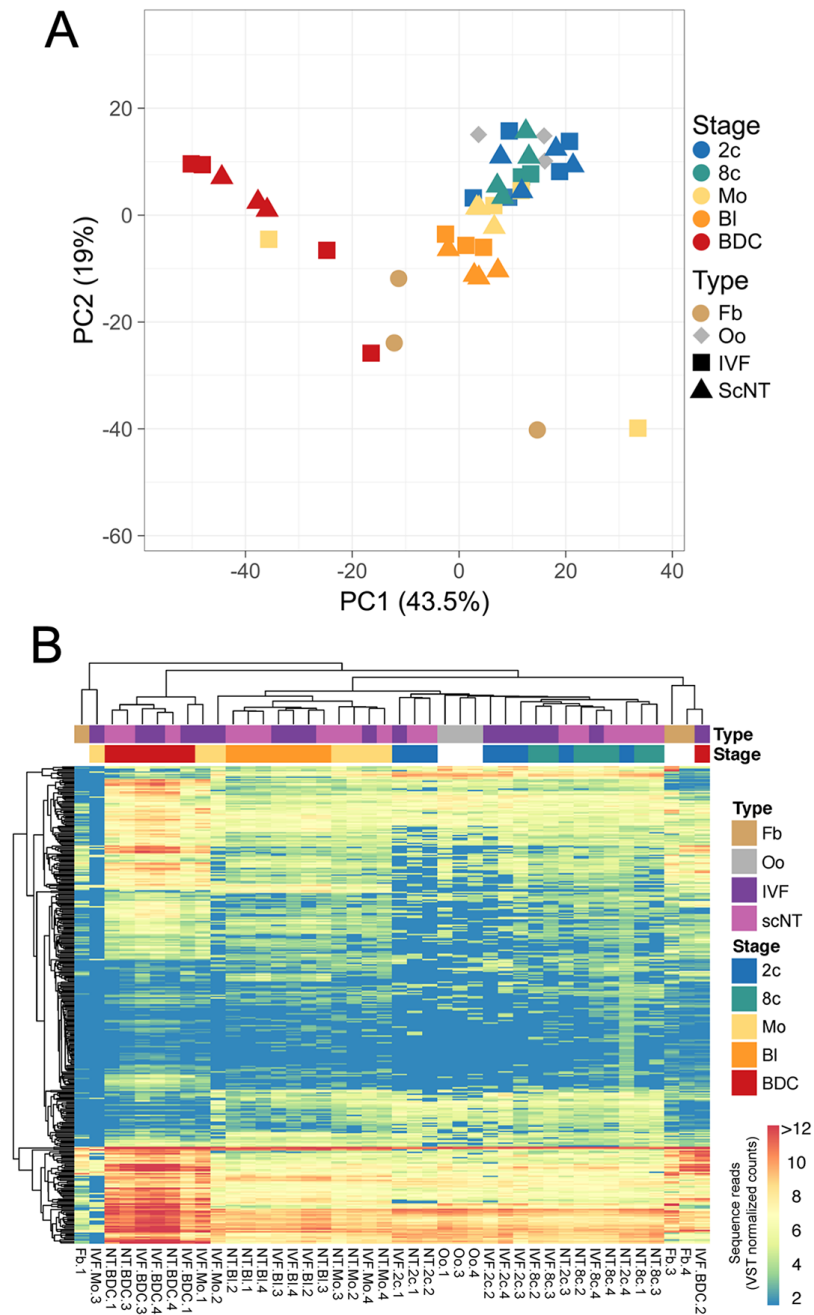**Figure 3.7.** Patterns of tRNA expression in bovine oocytes and donor fibroblast cells as well as early developing embryos and blastocyst-derived cells generated via IVF or scNT. (A) Principal components analysis of bovine tRNAs using the standard singular value decomposition method with imputation. (B) Unsupervised, bi-directional hierarchical cluster analysis (Euclidean distance method) of variance stabilized normalized read counts for all identified tRNAs. Data were categorized by sample type (Fb, fibroblast; Oo, oocyte; IVF, *in vitro*-fertilized embryo; scNT, scNT embryo) and stage (2c, 2-cell; 8c, 8-cell; Mo, morula; Bl, blastocyst; BDC, blastocyst-derived cells).
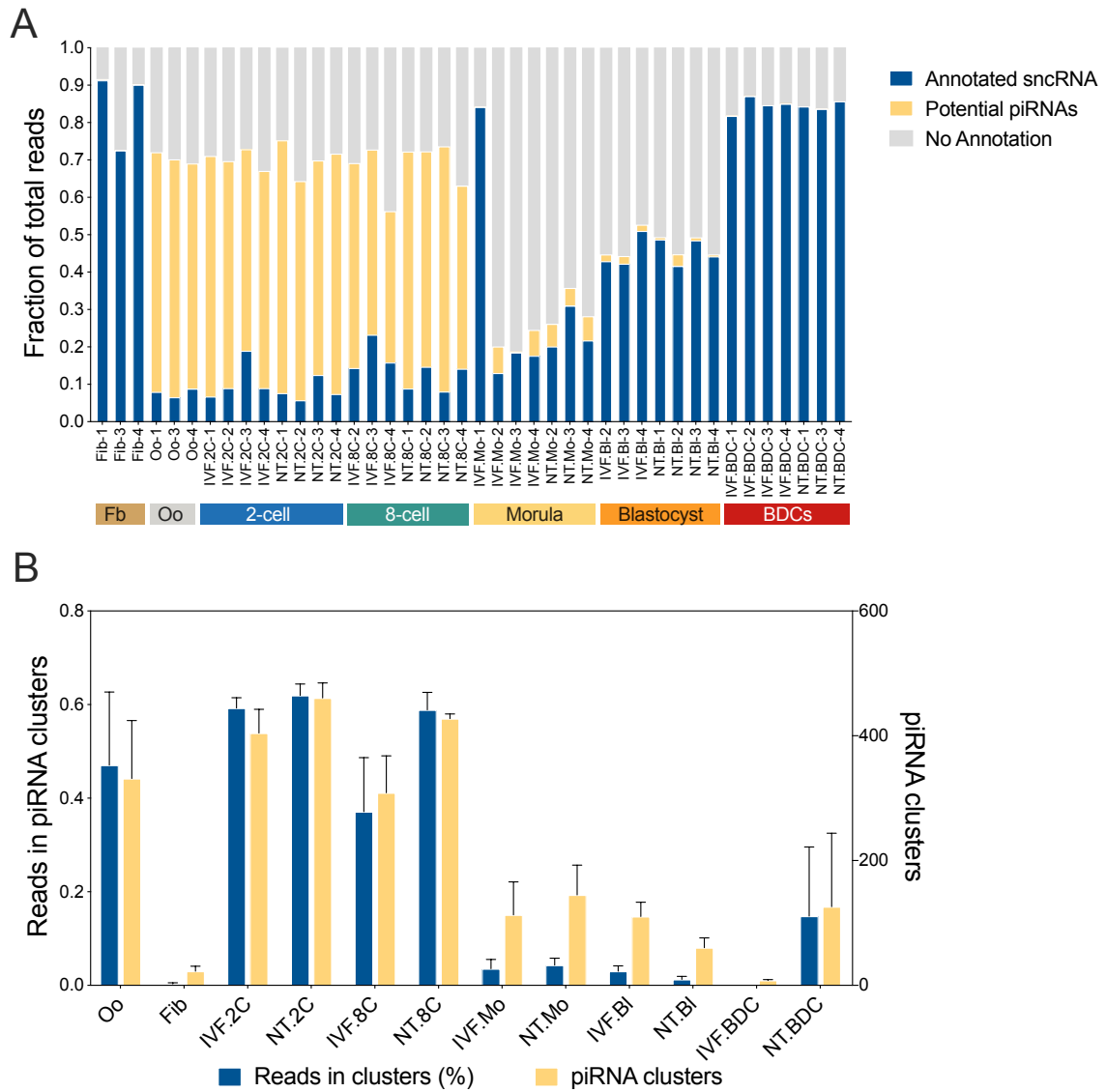
Figure 3.8. Annotation of piRNA-like RNAs (pilRNAs). (A) Fraction of sequence reads mapped as potential piRNAs compared to reads mapped to other sncRNA classes or not annotated by sample type. (B) Number of sequencing reads mapped to piRNA clusters (left axis) and number of piRNA clusters identified (right axis). Oo, oocyte; Fib, fibroblast; 2c, 2-cell; 8c, 8-cell; Mo, morula; Bl, blastocyst; BDC, blastocyst-derived cells; IVF, *in vitro* fertilized; NT, somatic cell nuclear transfer.
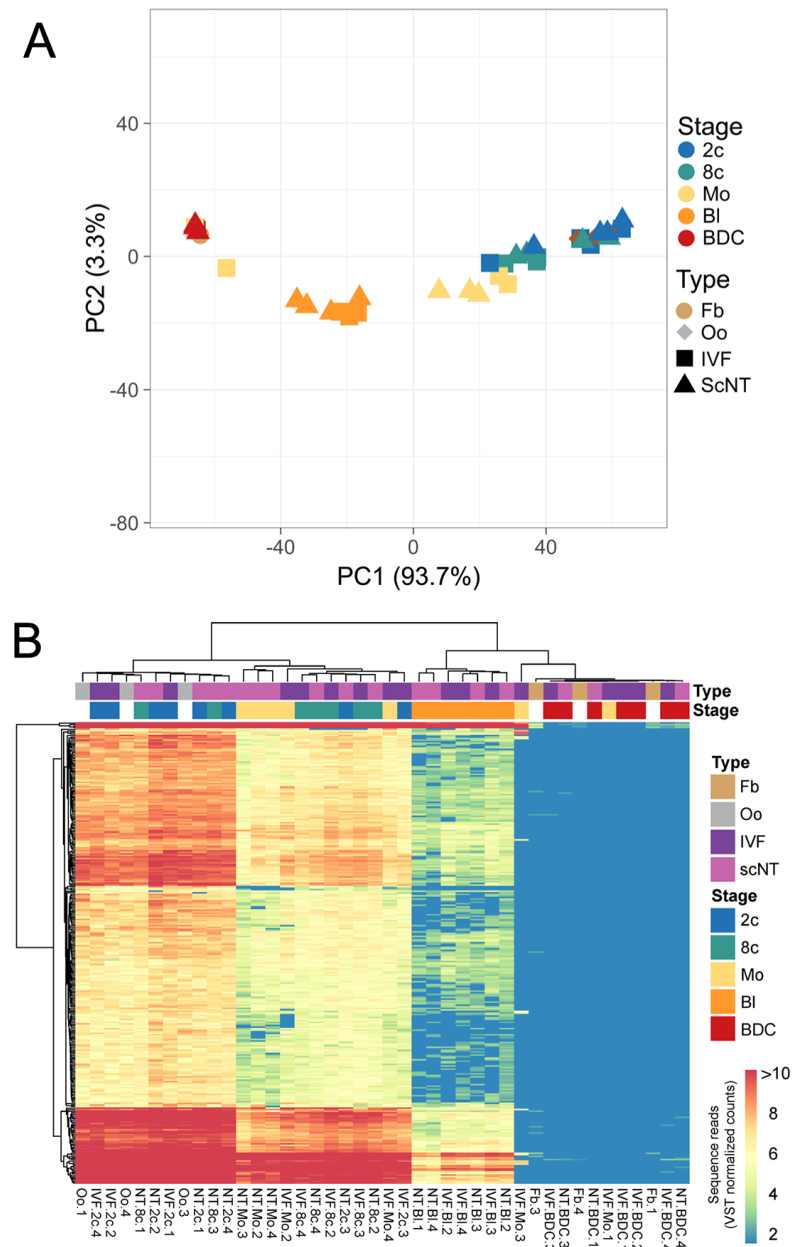
Figure 3.9. Patterns of piRNA expression in bovine oocytes and donor fibroblast cells as well as early developing embryos and blastocyst-derived cells generated via IVF or scNT. (A) Principal components analysis of bovine piRNAs using the standard singular value decomposition method with imputation. (B) Unsupervised, bi-directional hierarchical cluster analysis (Euclidean distance method) of variance stabilized normalized read counts for all identified piRNAs. Data were categorized by sample type (Fb, fibroblast; Oo, oocyte; IVF, *in vitro*-fertilized embryo; scNT, scNT embryo) and stage (2c, 2-cell; 8c, 8-cell; Mo, morula; Bl, blastocyst; BDC, blastocyst-derived cells).
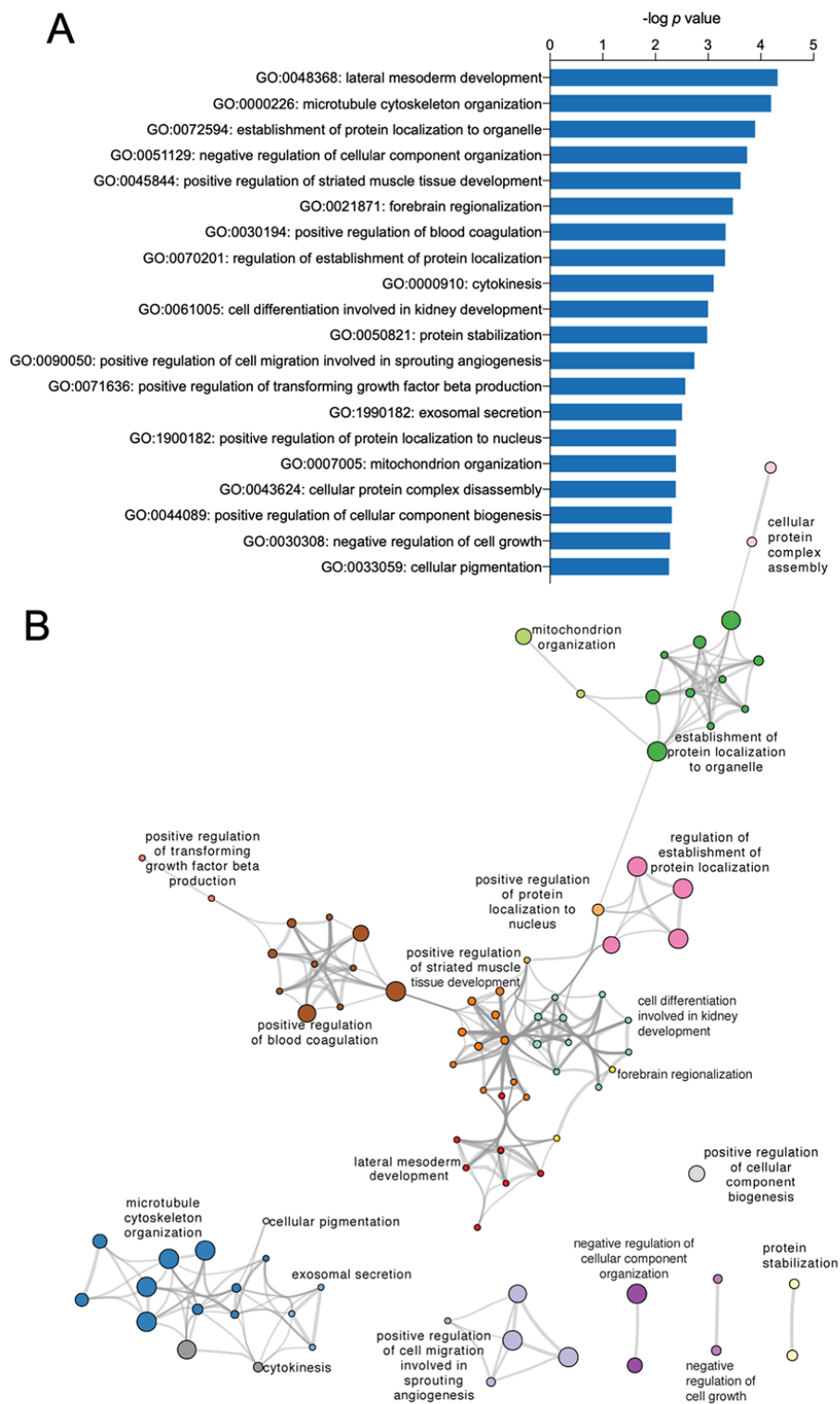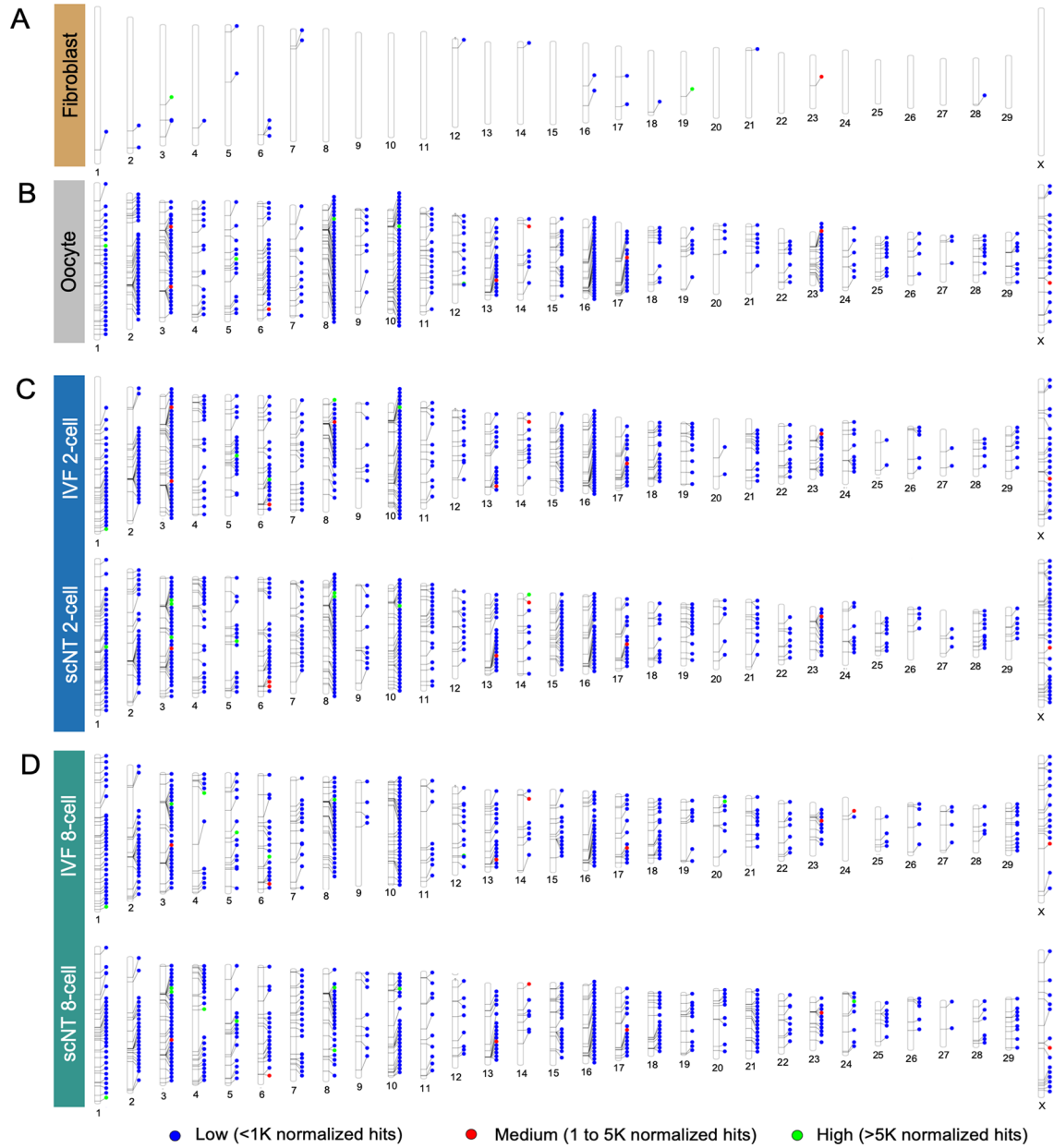
Figure 3.10. Expression of pilRNAs across chromosomes by sample type. (A) Fibroblast donor cells. (B) MII oocytes. (C) IVF and scNT 2-cell embryos. (D) IVF and scNT 8-cell embryos. Chromosome diagrams for the *Bos taurus* genome were created using PhenoGram [140] with each pilRNA cluster location marked by a circle colored according to expression level.
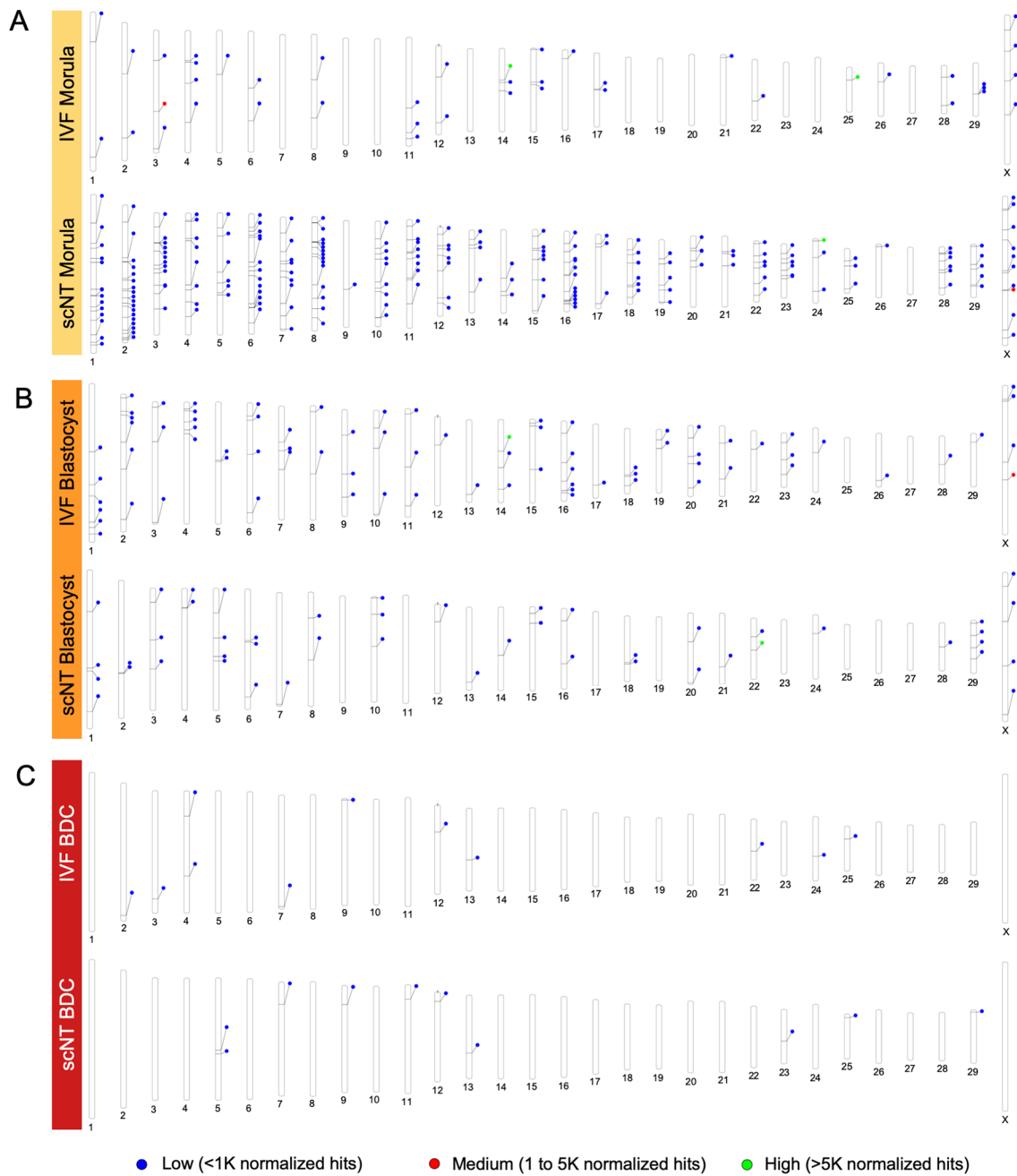
Figure 3.11. Expression of pilRNAs across chromosomes by sample type. (A) Morula IVF and scNT embryos. (B) Blastocyst IVF and scNT embryos. (C) Blastocyst-derived cells (BDCs) from IVF or scNT embryos. Chromosome diagrams for the *Bos taurus* genome were created using PhenoGram [140] with each pilRNA cluster location marked by a circle colored according to expression level.
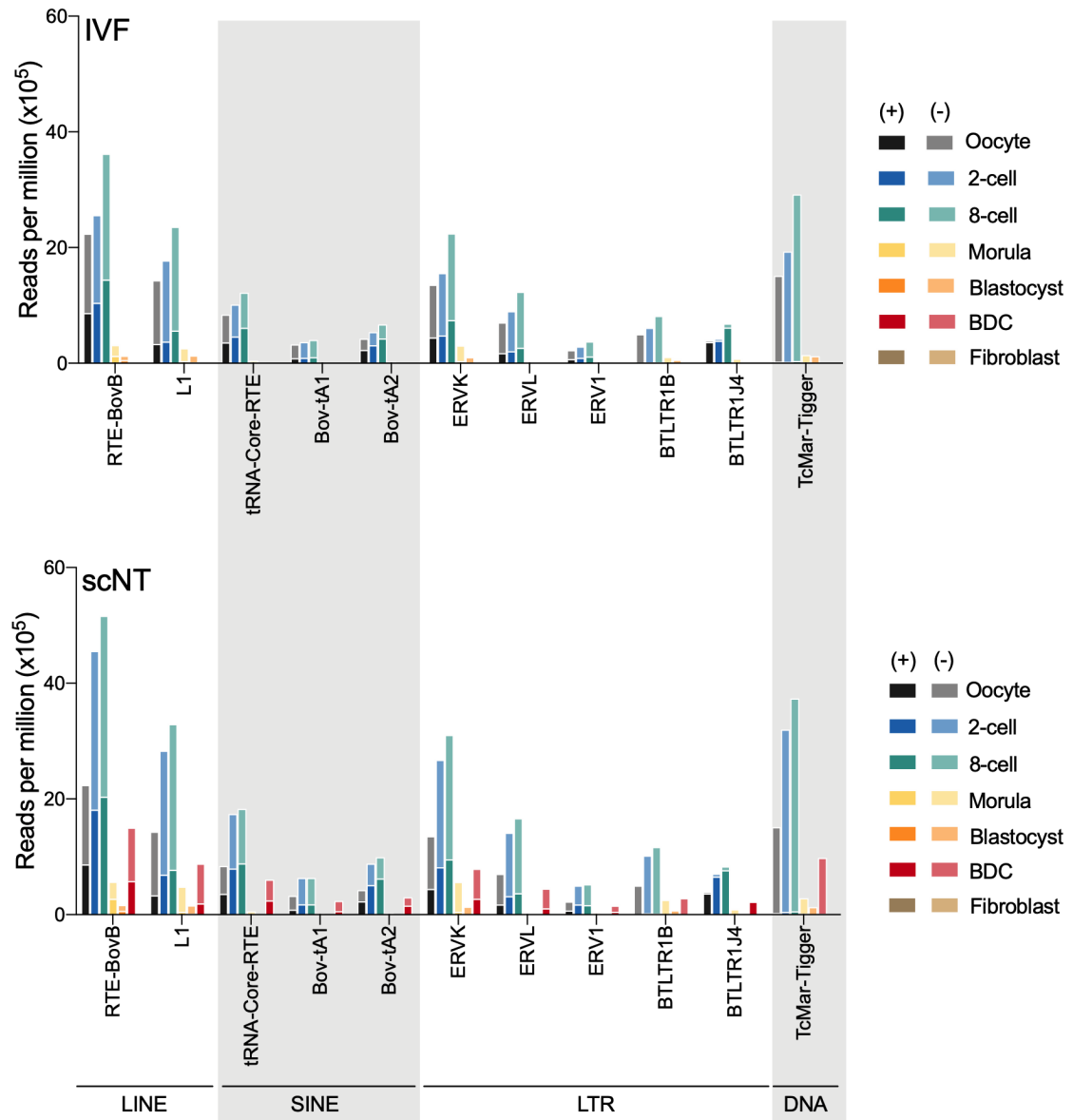
Figure 3.12. Mapping of piRNAs to TE sites in the *Bos taurus* genome. Values shown are the reads per million for the most common TE sites for putative pilRNAs on the sense (+) or antisense (–) DNA strand for IVF embryos (A) or scNT embryos (B). LINE, long interspersed nuclear elements; SINE, short interspersed nuclear elements; LTR, long terminal repeat; DNA, DNA transposon.

Table 3.1. Differentially abundant miRNAs

Comparison	All miRNAs	Cattle miRNAs	Cattle miRNA IDs
IVF vs. scNT			
2-cell	1	0	
8-cell	0	0	
Morula	6	3	miR-34a, miR-345-5p, miR-2340-3p (14T→A)
Blastocyst	1	0	
BDCs	26	24	miR-202-5p, miR-376c-3p(6A→G), miR-3432a-2-5p, miR-204-5p, miR-204, miR-3432b-5p(2T→G), miR-376b-3p, miR-221-3p, miR-202, miR-221, miR-380-3p, miR-133a-2-3p, miR-369-3p, miR-17-5p(1C→A), miR-3432a, miR-199b, miR-376c-3p, miR-18a-5p(20A→T), miR-376e, miR-22-3p, miR-222-3p, miR-409a-3p, miR-451, and miR-1260b-5p(9A→G)
Developmental stage			
IVF embryos			
Oo vs. 2c	2	0	
Fb vs. 2c	355	191	See Supplementary File 3.3; Appendix H
2C vs. 8c	1	0	
8c vs. Mo	29	22	miR-378, miR-378-3p, miR-378d-3p(1T→A), miR-2340-3p(14T→A), miR-378c-5p, miR-378-5p, miR-145, miR-2285aj-5p(14T→C), miR-378b-3p(4T→G), miR-34a-5p, miR-184, miR-199a-3p, miR-34a, miR-155-5p, miR-450a, miR-345-5p, miR-378c-5p, miR-199a-2-5p, miR-320a, miR-199b-3p, miR-214-3p, miR-339b,
Mo vs. Bl	25	19	miR-30e-5p, miR-143-3p, miR-28-3p, miR-145, miR-2340-3p(14T→A), miR-196a, miR-30f-5p(9C→T), miR-151-5p, miR-138-1-5p, ,miR-199b-3p, miR-93-5p,miR-345-5p,miR-30c,miR-146b, miR-339-5p, miR-199a-2-5p, miR-29b, miR-29e-3p(6T→C), miR-26a,
Bl vs. BDC	329	160	See Supplementary File 3.3; Appendix H
scNT embryos			
Oo vs. 2c	3	0	
Fb vs. 2c	365	202	See Supplementary File 3.3; Appendix H
2C vs. 8c	0	0	
8c vs. Mo	22	19	miR-378, miR-19b-3p, miR-378-5p, miR-6119-5p, miR-378-3p, miR-7-3-3p, miR-7-3-3p(25C→A), miR-138-2-5p(23G→A), miR-7, miR-7858-5p, miR-378c-5p, miR-27b-3p(19T→C), miR-138-2-5p(25C→A), miR-378d-3p(1T→A), miR-107-3p, miR-215-5p(19A→C), miR-30e-5p, miR-106b-5p, miR-302a-3p

Table 3.1. Differentially abundant miRNAs

Comparison	All miRNAs	Cattle miRNAs	Cattle miRNA IDs
Mo vs. Bl	56	33	miR-30e-5p, miR-497-5p, miR-125b-1-5p, miR-138-1-5p, miR-138-2-5p, miR-376c-3p, miR-152-3p, miR-11986c-5p(5C→T), miR-148b-3p(10C→G), miR-105a, miR-138-1-5p,miR-138-2-5p, miR-12058-5p(11T→A), miR-28-3p, miR-507b, miR-99a-5p, miR-138-1-5p(25C→A), miR-199a-3p, miR-15a, miR-30f-5p(9C→T), miR-2285cr-1-3p(3T→C), miR-10b-5p, miR-11986c-5p(15G→A), miR-146b, miR-122-5p, miR-218-1-5p, miR-34a-5p, miR-181a-1-5p, miR-324, miR-145-5p, miR-195-5p, miR-507-3p, miR-125b, miR-451
Bl vs. BDC	384	201	See Supplementary File 3.3; Appendix H

Note: Complete results of DESeq2 differential expression analyses for miRNAs are provided in Supplementary File 3.3; Appendix H. Abbreviations are: Oo, oocyte; 2c, 2-cell; 8c, 8-cell; Mo, morula; Bl, blastocyst; BDC, blastocyst-derived cells; Fb, fibroblast; IVF, in vitro-fertilized; scNT, somatic cell nuclear transfer; miRNA, microRNA; ID, identification number.

Table 3.2. Differentially abundant tRNA fragments

Comparison	Number
IVF vs. scNT	
2-cell	0
8-cell	0
Morula	4
Blastocyst	0
Developmental stage	
IVF embryos	
Oo vs. 2c	3
Fb vs. 2c	379
2c vs. 8c	10
8c vs. Mo	22
Mo vs. Bl	26
Bl vs. BDC	192
scNT embryos	
Oo vs. 2c	0
Fb vs. 2c	393
2C vs. 8c	0
8c vs. Mo	21
Mo vs. Bl	139
Bl vs. BDC	129

Note: Complete results of DESeq2 differential expression analyses for tRFs are provided in Supplementary File 3.3; Appendix H. Abbreviations are: Oo, oocyte; 2c, 2-cell; 8c, 8-cell; Mo, morula; Bl, blastocyst; BDC, blastocyst-derived cells; Fb, fibroblast; IVF, in vitro-fertilized; scNT, somatic cell nuclear transfer; miRNA, microRNA; ID, identification number.

Table 3.3. Differentially abundant piRNAs

Comparison	Number
IVF vs. scNT	
2-cell	0
8-cell	0
Morula	4
Blastocyst	0
Developmental stage	
IVF embryos	
Oo vs. 2c	1
Fb vs. 2c	4,586
2C vs. 8c	0
8c vs. Mo	49
Mo vs. Bl	70
Bl vs. BDC	1,766
scNT embryos	
Oo vs. 2c	0
Fb vs. 2c	4,672
2C vs. 8c	0
8c vs. Mo	24
Mo vs. Bl	0
Bl vs. BDC	717

Note: Complete results of DESeq2 differential expression analyses for piRNAs are provided in Supplementary File 3.3; Appendix H. Abbreviations are: Oo, oocyte; 2c, 2-cell; 8c, 8-cell; Mo, morula; Bl, blastocyst; BDC, blastocyst-derived cells; Fb, fibroblast; IVF, in vitro-fertilized; scNT, somatic cell nuclear transfer; miRNA, microRNA; ID, identification number.
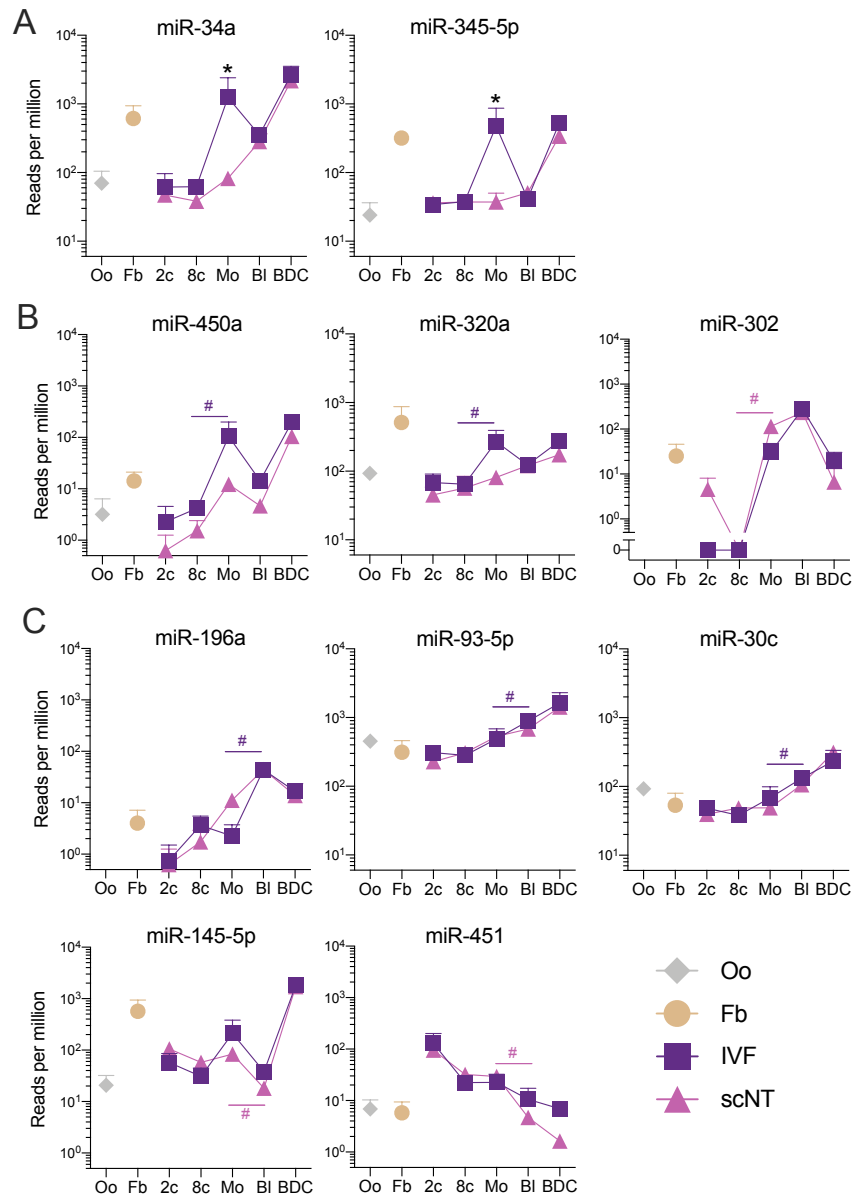
Supplementary Figure 3.S1. Dynamic expression of selected miRNAs of interest in early developing IVF and scNT embryos, oocytes and fibroblasts. Values shown are the normalized reads per million on a $\log_{10}$ scale. Selected miRNAs are shown to represent (A) miRNAs differentially expressed at the morula stage between IVF and scNT embryos, (B) miRNAs differentially expressed in morula vs. 8-cell IVF or scNT embryos, (C) miRNAs differentially expressed in blastocyst vs. morula IVF or ScNT embryos. * p <0.05 for IVF vs. scNT and # p <0.05 comparing by developmental stage as indicated and colored according to embryo type (purple, IVF; pink, scNT) as determined by DESeq2 analysis. Fb, fibroblast; Oo, oocyte; IVF, *in vitro*-fertilized embryo; scNT, scNT embryo; 2c, 2-cell; 8c, 8-cell; Mo, morula; Bl, blastocyst; BDC, blastocyst-derived cells.
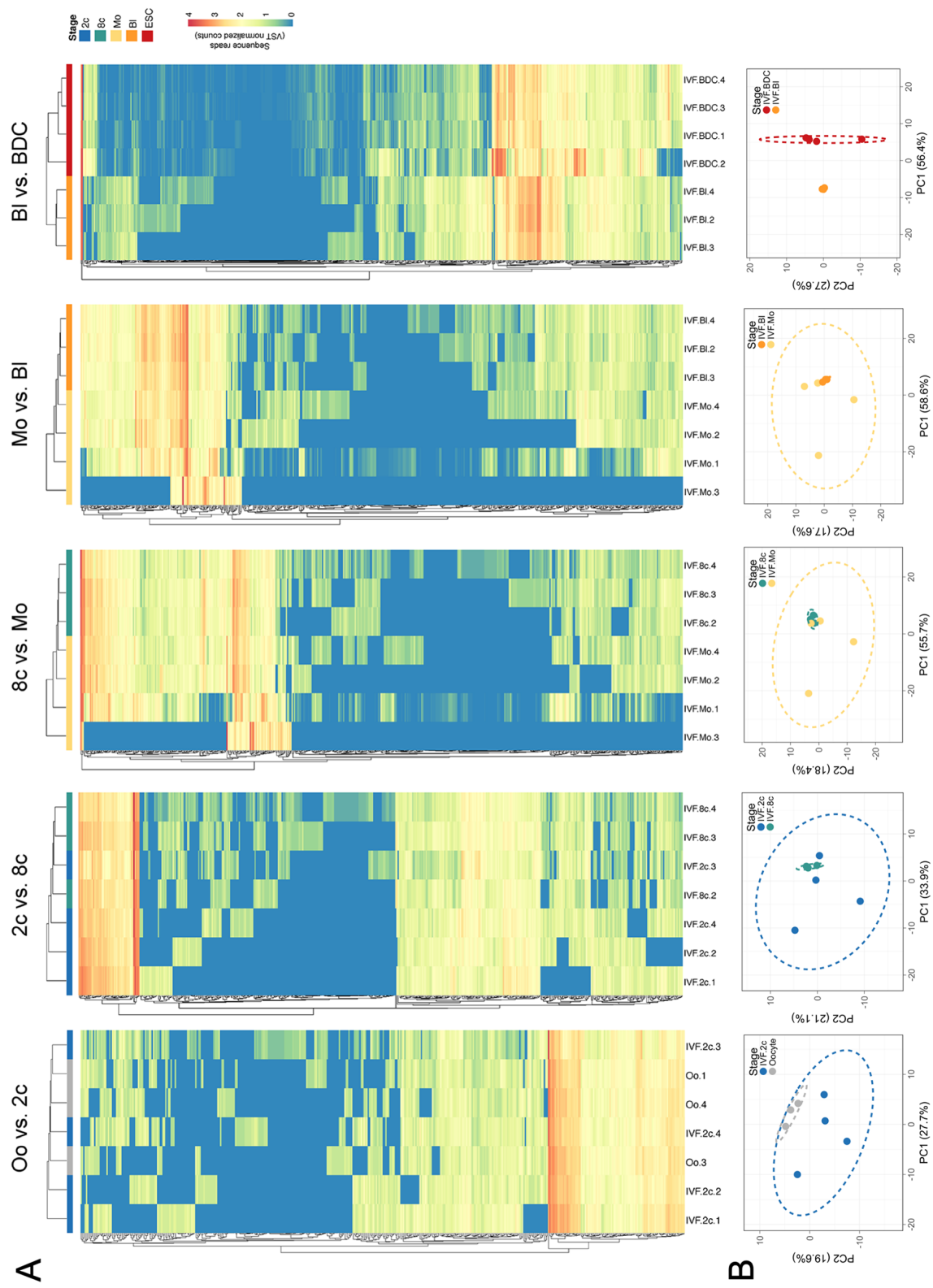
Supplementary Figure 3.S2. (prior page) Patterns of miRNA expression in IVF embryos with comparisons by developmental stage. (A) Unsupervised, bi-directional hierarchical cluster analysis (Euclidean distance method) of variance stabilized normalized read counts for all identified miRNAs present in either sample included in the heat map. Data were categorized by sample type (Fb, fibroblast; Oo, oocyte; IVF, *in vitro*-fertilized embryo; scNT, scNT embryo) and stage (2c, 2-cell; 8c, 8-cell; Mo, morula; Bl, blastocyst; BDC, blastocyst-derived cells). (B) Principal components analysis of bovine miRNAs present in either sample included, using the standard singular value decomposition method with imputation.
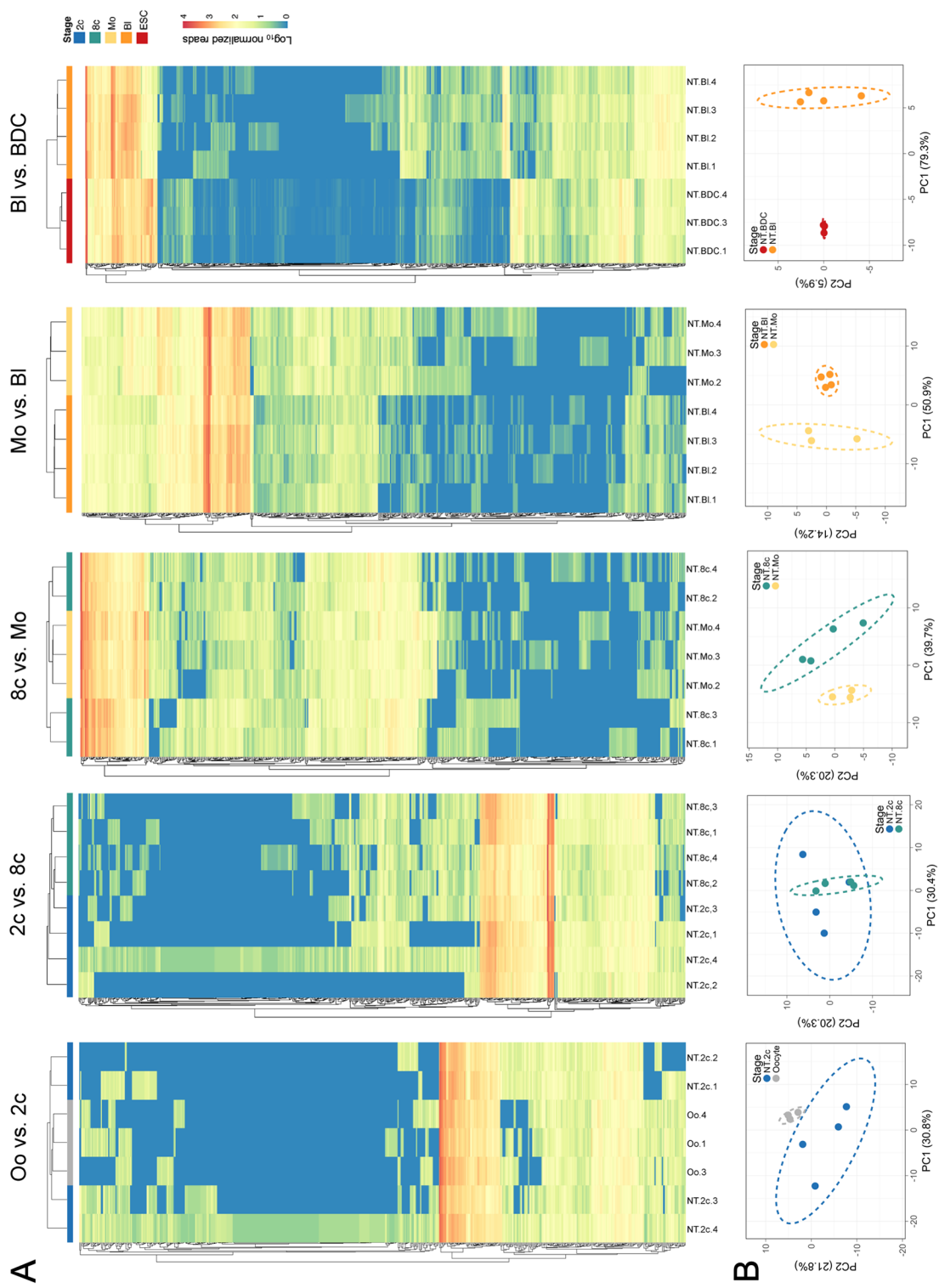
Supplementary Figure 3.S3. (prior page) Patterns of miRNA expression in scNT embryos with comparisons by developmental stage. (A) Unsupervised, bi-directional hierarchical cluster analysis (Euclidean distance method) of variance stabilized normalized read counts for all identified miRNAs present in either sample included in the heat map. Data were categorized by sample type (Fb, fibroblast; Oo, oocyte; IVF, *in vitro*-fertilized embryo; scNT, scNT embryo) and stage (2c, 2-cell; 8c, 8-cell; Mo, morula; Bl, blastocyst; BDC, blastocyst-derived cells). (B) Principal components analysis of bovine miRNAs present in either sample included, using the standard singular value decomposition method with imputation.

Supplementary Figure 3.S4. (prior page) Patterns of miRNA expression for comparisons of IVF and scNT embryos at each developmental stage. (A) Unsupervised, bi-directional hierarchical cluster analysis (Euclidean distance method) of variance stabilized normalized read counts for all identified miRNAs present in either sample included in the heat map. Data were categorized by sample type (Fb, fibroblast; Oo, oocyte; IVF, *in vitro*-fertilized embryo; scNT, scNT embryo) and stage (2c, 2-cell; 8c, 8-cell; Mo, morula; Bl, blastocyst; BDC, blastocyst-derived cells). (B) Principal components analysis of bovine miRNAs present in either sample included, using the standard singular value decomposition method with imputation.

A IVF: 8c vs. Mo

B

C IVF: Mo vs. Bl

D

Supplementary Figure 3.S5. Top 20 ontology and pathway terms and network clustering associated with predicted mRNA targets of over or under expressed miRNAs in IVF embryos. (A-B) Bovine IVF 8-cell vs. morula stage embryos and (C-D) bovine IVF morula vs. blastocyst stage embryos. For panels A and C, predicted mRNA targets (TargetScan total context ++ score <-0.35) of differentially expressed miRNAs ($|\log_2 R|>1$ and *p*-value <0.05) were subject to enrichment analysis using Metascape (Minimum overlap of 3, minimum enrichment 1.5 and *p*<0.01). Each heat map depicts the top 20 enriched summary terms identified for Gene Ontology biological processes, with the length of the bar determined by the $-\log_{10}$ *p*-value. For panels B and D, Cytoscape network analysis of gene ontology biological process terms for predicted mRNA targets of differentially expressed miRNAs. Each term is represented by a circle node, for which the size is proportional to the enrichment score and the color represents the enrichment pathway term. Terms with a similarity score >0.3 are linked by an edge (the thickness of the edge represents the similarity score) using force-directed network layout and bundled edges for clarity. Each cluster is labeled by the term with the highest *p*-value in the cluster group. Complete results of the Metascape analysis are provided in Supplementary File 3.5; Appendix J.

Supplementary Figure 3.S6. Top 20 ontology and pathway terms and network clustering associated with predicted mRNA targets of over or under expressed miRNAs in scNT embryos. (A-B) Bovine scNT 8-cell vs. morula stage embryos and (C-D) bovine scNT morula vs. blastocyst stage embryos. For panels A and C, predicted mRNA targets (TargetScan total context ++ score <-0.35) of differentially expressed miRNAs ($|\log_2 R|>1$ and p-value <0.05) were subject to enrichment analysis using Metascape (Minimum overlap of 3, minimum enrichment 1.5 and $p<0.01$). Each heat map depicts the top 20 enriched summary terms identified for Gene Ontology biological processes, with the length of the bar determined by the $-\log_{10} p$-value. For panels B and D, Cytoscape network analysis of gene ontology biological process terms for predicted mRNA targets of differentially expressed miRNAs. Each term is represented by a circle node, for which the size is proportional to the enrichment score and the color represents the enrichment pathway term. Terms with a similarity score >0.3 are linked by an edge (the thickness of the edge represents the similarity score) using force-directed network layout and bundled edges for clarity. Each cluster is labeled by the term with the highest p-value in the cluster group. Complete results of the Metascape analysis are provided in Supplementary File 3.5; Appendix J.

Supplementary Figure 3.S7. Identification of piRNA-like RNAs (pilRNAs) by ping-pong signature and 1U bias. Identification of pilRNAs by ping-pong signature and 1′U bias. Line graphs represent the number of pairs of reads with a 5′ to 5′ overlap in the datasets. The peak at 10 nt is characteristic of piRNAs produced by the ping-pong cycle. The sequence logos below represent nucleotide biases at each pilRNA position for reads 32 nt in length (base position number below each logo). Oo, oocyte; 8c, 8-cell stage embryo; Bl, blastocyst stage embryo.
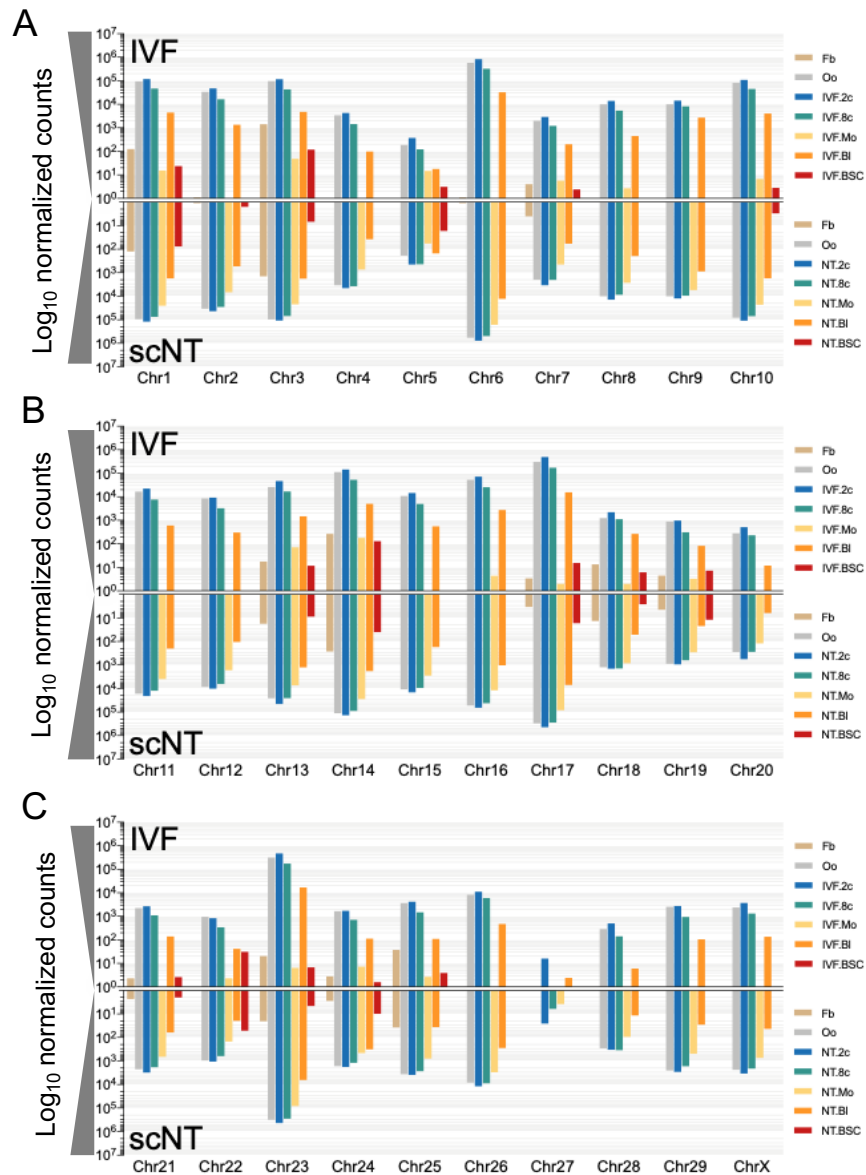
Supplementary Figure 3.S8. Potential piRNAs mapped to the *Bos Taurus* genome. Samples include fibroblasts, oocytes, IVF (above the x-axis) and scNT (below x-axis) two cell, eight cell, morula and blastocyst staged embryos. Values are the log$_{10}$ normalized counts for each sample type.

CHAPTER 4

MAPPING TRANSCRIPTIONAL CHANGES ASSOCIATED WITH THE

MATERNAL-TO-EMBRYONIC TRANSITION IN BOVINE EMBRYOS TO SMALL

NON-CODING RNA PROFILES IN BOTH IN VITRO-FERTILIZED AND SOMATIC CELL

NUCLEAR TRANSFER CATTLE EMBRYOS

Abstract

Background

The efficiency rate for the production of high quality embryos by somatic cell nuclear transfer (scNT) is far below that for *in vitro* fertilized (IVF) embryos, likely due to an accumulation of errors in genome reprogramming that impairs proper development. Moreover, errors associated with genome reprogramming in cloned embryos may also encompass dysregulation of expression of microRNAs (miRNA), an important class of non-coding RNAs that suppress translation or reduce stability of coding messenger RNA (mRNA). Thus, the objectives of this study were to determine the dynamics of mRNA expression in early developing scNT and IVF embryos in the context of the maternal-to-embryonic transition (MET) and to correlate apparent transcriptional dysregulation in cloned embryos with miRNA expression profiles. RNA sequencing was performed using cattle embryos produced via IVF or scNT at the 2-cell, 8-cell, morula and blastocyst developmental stages and using MII oocytes, donor cell fibroblasts and cells derived from either IVF or scNT blastocysts (BDCs). Sequencing data were analyzed by DESeq2 to identify transcripts differentially expressed in cloned embryos compared to their stage-matched IVF controls or to identify transcripts differentially expressed by developmental stage.

Results

Analysis of mRNA expression data revealed large-scale differences between scNT and IVF embryos at each developmental stage examined, with the greatest number of differentially expressed transcripts detected at the 8-cell and morula stages. Interestingly, those altered transcripts in 8-cell scNT embryos were associated with biological functions critical for the MET, such as mRNA processing and metabolism, ncRNA transcription and metabolism, methylation and chromatin modification. For two miRNAs previously identified as differentially expressed in scNT morulae, miR-34a and miR-345, negative correlations with some predicted mRNA targets were apparent, as would be expected if these miRNAs targeted the transcripts for repression or down-regulation, though these negative correlations were not widespread among all predicted targets.

Conclusion

Large-scale aberrations in expression of mRNAs were evident during the MET in cattle scNT embryos. However, this apparent dysregulation of gene expression in scNT embryos was not consistently correlated with aberrations in miRNA expression, suggesting that other mechanisms controlling gene expression may be at play.

Introduction

Somatic cell nuclear transfer (scNT) is a well-established method of animal cloning in livestock species. Using this method, genetic material is removed from a donor oocyte and replaced with a somatic donor cell's DNA. However, compared to *in vitro* fertilization technology, scNT embryos have low overall success rate, regardless of species [1, 2]. Improvement of scNT would make this assisted reproductive technology more accessible and allow for more efficient genetic improvement of livestock genetics. A possible cause of the low efficiency rates for development of scNT embryos is errors in the epigenetic reprogramming of

the donor somatic genome. Functioning as a second layer of coding information, the epigenome controls gene expression via changes in DNA structure that do not alter the genetic code, such as DNA methylation or histone modifications. In scNT embryos, the differentiated donor DNA must be reprogrammed into the pluripotent state necessary for successful embryo development. Errors in reprogramming that affect epigenome modifiers may lead to inappropriate gene expression and place high demands on the developing embryo [1]. Accordingly, inappropriate DNA methylation patterns [3-6] and histone modifications [7, 8] have been identified in scNT embryos, including blastocysts and aborted scNT fetuses from two to six months of gestation in cattle, and at the MZT (zygote) and blastocysts in mice. The significant cellular stress caused by extensive aberrant gene expression due to inappropriate genome reprogramming could be responsible, in part, for the low developmental rates of scNT embryos [9]. Of particular importance, proper genome reprogramming for genes involved in pluripotency is necessary for successful development to occur [10-12]. Interestingly, researchers have noted that patterns of epigenome programming in scNT embryos tend to resemble those of the donor cell, indicative of incomplete genome reprogramming in scNT embryos [13, 14]. It is likely that this incomplete reprogramming drives inappropriate gene expression and causes lower embryo development. Incomplete reprogramming would impact non-coding genes as well, such as those coding sncNRA, which could further impact aberrant gene expresison.

The maternal-to-embryonic transition (MET) is a highly orchestrated shift in genetic control from the maternal genome to the embryonic genome that may be especially sensitive to incomplete reprogramming in scNT embryos. In order for embryonic genome activation (EGA) to be successful, the maternally-deposited transcripts from the oocyte must be completely degraded to prepare the embryo for further development and tissue differentiation [15]. The shift in transcripts that occurs at the MET is quite profound, as maternal transcripts that may have been present in the oocyte and early embryo for weeks or months are eliminated in a matter of hours.

Three patterns of transcript changes are seen, including the destruction of oocyte-specific

transcripts that are not subsequently expressed in the embryo, the replacement of maternal

transcripts with embryonic transcripts (for ubiquitously expressed genes), and the *de novo*

transcription of embryonic transcripts that are not present in the oocyte [15]. During this time of

early development, transcripts can be regulated via changes in mRNA stability, translation, and

location. At the MET, mass degradation occurs with 30-40% of transcripts being completely

degraded and 60% being significantly degraded [16], driven by both embryonic and maternal

pathways. The maternal pathway involves several different degradation mechanisms, with the

primary of these being mediated by protein SMAD4, which acts to degrade the maternal

transcripts that would repress the embryonic genome [17]. Alternatively, embryonic degradation

machinery functions to rapidly degrade maternal transcripts still present [18]. Once the massive

turnover of transcripts is complete, embryonic gene expression actively drives embryonic

development allowing for differentiation to occur.

In bovine embryos, the major wave of the EGA occurs at the 8-16 cell stage, although

small amounts of transcriptional activation have been seen as early as the embryonic stage [19].

Graf and colleagues [19] determined that genes activated before the 4-cell stage in IVF embryos

were associated with RNA processing, translation, and transport biological functions; an uptick in

expression of these genes is likely necessary to prepare the embryo for genome activation and

active transcription and translation. For example, *KLF4* encodes a protein belonging to the

Kruppel family of transcription factors, and this gene was activated at the 4-cell stage. Similarly,

genes activated at the 8-cell developmental stage had functions associated with transcription and

translation, such as *EIF3*, which is required for the initiation of protein synthesis [19].

Researchers have shown that embryos generated via cloning harbored aberrations in their

populations of mRNA molecules, notably at the blastocyst developmental stage or later, with

expression patterns suggesting incomplete or faulty reprogramming of the donor genome [20, 21].

Moreover, the use of trichostatin A (TSA), a potent inhibitor of histone deacetylase, improved

development of scNT embryos. However, TSA did not appear to correct all reprograming errors,

suggesting that other mechanisms contribute to reprogramming errors in scNT embryos [22].

Prior to the EGA, transcript abundance is controlled via precise degradation pathways in early

developing embryos. Small non-coding RNAs (sncRNAs) function through RNA interference

(RNAi), the process by which ncRNAs block the translation or reduce the stability of the coding

mRNA. In mammals, microRNAs (miRNAs), piwi-interacting RNAs (piRNAs) and small

endogenous interfering RNAs (siRNAs) have been established as functioning in RNAi during

early development [23]. While the biogenesis and targeting mechanisms between these three

sncRNAs differ significantly, all can participate in RNAi through binding and degrading either

transposable elements (TEs) or mRNA [24]. MiRNAs act to silence mRNA expression through

binding interaction with the 3' UTR blocking translation and decreasing stability of target

transcripts [25]. MiRNAs are short, 22 nt single stranded RNAs that are transcribed from the

DNA by RNA polymerase II into primary-miRNAs (pri-miRNAs), which are then cleaved by

Drosha and its cofactor DGCR8. The cleavage results in precursor miRNAs (pre-miRNAs) that

are transported by Exportin proteins into the cytoplasm where the Dicer complex cleaves them

into double stranded, non-hairpin, miRNA molecules. AGO then unwinds and loads the miRNA

duplex into the miRNA-induced silencing complex (miRISC) [26]. The miRISC functions to

direct the miRNA to its target transcript. In cases of perfect complementary binding AGO will

mediate the target degradation [24], while in cases of non-complementary binding, miRNA

represses translation by inhibiting translation at the initiation step [25, 27]. MiRNAs have been

shown to be highly conserved, powerful regulators of gene expression, functioning in the

majority of investigated cell functions.

As the oocyte matures, it becomes transcriptionally silent, a developmental state that persists through early development until the EGA. The transcriptional silence makes the early embryo a unique window of development during which post-transcriptional machinery, such as miRNAs, may dominate the regulatory network. MiRNAs have been shown to contribute to the zygotic degradation pathway of maternal transcripts in Zebrafish, for which a single miRNA (miR-430) mediated the decay of hundreds of maternal mRNAs [28]. A similar mechanism of miRNA-mediated decay has also been found in *C. Elegans* by both maternal and zygotic miRNAs [29], in *Xenopus* by a miR-430 orthologue (miR-427) targeting mRNAs [30], and in *Drosophila* in which miR-309 degrades hundreds of mRNAs at the EGA [31]. Of note, miRNAs have not yet been shown to contribute to the embryonic degradation pathway in mammals, which may be due to heavy use of mouse models, as mice have a rodent unique sncRNA pathway in the oocyte and early embryo. In mice, piRNA proteins appear to be divergent from those in humans and cattle [32], and in mice and rats, a rodent oocyte-specific Dicer isoform preferentially loads siRNAs over miRNAs [33]. These species divergent pathways mean that further investigation in other mammalian species of miRNAs function in the MET is needed.

While most investigators have focused on aberrations in DNA methylation or histone marks to explain abnormalities in the transcriptome of scNT embryos, it is possible that deviations in populations of sncRNAs may also contribute to their poor development. Prior studies found that the donor cell miRNA expression pattern was persistent in the scNT bovine embryo or embryonic tissues [34, 35]. In addition to the limited studies that have examined miRNA populations in bovine embryos, researchers demonstrated that manipulation of specific miRNAs impacted rates of development in embryos [36, 37], including the improvement of scNT embryo rates [38-43]. Dysregulation of miRNA populations – and their transcript targets – necessary for a successful MET could contribute to high embryonic stress and death rates typical for production of scNT embryos.

While others have shown that scNT embryos harbor distinct transcriptomes compared to high quality embryos (IVF or *in vivo* produced), many of these studies were limited in that none have explored the dynamics of gene expression over the course of early embryo development through the MET [21, 44]. No prior studies have employed a discovery-based RNAseq approach to examine the entire population of miRNAs, and only a single recent study examined the entirety of the mRNA transcriptome through the MZT in scNT embryos [45]. Moreover, it is not known how aberrations in the transcriptome of scNT embryos correlate with populations of miRNAs during early embryo development. Thus, in this study, we sought to examine the dynamic changes in the transcriptome in scNT embryos through the MET and to examine those variations in context of the population of miRNAs in the same embryo samples. We hypothesized that transcripts derived from the maternal genome that need to be degraded for successful MET would be aberrantly expressed in the scNT embryos at the 8-cell stage. In addition, we hypothesized that samples found to have unique changes in miRNA populations would also show a distinct transcriptome profile in transcripts predicted to be targets of those unique miRNAs. An unbiased, discovery-based approach was employed using RNAseq of both small (<200 nt in length) and large (>200 nt in length) RNA populations in bovine oocytes, fibroblast donor cells, 2-cell embryos, 8-cell embryos, morula embryos, blastocyst embryos, and blastocyst-derived cells (BDCs) produced using both IVF and scNT methods.

Methods

Oocyte collection and in vitro maturation

Bovine ovaries were collected from a local abattoir (JBS, Hyrum, UT) and transported in a cooler containing 0.9% saline solution to the laboratory. The cumulus-oocyte complexes (COCs) were then aspirated from 3-8 mm follicles by using an 18-gauge needle and vacuum system. Only compact COCs with homogenous ooplasm and intact layers of cumulus cells were

used for scNT or IVF. Following aspiration, COCs were cultured at 39 °C with 5% CO_2 for 22 to 24 hr. The oocytes were cultured in TCM199 maturation medium with Earle's salts, L-glutamine, and sodium bicarbonate (Hyclone, Logan, Utah) supplemented with 10% fetal bovine serum (FBS), 0.05 mg/ml bovine follicle stimulating hormone (Sioux Biochemicals, Sioux city, IA), 5 mg/ml bovine luteinizing hormone (Sioux Biochemicals), 100 U/ml penicillin, and 100 mg/ml streptomycin.

In vitro fertilization

Following 22 to 24 hr of maturation, MII oocytes were fertilized using the laboratory's standard *in vitro* fertilization (IVF) protocol [46]. Briefly, one straw of cryopreserved bovine semen obtained from a Holstein bull (Hoffman AI, Logan UT) was removed from the liquid nitrogen tank and placed into a 35 °C water bath to thaw. Live sperm were isolated by centrifugation through a 45%/90% percoll gradient, suspended (final concentration $1x10^6$/ml) in Tyrode's albumin lactate pyruvate containing 10 μg/ml heparin and used to fertilize the mature oocytes. Twenty to 22 hr post-IVF, cumulus cells were removed by vortexing, and fertilized zygotes were washed in phosphate buffered saline with 0.32 mM sodium pyruvate, 5.55 mM glucose, and 3 mg/ml bovine serum albumin (PB1+). After washing, zygotes were cultured on a monolayer of bovine cumulus cells in 50 μl of synthetic oviductal fluid (SOFaa) with 3% FBS overlaid with mineral oil. The embryos were cultured at 39 °C in a humidified atmosphere with 5% CO_2. Half of the SOFaa medium was removed and replaced with fresh, equilibrated medium every other day starting the day after *in vitro* culture. Historically, production of IVF embryos via these methods followed by blastocyst stage embryo transfer to cattle recipients has generated successful pregnancies at an approximate rate of 50% by our research group (unpublished observations). The timing of embryo collection post-fertilization was 24-36 hr for 2-cell stage, 48-60 hr for 8-cell stage, 156-168 hr for morula stage, and 192-204 hr for blastocyst stage.

Somatic cell nuclear transfer

Primary bovine fibroblast cultures were established from ear biopsy tissues from a Brahma Spanish bull cross using well-established procedures. Frozen/thawed cells were grown to 80-100% confluence then passaged, with cells from passages 4-5 used as nuclear donors. Three days prior to a cloning session, donor fibroblast cells were thawed and propagated, and incubated in DMEM media supplemented with 15% (v/v) FBS at 39 °C with 5% CO_2. Bovine fibroblasts at 80-90% confluence were serum starved by replacing culture media with DMEM media containing 0.5% (v/v) FBS 24 hr prior to scNT. Oocytes were matured for 18-20 hr, then denuded using 100 μl of 1% (v/v) hyaluronidase, incubated for 5 min at 39˚C with 5% CO_2, followed by gentle pipetting. The removed cumulus cells were centrifuged at 500×*g* for 5 min at room temperature, and the supernatant was removed. Cumulus cells were then cultured in SOFaa under oil for use in embryo culture. The denuded oocytes were then rinsed using PB1+, and oocytes with polar bodies were selected and separated. ScNT was performed according to established protocols [47-49]. Briefly, oocytes with polar bodies were incubated in 0.6 μg/ml demecolcine for 30-40 min and a metaphase plate and a polar body were removed using a beveled pipette. Fibroblast cells were covered in 0.25% trypsin for 1 min, and then the trypsin was removed and fibroblasts were incubated for anther 6 min at 39˚C with 5% CO_2. Once fibroblast cells were detached, they were rinsed in warm medium and centrifuged at 150xg for 6 min. The supernatant was removed, and the cellular pellet was resuspended in 100 μl Hepes SOF medium [50]. One fibroblast cell was injected into the perivitelline space of the recipient oocyte, and fused using one direct pulse of 1.2 kV/cm for 22 μs by an Electro Cell Manipulator 2001 (BTX, San Diego, CA) in 0.28 M sorbitol, 0.05% (w/v) BSA, 0.1 mM $CaCl_2$, 0.5 mM $MgCl_2$, and 0.5 mM Hepes. Following fusion, embryos were washed through Hepes SOF, and incubated in embryo culture medium for 1 hr. Activation was then performed, with successfully fused

embryos cultured in 5μM ionomycin for 5 min, followed by 4 hr of incubation in activation

medium composed of SOFaa medium with 10μg/ml cyclohexamide and 1 mM 6-dimethylamino-

pyridine at 39˚C with 5% CO_2. Following activation, embryos were cultured on a monolayer of

bovine cumulus cells in 50 µl drops of SOFaa with 3% (v/v) FBS overlaid with mineral oil. The

embryos were cultured at 39 °C in a humidified atmosphere with 5% CO_2. Half of the SOFaa

medium was removed and replaced with fresh, equilibrated medium every other day starting the

day after *in vitro* culture. The timing of embryo collection was the same as for IVF embryos as

described above.

Isolation of blastocyst-derived cells

We followed methods previously described [51] to generate putative embryonic stem

cells (ESCs) from IVF and scNT embryos for comparison. Briefly, bovine expanded blastocysts

produced by IVF and scNT were used for BDC isolation. Zona pellucida-free blastocysts were

placed into 4-well culture dishes onto feeder layers of mitomycin C-treated mouse primary

embryonic fibroblasts. Feeder layer preparation was performed as previously reported [52].

These cells were cultured in DMEM supplemented with 0.2% (w/v) LIF, 1% (w/v) bFGF, and

30% (v/v) Knock out Serum Replacement (ThermoFisher Scientific, Waltham, MA). When

propagation was needed, cultured cell colonies were mechanically propagated using a small metal

blade. Germ line competent ESCs have only been reported for mouse and rat [53]. However,

bovine putative ESC-like cells expressing pluripotency markers have been reported by several

groups [51, 54-56]. Though the objective here was to generate putative ESC-like cells,

transcriptional profiling revealed that these cells did not consistently express the array of gene

markers typical of pluripotent stem cells (data available with Chapter 4). Thus, we instead refer

to these cells as blastocyst-derived cells (BDCs) herein.

Tissue collection

The same samples were used to isolate the small RNA fraction for profiling sncRNAs and the large RNA fraction for profiling mRNAs so that direct comparisons of the two RNA profiles could be made. Four pools of 20 mature oocytes were collected at 22 hr after maturation. To remove cumulus cells, the mature COCs were treated with 10 mM hyaluronidase for 5 min, followed by repeatedly pipetting the mixture until cumulus cells were removed. After visual inspection for complete removal of cumulus cells and the presence of a polar body, denuded mature metaphase II (MII) oocytes were washed through four droplets of PB1+. Oocytes were then snap frozen in cryotubes containing RNA/DNA shield (Zymo, Irvine, California), and stored at -80 °C until RNA isolation. Fibroblasts prepared for scNT were collected after centrifugation and pelleting, and were snap-frozen in 100ul of RNA/DNA shield (Zymo, Irvine, California). Four pools of 20 2-cell, 8-cell, morula, and blastocyst stage embryos were collected at 2, 3, 6, and 8 days post-IVF or scNT, washed in PBS, and stored at -80 °C.

Large RNA isolation and sequencing

RNA was isolated based on size using the RNA Clean & Concentrator™ 5 kit (Zymo) from three pools each of oocytes, 2-cell stage embryos, 8-cell stage embryos, morula stage embryos, or blastocyst stage embryos according to the manufacturer's protocol for purification of small RNA (<200 nt) and large RNA (>200 nt) as separate fractions. Small RNA sequencing was performed as outlined in Chapter 3. The large RNA fraction was used to generate sequencing libraries using the Ovation SoLo RNA-Seq System (NuGen Technologies; San Carlos, CA) as directed by the manufacturer's protocol for total RNA input. Briefly, first strand cDNA synthesis, cDNA processing, second strand synthesis, and end repair were conducted according to manufacturer's protocols, and samples were then stored at -20°C overnight. Then, adaptor ligation and purification were performed, followed by PCR amplification of the sequence

libraries (optimized to 16 cycles per instructions) and bead purification (twice) of the resulting

samples. Library DNA concentrations were quantified using the Qubit dsDNA HS Assay

(ThermoFisher Scientific). Because DNA concentrations were low, the PCR amplification step

was repeated once more with an additional 19 cycles to improve yield. Next, InDA-C treatment

was conducted, a third round of PCR was performed, and the resulting libraries were bead-

purified according to the manufacturer's protocol. Libraries were sequenced on the Illumina

NextSeq using single-end 75 bp reads by the USU Genomics Core Facility.

Data processing and analyses

Methods for processing and analysis of sncRNA sequencing data are provided in Chapter

3. For the large RNAseq data, high quality sequencing data were confirmed with FastQC [57],

and visualized across samples with MultiQC [58]. A STAR index was generated for read

alignment using *Bos taurus* UMD3.1 using the parameter --sjdbOverhang 74 [59]. Reads were

aligned with default parameters including --quantMode GeneCounts to generate feature count

tables (Supplementary File 4.1; Appendix L). Count tables were imported to R where the DESeq2

R package [60] was used to identify differentially expressed genes with a Benjamini–Hochberg

false discovery rate (FDR) adjusted *p*-value < 0.05. Comparisons were made for IVF vs. scNT

embryos at each developmental stage (2-cell, 8-cell, morula, blastocyst, and BDC) as well as

between developmental stages for each embryo type (oocyte vs. 2-cell, 2-cell vs. 8-cell, 8-cell vs.

morula, morula vs. blastocyst, and blastocyst vs. BDC). BioVenn was used to create proportional

Venn diagrams depicting the number of transcripts mapped to the cattle genome for each sample

[61]. Unsupervised, bidirectional hierarchical cluster analyses (Euclidean distance method with

average linkage) and principal component analyses were performed using ClustVis (large data

version) [62], but due to the large number of transcripts detected (over 20,000), only 2,400 genes

were randomly selected to include for hierarchical clustering and principal components analyses.

Lists of all differentially expressed genes (DEGs) (Supplementary File 4.2; Appendix M) were then subject to gene set enrichment analysis using Metascape and the Gene Ontology (GO) biological process database with the following parameters: minimum overlap of 3 genes in a category, minimum enrichment of 1.5, and p <0.05 [63]. Ontology analyses were performed using DEG sets for comparison between IVF and scNT embryos at each developmental stage and for comparisons by developmental stage within each embryo group. Resulting GO biological process terms were then clustered by semantic similarity (kappa score >0.3) and represented as networks in Cytoscape [64] (Supplementary File 4.3; Appendix N). Clusters of related terms were named according to the term within the cluster with the lowest p value; thus, clusters with different labels may have some similar terms within the cluster.

In Chapter 3, we identified a few differentially expressed miRNAs in scNT versus IVF embryos, or by sequentially comparing developmental stages. Here, one objective was to compare patterns of expression for those differentially expressed miRNAs to the relative abundance of putative target transcripts in IVF or scNT embryos. Thus, the list of mRNAs was filtered to include only predicted targets of those select miRNAs using TargetScan (release 7.2) [65] with the *Bos taurus* miRNA database (miRBase release 22). TargetScan predicts biological targets of miRNAs by searching for the presence of a conserved 8-mer, 7-mer, and 6-mer sites matching to the seed region of each miRNA. Then, correlation analyses were performed using cor and corrplot functions in R (www.R-project.org). A correlation was considered significant when the Spearman ρ > 0.5 or <-0.5 and p < 0.05. Last, networks were constructed linking those differentially expressed miRNAs and predicted mRNA targets using Cytoscape.

Results

Expression profiles of mRNAs in IVF and scNT embryos

Figure 4.1 illustrates the overlap in the number of sequences mapped to the cattle genome for each developmental stage (panel A) or comparing IVF and scNT embryos to either oocytes or donor fibroblast cells (panel B). The vast majority of transcripts sequenced were shared by each developmental stage, with an evident increase in the number of unique transcripts detected after the EGA in morula and blastocyst stage IVF and scNT embryos. Also, approximately 4000 fewer mRNAs were identified in scNT morula and blastocyst embryos (approximately 14,800 detected) than their IVF counterparts (more than 19,000 detected). More transcripts were in common with oocytes for each of the developmental stages than for fibroblasts, for both IVF and scNT embryos (Figure 4.1B). However, we also noted that the set of mRNAs identified in BDCs generated from both IVF and scNT embryos had greater overlap with fibroblasts than did the developing embryos.

Complete results for all DESeq2 differential expression analyses are provided in Supplementary File 4.2, and plots for individual genes of interest are provided in Supplementary Figure 4.S1. When comparing scNT embryos to their IVF controls, substantially more DEGs were identified in 8-cell and morula stage embryos, 1539 and 2140, respectively, than in earlier or later developmental stages with fewer than 300 DEGs (Table 4.1, Figure 4.2A). Also, for each type of embryo (IVF or scNT), we identified DEGs by progressively comparing developmental stages (Table 4.1, Figure 4.2B). Relatively few DEGs were identified when comparing 2-cell embryos to oocytes, whereas an increase in transcriptional activity was clearly evident by the 8-cell stage with nearly 3000 DEGs identified in 8-cell vs. 2-cell IVF embryos. Interestingly, transcriptional activity was reduced in scNT embryos at this transition point, with only 1092 DEGs identified when comparing 8-cell to 2-cell scNT embryos. The greatest boost in gene expression in developing embryos was noted when comparing morula to 8-cell staged embryos,

with 7106 and 10,942 DEGs for IVF and scNT embryos, respectively. The markedly greater

number of DEGs identified in scNT embryos at this later developmental transition (8-cell to

morula), coupled with the relatively low number in the prior transition (2-cell to 8-cell) suggests

that transcriptional activation may have been delayed in the scNT embryos. Indicative of

continuing transcriptional activity with further development, the morula to blastocyst transition

included 6,678 and 5,859 DEGs for IVF and scNT embryos, respectively. Last, when comparing

blastocyst stage embryos to the transcriptional profiles of the BDCs, substantial transcriptional

differences were noted with 8814 and 8623 DEGs identified for BDCs obtained from IVF or

scNT embryos, respectively.

Principal component analysis of the transcriptome data showed that samples clustered

clearly by developmental stage, although there was a clear segregation of samples by embryo

type at the morula stage (Figure 4.3A). Furthermore, unsupervised, bidirectional hierarchical

cluster analysis showed that samples were clearly grouped primarily by developmental stage,

with 2-cell and 8-cell embryos clustered in the same tree as oocytes, and morula and blastocyst

embryos clustered as separate sub-trees very distant from earlier developing embryos (Figure

4.3B). Fibroblasts and BDCs were also clustered, but the fibroblast's sub-trees clustered

distinctly from other later-development stages. Interestingly, within the developmental stages,

the transcriptional profiles of IVF and scNT embryos were sufficiently distinct that the samples

clustered by type, with IVF and scNT embryos clustered separately for the morula and blastocyst

stages and for BDCs (Figure 4.3B; Supplementary Figure 4.S2). However, IVF and scNT

transcriptional profiles for 2-cell and 8-cell embryos were not sufficiently different to cause

complete segregation by embryo type for those stages.

Principal components and hierarchical clustering were also performed for each embryo

type to compare transcriptional profiles progressively by developmental stage (Supplementary

Figures 4.S3 to S4). For IVF embryos, distinct transcriptional profiles were evident by the 8-cell

stage, at which point the pattern of gene expression was clearly distinct from that for morula stage

embryos. This pattern persisted through the blastocyst stage and was also evident when

comparing blastocysts to BDCs and fibroblast cells (Supplementary Figure 4.S3). Interestingly,

for scNT embryos, embryos segregated by developmental stage for all stages examined, from 2-

cell through blastocyst and for comparisons of blastocyst to BDCs and fibroblast donor cells

(Supplementary Figure 4.S4).

Ontology analysis of mRNA populations in IVF and scNT embryos

Results of biological process gene ontology analysis of mRNAs differentially expressed

in scNT versus IVF embryos at each developmental stage are summarized in Figures 4.4-4.8 and

in Supplementary File 4.3. Note that bar charts represent the top 20 terms (clustered by kappa

score for similarity) for each data set. At the 2-cell stage, the most significant terms and highly

enriched terms included "extracellular structure organization", "plasma lipoprotein particle

clearance", and "smooth muscle cell proliferation" (Figure 4.4). Notably, at the 8-cell stage, clear

trends in biological processes dysregulated in scNT embryos compared to IVF were evident, with

terms related to epigenetic control of gene expression highly enriched, including "covalent

chromatin modification" and "methylation" (Figure 4.5). Terms related to small or non-coding

RNA were also enriched at this developmental stage only, including "ncRNA metabolic process",

"ncRNA transcription", and "tRNA processing". Also, terms related to RNA metabolism and

functions were also highly enriched, including "regulation of mRNA processing" (and variations

thereof), "RNA localization", and "translation". Last, two terms related to development were

associated with transcriptional dysregulation in scNT embryos, "chordate embryonic

development" and "stem cell population maintenance." Of note, the enrichment scores and

significance p values were the highest and lowest for comparison of scNT to IVF embryos at the

8-cell developmental stage.

After the EGA, the collection of biological processes associated with DEGs for scNT and IVF morulae was distinct from the prior developmental stages, with enriched terms including those associated with early nervous system development enriched (e.g., "dendrite development", "cell surface receptor signaling pathway involved in cell-cell signaling", "positive regulation of synaptic transmission", and "signal release") and cellular organization ("actin cytoskeleton organization" and "mitochondrion organization") (Figure 4.6). At the blastocyst stage, top biological process terms for DEGs in scNT embryos included "synapsis", "regulation of receptor localization to synapse", and "chloride transport" (Figure 4.7). Last, the array of terms associated with differential expression in BDCs derived from scNT embryos compared to IVF included many associated with metabolic processes, such as "aspartate family amino acid process", "small molecule catabolic process", and "phosphatidylinositol metabolic process" (Figure 4.8) as well as a few terms connected to development ("regulation of cell morphogenesis", "head development", and "regulation of collateral sprouting").

Ontology analyses were also performed for sets of DEGs for progressive comparison by developmental stage (Supplemental Figures 4.S5 to S8). The most interesting feature of these analyses in both IVF and scNT embryos was the identification of terms with functions relevant to embryonic genome activation at the 8-cell developmental stage (Supplementary Figures 4.S6 and 4.S8), including "mRNA processing", "RNA localization", and "ncRNA metabolic process" as examples. Also enriched at this stage were terms associated with "cellular response to stress", "DNA repair" and regulation of the cell cycle checkpoint (*i.e.*, terms clustered with "signal transduction by p53 class mediator" in scNT embryos or "response to cellular stress" in IVF embryos). For comparisons beyond the 8-cell developmental stage, no significant enrichment in ontology terms was noted for differentially expressed genes in either IVF or scNT embryos.

Correlation of mRNA and miRNA expression in IVF and scNT embryos

As outlined in Chapter 3, several miRNAs were identified as differentially expressed in scNT versus IVF cattle embryos. Here, we selected miR-34a and miR-345 for further exploration, as these miRNAs have putative roles in early embryogenesis and they were significantly over-expressed in scNT morula stage embryos compared to IVF. Relative expression of these miRNAs was compared to a subset of mRNAs identified as predicted targets as described above. Correlation plots were generated for the full set of predicted targets for each miRNA and for a subset of transcripts that were identified as differentially expressed in scNT versus IVF embryos at the morula stage. MiR-34a was significantly negatively correlated with 24 of the 57 predicted target mRNAs that were detected by this RNAseq experiment and positively correlated with 17 other mRNAs (Supplementary Figure 4.S9). The strongest negative relationship in expression was evident for genes *ELMOD1*, *TMED8*, *PTPN4*, *TMCC3*, *MLLT1*, *SIDT2*, *PNOC*, *RSPO4*, *SIX3*, *UHRF2*, and *EVI5L*. When further refining this list to include only those transcripts differentially expressed in scNT morulae, clear inverse relationship in expression evident for *RALGDS*, *AMER1*, *SIPA1*, *EVI5L*, and *HCFC2* (Figure 4.9; Supplementary Figure 4.S10). *WNT2B* was the only predicted target of miR-34a that was differentially expressed in scNT morula embryos but not negatively correlated with miR-34a expression. A similar analysis was performed for miR-345, however a clear trend toward negative correlations for this set of 104 target transcripts was not apparent (Supplementary Figures 4.S11). When the target transcript list was filtered to include only those mRNAs differentially expressed in scNT morula stage embryos, significant negative correlations were noted for *SENP8*, *TINF2*, *CREBL2*, *NOSTRIN*, *VSX2*, *GNB5*, *FHAD1*, and *PDK2* (Figure 4.9; Supplementary Figure 4.S12). Alternatively, no correlation was evident for miR-345 and expression of HIGD1A or F3, and a positive correlation was noted for *EIF3L*, suggesting that these mRNAs were not degraded by miR-345.

Network analysis of differentially expressed miRNAs and
predicted mRNA targets

To determine whether predicted mRNA targets were inversely regulated by their miRNA

families, a Cytoscape network was created, where differentially expressed miRNAs in IVF and

scNT morula stage embryos are connected to their predicted mRNA targets. Of note, very few

mRNA transcripts were targeted by both miR-34a and miR-345 (Figure 4.10). As was noted for

the correlation analyses, these network visualizations show that the pattern of expression of

predicted mRNA targets is fairly mixed with no clear trend for decreased mRNA expression in

the context of elevated miRNA expression. Some exceptions are noted (medium- to dark-purple

mRNA nodes connected to yellow miRNA nodes), as was evident in the correlation analyses

(Supplementary File 4.4; Appendix O) contains resulting networks for changes in miRNA and

mRNA within type for stage transitions 8-cell to morula stage embryos and morula to blastocyst

stage embryos, as these were the only stage transitions with significantly different miRNA

expression profiles. These networks were all markedly more complex with many more

overlapping miRNA targets among the differentially expressed miRNAs, as more differences in

miRNA expression were noted when comparing by developmental stage than when comparing

scNT versus IVF embryo types. (No differentially expressed miRNAs were identified for oocyte

vs. 2-cell or for 2-cell vs. 8-cell stage comparisons for either IVF or scNT embryo types, and thus

networks were not generated.) When comparing 8-cell to morula stage IVF embryos, most

differentially expressed miRNAs were over-expressed in morulae (Supplementary File 4.4;

Appendix O); of those, a higher fraction of down-regulated transcripts was connected to miR-

214. However, each subnetwork was composed of a mix of up- and down-regulated transcripts

with no overwhelming trend toward suppression of mRNAs by an over-expressed miRNA or vice

versa. This pattern was also evident when examining the miRNA-to-mRNA network for IVF

blastocyst embryos (Supplementary File 4.4; Appendix O). Networks for scNT embryos

transitioning from 8-cell to morula were also mixed in their expression profiles, though a greater proportion of predicted targets for some miRNAs were inversely expressed, such as for miR-378d or miR-138 (Supplementary File 4.4; Appendix O). Last, in the comparison of scNT morula- to blastocyst stage embryos, many more miRNAs were under-expressed in the blastocyst embryos (Supplementary File 4.4; Appendix O). However, the networks do not show a clear pattern of corresponding over-expression for most mRNA targets. Rather, expression values for each of the subnetworks of miRNA-to-mRNA targets appear fairly heterogeneous. In summary, for each of these networks, only a fraction of the predicted mRNA targets depicts the anticipated inverse expression pattern for the differentially expressed miRNA.

Discussion

The present study is the first to employ a discovery-based, RNA sequencing approach to compare both sncRNAs and mRNAs transcriptional profiles of IVF and scNT embryos during the MET in cattle. In addition, sncRNA sequencing was employed on the same samples, demonstrating that small quantities of biological sample can be successfully utilized for both sncRNA and mRNA sequencing approaches with the production of reasonable numbers of scNT embryos. Our analysis of mRNA expression data revealed large-scale differences between IVF and scNT embryos at each developmental stage examined. Aberrant expression of genes that participate in developmentally important pathways could restrict the developmental competency of scNT embryos. However, the overall transcriptional profiles primarily reflected the developmental stage with pre-MET embryos, post-MET embryos, and more differentiated cells harboring distinct patterns of mRNA expression. Though a core set of transcripts was shared between differentiated fibroblasts and all embryo types, the clustering and principal components analyses did not suggest that mRNA profiles of scNT embryos were broadly similar to their donor cell fibroblasts. Previously, we determined that three miRNAs were differentially

expressed in morula staged IVF and scNT embryos. Interestingly, results of this study revealed the largest number of differentially expressed transcripts in IVF and clones also at the morula stage. Taken together, these observations suggest that the greatest deviations in RNA expression were manifest shortly after the EGA in cattle. By sequencing both large and small fractions of the population of RNA molecules, this study provided us the ability to correlate expression patterns of miRNAs and their putative mRNA targets in the same embryo samples. For two miRNAs identified as differentially expressed in morulae, miR-34a and miR-345, negative correlations with some predicted mRNA targets were apparent, as would be expected if these miRNAs targeted the transcripts for repression or down-regulation. However, when considering the whole population of potential targets, these negative correlations were not widespread, suggesting other mechanisms controlling gene expression were likely at play .

Differences in transcript abundance detected at the 2-cell stage reflect populations present in the oocyte and transcripts originating from either the sperm for IVF embryos or donor cell for scNT embryos. At the 2-cell stage, differentially expressed transcripts were associated with key functions in early development, as exemplified by *FGF7*, which was expressed at higher levels in scNT embryos and contributes to embryonic development, cell growth, and morphogenesis [66]. *PRRX1* was also more highly expressed in scNT embryos and is a member of the homeobox family, acting as a transcription co-activator for growth and differentiation factor response. Interestingly, there is evidence for post-transcriptional regulation of *PRRX1* gene in embryos [67]. Genes that were under-expressed in scNT 2-cell stage embryos included *CATSBER*, a gene required for sperm motility transcripts of which were likely provided by sperm in the IVF embryos [68]. Collectively, differentially expressed transcripts in 2-cell embryos were associated with biological functions related to cell proliferation, division, and growth. However, as some cyclins and cyclin dependent kinases appear to be dispensable at this stage, it is possible that changes in these pathways in scNT embryos do not negatively impact development until later in

development [69]. Differentially expressed transcripts in scNT embryos were also associated with ontology terms "smooth muscle proliferation" and "epithelial cell proliferation", which may be derived from the differentiated fibroblast nuclear donor causing aberrant expression in scNT embryos. In addition, enriched terms included positive regulation of *PI3K* signaling, which is a key pathway to control cellular differentiation in the embryo; loss of this pathway is embryonic lethal [70].

In cattle, the EGA occurs at the 8-cell stage, and abnormalities in gene expression during this critical transition at could impair the developmental competence of scNT embryos. Several developmentally important genes were up-regulated in scNT 8-cell stage embryos, including *YAP1*, which restricts proliferation and promotes apoptosis [71], as well as continued over-expression of *FGF7*. Of particular interest was the apparent over-expression of *DICER1* and *DROSHA* in scNT 8-cell embryos, as these genes code for key biogenesis proteins involved in processing miRNA. Thus, one may expect excess miRNA biogenesis in 8-cell scNT embryos, although we did not detect a significant over-abundance of any specific miRNA at this stage (Chapter 3). However, it is possible that the lack of differential miRNA expression was due to the apparent down regulation of *XPO1* in the scNT embryos, as this transcript codes for EXPORTIN 1. Although EXPORTIN 5 is the primary Exportin protein responsible for pre-miRNA export into the cytoplasm, EXPORTIN 1 promotes the processing of pri-miRNA to pre-miRNA in *C. Elegans* and *Drosophila* [72]. *NOBOX* was also expressed at higher levels in scNT embryos compared to their IVF counterparts at the 8-cell stage; this gene is oocyte-specific and degraded by the miRNA miR-196a [73]. That *NOBOX* was higher in 8-cell embryos but miR-196a was not correspondingly decreased, suggests that the abundance of this miRNA may have been insufficient to effectively target *NOBOX* transcripts, or possibly other miRNAs may cooperate in its down regulation. (Incidentally, the TargetScan database does not map any known cattle miRNAs as targeting *NOBOX* transcripts.) *DPPA2* (*developmental pluripotency associated 2*)

expression was decreased in 8-cell scNT embryos; this developmentally-important gene was shown to be necessary for the activation of the first group of genes activated in EGA in mice [74]. It is possible that down-regulation of *DPPA2* in scNT embryos would delay the onset of EGA and may interfere with proper genome reprogramming. Interestingly, down-regulated genes in scNT embryos at the 8-cell stage also included *NANOG*, a key pluripotency factor [75]. *NANOG* is critical for maintenance of pluripotency in stem cells and represses differentiation of early embryonic cells in a multifaceted manner, as *NANOG* is both an activator and repressor of multiple gene targets of pluripotency and differentiation [76]. Though *OCT4* has been shown to be required for *NANOG* expression in bovine embryos [77], expression of this gene was not significantly different in IVF and scNT embryos at the 8-cell stage, suggesting other mechanisms may be responsible for altering *NANOG* expression. Aberrant expression of this key transcriptional factor could greatly compromise the developmental competence of scNT embryos, especially occurring at the MET when embryonic transcripts must be regulated carefully. Also of interest was the apparent over-expression of *NOTCH2* in 8-cell scNT embryos, as this gene is involved in cell fate determination in early embryogenesis.

In a previous study using RNAseq to explore transcriptional changes associated with EGA in bovine embryos, Graf et al. determined that embryonic genes activated at the 8-cell stage were involved in "DNA dependent transcription", "purine nucleotide metabolic processes", "translational initiation", and "RNA metabolic process" by ontology analysis [19]. These annotated biological processes correlate with the major events of the MET, such as transcriptional initiation and translation, and the degradation of maternally stored transcripts and proteins. Results from this RNAseq comparison of IVF and scNT embryos suggest major dysregulation for genes critical for embryonic genome activation, including "mRNA processing", "ncRNA metabolic process", "RNA localization and regulation of mRNA metabolic process." Of note, differentially expressed genes in scNT 8-cell embryos were associated with "ncRNA metabolic

process", "tRNA processing", and "ncRNA transcription," which would suggest potential aberrant control of the production of non-coding RNA molecules, including sncRNAs. However, RNAseq of the population of sncRNAs in scNT embryos did not reveal large scale, consistent differences in relative abundance of miRNA, tRNA fragments or piRNAs at this developmental stage. If such aberrations in the population of sncRNAs did occur, they may have been random in nature. Also, aberrant gene expression in scNT 8-cell embryos was associated with dysregulation of mRNA processing, mRNA metabolism, and RNA localization, an observation that aligns well with the very large shift in transcript abundance observed subsequently in morula stage scNT embryos.

Several transcripts that were noted as under-expressed in 8-cell scNT embryos compared to stage-matched IVF controls were noted as over-expressed in scNT embryos at the morula developmental stage, pointing to dynamic dysregulation of their expression. These transcripts included *EIF3D*, which functions to initiate translation, and *EI24*, which is involved in P53-mediated apoptosis. Another example includes *XPO1* (exportin), which was down-regulated in 8-cell scNT embryos, but over-expressed in clones at the morula stage. Conversely, the inverse expression pattern was noted for the miRNA biogenesis genes *DICER* and *DROSHA*, which were higher in 8-cell scNT embryos, but lower in scNT morulae compared to their stage-matched IVF controls. This pattern of expression may explain the apparent differential expression of some miRNAs in scNT morula embryos. Moreover, although expression of *NOBOX* was higher in 8-cell scNT embryos, this transcript was significantly lower in morula scNT embryos, perhaps as a consequence of elevated (though not significantly so) expression of miR-196a.

Some genes down regulated in morula stage scNT embryos are involved in epigenetic control of gene expression, including histone modifiers, chromatin-binding proteins, and control of DNA methylation (e.g., *KDM6A*, *SUV39H1*, *SUV39H2*, *KDM1B*, *HDAC7*, *MBD1*, *MBD5*, *METTL3*, and *TRIM28*). *TRIM28* is required for genomic imprinting through the establishment

and maintenance of DNA methylation marks, even after genome reprogramming [78], and loss of

TRIM28 in bovine scNT embryos causes loss of methylation in imprinted genes [79]. Errors in

epigenetic coding of imprinted genes are a known problem in scNT embryos and could cause

lower developmental rates as well as often observed placental abnormalities [77]. *TDRD1* and

TDRD5 encode transcripts that repress transposable elements and form complexes with piRNA

and PIWI proteins in spermatogenesis [80]; interestingly, expression of these genes was down

regulated in scNT morula staged embryos. Importantly, higher retrotransposon expression has

been observed in scNT embryos, and loss of these transcripts may contribute to that issue [81].

As the largest differences in expression of piRNAs were noted when comparing IVF and scNT

embryos at the morula stage, the changes in expression of *TDRD1* and *TDRD5* transcripts are of

key functional interest and warrant further investigation. Lastly, Graf et al. observed that genes

with functions connected to basic metabolic processes were activated shortly after the EGA in 16-

cell cattle embryos [19]. Of note, our analyses revealed that differential expression of transcripts

in scNT embryos at the morula stage included many genes connected to cellular processes such as

"small GTPase mediated signal transduction", "actin cytoskeleton organization", and

"mitochondrial organization." Aberrant regulation of key metabolic and cellular function

pathways in scNT embryos after the MET could contribute to the high loss of scNT embryos

during early development and further impact successful implantation and placentation of these

embryos following transfer to recipient animals.

While a remarkable number of transcripts were differentially expressed in morula stage

scNT embryos, many fewer were noted by the blastocyst stage of development. This difference

could be attributed to an element of natural selection, such that embryos harboring large scale

aberrations in gene expression were not sufficiently competent to develop to the blastocyst stage

and instead arrested their development. Functions of those genes that were significantly different

in scNT blastocysts appeared connected to general functions, such as "cyclic nucleotide

metabolic process", "synapsis" and "chloride transport." Overall, the relatively low enrichment

values for differentially expressed genes in blastocyst stage scNT embryos suggest that those that

progress to this developmental stage harbor fewer overall differences compared to their stage-

matched IVF counterparts than did earlier stages evaluated. However, these persistent

dysregulated pathways may yet contribute to embryo loss post embryo transfer. Interestingly,

HAND1 was down regulated in scNT blastocyst stage embryos; this transcriptional factor is

required for placental development through promotion of differentiation in trophoblast giant cells

[82]. *HAND1* has been shown previously to be reduced in scNT bovine embryos at day 17 [83],

and may be connected to some of the placental aberrations commonly found in scNT pregnancies.

Our main objective in collecting sncRNA and mRNA sequencing data from the same

embryo samples was to correlate differences in expression of specific miRNAs in scNT embryos

with abundance of their predicted mRNA transcripts over the course of early development. Our

approach here was similar to that as we employed previously for sncRNA sequencing data and

published transcriptome data for mRNA expression in 8-cell and blastocyst embryos [84].

However, as discussed above, relatively few significantly different miRNAs were identified in

scNT embryos, with only three differentially expressed miRNAs noted at the morula stage.

Interaction networks were constructed for these miRNAs and their TargetScan predicted targets.

Within the morula stage embryo network, we noted that miR-34a had the most mRNA targets.

MiR-34a has been demonstrated to control cell fate and differentiation, as knock down of its

expression in pluripotent stem cells expanded cell fate potential [85]. Even though miR-34a was

under-expressed in scNT morulae, a broad pattern of over-expressed mRNA targets was not

dominant at this developmental stage. However, the expression of a subset of these predicted

targets was indeed strongly negatively correlated with miR-34a over the entire window of embryo

development studied (2-cell to blastocyst stage). Of these gene targets, many had regulatory

functions in cancer (e.g., *AMER1*, *RALGDS*, and *SIPA1*) suggesting possible function in

controlling pluripotency and cell fate [86-88]. A smaller network of predicted targets was identified for miR-345, and a similar trend was noted with the expression of only a fraction of the targets inversely correlated with this miRNA. To date, no evidence specifically points to a role for miR-345 in early development, although this miRNA is known to suppress proliferation, migration and invasion in cancer cells [89]. We also considered the pattern of expression for miR-345 over the full window of early development and correlated expression with predicted gene targets. Of those genes with the expected inverse expression pattern from 2-cell to blastocyst stage, *PDK1* is particularly interesting as this gene codes for a master kinase that is required for normal development, [90] and promotes somatic reprogramming efficiency through functions in self-renewal and differentiation in embryonic stem cells [91]. A similar pattern was evident for *NOSTRIN*, which modulates the differentiation of trophoblast giant cells as they begin to differentiate from trophoblast stem cells [92]. While miR-2340 was identified as significantly differentially expressed in morula scNT embryos, predicted targets of this miRNA overlapped entirely with miR-34a and miR-345; also, this miRNA does not have a known homologue in humans or mice. Moreover, although its expression was statistically significant, we noted that response was largely driven by a single sample.

Overall, the miRNA-mRNA networks did not suggest broad inverse patterns of expression as would be predicted given the role of miRNAs to degrade mRNA transcripts. This pattern was evident when considering interaction networks for differentially expressed transcripts in scNT embryos compared to their IVF counterparts at specific developmental stages as described above, or when comparing differentially expressed transcripts at progressive developmental stages (e.g., morula vs. 8-cell IVF embryos). However, interpretation of RNAi mediated regulation is complicated by a lack of one-to-one regulation of miRNAs and mRNAs. As seen in the networks, multiple miRNAs can target a single mRNA transcript, and individual miRNAs may have hundreds of mRNA targets. The complexity of how miRNAs interact with

targets was exemplified by the network generated for transcripts at the morula-to-blastocyst transition, as the highly intricate map shows overlap in many targeted transcripts between three or four miRNAs. The lack of a dominant pattern of negative correlation between miRNAs and predicted target mRNAs may be explained by the influence of other regulatory factors, such as RNA-binding proteins, which can allow some mRNA targets to escape repression through mediation of binding between miRNAs and the 3' untranslated region (UTR). As an example, Dead end homologue 1 (DND1) binds to mRNAs containing an AU-rich element in the 3' untranslated region, blocking miRNA binding to this site. DND1 is abundant in pig oocytes and early developing embryos, and is down-regulated at the MET [93]. DND1 may function in early development to protect developmentally important genes from degradation before the activation of the embryonic genome, as DND1 is functionally capable of binding pluripotency genes vital during early development [93]. It likely that miRNA targeting of transcripts for degradation is a combination of degradation signals from miRNA and interactions with RNA-binding proteins, as opposed to a simple one-to-one matching process [94]. Also, it is possible that the TargetScan prediction algorithm is not accurately predicting mRNA targets for these mRNAs, or that multiple miRNA species are necessary to degrade these transcripts. Lastly, the lack of broad-scale negative correlations may be attributed to the timing for comparisons. Herein, the network for miRNA-mRNA interactions was visualized using time-matched data from the same developmental stage. However, this approach would not fully capture delayed degradation of mRNA targets, which has been detected up to 32 hours after miRNA induction [95].

Of particular interest was the vast scale of transcriptional changes associated with the later developmental transitions from 8-cell to morula and morula to blastocyst stages, with thousands of DEGs identified, even while controlling for type I errors. We explored these lists of DEGs in context of the report by Graf et al. [19], who applied RNAseq methods to determine the timing of EGA in cattle IVF embryos. Comparing the lists of genes that were induced during

embryo development, our numbers are not out of line with those reported by Graf et al., although their stage comparisons were slightly different. For example, 1,937 genes were induced over the transition from 2-cell to 8-cell stage in this study compared to 1,296 genes for the transition from 4-cell to 8-cell stage in Graf et al. Similarly, 2,980 genes were induced when transitioning from 8-cell to morula stage in IVF embryos in this study as compared to 2,943 genes for the transition from 8-cell to 16 cell in Graf et al. Interestingly, we did identify many more up-regulated transcripts when comparing oocytes to 2-cell embryos in this study than did Graf and coworkers [19], though our study design did not facilitate distinction of maternal transcripts from those newly produced by the embryo. Thus, our data from the 2-cell stage also captures maternal genes that had undergone polyadenylation and been protected from active degradation pathways in very early development, likely inflating the number of over-expressed genes.

There were several limitations in this study that should be considered. Due to the small amount of genetic material in the developing embryos, pools of embryos were necessary for RNA sequencing. The use of pooled samples can mask the biological variability of individual embryos, and, conversely, an extreme outlier can shift the RNA population of the pool. In our studies, marked variability was noted for sncRNA profiles, though less so for the mRNA profiles of the same samples. Embryos were staged based on their morphology, though others have shown that blastocyst stage embryos with similar morphology can harbor distinct transcriptomes [96]. Populations of sncRNAs could also differ for embryos with similar morphologies. Another possibility to consider is that the sex of the IVF embryos contributed to some variation in gene expression after the EGA. Others have reported sex-based differences in expression of both miRNA and mRNA, with up to one third of expressed transcripts demonstrating sex-specific transcriptional regulation in bovine preimplantation embryos [97]. Also, bovine blastocysts showed sexual dimorphism in secreted miRNAs [98]. While IVF embryos were created using a single bull's semen, the semen was not sex sorted. Thus, the sample pools for IVF embryos

contained both males and females, whereas the pools of scNT were exclusively male. Moreover, IVF embryo samples would also be impacted by sperm-specific transcripts and sncRNAs. Last, culture conditions could have increased variability within embryos as well, as cumulus cell transcriptomes and secreted factors can vary from oocyte to oocyte [99, 100].

In conclusion, major differences were observed in the transcriptomes of scNT and IVF embryos at every stage through the MET. Importantly, we identified a cohort of genes differentially expressed in scNT embryos during the EGA that function in epigenetic control of gene expression, processing of sncRNA and processing of mRNAs. To our knowledge, our group is the first to assess the dynamics of sncRNA and mRNA populations derived from the same scNT embryos through the MET. When comparing the dynamics of specific miRNAs and populations of predicted targets, negative correlations were not widespread, suggesting other mechanisms controlling gene expression were likely responsible for the large-scale aberrant transcript expression in scNT embryos. Such mechanisms may involve abnormal patterns of DNA methylation or chromatin modification, as others have noted in scNT embryos. Moreover, expression of sncRNA loci in scNT embryos may also be impacted by poor reprogramming of these epigenetic marks, although results from this work suggest that such changes in miRNA expression may be more random than for transcripts, as we identified many fewer significantly differentially expressed miRNAs than mRNAs within the same RNA population. Additionally, further study is needed to better explore the complex network of miRNA-mRNA interactions, given that an individual miRNA may have many hundreds of gene targets and a transcript may be targeted by multiple miRNAs. Also, future studies should consider the added complexity of RNA binding proteins that restrict access to transcripts and block miRNA-mediated repression or degradation.

References

1.   Wilmut I, Beaujean N, de Sousa PA, Dinnyes A, King TJ, Paterson LA, Wells DN, Young LE. Somatic cell nuclear transfer. Nature. 2002;419(6907):583-586.

2.   Solter D. Mammalian cloning: advances and limitations. Nat Rev Genet. 2000;1(3):199-207.

3.   Cezar GG, Bartolomei MS, Forsberg EJ, First NL, Bishop MD, Eilertsen KJ. Genome-wide epigenetic alterations in cloned bovine fetuses. Biol Reprod. 2003;68(3):1009-1014.

4.   Kang YK, Park JS, Koo DB, Choi YH, Kim SU, Lee KK, Han YM. Limited demethylation leaves mosaic-type methylation states in cloned bovine pre-implantation embryos. EMBO J. 2002;21(5):1092-1100.

5.   Mann MR, Chung YG, Nolen LD, Verona RI, Latham KE, Bartolomei MS. Disruption of imprinted gene methylation and expression in cloned preimplantation stage mouse embryos. Biol Reprod. 2003;69(3):902-914.

6.   Han YM, Kang YK, Koo DB, Lee KK. Nuclear reprogramming of cloned embryos produced in vitro. Theriogenology. 2003;59(1):33-44.

7.   Kang YK, Koo DB, Park JS, Choi YH, Chung AS, Lee KK, Han YM. Aberrant methylation of donor genome in cloned bovine embryos. Nat Genet. 2001;28(2):173-177.

8.   Matoba S, Liu Y, Lu F, Iwabuchi KA, Shen L, Inoue A, Zhang Y. Embryonic development following somatic cell nuclear transfer impeded by persisting histone methylation. Cell. 2014;159(4):884-895.

9.   Matoba S, Zhang Y. Somatic cell nuclear transfer reprogramming: mechanisms and applications. Cell Stem Cell. 2018;23(4):471-485.

10.   Bourc'his D, Le Bourhis D, Patin D, Niveleau A, Comizzoli P, Renard JP, Viegas-Pequignot E: Delayed and incomplete reprogramming of chromosome methylation patterns in bovine cloned embryos. Curr Biol. 2001;11(19):1542-1546.

11.   Pesce M, Scholer HR. Oct-4: gatekeeper in the beginnings of mammalian development. Stem Cells. 2001;19(4):271-278.

12.   Yamazaki Y, Fujita TC, Low EW, Alarcon VB, Yanagimachi R, Marikawa Y. Gradual DNA demethylation of the Oct4 promoter in cloned mouse embryos. Mol Reprod Dev. 2006;73(2):180-188.

13.   Santos F, Zakhartchenko V, Stojkovic M, Peters A, Jenuwein T, Wolf E, Reik W, Dean W. Epigenetic marking correlates with developmental potential in cloned bovine preimplantation embryos. Curr Biol. 2003;13(13):1116-1121.

14.   Shi WaH T. Aberrant methylation patterns at the two-cell stage as an indicator of early developmental failure. Mol Reprod Dev. 2002;63(3):329-334.

15. Schultz RM. The molecular foundations of the maternal to zygotic transition in the preimplantation embryo. Hum Reprod Update. 2002;8(4):323-331.

16. Thomsen S, Anders S, Janga SC, Huber W, Alonso CR. Genome-wide analysis of mRNA decay patterns during early Drosophila development. Genome Biol. 2010;11(9):R93.

17. Gotze M, Wahle E. Smaug destroys a huge treasure. Genome Biol. 2014;15(1):101.

18. Benoit B, He CH, Zhang F, Votruba SM, Tadros W, Westwood JT, Smibert CA, Lipshitz HD, Theurkauf WE. An essential role for the RNA-binding protein Smaug during the Drosophila maternal-to-zygotic transition. Development. 2009;136(6):923-932.

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

20. Betsha S, Hoelker M, Salilew-Wondim D, Held E, Rings F, Grosse-Brinkhause C, Cinar MU, Havlicek V, Besenfelder U, Tholen E et al. Transcriptome profile of bovine elongated conceptus obtained from SCNT and IVP pregnancies. Mol Reprod Dev. 2013;80(4):315-333.

21. Min B, Cho S, Park JS, Lee YG, Kim N, Kang YK. Transcriptomic features of bovine blastocysts derived by somatic cell nuclear transfer. G3 (Bethesda). 2015;5(12):2527-2538.

22. Hosseini SM, Dufort I, Nieminen J, Moulavi F, Ghanaei HR, Hajian M, Jafarpour F, Forouzanfar M, Gourbai H, Shahverdi AH et al. Epigenetic modification with trichostatin A does not correct specific errors of somatic cell nuclear transfer at the transcriptomic level; highlighting the non-random nature of oocyte-mediated reprogramming errors. BMC Genomics. 2016;17:16.

23. Ohnishi Y, Totoki Y, Toyoda A, Watanabe T, Yamamoto Y, Tokunaga K, Sakaki Y, Sasaki H, Hohjoh H. Small RNA class transition from siRNA/piRNA to miRNA during pre-implantation mouse development. Nucleic Acids Res. 2010;38(15):5141-5151.

24. Carthew RW, Sontheimer EJ. Origins and mechanisms of miRNAs and siRNAs. Cell. 2009;136(4):642-655.

25. Fabian MR, Sonenberg N, Filipowicz W. Regulation of mRNA translation and stability by microRNAs. Annu Rev Biochem. 2010;79:351-379.

26. Clarke S, Sunkar, R. Biogenesis of Mammalian miRNA. In: MicroRNAs as Tools in Biopharmaceutical Production. Springer Science. 2012. p. 15-27

27. Valencia-Sanchez MA, Liu J, Hannon GJ, Parker R. Control of translation and mRNA degradation by miRNAs and siRNAs. Genes Dev. 2006;20(5):515-524.

28. Giraldez AJ, Mishima Y, Rihel J, Grocock RJ, Van Dongen S, Inoue K, Enright AJ, Schier AF. Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. Science 2006;312(5770):75-79.

29. Wu E, Thivierge C, Flamand M, Mathonnet G, Vashisht AA, Wohlschlegel J, Fabian MR, Sonenberg N, Duchaine TF. Pervasive and cooperative deadenylation of 3'UTRs by embryonic microRNA families. Mol Cell. 2010;40(4):558-570.

30. Lund E, Liu M, Hartley RS, Sheets MD, Dahlberg JE. Deadenylation of maternal mRNAs mediated by miR-427 in Xenopus laevis embryos. RNA. 2009;15(12):2351-2363.

31. Bushati N, Stark A, Brennecke J, Cohen SM. Temporal reciprocity of miRNAs and their targets during the maternal-to-zygotic transition in Drosophila. Curr Biol. 2008;18(7):501-506.

32. Roovers EF, Rosenkranz D, Mahdipour M, Han CT, He N, Chuva de Sousa Lopes SM, van der Westerlaken LA, Zischler H, Butter F, Roelen BA et al. Piwi proteins and piRNAs in mammalian oocytes and early embryos. Cell Rep. 2015;10(12):2069-2082.

33. Much C, Auchynnikava T, Pavlinic D, Buness A, Rappsilber J, Benes V, Allshire R, O'Carroll D. Endogenous mouse Dicer is an exclusively cytoplasmic protein. PLoS Genet. 2016;12(6):e1006095.

34. Hossain MM, Tesfaye D, Salilew-Wondim D, Held E, Proll MJ, Rings F, Kirfel G, Looft C, Tholen E, Uddin J et al. Massive deregulation of miRNAs from nuclear reprogramming errors during trophoblast differentiation for placentogenesis in cloned pregnancy. BMC Genomics. 2014;15:43.

35. Castro FO, Sharbati S, Rodriguez-Alvarez LL, Cox JF, Hultschig C, Einspanier R. MicroRNA expression profiling of elongated cloned and in vitro-fertilized bovine embryos. Theriogenology. 2010;73(1):71-85.

36. Kropp J, Khatib H. Characterization of microRNA in bovine in vitro culture media associated with embryo quality and development. J Dairy Sci. 2015;98(9):6552-6563.

37. Lin X, Beckers E, Mc Cafferty S, Gansemans Y, Joanna Szymanska K, Chaitanya Pavani K, Catani JP, Van Nieuwerburgh F, Deforce D, De Sutter P et al. Bovine embryo-secreted microRNA-30c is a potential non-invasive biomarker for hampered preimplantation developmental competence. Front Genet. 2019;10:315.

38. Liang S, Nie ZW, Guo J, Niu YJ, Shin KT, Ock SA, Cui XS. Overexpression of MicroRNA-29b decreases expression of DNA methyltransferases and improves quality of the blastocysts derived from somatic cell nuclear transfer in cattle. Microsc Microanal. 2018;24(1):29-37.

39. Zhang J, Qu P, Zhou C, Liu X, Ma X, Wang M, Wang Y, Su J, Liu J, Zhang Y. MicroRNA-125b is a key epigenetic regulatory factor that promotes nuclear transfer reprogramming. J Biol Chem. 2017;292(38):15916-15926.

40. Wang P, Li X, Cao L, Huang S, Li H, Zhang Y, Yang T, Jiang J, Shi D. MicroRNA-148a overexpression improves the early development of porcine somatic cell nuclear transfer embryos. PLoS One. 2017;12(6):e0180535.

41. Rashmi, Sah S, Shyam S, Singh MK, Palta P. Treatment of buffalo (Bubalus bubalis) SCNT embryos with microRNA-21 mimic improves their quality and alters gene expression but does not affect their developmental competence. Theriogenology. 2019;126:8-16.

42. Chang HY, Xie RX, Zhang L, Fu LZ, Zhang CT, Chen HH, Wang ZQ, Zhang Y, Quan FS. Overexpression of miR-101-2 in donor cells improves the early development of Holstein cow somatic cell nuclear transfer embryos. J Dairy Sci. 2019;102(5):4662-4673.

43. Wang M, Gao Y, Qu P, Qing S, Qiao F, Zhang Y, Mager J, Wang Y. Sperm-borne miR-449b influences cleavage, epigenetic reprogramming and apoptosis of SCNT embryos in bovine. Sci Rep. 2017;7:13403.

44. Gao G, Wang S, Zhang J, Su G, Zheng Z, Bai C, Yang L, Wei Z, Wang X, Liu X et al. Transcriptome-wide analysis of the SCNT bovine abnormal placenta during mid- to late gestation. Sci Rep. 2019;9(1):20035.

45. Zhou C, Zhang J, Zhang M, Wang D, Ma Y, Wang Y, Wang Y, Huang Y, Zhang Y. Transcriptional memory inherited from donor cells is a developmental defect of bovine cloned embryos. FASEB J. 2020;34(1):1637-1651.

46. Funston RN, Nauta WJ, Seidel GE. Culture of bovine embryos in buffalo rat liver cell-conditioned media or with leukemia inhibitory factor. J Anim Sci. 1997;75(5):1332-1336.

47. Aston KI, Li GP, Hicks BA, Sessions BR, Pate BJ, Hammon D, Bunch TD, White KL. Effect of the time interval between fusion and activation on nuclear state and development in vitro and in vivo of bovine somatic cell nuclear transfer embryos. Reproduction. 2006;131(1):45-51.

48. Aston KI, Li GP, Hicks BA, Sessions BR, Pate BJ, Hammon DS, Bunch TD, White KL. The developmental competence of bovine nuclear transfer embryos derived from cow versus heifer cytoplasts. Anim Reprod Sci. 2006;95(3-4):234-243.

49. Li GP, Bunch TD, White KL, Aston KI, Meerdo LN, Pate BJ, Sessions BR. Development, chromosomal composition, and cell allocation of bovine cloned blastocyst derived from chemically assisted enucleation and cultured in conditioned media. Mol Reprod Dev. 2004;68(2):189-197.

50. Gandhi AP, Lane M, Gardner DK, Krisher RL. A single medium supports development of bovine embryos throughout maturation, fertilization and culture. Hum Reprod. 2000;15(2):395-401.

51. Polejaeva IA. White KL, Ellis LC, Reed WA. Isolation and long-term culture of mink and bovine embryonic stem-like cells. Theriogenology. 1995;43:300.

52.   Polejaeva IA, Reed WA, Bunch TD, Ellis LC, White KL. Prolactin-induced termination of obligate diapause of mink (Mustela vison) blastocysts in vitro and subsequent establishment of embryonic stem-like cells. J Reprod Fertil. 1997;109(2):229-236.

53.   Polejaeva I, Mitalipov S. Stem cell potency and the ability to contribute to chimeric organisms. Reproduction. 2013;145(3):R81-88.

54.   Malaver-Ortega LF, Sumer H, Liu J, Verma PJ. The state of the art for pluripotent stem cells derivation in domestic ungulates. Theriogenology. 2012;78(8):1749-62

55.   Saito S, Strelchenko N, Niemann H. Bovine embryonic stem cell-like cell lines cultured over several passages. Roux Arch Dev Biol. 1992;201(3):134-141.

56.   Cibelli JB, Stice SL, Golueke PJ, Kane JJ, Jerry J, Blackwell C, Ponce de Leon FA, Robl JM. Transgenic bovine chimeric offspring produced from somatic cell-derived stem-like cells. Nat Biotechnol. 1998;16(7):642-646.

57.   FastQC: a quality control tool for high throughput sequence data In: Babraham Bioinformatics. 2019. http://www.bioinformatics.babraham.ac.uk/projects/fastqc. Accessed July 2019.

58.   Ewels P, Magnusson M, Lundin S, Kaller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics. 2016;32(19):3047-3048.

59.   Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29(1):15-21.

60.   Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.

61.   Hulsen T, de Vlieg J, Alkema W. BioVenn - a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. BMC Genomics. 2008;9:488.

62.   Metsalu T, Vilo J. ClustVis: a web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. Nucleic Acids Res. 2015;43(W1):W566-570.

63.   Tripathi S, Pohl MO, Zhou Y, Rodriguez-Frandsen A, Wang G, Stein DA, Moulton HM, DeJesus P, Che J, Mulder LC et al. Meta- and orthogonal integration of influenza "OMICs" data defines a role for UBR4 in virus budding. Cell Host Microbe. 2015;18(6):723-735.

64.   Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13(11):2498-2504.

65. Agarwal V, Bell GW, Nam JW, Bartel DP. Predicting effective microRNA target sites in mammalian mRNAs. Elife. 2015;4.

66. Beenken A, Mohammadi M. The FGF family: biology, pathophysiology and therapy. Nat Rev Drug Discov 2009;8(3):235-253.

67. Chesterman ES, Gainey GD, Varn AC, Peterson RE, Kern MJ. Investigation of Prx1 protein expression provides evidence for conservation of cardiac-specific posttranscriptional regulation in vertebrates. Dev Dyn. 2001;222(3):459-470.

68. Chung JJ, Navarro B, Krapivinsky G, Krapivinsky L, Clapham DE. A novel gene required for male fertility and functional CATSPER channel formation in spermatozoa. Nat Commun. 2011;2:153.

69. Artus J, Babinet C, Cohen-Tannoudji M. The cell cycle of early mammalian embryos: lessons from genetic mouse models. Cell Cycle. 2006;5(5):499-502.

70. Yu JS, Cui W. Proliferation, survival and metabolism: the role of PI3K/AKT/mTOR signalling in pluripotency and cell fate determination. Development. 2016;143(17):3050-3060.

71. Liu Y, He K, Hu Y, Guo X, Wang D, Shi W, Li J, Song J. YAP modulates TGF-beta1-induced simultaneous apoptosis and EMT through upregulation of the EGF receptor. Sci Rep. 2017;7:45523.

72. Muqbil I, Bao B, Abou-Samra AB, Mohammad RM, Azmi AS. Nuclear export mediated regulation of microRNAs: potential target for drug intervention. Curr Drug Targets. 2013;14(10):1094-1100.

73. Tripurani SK, Lee KB, Wee G, Smith GW, Yao J. MicroRNA-196a regulates bovine newborn ovary homeobox gene (NOBOX) expression during early embryogenesis. BMC Dev Biol. 2011;11:25.

74. De Iaco A, Coudray A, Duc J, Trono D. DPPA2 and DPPA4 are necessary to establish a 2C-like state in mouse embryonic stem cells. EMBO Rep. 2019;20(5).

75. Pan G, Thomson JA. Nanog and transcriptional networks in embryonic stem cell pluripotency. Cell Res. 2007;17(1):42-49.

76. Boyer LA, Lee TI, Cole MF, Johnstone SE, Levine SS, Zucker JP, Guenther MG, Kumar RM, Murray HL, Jenner RG et al. Core transcriptional regulatory circuitry in human embryonic stem cells. Cell. 2005;122(6):947-956.

77. Simmet K, Zakhartchenko V, Philippou-Massier J, Blum H, Klymiuk N, Wolf E. OCT4/POU5F1 is required for NANOG expression in bovine blastocysts. Proc Natl Acad Sci USA. 2018;115(11):2770-2775.

78. Alexander KA, Wang X, Shibata M, Clark AG, Garcia-Garcia MJ: TRIM28 controls genomic imprinting through distinct mechanisms during and after early genome-wide reprogramming. Cell Rep 2015, 13(6):1194-1205.

79. Ma X, Zhang S, Zhang M, Zhu Y, Ma P, Yang S, Su L, Li Z, Lv W, Luan W. TRIM28 down-regulation on methylation imprints in bovine preimplantation embryos. Zygote. 2018;26(6):449-456.

80. Ding D, Liu J, Midic U, Wu Y, Dong K, Melnick A, Latham KE, Chen C. TDRD5 binds piRNA precursors and selectively enhances pachytene piRNA processing in mice. Nat Commun. 2018;9(1):127.

81. Bui LC, Evsikov AV, Khan DR, Archilla C, Peynot N, Henaut A, Le Bourhis D, Vignon X, Renard JP, Duranthon V. Retrotransposon expression as a defining event of genome reprogramming in fertilized and cloned bovine embryos. Reproduction. 2009;138(2):289-299.

82. Cross JC, Flannery ML, Blanar MA, Steingrimsson E, Jenkins NA, Copeland NG, Rutter WJ, Werb Z. Hxt encodes a basic helix-loop-helix transcription factor that regulates trophoblast cell development. Development. 1995;121(8):2513-2523.

83. Arnold DR, Bordignon V, Lefebvre R, Murphy BD, Smith LC. Somatic cell nuclear transfer alters peri-implantation trophoblast differentiation in bovine embryos. Reproduction. 2006;132(2):279-290.

84. Cuthbert JM, Russell SJ, White KL, Benninghoff AD. The maternal-to-zygotic transition in bovine in vitro-fertilized embryos is associated with marked changes in small non-coding RNAsdagger. Biol Repro. 2019;100(2):331-350.

85. Choi YJ, Lin CP, Risso D, Chen S, Kim TA, Tan MH, Li JB, Wu Y, Chen C, Xuan Z et al. Deficiency of microRNA miR-34a expands cell fate potential in pluripotent stem cells. Science. 2017;355(6325).

86. Moghadam AR, Patrad E, Tafsiri E, Peng W, Fangman B, Pluard TJ, Accurso A, Salacz M, Shah K, Ricke B et al. Ral signaling pathway in health and cancer. Cancer Med. 2017;6(12):2998-3013.

87. Sanz-Pamplona R, Lopez-Doriga A, Pare-Brunet L, Lazaro K, Bellido F, Alonso MH, Ausso S, Guino E, Beltran S, Castro-Giner F et al. Exome sequencing reveals AMER1 as a frequently mutated gene in colorectal cancer. Clin Cancer Res. 2015;21(20):4709-4718.

88. Ji K, Ye L, Toms AM, Hargest R, Martin TA, Ruge F, Ji J, Jiang WG. Expression of signal-induced proliferation-associated gene 1 (SIPA1), a RapGTPase-activating protein, is increased in colorectal cancer and has diverse effects on functions of colorectal cancer cells. Cancer Genomics Proteomics. 2012;9(5):321-327.

89. Chen QG, Zhou W, Han T, Du SQ, Li ZH, Zhang Z, Shan GY, Kong CZ. MiR-345 suppresses proliferation, migration and invasion by targeting Smad1 in human prostate cancer. J Cancer Res Clin Oncol, 2016;142(1):213-224.

90. Lawlor MA, Mora A, Ashby PR, Williams MR, Murray-Tait V, Malone L, Prescott AR, Lucocq JM, Alessi DR: Essential role of PDK1 in regulating cell size and development in mice. EMBO J. 2002;21(14):3728-3738.

91. Wang XQ, Lo CM, Chen L, Ngan ES, Xu A, Poon RY. CDK1-PDK1-PI3K/Akt signaling pathway regulates embryonic and induced pluripotency. Cell Death Differ 2017;24(1):38-48.

92. Chakraborty S, Ain R. NOSTRIN: A novel modulator of trophoblast giant cell differentiation. Stem Cell Res. 2018;31:135-146.

93. Rivera RM, Ross JW. Epigenetics in fertilization and preimplantation embryo development. Prog Biophys Mol Biol. 2013;113(3):423-432.

94. Alonso CR. A complex 'mRNA degradation code' controls gene expression during animal development. Trends Genet. 2012;28(2):78-88.

95. Eichhorn SW, Guo H, McGeary SE, Rodriguez-Mias RA, Shin C, Baek D, Hsu SH, Ghoshal K, Villen J, Bartel DP. mRNA destabilization is the dominant effect of mammalian microRNAs by the time substantial repression ensues. Mol Cell. 2014;56(1):104-115.

96. Driver AM, Penagaricano F, Huang W, Ahmad KR, Hackbart KS, Wiltbank MC, Khatib H. RNA-Seq analysis uncovers transcriptomic variations between morphologically similar in vivo- and in vitro-derived bovine blastocysts. BMC Genomics. 2012;13:118.

97. Bermejo-Alvarez P, Rizos D, Rath D, Lonergan P, Gutierrez-Adan A. Sex determines the expression level of one third of the actively expressed genes in bovine blastocysts. Proc Natl Acad Sci USA. 2010;107(8):3394-3399.

98. Gross N, Kropp J, Khatib H. Sexual dimorphism of miRNAs secreted by bovine in vitro-produced embryos. Front Genet. 2017;8:39.

99. Hamel M, Dufort I, Robert C, Gravel C, Leveille MC, Leader A, Sirard MA. Identification of differentially expressed markers in human follicular cells associated with competent oocytes. Hum Reprod. 2008;23(5):1118-1127.

100. Huang Z, Wells D. The human oocyte and cumulus cells relationship: new insights from the cumulus cell transcriptome. Mol Hum Reprod. 2010;16(10):715-725.
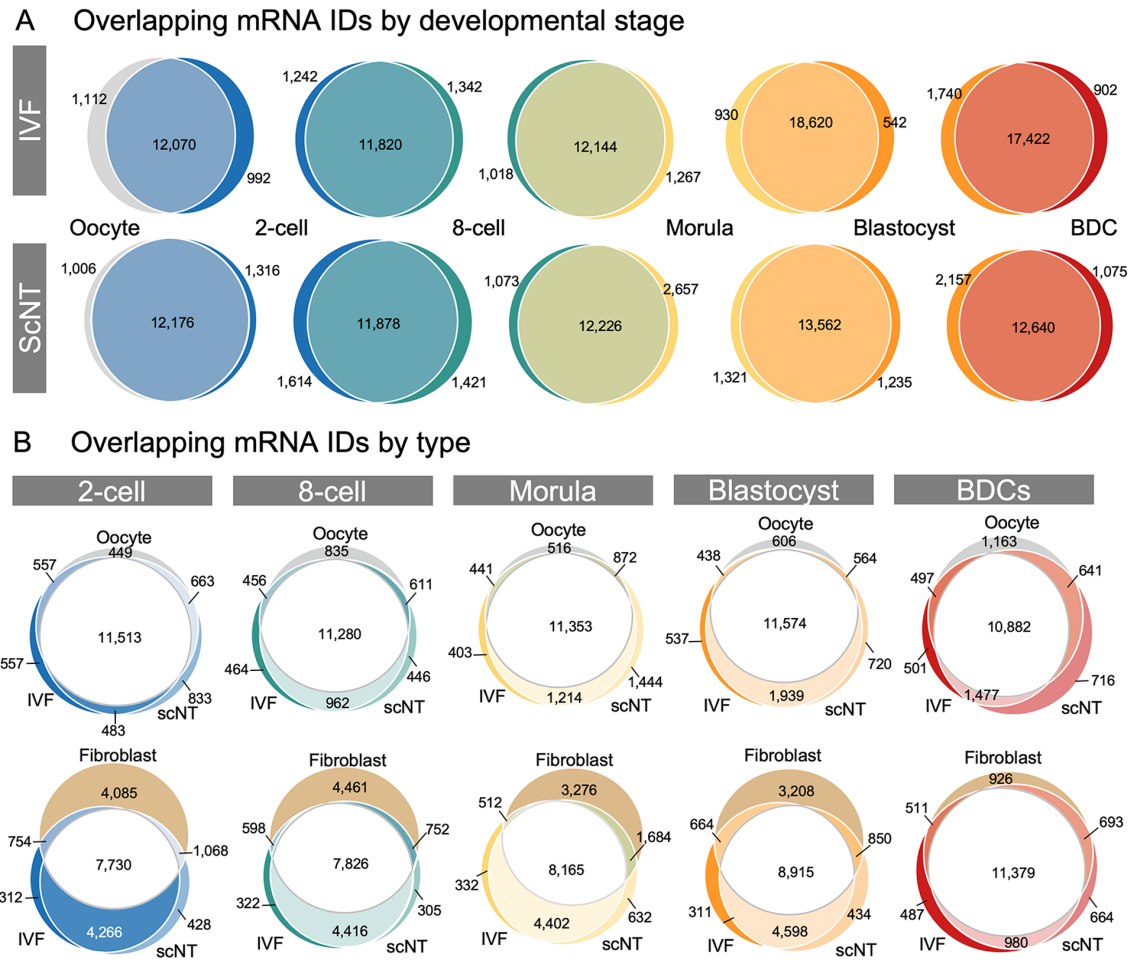
Figure 4.1. Overlap in detection of all mapped mRNAs in bovine oocytes and donor fibroblast cells as well as early developing embryos and blastocyst-derived cells generated via IVF or scNT. Proportional Venn diagrams depicting mapped to the *Bos Taurus* transcriptome, including unique and overlapping transcript IDs. (A) mRNAs identified by sequential comparison of developmental stages for IVF or scNT embryos. (B) mRNAs identified in IVF, scNT and either oocytes or donor fibroblast cells at each developmental stage.
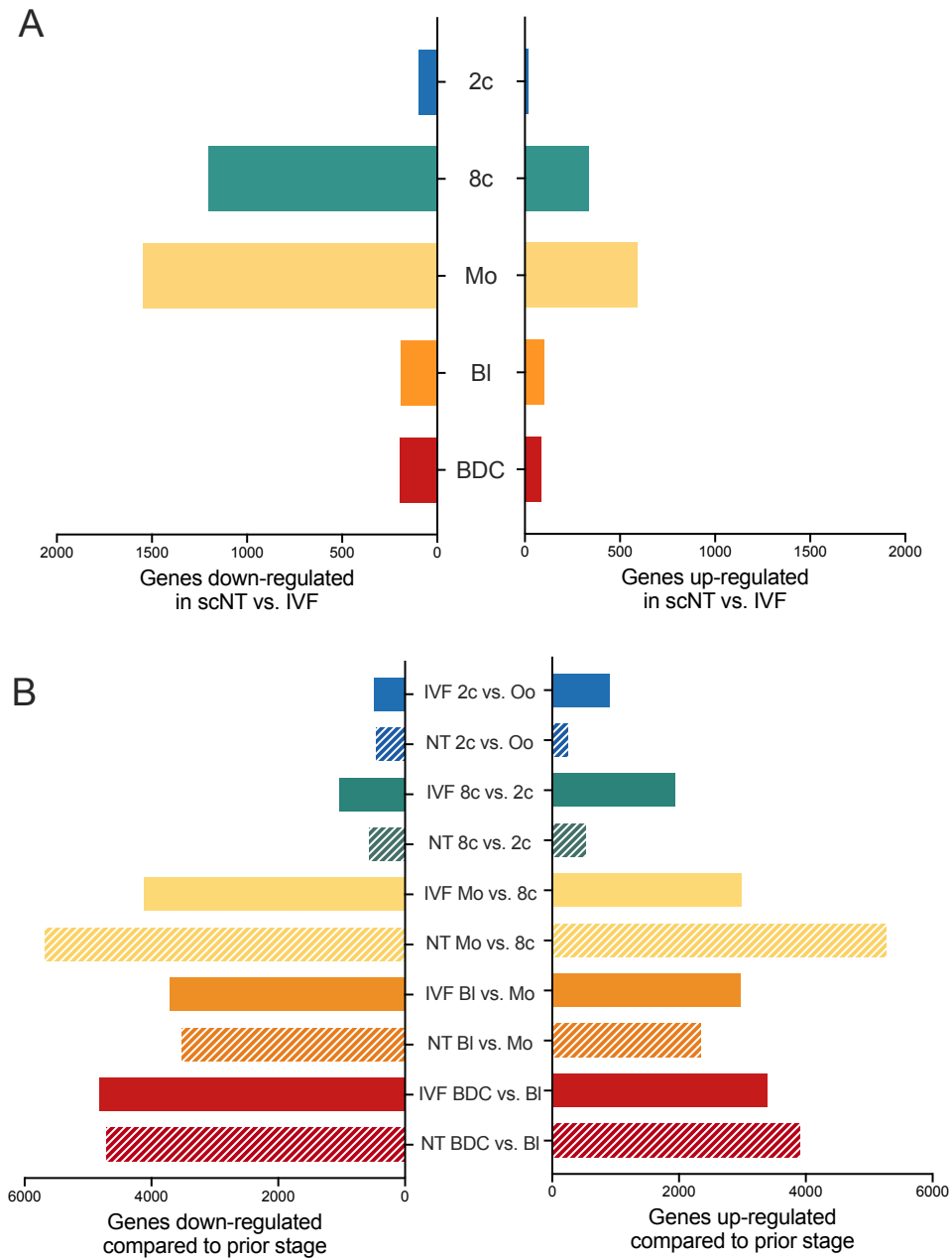
Figure 4.2. Distribution of the number of significantly different genes by embryo type or developmental stage. (A) Number of significantly different genes down-regulated (left) or up-regulated (right) in scNT embryos as compared to their stage-matched IVF counterparts. (B) Number of genes down-regulated (left) or up-regulated (right) in IVF or scNT embryos by sequential comparison of developmental stage. A significant difference in gene expression was inferred when the FDR adjusted *p*-value <0.05.
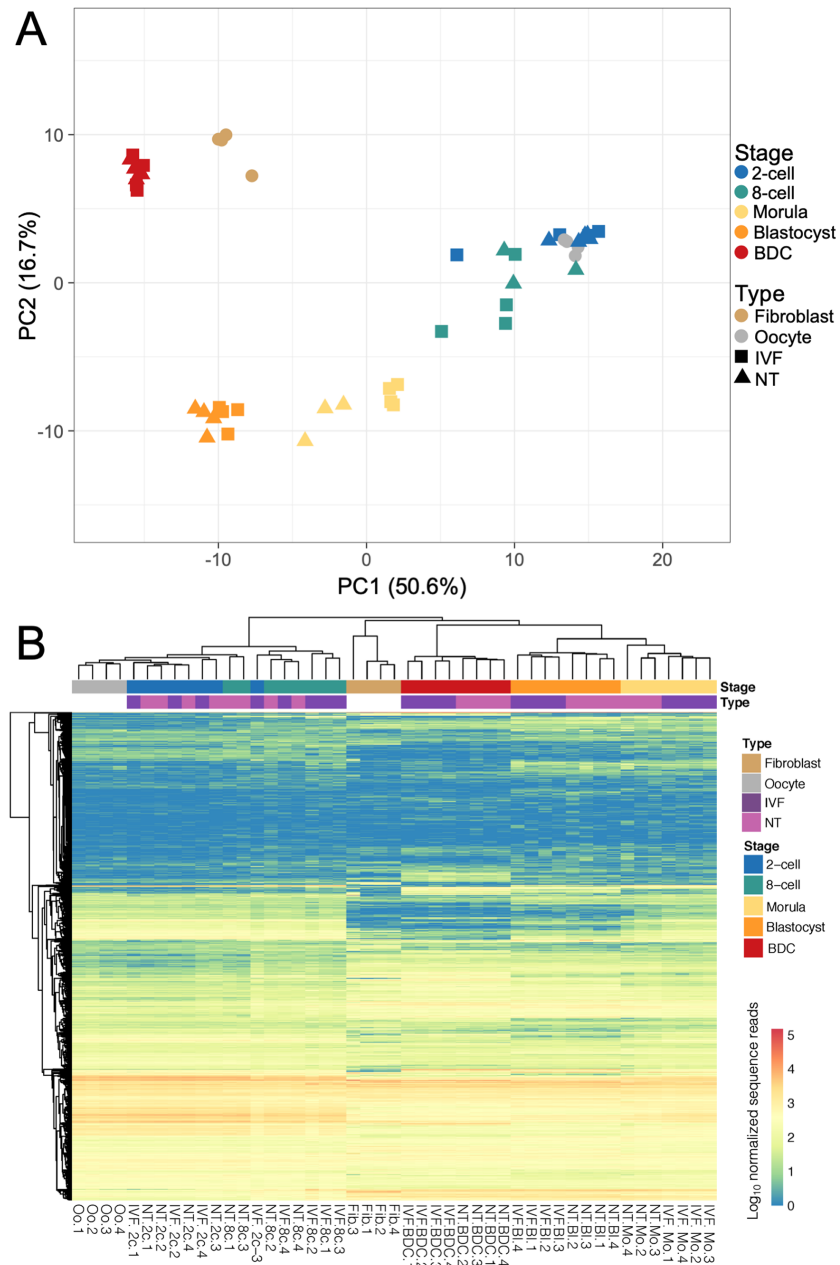
Figure 4.3. Patterns of mRNA expression in bovine oocytes and donor fibroblast cells as well as early developing embryos and blastocyst-derived cells generated via IVF or scNT. (A) Principal components analysis of bovine miRNAs using the standard singular value decomposition method with imputation. (B) Unsupervised, bi-directional hierarchical cluster analysis (Euclidean distance method) of variance stabilized normalized read counts for a subset of 2,400 randomly chosen mRNAs. Data were categorized by sample type (Fb, fibroblast; Oo, oocyte; IVF, *in vitro*-fertilized embryo; scNT, scNT embryo) and stage (2c, 2-cell; 8c, 8-cell; Mo, morula; Bl, blastocyst; BDC, blastocyst-derived cells).

Figure 4.4. Significant biological process ontology terms and network clustering associated with significantly different genes in bovine scNT versus IVF 2-cell stage embryos. (A) Bar chart depicts the top 20 enriched summary terms identified for ontology biological processes. (B) Cytoscape network depicting biological process ontology terms represented as nodes connected with terms with shared differentially expressed genes. The size of the node is proportional to the term's enrichment score. Terms with a similarity score >0.3 are similarly colored and linked by an edge (the thickness of the edge represents the similarity score) using force-directed network layout and bundled edges for clarity. Each cluster is labeled by the term with the lowest *p*-value within the cluster group. Complete results of the Metascape analysis are provided in Supplementary File 4.3.

Figure 4.5. Significant biological process ontology terms and network clustering associated with significantly different genes in bovine scNT versus IVF 8-cell stage embryos. (A) Bar chart depicts the top 20 enriched summary terms identified for ontology biological processes. (B) Cytoscape network depicting biological process ontology terms represented as nodes connected with terms with shared differentially expressed genes. The size of the node is proportional to the term's enrichment score. Terms with a similarity score >0.3 are similarly colored and linked by an edge (the thickness of the edge represents the similarity score) using force-directed network layout and bundled edges for clarity. Each cluster is labeled by the term with the lowest *p*-value within the cluster group. Complete results of the Metascape analysis are provided in Supplementary File 4.3.

A
-log *p* value

GO:0007264: small GTPase mediated signal transduction
GO:0120039: plasma membrane bounded cell projection morphogenesis
GO:0030036: actin cytoskeleton organization
GO:0009117: nucleotide metabolic process
GO:0007005: mitochondrion organization
GO:0044257: cellular protein catabolic process
GO:0140053: mitochondrial gene expression
GO:0016358: dendrite development
GO:0051098: regulation of binding
GO:0023061: signal release
GO:0071407: cellular response to organic cyclic compound
GO:0051345: positive regulation of hydrolase activity
GO:1905114: cell surface receptor signaling pathway involved in cell-cell signaling
GO:0048871: multicellular organismal homeostasis
GO:0016570: histone modification
GO:0043043: peptide biosynthetic process
GO:0007276: gamete generation
GO:0050806: positive regulation of synaptic transmission
GO:0010942: positive regulation of cell death
GO:0045859: regulation of protein kinase activity

B

signal release

cell surface receptor signaling pathway involved in cell-cell signaling

positive regulation of cell death

positive regulation of hydrolase activity

small GTPase mediated signal transduction

dendrite development

positive regulation of synaptic transmission

plasma embrane bounded cell projection morphogenesis

regulation of protein kinase activity

cellular protein catabolic process

nucleotide metabolic process

histone modification

actin cytoskeleton organization

peptide biosynthetic process

mitochondrion organization

mitochondrial gene expression

cellular response to organic cyclic compound

gamete generation

multicellular organismal homeostasis
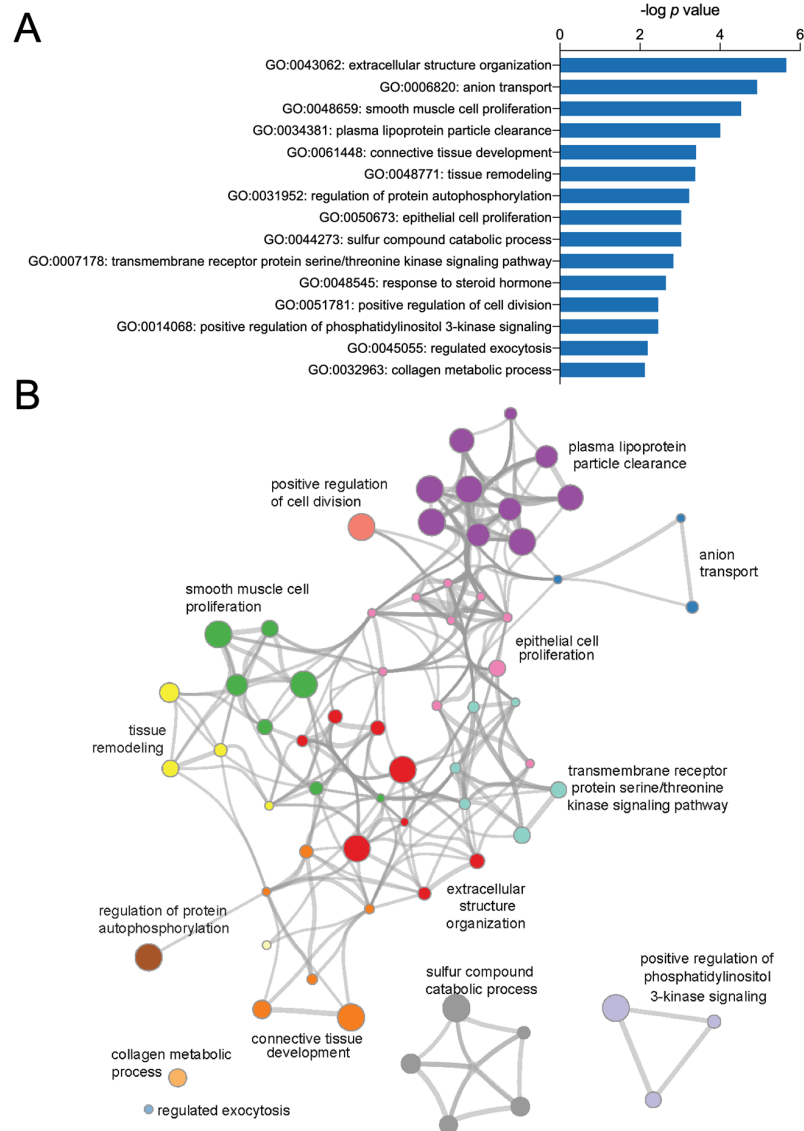
regulation of binding

Figure 4.6. (prior page) Significant biological process ontology terms and network clustering associated with significantly different genes in bovine scNT versus IVF morula stage embryos. (A) Bar chart depicts the top 20 enriched summary terms identified for ontology biological processes. (B) Cytoscape network depicting biological process ontology terms represented as nodes connected with terms with shared differentially expressed genes. The size of the node is proportional to the term's enrichment score. Terms with a similarity score >0.3 are similarly colored and linked by an edge (the thickness of the edge represents the similarity score) using force-directed network layout and bundled edges for clarity. Each cluster is labeled by the term with the lowest *p*-value within the cluster group. Complete results of the Metascape analysis are provided in Supplementary File 4.3.
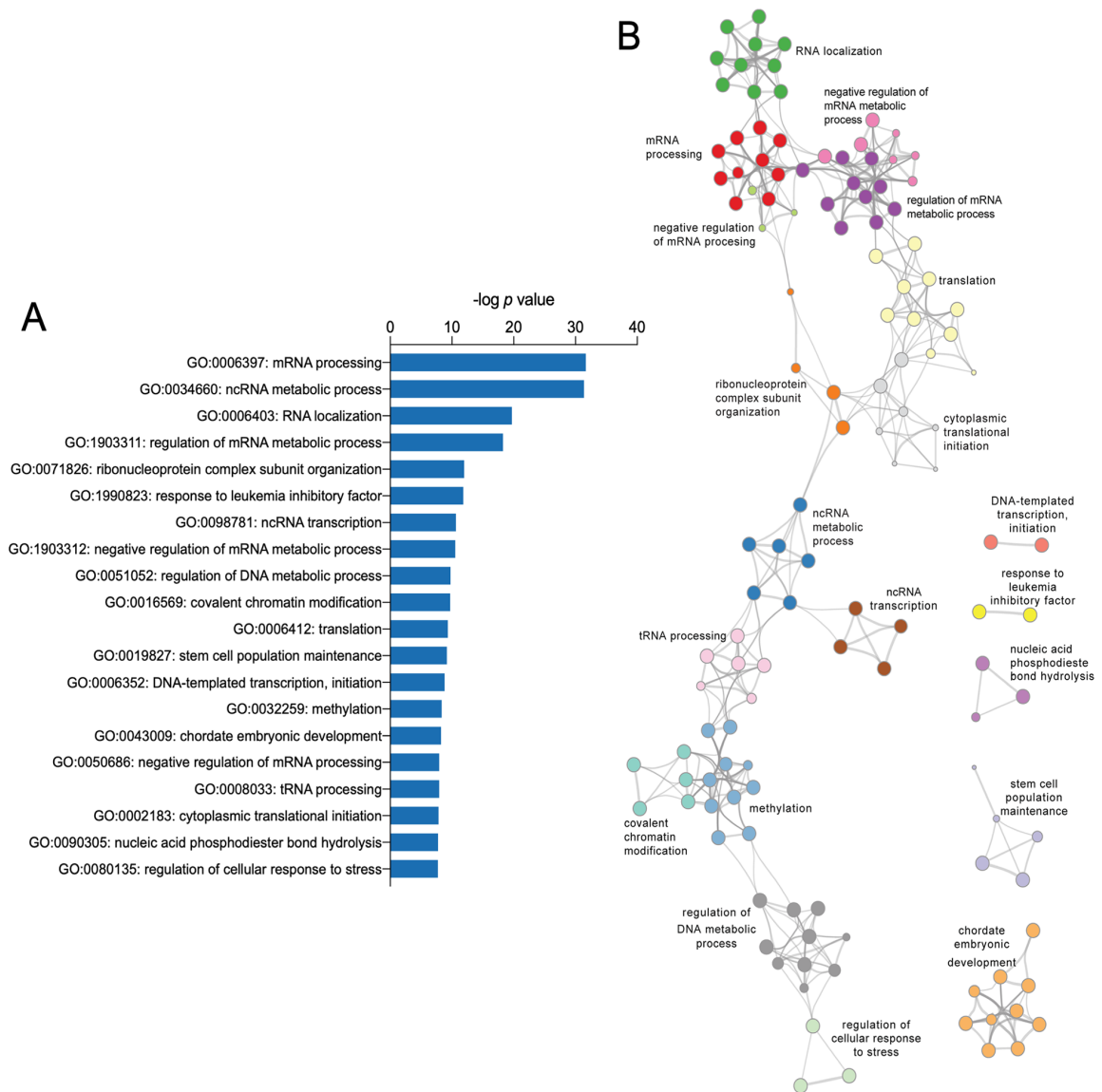
Figure 4.7. Significant biological process ontology terms and network clustering associated with significantly different genes in bovine scNT versus IVF blastocyst stage embryos. (A) Bar chart depicts all enriched summary terms identified for ontology biological processes. (B) Cytoscape network depicting biological process ontology terms represented as nodes connected with terms with shared differentially expressed genes. The size of the node is proportional to the term's enrichment score. Terms with a similarity score >0.3 are similarly colored and linked by an edge (the thickness of the edge represents the similarity score) using force-directed network layout and bundled edges for clarity. Each cluster is labeled by the term with the lowest p-value within the cluster group. Complete results of the Metascape analysis are provided in Supplementary File 4.3.
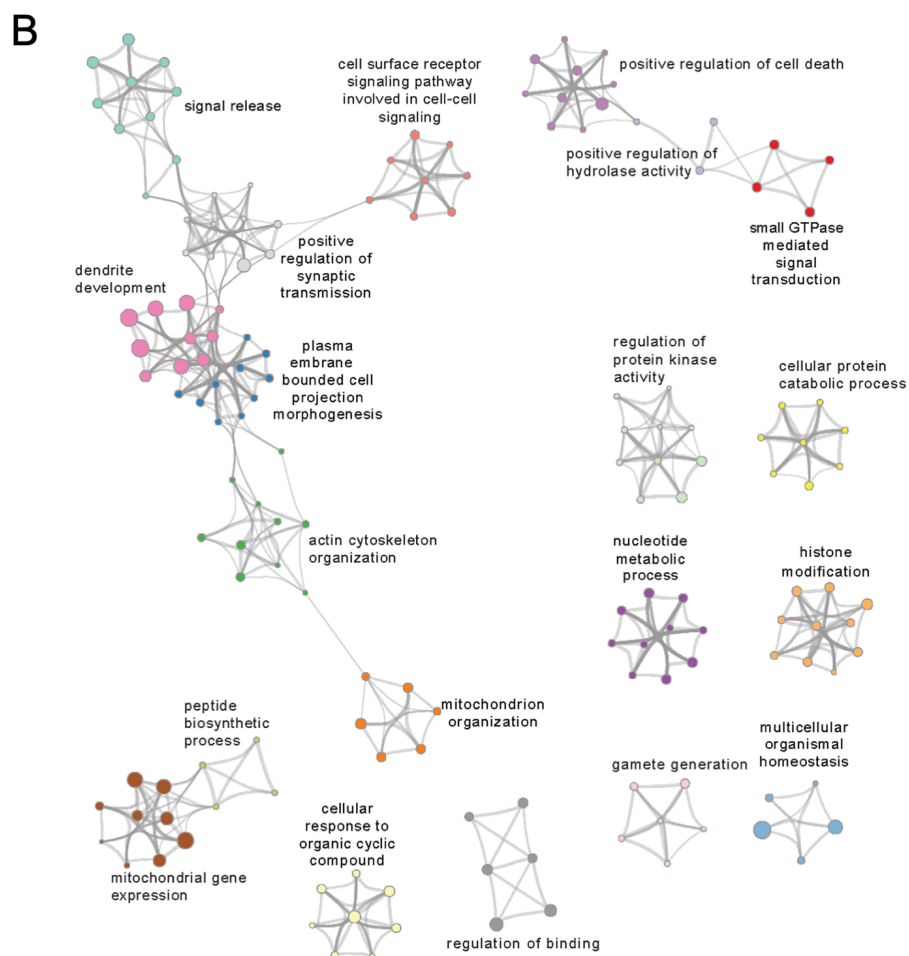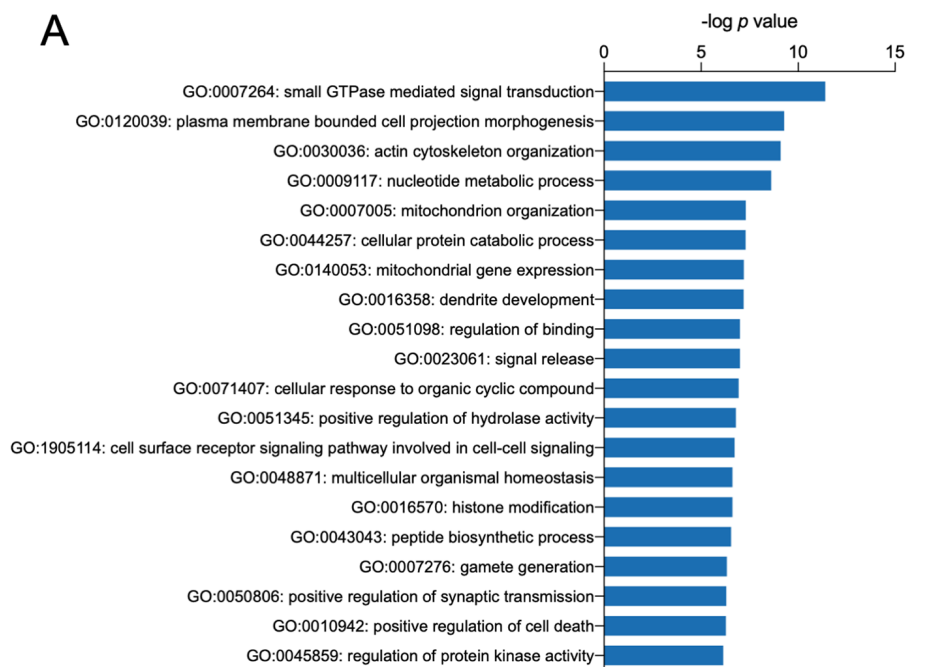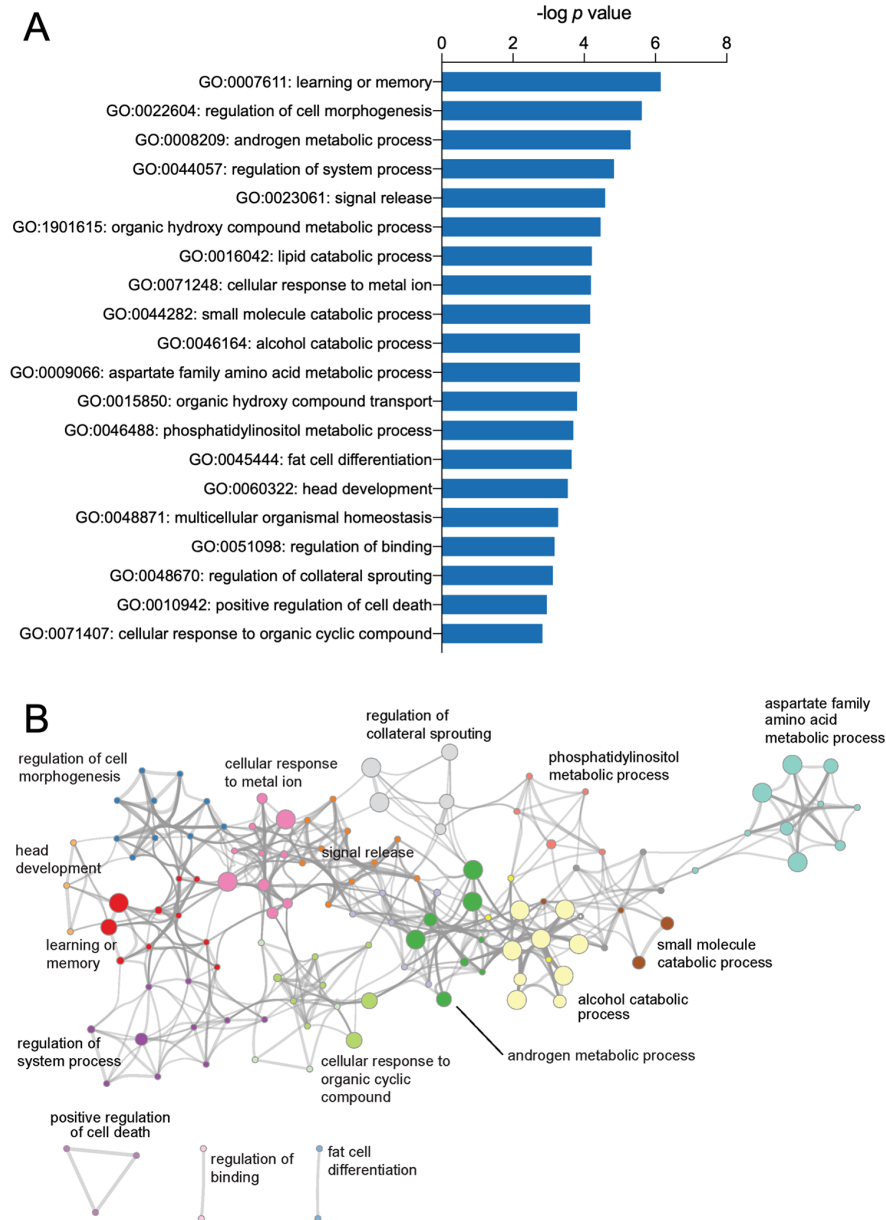
Figure 4.8. Significant biological process ontology terms and network clustering associated with significantly different genes in bovine BDCs derived from scNT versus IVF embryos. (A) Bar chart depicts the top 20 enriched summary terms identified for ontology biological processes. (B) Cytoscape network depicting biological process ontology terms represented as nodes connected with terms with shared differentially expressed genes. The size of the node is proportional to the term's enrichment score. Terms with a similarity score >0.3 are similarly colored and linked by an edge (the thickness of the edge represents the similarity score) using force-directed network layout and bundled edges for clarity. Each cluster is labeled by the term with the lowest p-value within the cluster group. Complete results of the Metascape analysis are provided in Supplementary File 4.3.
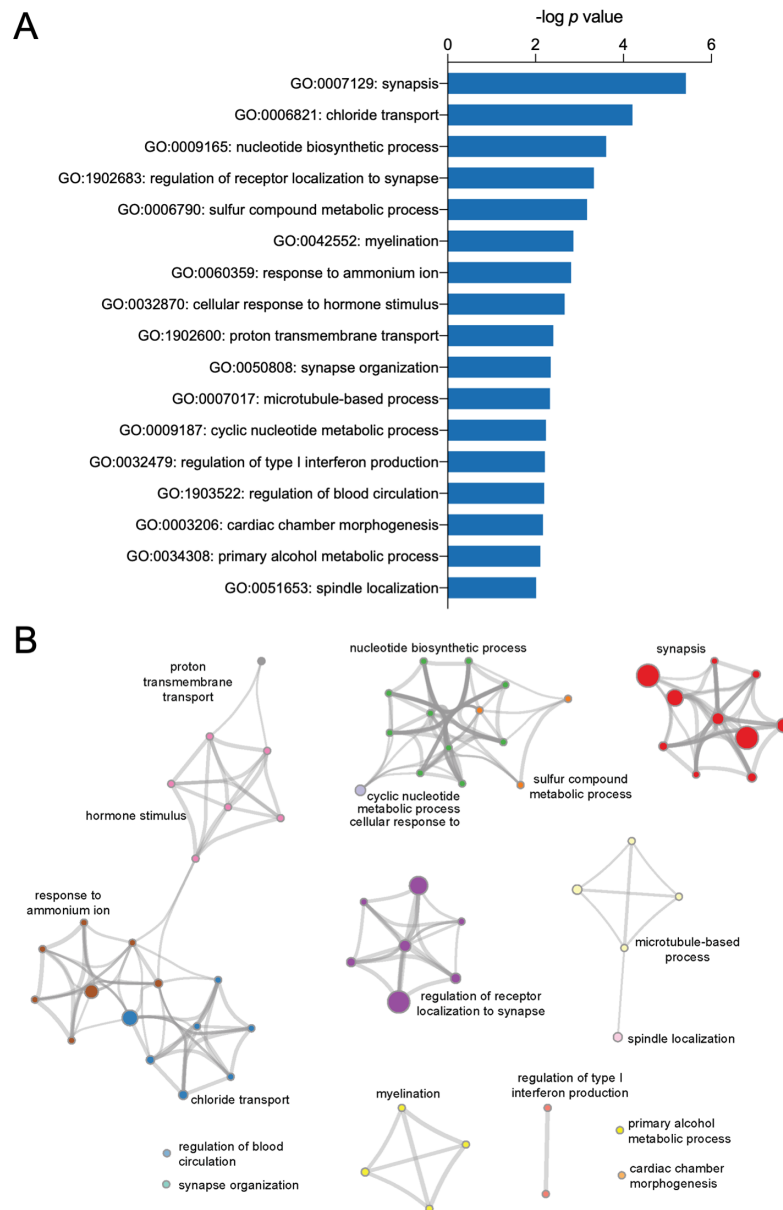
Figure 4.9. Spearman correlation plots for miRNAs and select predicted mRNA target genes. Significant correlation values ($p < 0.05$) are represented by colored dots according to the scale, with positive correlations in blue and negative correlations in red. Only mRNA predicted targets that were differentially expressed in scNT and IVF morula staged embryos are shown, although the correlation analyses used all expression data for all developmental time points.

Log$_2$ Fold Change

-3.0   0   3.0
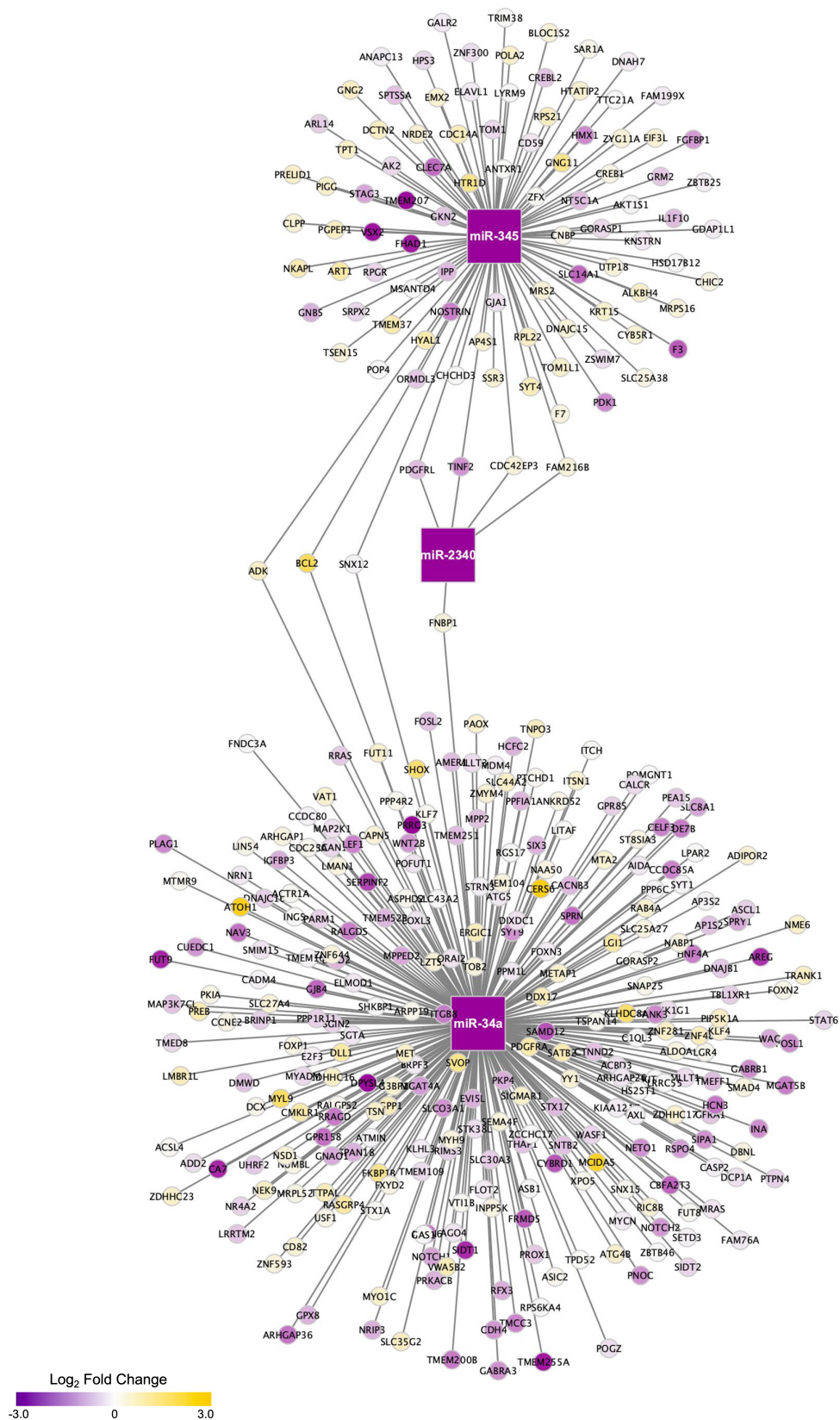
Figure 4.10. (prior page) Network of differentially expressed miRNAs in morula stage embryos and the expression of predicted mRNA targets. A network of miRNA differentially expressed in scNT morula stage embryos as compared of IVF morula stage embryos, and their expressed predicted mRNA targets (TargetScan total context ++ score <-0.35) was created using Cytoscape with a force-directed layout. Only miRNAs within families annotated by TargetScan as conserved or broadly conserved were included in the network analysis. MiRNAs are represented by squares and colored according to their relative abundance in scNT morula stage embryos as compared to IVF morula stage embryos, and target mRNAs are shown as circles and colored by their expression in scNT morula stage embryos as compared to IVF morula stage embryos using RNAseq data collected from the same samples (values are $\log_2$ fold change of normalized reads). Edges between miR and mRNA nodes indicate that the connected mRNA is a predicted target of the miRNA in *Bos taurus*.

Table 4.1. Differentially abundant mRNA transcripts

Comparison	Cattle mRNAs	Top 20 mRNA IDs
IVF vs. NT		
2-cell	113	*ASPN, ECM2, MEDAG, PRRX1, FGF7, COL6A3, BGN, S100A4, LOC534578, BEX2, CLCA3, LOC616308, LOC511494, ZSCAN4, SLC6A6, TMC6, LOC520023, CYP11A1, SLC35D2, KCNE4*
8-cell	1,539	*CALML5, GALR1, CACNA2D4, ATP6V0D2, CD48, SDS, SLC28A2, STC2, CXCL2, IGSF5, TACR3, GPR171, TENT5D, KRT15, LOC781486, GRO1, LOC789301, C15H11orf34, LOC786811, INSL3*
Morula	2,140	*ENPP6, LOC507550, BEST3, MZB1, FADS6, SRRM4, KRT73, DNER, GRIA3, RBFOX1, RBP7, GPR34, KLHL31, CSN3, ADAMDEC1, PRRX1, KRT1, NAPRT, DPT, C16H1orf158*
Blastocyst	289	*GIMAP4, GSTA3, GABRB3, AGMO, WNT11, ACTL8, LOC524236, CXCL14, LOC614881, ST6GALNAC1, NPSR1, LOC619094, GABRB1, FAM216B, CSN3, C16H1orf158, BCL11B, BHMT, CD8B, TPH1*
BDC	283	*GRIN2B, LOC781494, DCDC1, LOC617219, LOC785803, CRHBP, OXT, FAM216B, NETO1, FGF12, WSCD1, TMEM176B, TMEM176A, MRO, IGSF11, ZBTB16, DMRTA1, LIN7A, ASIP, CYP19A1*
Developmental Stage		
IVF embryos		
Oo vs. 2c	1,400	*PTGES, LOC511494, LOC616308, APOA1, BCHE, LOC281376, COL11A1, LOC789138, VILL, XCL1, TAGLN, LOC100299783, MSC, ATP1A2, ADPRHL1, COL5A2, LOC520034, CLCA3, KLF17, TMEM26*
2C vs. 8c	2,977	*TFAP2C, RET, MIR371, ZFP42, MYOD1, CT55, LOC782781, TENT5D, GALR1, ESRRB, SLC34A2, CNR2, KCNK5, FOLR1, MEDAG, LOC617141, SCTR, ETNPPL, NANOG, INSL3*
8c vs. Mo	7,106	*INSL3, MMP3, BRB, MPZ, CHI3L2, GDPD2, CRCT1, IL27, SPIC, PLAC8, NUPR1, FCGR2B, LOC520336, SUSD2, PECAM1, S100A5, WNT6, SMPDL3A, LOC528262, MAG*

Mo vs. Bl	6,678	*NLRP9, LBX1, RGS2, WEE2, SOX5, PRKG2, LRMP, UBASH3A, OXT, KRT73, IL18RAP, FAM216B, POPDC3, LOC783399, CLDN8, PAG2, CCL24, PRSS8, RBBP8NL, LOC508879*
Bl vs. BDC	8,814	*PAG11, GJB6, TKDP4, S100G, NUDT11, FGF12, RBM24, NETO1, CRHBP, WSCD1, INSL3, OXT, LOC516378, POU5F1, SPIC, BRB, ZAP70, DPPA2, OTX2, LOC528262*
2C. vs Fib	12,471	*CCN3, ESM1, S100A4, CCN5, BGN, PRRX2, COL6A3, BTG4, NEFL, SALL4, KCNN3, KPNA7, EIF4E1B, NLRP8, MOS, LOC100301263, WEE2, ACCSL, BMP15, NLRP14*

scNT embryos

Oo vs. 2c	696	*TFCP2L1, TLL2, A4GALT, BEX2, MBD3L3, PRRX1, S100A4, BGN, NOV, AEBP1, FGF7, COL3A1, COL6A3, LOC281376, COL11A1, EMP1, THBD, ECM2, TFPI2, OMD*
2C vs. 8c	1,092	*BEX2, KLHL31, LOC534578, LOC616308, CLCA3, FOLR1, LOC789138, LOC617709, KLF17, LOC525100, LOC100139585, LOC520023, LOC782781, LOC616911, BECN2, LOC525101, LOC530538, MBD3L3, LOC520034, ZSCAN4*
8c vs. Mo	10,942	*CYP4A22, B3GNT6, FGFR4, CHI3L2, CLDN6, CSN3, EIF1AY, LOC615989, SLC39A2, RNASE1, SPIC, LOC785540, BRB, S100A5, LOC100847738, KRT1, LOC528262, CD48, ATP6V0D2, KLHL31*
Mo vs. Bl	5,859	*C26H10orf90, LOC100139585, POPDC3, GABRB1, CHRM3, KLHL31, CCL17, PAG2, CLDN8, C9H6orf58, CCL22, PRSS8, MYL7, PRSS22, CLDN23, HNF4A, ZBTB42, SULT1E1, GPR34, SLC9A3*
Bl vs. BDC	8,623	*MMP9, TKDP4, GABRB1, S100G, IGSF11, NUDT11, TMEM176B, LIN7A, TMEM176A, ZBTB16, MRO, RBM24, DMRTA1, INSL3, LOC516378, SPIC, DPPA2, CFAP54*
2C vs. Fib	13,049	*BTG4, NEFL, LOC100848540, KCNN3, MBD3L3, SALL4, KPNA7, LOC101904481, EIF4E1B, MOS, WEE2, NLRP8, ACCSL, NLRP14, LOC100301263, DPPA3, BMP15, LHX8, NPM2, HHIP*

Complete results of DESeq2 differential expression analyses for miRNAs are provided in Supplementary File 4.2. Abbreviations are: Oo, oocyte; 2c, 2-cell; 8c, 8-cell; Mo, morula; Bl, blastocyst; BDC, blastocyst-derived cells; Fb, fibroblast; IVF, in vitro-fertilized; scNT, somatic cell nuclear transfer; miRNA, microRNA; ID, identification number.
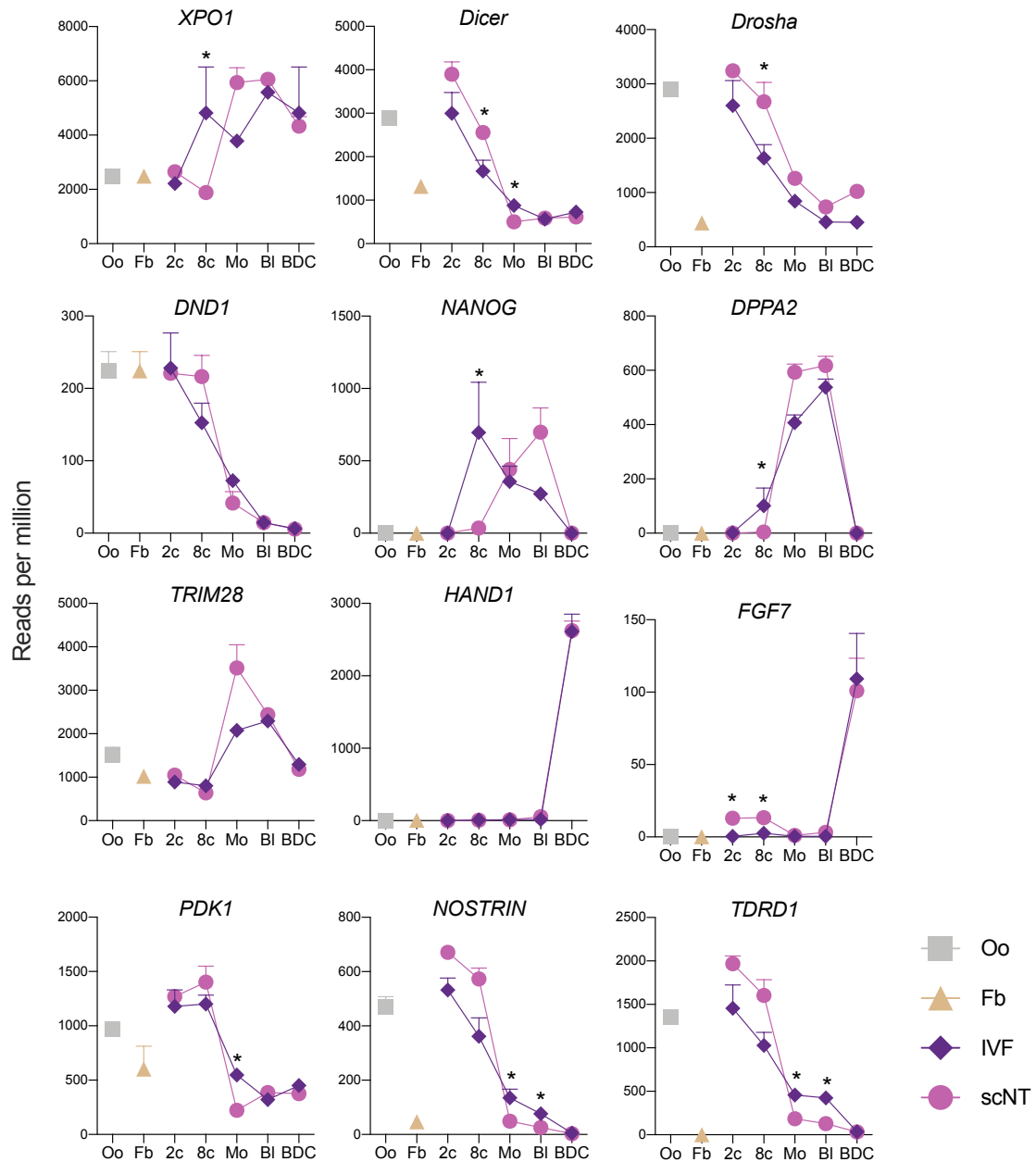
Supplementary Figure 4.S1. Dynamic expression of selected mRNAs of interest in early developing IVF and scNT embryos, oocytes and fibroblasts. Values shown are the normalized reads per million. * *p* <0.05 for IVF vs. scNT as determined by DESeq2 analysis. Differential expression by embryo stage is summarized in Supplementary File 4.2. Fb, fibroblast; Oo, oocyte; IVF, *in vitro*-fertilized embryo; scNT, scNT embryo; 2c, 2-cell; 8c, 8-cell; Mo, morula; Bl, blastocyst; BDC, blastocyst-derived cells.
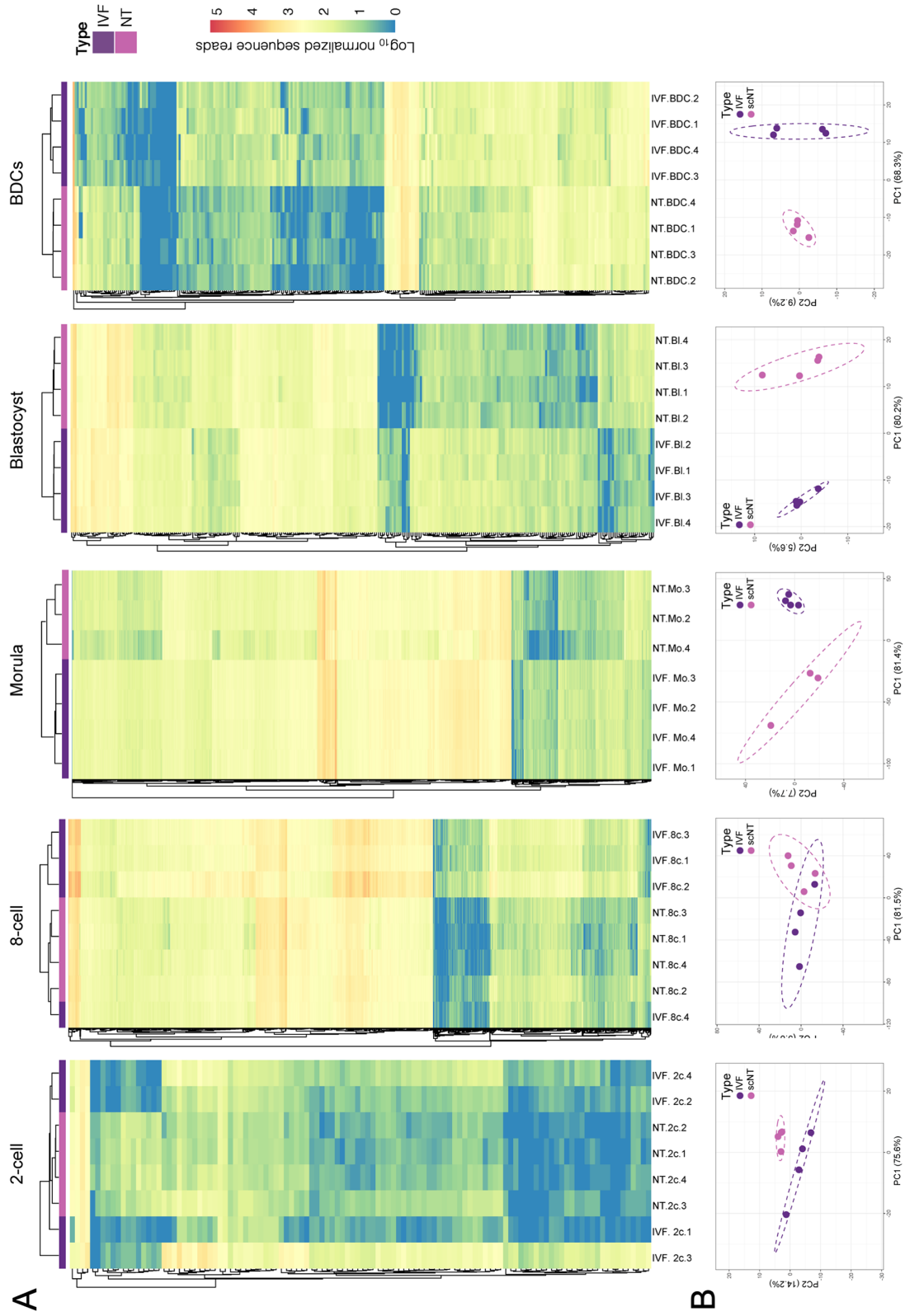
Supplementary Figure 4.S2. (prior page) Patterns of mRNA expression in IVF and scNT embryos at each developmental stage. (A) Unsupervised, bi-directional hierarchical cluster analysis (Euclidean distance method) of variance stabilized normalized read counts for mRNAs identified as significantly different between embryos of the same stage and different types. Data were categorized by sample type (IVF, *in vitro*-fertilized embryo; scNT, scNT embryo) and stage (2c, 2-cell; 8c, 8-cell; Mo, morula; Bl, blastocyst; BDC, blastocyst-derived cells). (B) Principal components analysis of bovine mRNAs significantly different for the stage transition, using the standard singular value decomposition method with imputation.

Supplementary Figure 4.S3. Patterns of mRNA expression in IVF embryos with comparisons by developmental stage. (A) Unsupervised, bi-directional hierarchical cluster analysis (Euclidean distance method) of variance stabilized normalized read counts for mRNAs identified as significantly different between embryos of the same stage and different types. Data were categorized by sample type (IVF, *in vitro*-fertilized embryo; scNT, scNT embryo) and stage (2c, 2-cell; 8c, 8-cell; Mo, morula; Bl, blastocyst; BDC, blastocyst-derived cells). (B) Principal components analysis of bovine mRNAs significantly different for the stage transition, using the standard singular value decomposition method with imputation.
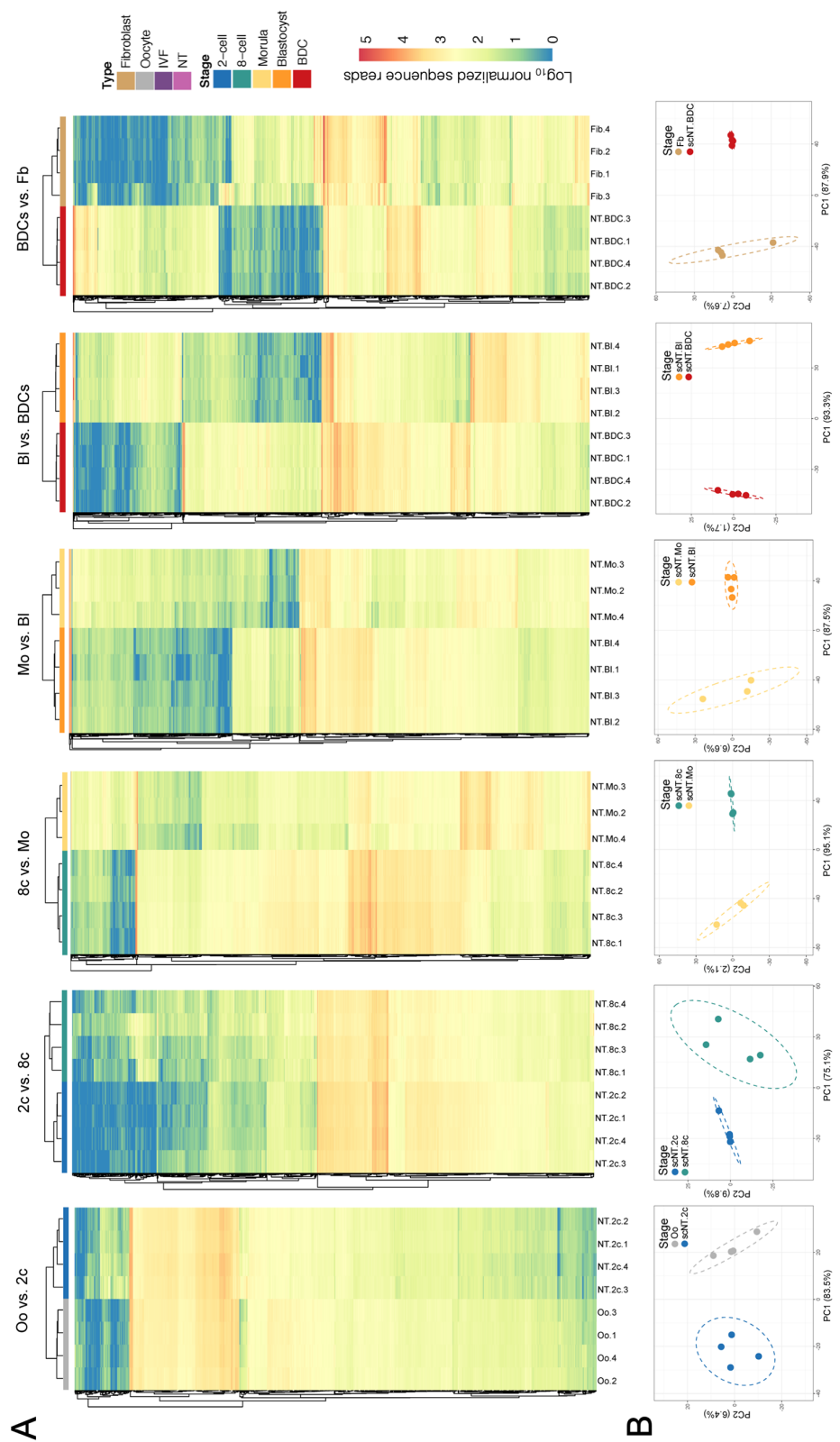
Supplementary Figure 4.S4. (prior page) Patterns of mRNA expression in scNT embryos with comparisons by developmental stage. (A) Unsupervised, bi-directional hierarchical cluster analysis (Euclidean distance method) of variance stabilized normalized read counts for mRNAs identified as significantly different between embryos of the same type and subsequence stages. Data were categorized by sample type (Fb, fibroblast; Oo, oocyte; IVF, *in vitro*-fertilized embryo; scNT, scNT embryo) and stage (2c, 2-cell; 8c, 8-cell; Mo, morula; Bl, blastocyst; BDC, b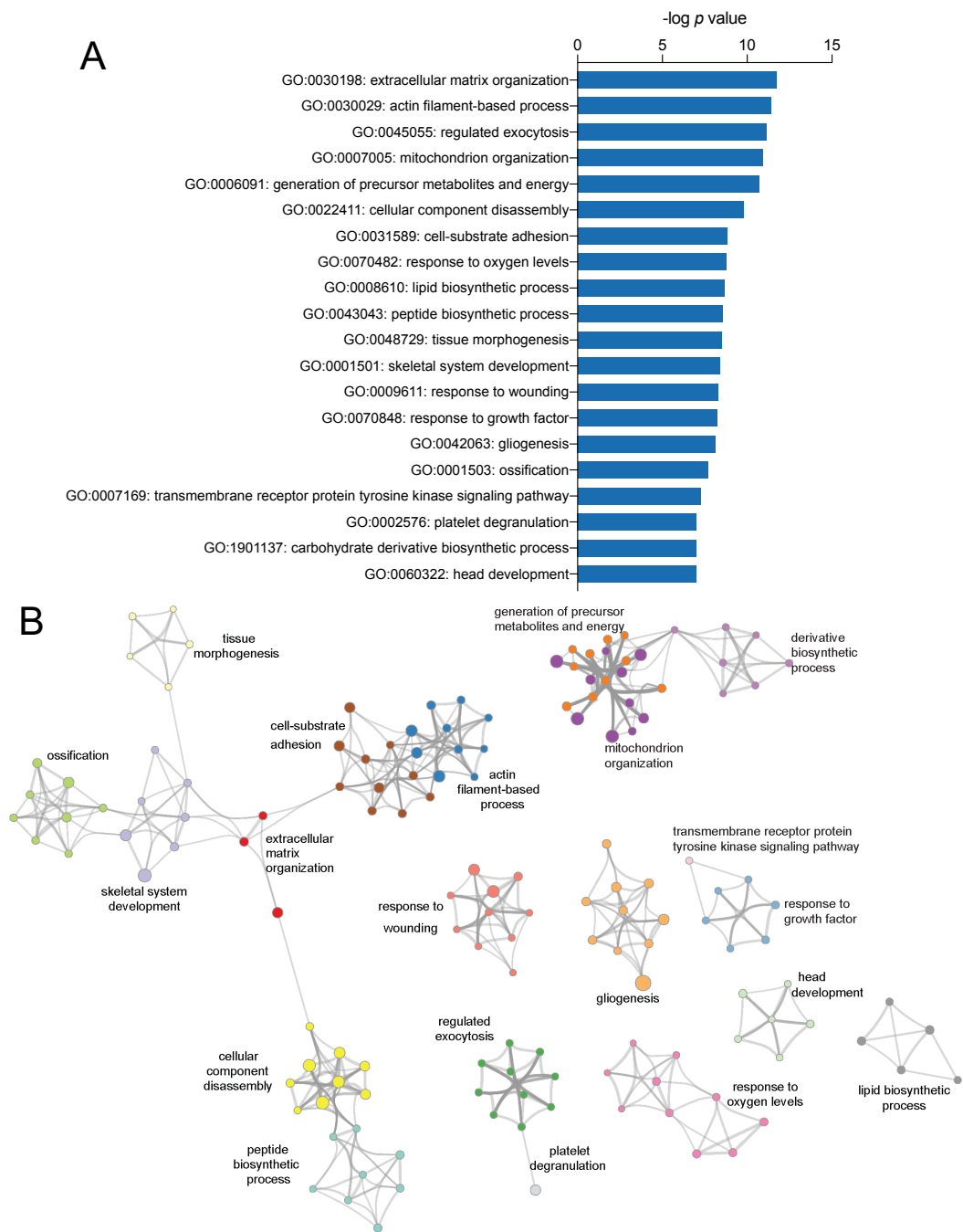lastocyst-derived cells). (B) Principal components analysis of bovine mRNAs significantly different for the stage transition, using the standard singular value decomposition method with imputation.

A

-log *p* value

GO:0030198: extracellular matrix organization
GO:0030029: actin filament-based process
GO:0045055: regulated exocytosis
GO:0007005: mitochondrion organization
GO:0006091: generation of precursor metabolites and energy
GO:0022411: cellular component disassembly
GO:0031589: cell-substrate adhesion
GO:0070482: response to oxygen levels
GO:0008610: lipid biosynthetic process
GO:0043043: peptide biosynthetic process
GO:0048729: tissue morphogenesis
GO:0001501: skeletal system development
GO:0009611: response to wounding
GO:0070848: response to growth factor
GO:0042063: gliogenesis
GO:0001503: ossification
GO:0007169: transmembrane receptor protein tyrosine kinase signaling pathway
GO:0002576: platelet degranulation
GO:1901137: carbohydrate derivative biosynthetic process
GO:0060322: head development

B

Supplementary Figure 4.S5. Significant biological process ontology terms and network clustering associated with significantly different genes in bovine 2-cell IVF embryos compared to oocytes. (A) Bar chart depicts the top 20 enriched summary terms identified for ontology biological processes. (B) Cytoscape network depicting biological process ontology terms represented as nodes connected with terms with shared differentially expressed genes. The size of the node is proportional to the term's enrichment score. Terms with a similarity score >0.3 are similarly colored and linked by an edge (the thickness of the edge represents the similarity score) using force-directed network layout and bundled edges for clarity. Each cluster is labeled by the term with the lowest *p*-value within the cluster group. Complete results of the Metascape analysis are provided in Supplementary File 4.3.

-log *p* value

A

GO:0006397: mRNA processing
GO:0034660: ncRNA metabolic process
GO:0006403: RNA localization
GO:0010564: regulation of cell cycle process
GO:1903311: regulation of mRNA metabolic process
GO:0006281: DNA repair
GO:0016569: covalent chromatin modification
GO:0098781: ncRNA transcription
GO:0031124: mRNA 3'-end processing
GO:0090305: nucleic acid phosphodiester bond hydrolysis
GO:0080135: regulation of cellular response to stress
GO:0006412: translation
GO:0043009: chordate embryonic development
GO:0032984: protein-containing complex disassembly
GO:0033044: regulation of chromosome organization
GO:0051603: proteolysis involved in cellular protein catabolic process
GO:0009314: response to radiation
GO:0006354: DNA-templated transcription, elongation
GO:0006260: DNA replication
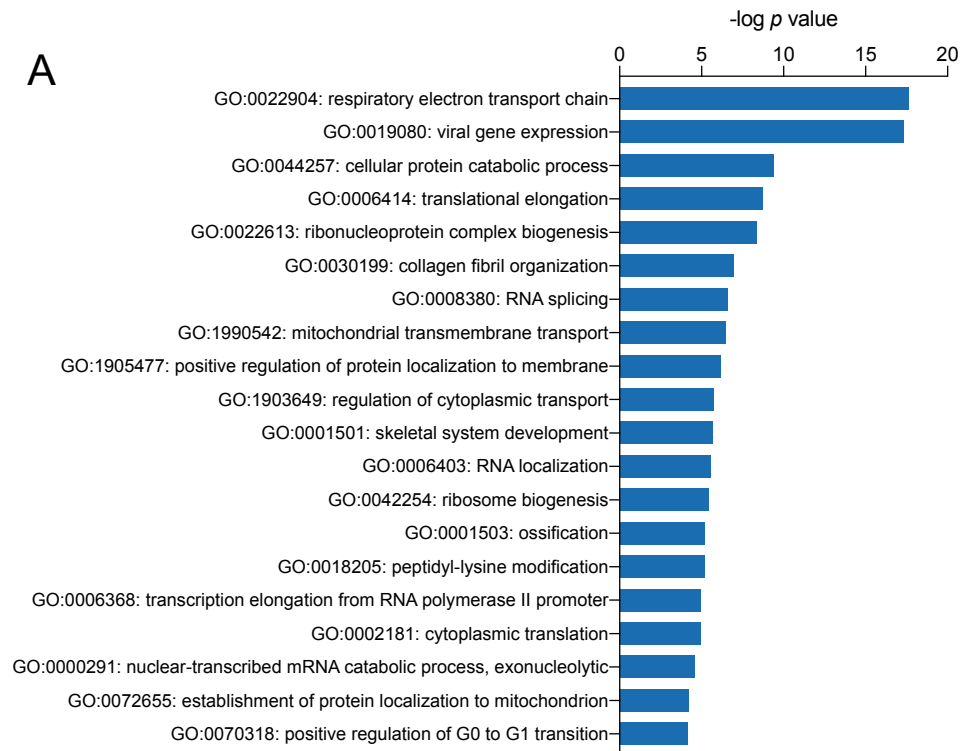GO:0016236: macroautophagy

B

Supplementary Figure 4.S6. Significant biological process ontology terms and network clustering associated with significantly different genes in bovine 8-cell versus 2-cell IVF embryos. (A) Bar chart depicts the top 20 enriched summary terms identified for ontology biological processes. (B) Cytoscape network depicting biological process ontology terms represented as nodes connected with terms with shared differentially expressed genes. The size of the node is proportional to the term's enrichment score. Terms with a similarity score >0.3 are similarly colored and linked by an edge (the thickness of the edge represents the similarity score) using force-directed network layout and bundled edges for clarity. Each cluster is labeled by the term with the lowest *p*-value within the cluster group. Complete results of the Metascape analysis are provided in Supplementary File 4.3.
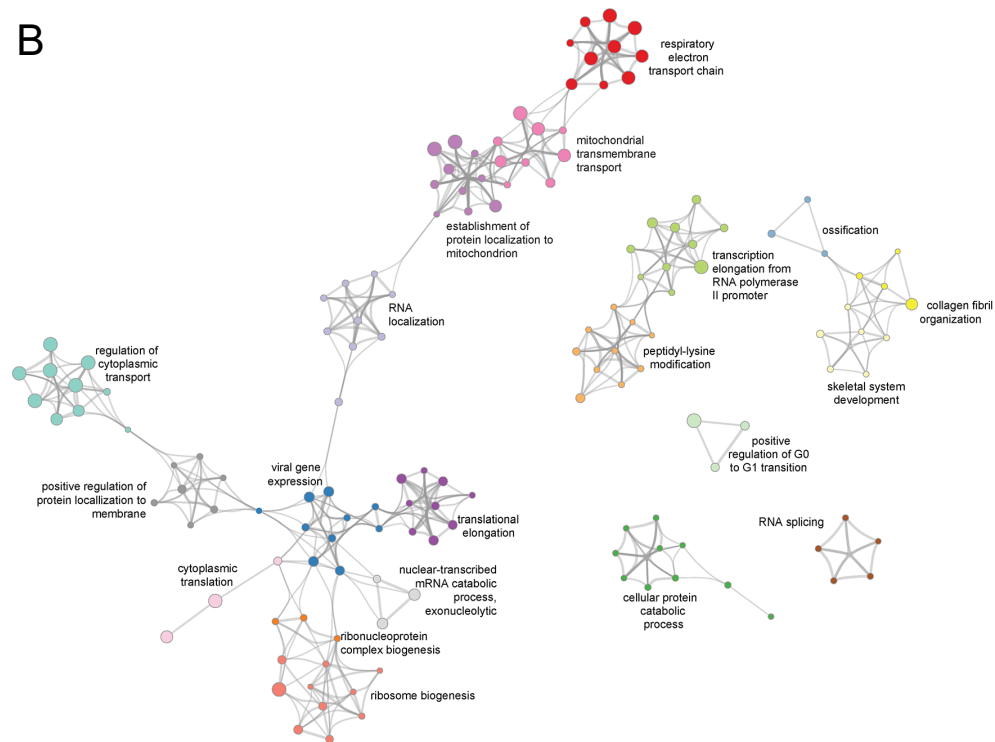
A

-log *p* value

GO:0022904: respiratory electron transport chain
GO:0019080: viral gene expression
GO:0044257: cellular protein catabolic process
GO:0006414: translational elongation
GO:0022613: ribonucleoprotein complex biogenesis
GO:0030199: collagen fibril organization
GO:0008380: RNA splicing
GO:1990542: mitochondrial transmembrane transport
GO:1905477: positive regulation of protein localization to membrane
GO:1903649: regulation of cytoplasmic transport
GO:0001501: skeletal system development
GO:0006403: RNA localization
GO:0042254: ribosome biogenesis
GO:0001503: ossification
GO:0018205: peptidyl-lysine modification
GO:0006368: transcription elongation from RNA polymerase II promoter
GO:0002181: cytoplasmic translation
GO:0000291: nuclear-transcribed mRNA catabolic process, exonucleolytic
GO:0072655: establishment of protein localization to mitochondrion
GO:0070318: positive regulation of G0 to G1 transition

B

Supplementary Figure 4.S7. Significant biological process ontology terms and network clustering associated with significantly different genes in bovine 2-cell scNT embryos compared to oocytes. (A) Bar chart depicts the top 20 enriched summary terms identified for ontology biological processes. (B) Cytoscape network depicting biological process ontology terms represented as nodes connected with terms with shared differentially expressed genes. The size of the node is proportional to the term's enrichment score. Terms with a similarity score >0.3 are similarly colored and linked by an edge (the thickness of the edge represents the similarity score) using force-directed network layout and bundled edges for clarity. Each cluster is labeled by the term with the lowest *p*-value within the cluster group. Complete results of the Metascape analysis are provided in Supplementary File 4.3.
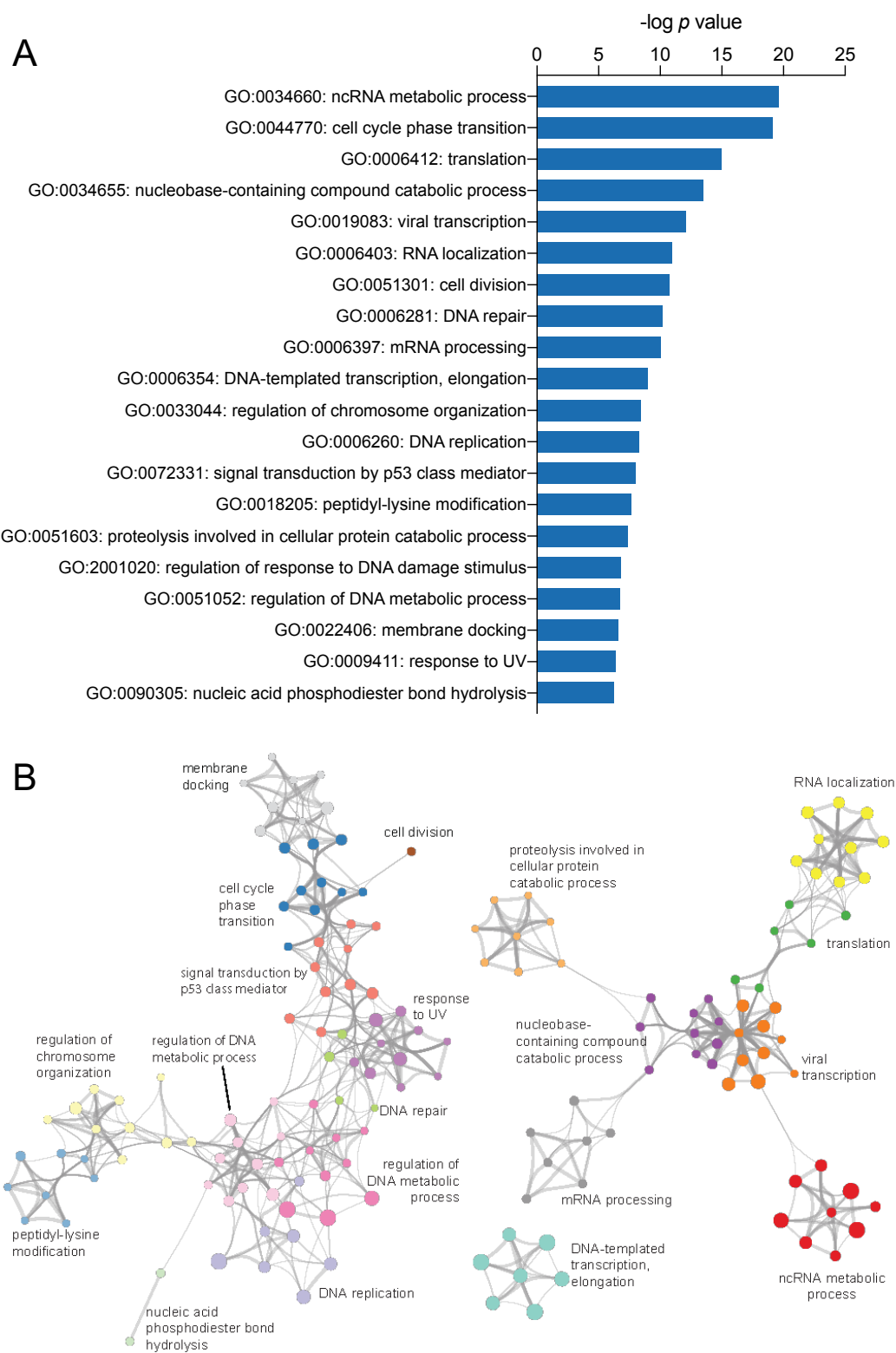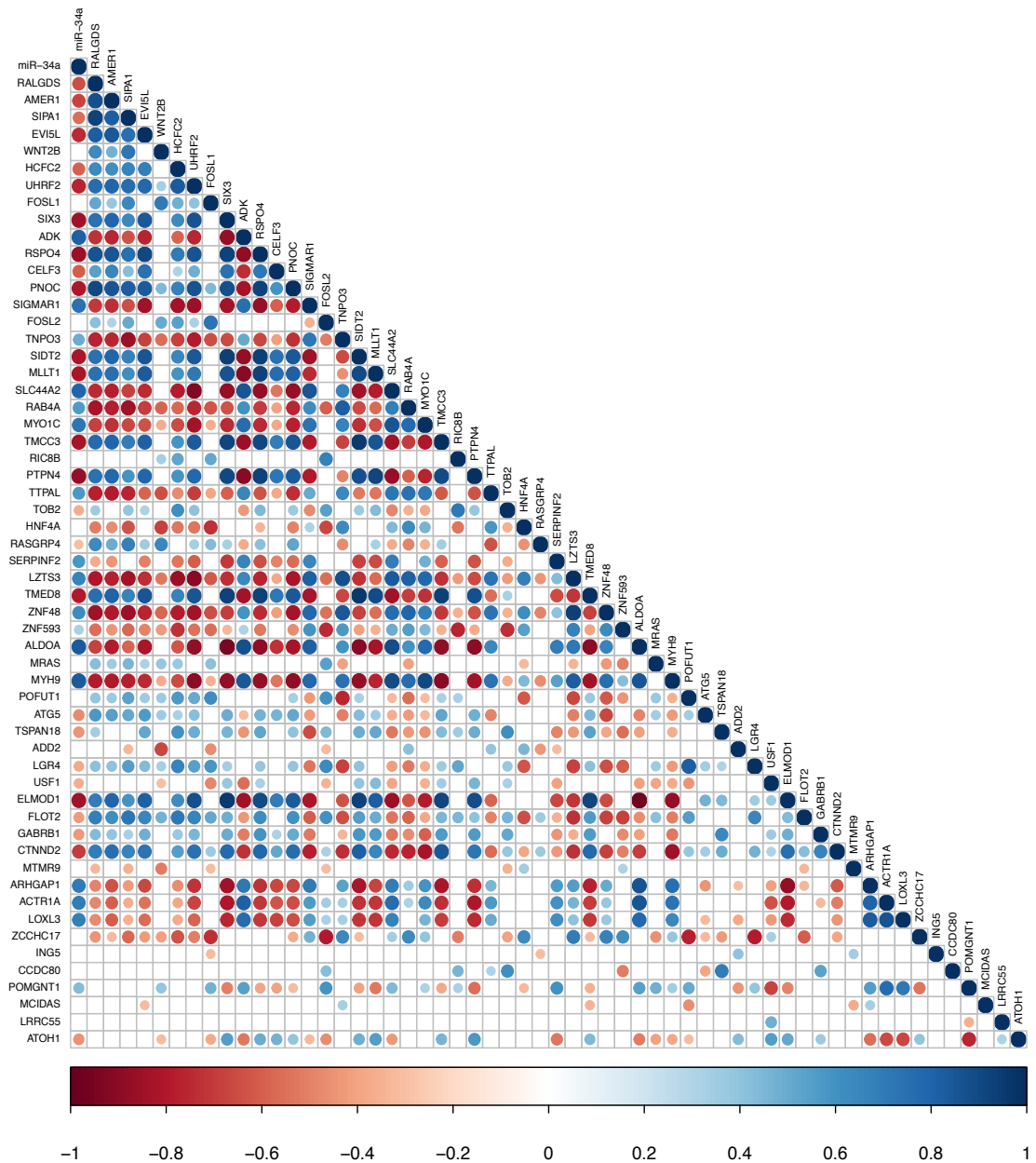
Supplementary Figure 4.S8. Significant biological process ontology terms and network clustering associated with significantly different genes in bovine 2-cell versus 8-cell scNT embryos. (A) Bar chart depicts the top 20 enriched summary terms identified for ontology biological processes. (B) Cytoscape network depicting biological process ontology terms represented as nodes connected with terms with shared differentially expressed genes. The size of the node is proportional to the term's enrichment score. Terms with a similarity score >0.3 are similarly colored and linked by an edge (the thickness of the edge represents the similarity score) using force-directed network layout and bundled edges for clarity. Each cluster is labeled by the term with the lowest *p*-value within the cluster group. Complete results of the Metascape analysis are provided in Supplementary File 4.3.
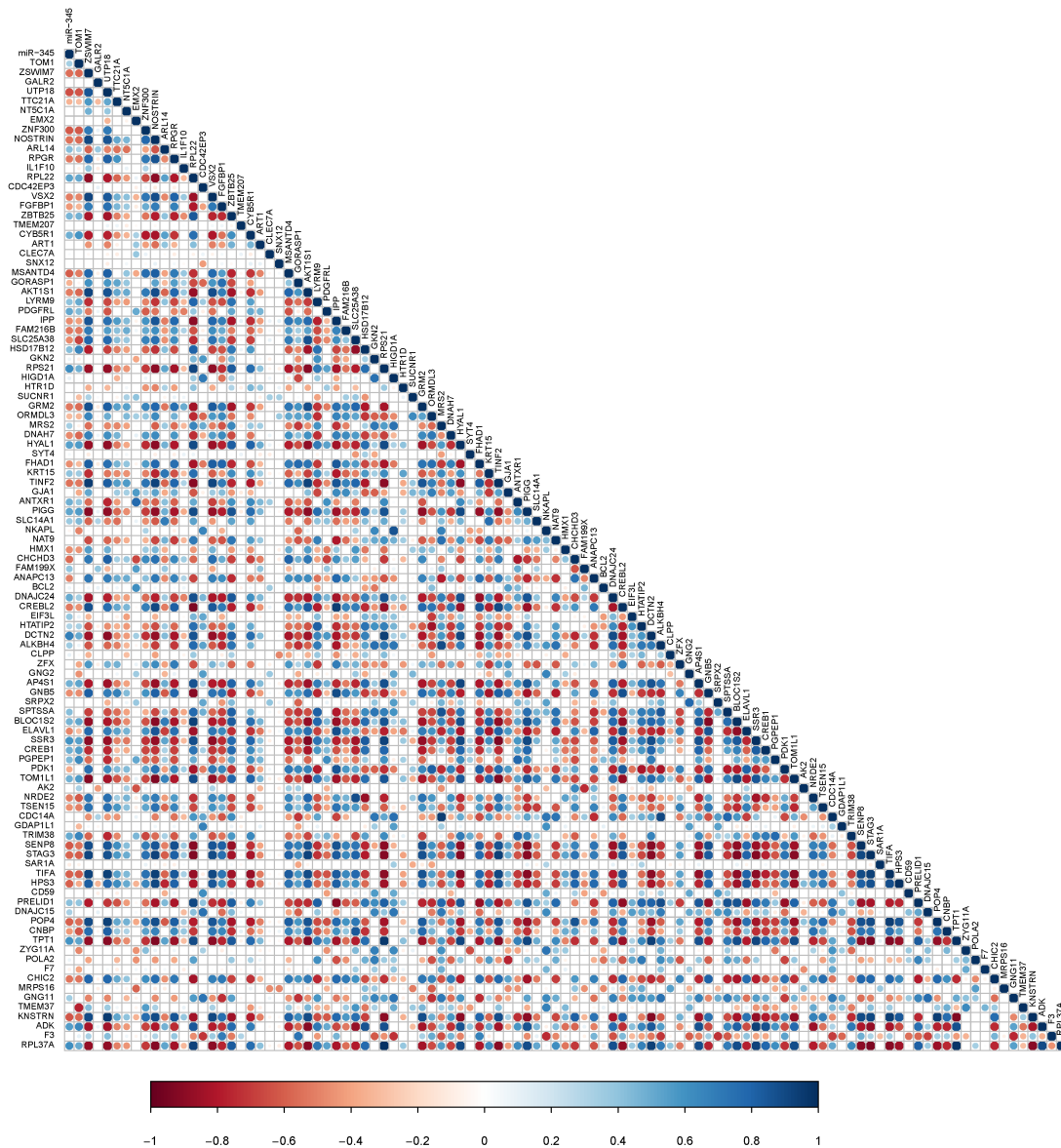
Supplementary Figure 4.S9. Spearman correlation plots for miR-34a and select predicted mRNA target genes. Significant correlation values ($p < 0.05$) are represented by colored dots according to the scale, with positive correlations in blue and negative correlations in red. Correlation analyses used all expression data for all developmental time points.
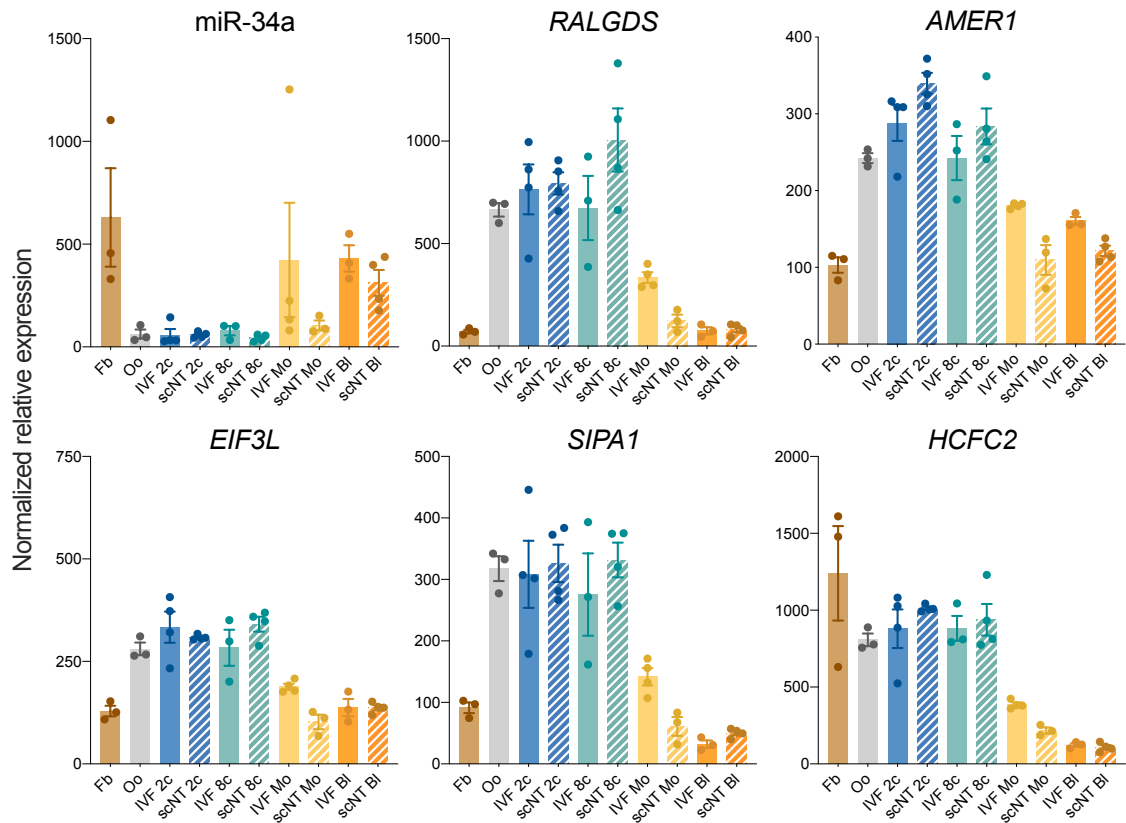
Supplementary Figure 4.S10. Comparison of expression for miR-34a and selected mRNA targets. Values shown are the normalized reads per sample for mRNAs, showing mean and +/- SEM for transcripts predicted to be targeted by miR-34a and demonstrating a negative correlation of expression in correlation plots. See Supplementary File 4.3 for statistical results for all comparisons of interest.
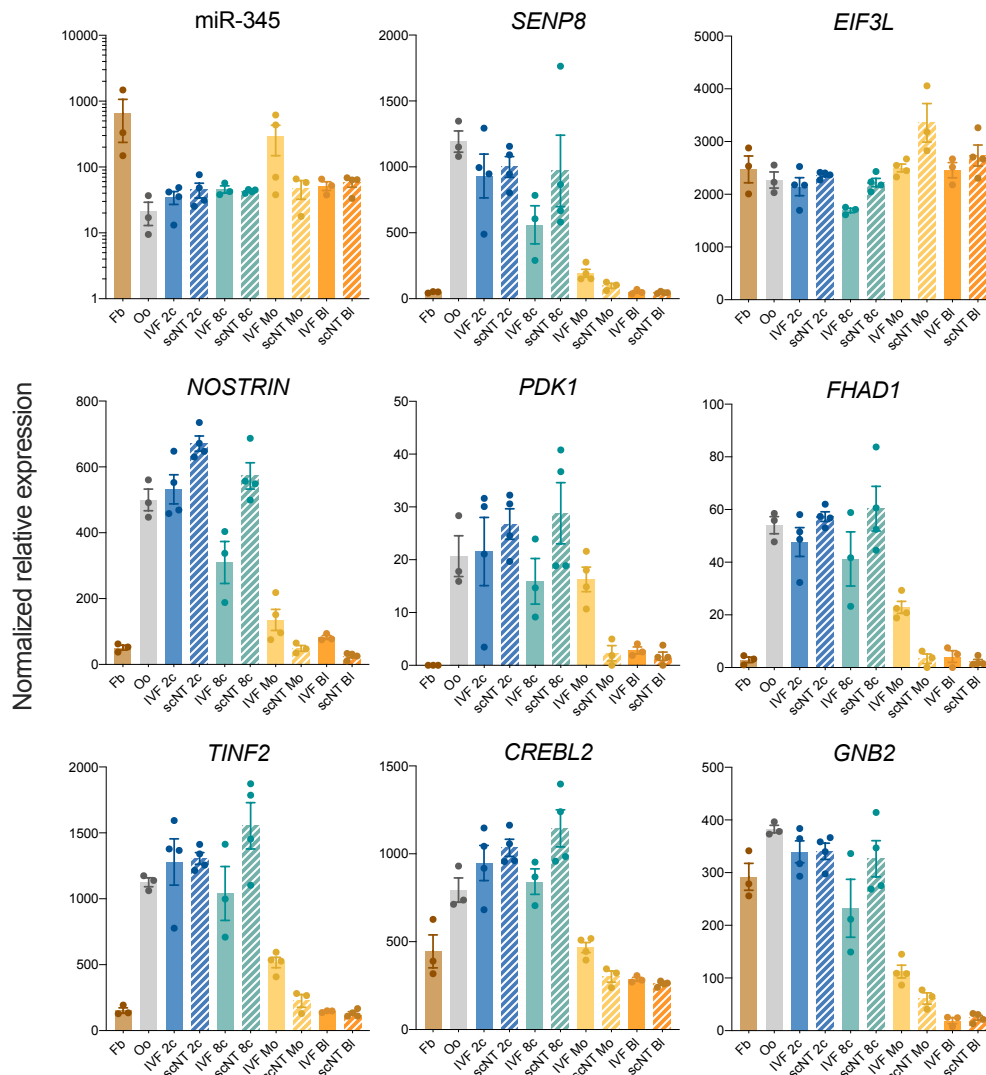
Supplementary Figure 4.S11. Spearman correlation plots for miR-345 and select predicted mRNA target genes. Significant correlation values ($p < 0.05$) are represented by colored dots according to the scale, with positive correlations in blue and negative correlations in red. Correlation analyses used all expression data for all developmental time points.

Supplementary Figure 4.S12. Comparison of expression for miR-345 and selected mRNA targets. Values are the normalized reads per sample for mRNAs, showing mean and +/- SEM for transcripts predicted to be targeted by miR-345 and demonstrating a negative correlation of expression in correlation plots. See Supplementary File 4.3 for statistical results for all comparisons of interest.

CHAPTER 5

CONCLUSIONS

Deficiencies in cloned embryo development are likely the result of incomplete or abnormal epigenetic reprogramming. Following somatic cell nuclear transfer (scNT), persistent DNA methylation [1, 2] and histone profiles [3, 4] of the somatic donor cell are seen in the resulting embryo. These aberrant epigenetic profiles are expected to lead to inappropriate gene expression, likely causing the low developmental rates observed in scNT embryos [5]. Small non-coding RNAs (sncRNAs) target and degrade messenger RNA (mRNA) through RNA interference (RNAi). These sncRNAs are dynamically expressed through early development, and it is possible that these regulatory molecules are also incompletely reprogrammed in scNT embryos. In Chapter 2, our first hypothesis was that miRNAs would be more abundant in 8-cell embryos at the start of the MZT and that abundant miRNAs would target maternal mRNAs with important developmental functions associated with epigenetic programming and development. In Chapter 3, we next we hypothesized that miRNAs found in the nuclear donor cell would be more abundant in scNT embryos as compared to embryos produced via *in vitro* fertilization (IVF). Also, we hypothesized that aberrant expression patterns of miRNAs in scNT embryos would occur as compared to IVF embryos, and the sncRNA profile of the somatic donor cells would persist in scNT embryos potentially due to incomplete reprogramming. Finally, in Chapter 4, we hypothesized that transcripts derived from the maternal genome that need to be degraded for successful MET would be aberrantly expressed in the scNT embryos at the 8-cell stage. In addition, we hypothesized that samples found to have unique changes in miRNA populations would also show a distinct transcriptome profile in transcripts predicted to be targets of those unique miRNAs. The experimental approach involved the use of sncRNA sequencing on a pool of IVF embryos from embryo developmental stages before, during, and after the MET. In a

second experiment, we also performed RNA sequencing of both small and large RNA fractions isolated from the same pool of samples, comprised of developmental stages before, during, and after the MET in addition to somatic nuclear donor cells (fibroblasts) and blastocyst-derived cells (BDCs). To examine the impact of specific miRNA errors on the transcriptome of the embryo through the MZT, we aimed to identify sncRNA molecules aberrantly expressed bovine scNT embryos, and to link those differences to abnormally expressed target transcripts.

We determined that pools of 20 embryos could be successfully used to isolate both large and small RNA, and to generate good quality sequencing data. We found that sncRNA populations were distinct at different stages of embryo development. Moreover, over the course of development, the profile of scRNAs changed accordingly to the level of differentiation, within the embryo type (IVF or scNT). However, no striking differences in the profile of sncRNAs were evident when comparing scNT to IVF embryos for microRNAs (miRNAs), piwi-interacting RNAs (piRNAs), and transfer RNA fragments (tRFs). However, a few statistically significant differences were noted for several sncRNAs at the morula stage of development when comparing scNT versus IVF embryos. Alternatively, large-scale differences in the transcriptome were evident at multiple developmental stages when comparing scNT embryos to their IVF counterparts, with the greatest number of differentially expressed genes noted at the morula stage, coinciding with the developmental stage with significant alternations in sncRNAs. Ontology analysis of these differentially expressed genes pointed to enriched biological process pathways related to RNA processing and ncRNA pathways, suggesting that aberrant gene expression in scNT embryos could negatively impact the EGA. Finally, through network mapping, potential associations were examined between differentially expressed miRNAs and their predicted target mRNAs. Overall, the expression patterns were not broadly inversely correlated as would be anticipated for widespread miRNA target binding and degradation of the predicted targets. This

observation suggests that the observed marked changes in mRNA abundance were not closely tied to changes miRNA abundance in these embryos.

As in any study, this work had limitations. The first limitation concerns the accuracy of prediction tools to tie miRNA changes to predicted target expression. Identifying miRNA targets can be difficult as a single miRNA molecule may target hundreds of mRNAs, and one mRNA may be targeted by multiple miRNAs [6]. Further complicating the miRNA-mRNA relationship, the miRNA-mRNA network is influenced by regulatory factors that contribute to post-transcriptional gene silencing. RNA-binding proteins mediate binding between miRNAs and the 3' untranslated region (UTR) in target transcripts, allowing those targets to escape repression. One such protein is dead end homologue 1 (DND1), which binds to mRNAs containing an AU-rich element in the 3′ UTR and blocks miRNA binding to this site. DND1 is abundant in early developing embryos, although its expression is down-regulated at the EGA, the stage at which maternal transcripts would require degradation. DND1 is also functionally capable of binding pluripotency genes, which are vital during early development, and may protect developmentally important transcripts from degradation before the embryonic stage at which they are functional [7]. It is likely that the effectiveness of miRNAs targeting transcripts for degradation is a combination of degradation signals from miRNA and interactions with RNA-binding proteins as opposed to a simple one-to-one matching process [6]. However, the complexity of the miRNA-mRNA binding code limits our ability to predict when a miRNA will repress expression of a target mRNA.

The difference noted for expression of sncRNA in Chapter 2 and Chapter 3 may be due to methodological differences in embryo sample collection and culture. Pools of oocytes or embryos were required to obtain sufficient amounts of small RNA and mRNA for sequencing. In our first study, pools of 40 embryos and oocytes were used. However, based on the robust results of the first study, the subsequent experiment used pools of only 20 embryos. However, this two-

fold variation in the number of embryos between pools in the two studies could mean that the variability among individual embryos was more effectively masked in the first study. Pooling can mask biological variability between individual embryos, making it difficult to identify embryos that may represent outliers. Staged embryos were collected based on consistency in morphology, however high quality morphology blastocyst stage embryos are known to express variable transcriptomes [8], and one may also expect similar differences in sncRNA expression. In this project, the same donor semen was used for all experiments, with the first study only requiring two weeks to generate the total population of IVF embryos sequenced. Because the second experiment required the generation of both IVF and scNT embryos, and additional embryo stages were needed, the production of embryos occurred over the span of two years. Given the short duration of collection for the first experiment, it is possible that the abattoir had received cattle from a restricted geographical area or from the same breed, which could lead to bias in the population of cattle from which oocytes were collected. Alternatively, over a two year span, a range of cattle breeds from many regions of the U.S. and Canada would be expected to be processed through the abattoir. Differences in breed, season of collection, and nutritional status could influence variability in the transcriptome and sncRNA populations of the oocyte, and thus the embryo, which would have been distributed broadly across samples in the project that collected samples over two years. Heat stress has been shown to cause low quality of oocytes, and therefore negatively impact embryo development [9], as well as to impact gene expression in heat-stressed oocytes [10]. Oocyte quality, and therefore the transcriptome and sncRNA populations of the resulting embryos, could be impacted by age, as pre-pubertal heifers and older cattle oocytes contain differences in quality and gene expression [11, 12]. Follicular and oocyte physiology, specifically oocyte gene expression and quality, are also likely impacted by nutritional status, as differences based on under-nutrition, high protein diets more typical of dairy cattle, and supplementation have been shown [13, 14]. Again, it is likely that over a 24-month

period, changes in transcriptome or sncRNA populations due to these environmental factors would not have an outsized impact on one pool over another, as samples were constructed by pooled together several different IVF and scNT sessions from a range of seasons and time frames. In addition to environmental differences, the sex of the embryo can contribute to both changes in the mRNA population [15], as well as differences in miRNA profiles [16]. While the changes in transcription and the sncRNA profile originating from genetic sex would not occur until the embryonic genome activation (EGA), they also would not be present in scNT embryos, which were all genetically the same male. The sexual dimorphism of miRNAs and mRNA may explain some of the variability seen in IVF embryos in comparison to scNT embryos, although sexual dimorphism would not be expected in embryos collected prior to EGA.

Culture conditions are another possible source of differences between the studies detailed in Chapters 2 and 3. The investigations described in Chapter 2 utilized Charles Rosenkrans 2 medium (CR2) in larger volumes (500 μl) to negate the need for oil incubation, while the studies of Chapter 3 utilized smaller volumes (50 μl) of synthetic oviductal fluid (SOFaa) medium under mineral oil. Embryos secrete factors to communicate with the maternal environment, and those factors may influence other embryos when co-cultured as groups *in vitro*. The resulting embryo-driven modification of the *in vitro* culture environment can be either beneficial or negative. Overall, group culture of embryos is beneficial for development, although poor-quality embryos may secrete molecules that negatively impact the development of the other co-cultured embryos [17-20]. Both human and bovine embryos secrete miRNAs into culture media, and specific miRNAs may be used as biomarkers of embryo quality and developmental competence [21, 22]. Thus, populations of sncRNA may also be impacted by group culture of embryos. While embryos were group cultured in both experiments for this dissertation, the embryo-to-media ratio was different, which has been shown to impact development rates [23]. A higher ratio of embryo volume to media volume (as in Chapters 3-4) would lead to greater concentrations of these

secreted factors in the medium, potentially leading to greater impact on embryo development and variability in sncRNA profiles. Additionally, in Chapters 3-4, the embryos were also co-cultured with cumulus cells that had been stripped from the donor oocytes. Cattle cumulus cells exhibit highly variable expression of miRNAs, although these varying patterns of miRNAs do not correlate with quality of the associated oocyte [24]. In addition, cumulus-oocyte-complexes have been shown to express specific populations of miRNA [25]. It is possible that when cumulus cells are co-cultured with embryos, high variability in cumulus cell miRNA expression could cause higher variability in miRNA expression of co-cultured embryos. In addition to sncRNA populations, cumulus cell mRNA transcriptomes vary from oocyte to oocyte, and specific mRNA expression can be correlated with improved oocyte quality [26, 27]. It is possible that with embryo and cumulus cell cross talk, co-culture of early embryos with cumulus cells originating from a large group of oocytes could impact the transcriptome or sncRNA populations of these embryos. The use of cumulus cells collected from each batch of matured oocytes introduces a culture variable that was not defined, and there may have been significant differences between each culture batch that contributed to some of the observed variation in sncRNA. Though the nuclear donor bibroblast cells were derived from the same primary cell culture line, cultured in the same cell medium under the same culture conditions, and consistently collected at passages 5-7, these cells showed higher variability than expected with respect to their sncRNA profiles. However, this pattern was not apparent for the fibroblast transcriptomes. The source of this sncRNA variability in the fibroblast samples needs to be established. Another open question is whether cell-to-cell variability of the nuclear donor cell impacts cloning rates, and what functional role such variability may play in the fibroblasts [28]. The lack of variability in the majority of other sample types suggests that the variability noted in the fibroblast samples was not technical noise, and the very tight clustering and similarities between the transcriptome data in sample types suggests that these samples were not abnormal. However, sncRNA may be more

variable then mRNA, and perhaps mechanisms mediating expression of these molecules are more sensitive in some cells types to differing culture conditions.

A follow-up project is in progress that utilizes the supplementation of scNT embryos with miRNA mimics to determine whether artificial induction of specific miRNAs that were determined to be deficient in scNT morula stage embryos will improve developmental competence and repress predicted target transcripts. MiRNA mimics are synthetic miRNAs that contain the same sequences as our miRNAs of interest, and these mimics should functionally act to degrade the same mRNA molecules as endogenously-produced miRNAs. We will quantify mRNA and protein levels of predicted targets of miR-34a (SATB2) and miR-345 (KLF5) to assess the function of these miRNA in early development. Observations from this follow-up study will provide important information as to whether these selected miRNAs have a functional role in early development and confirm whether they indeed bind to and degrade the predicted target transcripts.

While functional examination of differentially expressed miRNAs is of great interest for future study, the work presented in this dissertation provided novel techniques and findings that enhanced the field of somatic cell nuclear transfer. These findings provided a better understanding of the dynamics of sncRNA populations in IVF and scNT bovine embryos through the MET and highlighted some miRNAs that may be used to increase developmental competence of these embryos. The first publication from our group demonstrated that cattle miRNAs and other sncRNAs are dynamically regulated with large-scale population changes occurring through the MET [29]. Other researchers, who quantified only a fraction of the total miRNA population, also found dynamic changes in miRNA molecules in pre-implantation cattle embryos, including miRNA increases at the 8-call stage [30, 31]. While numerous studies have compared differences in specific miRNAs in scNT and IVF cattle embryos at more advanced stages of embryonic or fetal development [32, 33], those studies did not employ a global approach and, therefore, likely

missed many miRNAs of potential biological importance. Those earlier studies also failed to assess other populations of sncRNA and did not address dynamic changes in sncRNA through the MET. Most importantly, no other study has isolated sncRNA and mRNA from the same samples in order to link alterations in sncRNAs, specifically miRNAs, to differential transcript abundance. The identification of sncRNA populations, including tRNA fragments, piRNAs, and miRNAs, that were differentially expressed through development or that were aberrant in scNT bovine embryos provides valuable information for understanding early development. The results of this study produced good quality data, showing that in-depth global analysis of sncRNA populations and mRNA populations are possible from a much smaller pool of samples then has been used previously in cattle. The large amount of data produced herein should provide ample opportunity to pursue further assessment of sncRNA dynamics in early embryos as more tools become available to assess tRNA fragments and piRNA function in early embryos.

While the main focus of the project was to identify differences in sncRNA expression between scNT and IVF embryos, few miRNAs were found to differ significantly between IVF and scNT bovine embryos, and a high amount of sncRNA expression variability was seen within biological replicates. My dissertation research assessed average expression levels of sncRNA, and high variability prevented a clear understanding of sncRNA changes between IVF and scNT embryos and function. Gene expression is both signal and noise, and the idea that investigating variability can shed light on regulatory control has been gaining traction [34]. Examining biological variability itself can have high value in identifying genes that may demonstrate distinctive patterns of variable expression, which could clarify potential functions and regulation [35]. In support of this strategy, specific differences in protein variability have been found to be strongly correlated to the mode of regulation and function of that protein [28]. New technologies that allow quantification of single cell gene expression have been invaluable to identify specific genes that may direct cell fate transitions [36, 37]. Researchers can now identify genes that have

low expression variability, which may provide a marker for cell fate determination, as well as genes with high variability, which may provide valuable information about the environmental regulation of that gene [38-40]. A knowledge gap in developmental biology is the understanding of patterns in gene expression variability and how variable expression might contribute to function. Utilizing single cell technologies to examine variability in both transcripts and sncRNA could improve understanding of how sncRNAs may be acting as regulators within the parameters of an individual transcriptome in future studies. Studying variability can help identify key regulators and offer valuable insights into the regulations of the transcriptional programs that drive early development. Results could guide possible follow up studies that connect regulatory pathways in early development. The transcriptional silence of the early embryo provides a unique opportunity to examine how non-coding RNA mechanisms work together to achieve developmental competence. While the data presented in this dissertation do not point to widespread differences in sncRNA, they does provide an opportunity to better understand regulation and variable expression patterns of sncRNA that may be critical in early development.

In addition to the further assessment of biological variability in sncRNA and mRNA populations and how they may help us understand function in early development, further work is needed to understand the role of technological noise and variability in our results as well. The tools and protocols for sncRNA sequencing are much newer than for traditional mRNA sequencing, and thus, these approaches have be subject to less technological development, including optimization for applications using low inputs. In addition, library preparation of sncRNA cannot include an amplification step prior to barcoding because these RNA molecules are short and nucleotide differences would induce amplification bias. The inability to include an amplification step necessarily limits the number of reads that can be sequenced when using a small amount of starting material [41], as for early developing embryos. While the sncRNA library was sequenced on the Ion Torrent using newer library preparation technologies, the

mRNA library was prepared using a kit specifically designed for low biological input and sequenced using the newer technology of the Illumina Miseq. Differences in data quality produced via these two technologies have been shown [42], which may have impacted our ability to directly compare miRNA to predicted target mRNA in the most accurate way possible.

In addition, studies that aim to elucidate the complex miRNA-mRNA regulatory relationship in early embryonic development would be incredibly valuable. Further work is also needed to understand the timing of miRNA-mediated degradation of target mRNAs. Each of our stages assessed represents a snapshot in development, and whether or not the timing of this snapshot was accurate to capture miRNA-mediated mRNA degradation is unknown. Destabilization of mRNA explains 66% to more than 90% of miRNA-mediated repression regardless of the specific miRNA, the cell type, the cell growth conditions, or the translational state of the mRNA [43]. Due to the possibility of delayed degradation of target mRNAs, additional studies that utilize additional stages of development, such as the 4-cell stage, 16-cell stage, compact morula, and expanded blastocyst stage embryos, could be very useful for more closely examining the dynamic relationship between miRNA and mRNA. Because of these unknowns, the lack of negative correlation between miRNA and mRNA targets may be a reflection of the early embryo's specialized mRNA degradation pathways, and it is possible that functional binding, repression, and degradation are in fact occurring. Further functional follow up is needed to better understand the dynamics of miRNA binding and mRNA degradation in the context of the early embryo.

Luciferase binding assays could be used to identify the mRNA binding targets of specific miRNAs in the context of the developing embryo. However, luciferase binding assays would only demonstrate that the miRNA has the potential to bind the mRNA 3' UTR and would not reveal whether RNA binding proteins are blocking binding and degradation of targets [44, 45]. Once mRNA targets are identified via luciferase binding assay, the use of a miRNA mimic to

assess actual binding in an early embryo context would be appropriate. If the target transcript and associated protein of the miRNA mimic are reduced with mimic supplementation, it would be important to identify whether or not the same results are found in both IVF and scNT embryos. In any cases where the miRNA mimic does not successfully reduce transcripts or protein abundance, follow up work that utilizes an RNA-protein crosslink, RNA pull-down, or shotgun proteomics approach to isolate and assess any proteins that may be bound to the mRNA may help identify why miRNAs and their targets can both be present at high levels in the embryo and whether this is the result of RNA binding proteins [46]. In addition, continued work to identify biologically relevant targets outside of the bioinformatically predicted targets would be beneficial, as individual miRNAs can regulate hundreds of genes using a complex degradation code [47, 48].

In conclusion, this research provided valuable new information on changes in the sncRNA population and mRNA transcriptome in scNT embryos that may be contributing to the low rates of reproductive success. Although few individual miRNAs were found to be statistically different between scNT and IVF embryos and there was limited negative correlation in the relative abundance of these miRNAs and their putative mRNA targets, some interesting differences in the pattern of miRNA expression between stages of early developmental were found. To our knowledge, this research project is the first to utilize a global, unbiased approach to quantify sncRNAs and mRNAs from the same samples and to compare sncRNA expression between scNT and IVF embryos through the MET. Limitations in our studies included different culture conditions used in the two sncRNA sequencing investigations and the lack of functional testing of miRNAs. Future studies should focus on identifying patterns of variability within populations of sncRNAs to identify functionally significant sncRNAs of possible function, to identify which target mRNAs are bound by miRNAs in early development, and to explore the roles of RNA binding proteins in miRNA function in this opening period of development.

References

1.   Zhang S, Chen X, Wang F, An X, Tang B, Zhang X, Sun L, Li Z. Aberrant DNA methylation reprogramming in bovine SCNT preimplantation embryos. Sci Rep. 2016;6:30345.

2.   Simonsson S, Gurdon J. DNA demethylation is necessary for the epigenetic reprogramming of somatic cell nuclei. Nat Cell Biol. 2004;6(10):984-990.

3.   Liu X, Wang Y, Gao Y, Su J, Zhang J, Xing X, Zhou C, Yao K, An Q, Zhang Y. H3K9 demethylase KDM4E is an epigenetic regulator for bovine embryonic development and a defective factor for nuclear reprogramming. Development. 2018;145(4).

4.   Santos F, Zakhartchenko V, Stojkovic M, Peters A, Jenuwein T, Wolf E, Reik W, Dean W. Epigenetic marking correlates with developmental potential in cloned bovine preimplantation embryos. Curr Biol. 2003;13(13):1116-1121.

5.   Matoba S, Zhang Y. Somatic cell nuclear transfer reprogramming: mechanisms and applications. Cell Stem Cell. 2018;23(4):471-485.

6.   Alonso CR. A complex 'mRNA degradation code' controls gene expression during animal development. Trends Genet. 2012;28(2):78-88.

7.   Rivera RM, Ross JW. Epigenetics in fertilization and preimplantation embryo development. Prog Biophys Mol Biol. 2013;113(3):423-432.

8.   Driver AM, Penagaricano F, Huang W, Ahmad KR, Hackbart KS, Wiltbank MC, Khatib H. RNA-Seq analysis uncovers transcriptomic variations between morphologically similar in vivo- and in vitro-derived bovine blastocysts. BMC Genomics. 2012;13:118.

9.   Wolfenson D, Roth Z, Meidan R. Impaired reproduction in heat-stressed cattle: basic and applied aspects. Anim Reprod Sci. 2000;60-61:535-547.

10.   Argov N, Moallem U, Sklan D. Summer heat stress alters the mRNA expression of selective-uptake and endocytotic receptors in bovine ovarian cells. Theriogenology. 2005;64(7):1475-1489.

11.   Takeo S, Kawahara-Miki R, Goto H, Cao F, Kimura K, Monji Y, Kuwayama T, Iwata H. Age-associated changes in gene expression and developmental competence of bovine oocytes, and a possible countermeasure against age-associated events. Mol Reprod Dev. 2013;80(7):508-521.

12.   Krisher RL. Maternal age affects oocyte developmental potential at both ends of the age spectrum. Rerpro Dev Fert. 2018;31(1):1-9.

13.   Moussa M, Shu J, Zhang XH, Zeng F. Maternal control of oocyte quality in cattle "a review". Anim Reprod Sci. 2015;155:11-27.

14. Penagaricano F, Souza AH, Carvalho PD, Driver AM, Gambra R, Kropp J, Hackbart KS, Luchini D, Shaver RD, Wiltbank MC. Effect of maternal methionine supplementation on the transcriptome of bovine preimplantation embryos. PLoS One. 2013;8(8):e72302.

15. Lowe R, Gemma C, Rakyan VK, Holland ML. Sexually dimorphic gene expression emerges with embryonic genome activation and is dynamic throughout development. BMC Genomics. 2015;16:295.

16. Gross N, Kropp J, Khatib H. Sexual dimorphism of miRNAs secreted by bovine in vitro-produced embryos. Front Genet. 2017;8:39.

17. Tao T, Robichaud A, Mercier J, Ouellette R. Influence of group embryo culture strategies on the blastocyst development and pregnancy outcome. J Assist Reprod Genet. 2013;30(1):63-68.

18. Paria BC, Dey SK. Preimplantation embryo development in vitro: cooperative interactions among embryos and role of growth factors. Proc Natl Acad Sci USA. 1990;87(12):4756-4760.

19. Fujita T, Umeki H, Shimura H, Kugumiya K, Shiga K. Effect of group culture and embryo-culture conditioned medium on development of bovine embryos. J Reprod Dev. 2006;52(1):137-142.

20. Stokes PJ, Abeydeera LR, Leese HJ. Development of porcine embryos in vivo and in vitro; evidence for embryo 'cross talk' in vitro. Dev Biol. 2005;284(1):62-71.

21. Kropp J, Salih SM, Khatib H. Expression of microRNAs in bovine and human pre-implantation embryo culture media. Front Genet. 2014;5:91.

22. Kropp J, Khatib H. Characterization of microRNA in bovine in vitro culture media associated with embryo quality and development. J Dairy Sci. 2015;98(9):6552-6563.

23. Ebner T, Shebl O, Moser M, Mayer RB, Arzt W, Tews G. Group culture of human zygotes is superior to individual culture in terms of blastulation, implantation and life birth. Reprod Biomed Online. 2010;21(6):762-768.

24. Uhde K, van Tol HTA, Stout TAE, Roelen BAJ. MicroRNA expression in bovine cumulus cells in relation to oocyte quality. Noncoding RNA. 2017;3(1).

25. Assou S, Al-edani T, Haouzi D, Philippe N, Lecellier CH, Piquemal D, Commes T, Ait-Ahmed O, Dechaud H, Hamamah S. MicroRNAs: new candidates for the regulation of the human cumulus-oocyte complex. Hum Reprod. 2013;28(11):3038-3049.

26. Hamel M, Dufort I, Robert C, Gravel C, Leveille MC, Leader A, Sirard MA. Identification of differentially expressed markers in human follicular cells associated with competent oocytes. Hum Reprod. 2008;23(5):1118-1127.

27. Huang Z, Wells D. The human oocyte and cumulus cells relationship: new insights from the cumulus cell transcriptome. Mol Hum Reprod. 2010;16(10):715-725.

28.   Pelkmans L. Cell Biology. Using cell-to-cell variability--a new era in molecular biology. Science. 2012;336(6080):425-426.

29.   Cuthbert JM, Russell SJ, White KL, Benninghoff AD. The maternal-to-zygotic transition in bovine in vitro-fertilized embryos is associated with marked changes in small non-coding RNAs. Biol Reprod. 2019;100(2):331-350.

30.   Mondou E, Dufort I, Gohin M, Fournier E, Sirard MA. Analysis of microRNAs and their precursors in bovine early embryonic development. Mol Hum Reprod. 2012;18(9):425-434.

31.   Berg DK, Pfeffer PL. MicroRNA expression in bovine preimplantation embryos. Reprod Fertil Dev. 2018;30(3):546-554.

32.   Hossain MM, Tesfaye D, Salilew-Wondim D, Held E, Proll MJ, Rings F, Kirfel G, Looft C, Tholen E, Uddin J. Massive deregulation of miRNAs from nuclear reprogramming errors during trophoblast differentiation for placentogenesis in cloned pregnancy. BMC Genomics. 2014;15:43.

33.   Castro FO, Sharbati S, Rodriguez-Alvarez LL, Cox JF, Hultschig C, Einspanier R. MicroRNA expression profiling of elongated cloned and in vitro-fertilized bovine embryos. Theriogenology. 2010;73(1):71-85.

34.   Eldar A, Elowitz MB. Functional roles for noise in genetic circuits. Nature. 2010;467(7312):167-173.

35.   Padovan-Merhar O, Raj A. Using variability in gene expression as a tool for studying gene regulation. Wiley Interdiscip Rev Syst Biol Med. 2013;5(6):751-759.

36.   Kalmar T, Lim C, Hayward P, Munoz-Descalzo S, Nichols J, Garcia-Ojalvo J, Martinez Arias A. Regulated fluctuations in nanog expression mediate cell fate decisions in embryonic stem cells. PLoS Biol. 2009;7(7):e1000149.

37.   Maamar H, Raj A, Dubnau D. Noise in gene expression determines cell fate in Bacillus subtilis. Science. 2007;317(5837):526-529.

38.   Barkai N, Shilo BZ. Variability and robustness in biomolecular systems. Mol Cell. 2007;28(5):755-760.

39.   Wolf L, Silander OK, Nimwegen E. Expression noise facilitates the evolution of gene regulation. Elife. 2015;4.

40.   Moignard V, Gottgens B. Transcriptional mechanisms of cell fate decisions revealed by single cell expression profiling. Bioessays. 2014;36(4):419-426.

41.   Eminaga S, Christodoulou DC, Vigneault F, Church GM, Seidman JG. Quantification of microRNA expression with next-generation sequencing. Curr Protoc Mol Biol. 2013;Chapter 4:Unit 4 17.

42.   Quail MA, Smith M, Coupland P, Otto TD, Harris SR, Connor TR, Bertoni A, Swerdlow HP, Gu Y. A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. BMC Genomics. 2012;13:341.

43.   Eichhorn SW, Guo H, McGeary SE, Rodriguez-Mias RA, Shin C, Baek D, Hsu SH, Ghoshal K, Villen J, Bartel DP. mRNA destabilization is the dominant effect of mammalian microRNAs by the time substantial repression ensues. Mol Cell. 2014;56(1):104-115.

44.   Jin Y, Chen Z, Liu X, Zhou X. Evaluating the microRNA targeting sites by luciferase reporter gene assay. Methods Mol Biol. 2013;936:117-127.

45.   Campos-Melo D, Droppelmann CA, Volkening K, Strong MJ. Comprehensive luciferase-based reporter gene assay reveals previously masked up-regulatory effects of miRNAs. Int J Mol Sci. 2014;15(9):15592-15602.

46.   Theil K, Imami K, Rajewsky N. Identification of proteins and miRNAs that specifically bind an mRNA in vivo. Nat Commun. 2019;10(1):4205.

47.   Wolter JM, Kotagama K, Pierre-Bez AC, Firago M, Mangone M. 3'LIFE: a functional assay to detect miRNA targets in high-throughput. Nucleic Acids Res. 2014;42(17):e132.

48.   Cloonan N. Re-thinking miRNA-mRNA interactions: intertwining issues confound target discovery. Bioessays. 2015;37(4):379-388.

APPENDICES

Appendix A

March 19th, 2020

Jocelyn Cuthbert
4815 Old Main Hill, Logan UT 84321

To the Authors:

I am in the process of preparing my dissertation in the Animal, Dairy, and Veterinary Sciences department at Utah State University. I hope to complete my degree program in April, 2020.

The article "The Maternal-To-Zygotic Transition in Bovine in Vitro-Fertilized Embryos Is Associated With Marked Changes in Small Non-Coding RNAs", of which I am first author, appeared in the journal Biology of Reproduction in February, 2019 on pages 331-350 and reports an essential part of my research. As authors on this publication, I would like your permission to reprint it as a chapter in my dissertation, which may require some revision. Please note that USU sends every dissertation to USU Digital Commons to be made available for reproduction.

I will include acknowledgment to the article on the first page of the chapter, as shown below. Copyright and permission information will be included in a special appendix.
Please let me know if you would like a different acknowledgement.

Please indicate your approval of this request by signing in the space provided and attach any other form necessary to confirm permission. If you have any questions, please contact me at the phone number or email below. Thank you for your assistance.

Jocelyn Cuthbert
303-921-6748
jcuthber@aggiemail.usu.edu

I hereby give permission to Jocelyn Cuthbert to reprint the requested article in her dissertation, with the following acknowledgement:

Signed: _____(Dr. Abby Benninghoff)
Date: _____3/24/20_____

March 19th, 2020

Jocelyn Cuthbert
4815 Old Main Hill, Logan UT 84321

To the Authors:

I am in the process of preparing my dissertation in the Animal, Dairy, and Veterinary Sciences department at Utah State University. I hope to complete my degree program in April, 2020.

The article "The Maternal-To-Zygotic Transition in Bovine in Vitro-Fertilized Embryos Is Associated With Marked Changes in Small Non-Coding RNAs", of which I am first author, appeared in the journal Biology of Reproduction in February, 2019 on pages 331-350 and reports an essential part of my research. As authors on this publication, I would like your permission to reprint it as a chapter in my dissertation, which may require some revision. Please note that USU sends every dissertation to USU Digital Commons to be made available for reproduction.

I will include acknowledgment to the article on the first page of the chapter, as shown below. Copyright and permission information will be included in a special appendix.
Please let me know if you would like a different acknowledgement.

Please indicate your approval of this request by signing in the space provided and attach any other form necessary to confirm permission. If you have any questions, please contact me at the phone number or email below. Thank you for your assistance.

Jocelyn Cuthbert
303-921-6748
jcuthber@aggiemail.usu.edu

I hereby give permission to Jocelyn Cuthbert to reprint the requested article in her dissertation, with the following acknowledgement:

_____   Date: 24 March 2020
(Dr. Stewart Russell)

March 24th, 2020

Jocelyn Cuthbert
4815 Old Main Hill, Logan UT 84321

To the Authors:

I am in the process of preparing my dissertation in the Animal, Dairy, and Veterinary Sciences department at Utah State University. I hope to complete my degree program in April, 2020.

The article "The Maternal-To-Zygotic Transition in Bovine in Vitro-Fertilized Embryos Is Associated With Marked Changes in Small Non-Coding RNAs", of which I am first author, appeared in the journal Biology of Reproduction in February, 2019 on pages 331-350 and reports an essential part of my research. As authors on this publication, I would like your permission to reprint it as a chapter in my dissertation, which may require some revision. Please note that USU sends every dissertation to USU Digital Commons to be made available for reproduction.

I will include acknowledgment to the article on the first page of the chapter, as shown below. Copyright and permission information will be included in a special appendix.
Please let me know if you would like a different acknowledgement.

Please indicate your approval of this request by signing in the space provided and attach any other form necessary to confirm permission. If you have any questions, please contact me at the phone number or email below. Thank you for your assistance.

Jocelyn Cuthbert
303-921-6748
jcuthber@aggiemail.usu.edu

I hereby give permission to Jocelyn Cuthbert to reprint the requested article in her dissertation, with the following acknowledgement:

Signed: _____(Dr. Ken White)
Date: _____03/24/2020_____

Cuthbert - Co-author Permission Letter

Final Audit Report   2020-03-24

Created:	2020-03-24
By:	Judy Wilde (judy.wilde@usu.edu)
Status:	Signed
Transaction ID:	CBJCHBCAABAADOn-eHesDT-cTtYtHOqJN6ovbIg00a85

"Cuthbert - Co-author Permission Letter" History

Document created by Judy Wilde (judy.wilde@usu.edu)
2020-03-24 - 8:00:29 PM GMT- IP address: 67.45.32.126

Document emailed to Ken White (ken.white@usu.edu) for signature
2020-03-24 - 8:01:12 PM GMT

Email viewed by Ken White (ken.white@usu.edu)
2020-03-24 - 8:37:47 PM GMT- IP address: 129.123.207.90

Document e-signed by Ken White (ken.white@usu.edu)
Signature Date: 2020-03-24 - 8:37:57 PM GMT - Time Source: server- IP address: 129.123.207.90

Signed document emailed to Jocelyn Cuthbert (jcuthber@aggiemail.usu.edu), Judy Wilde (judy.wilde@usu.edu) and Ken White (ken.white@usu.edu)
2020-03-24 - 8:37:57 PM GMT

Appendix B

Date:	Aug 22, 2018
To:	"Abby Diane Benninghoff" abby.benninghoff@usu.edu
From:	"Biology of Reproduction" biolre.editorialoffice@oup.com
Subject:	Biology of Reproduction: Decision on BIOLRE-2018-0268R1

Re: BIOLRE-2018-0268R1 - The maternal-to-zygotic transition in bovine in vitro-fertilized embryos is associated with marked changes in small non-coding RNAs

Dear Dr. Benninghoff,

Congratulations! We are pleased to accept your manuscript for publication. The PDF version of the article will be published first in Advance Articles, which establishes its date of publication. Subsequently, your paper will be published in traditional form in the online version of Biology of Reproduction and be available in PDF format for traditional print.

If the research reported in this paper was funded entirely or in part by the U.S. National Institutes of Health (NIH) or an agency that requires public archiving and access and accepts deposit in PubMed Central (PMC) or PMC International, BOR will deposit the final, published manuscript in PMC in compliance with the NIH Enhanced Public Access Policy, and you need do nothing further. You will receive your official acceptance date from Oxford University Press once you have signed your license to publish. (N.B. If you are a UK-based author and are looking to comply with the HEFCE policy on open access in the Research Excellence Framework, you should use the official acceptance date when depositing in your repository.)

Also, beginning 12 months after the date of final publication, all articles published in Biology of Reproduction will be covered by a Creative Commons Attribution-NonCommercial 4.0 International license (CC BY-NC). Authors whose funding agency requires access earlier than the 12-month embargo must choose an open-access option. Upon final publication, Open Access articles will be publicly available on the journal site and, if required, deposited to PMC/I with a 6 month embargo and published with CC BY-NC license in accordance with funding agency policy. Articles may be shared or adapted--provided that the original source is credited and changes are indicated. Under the CC BY-NC license, articles and material therein cannot be used for commercial purposes.

COVER IMAGES: If you have an image that you would like to be considered as a potential cover image for the issue your article will appear in, please submit it to biolrepro@oup.com.

If you would like to provide a graphical abstract, please email it to the Editors-in-Chief.

Thank you for submitting this excellent paper to Biology of Reproduction.

Best regards,

Wei Yan and Hugh Clarke
Editors-in-Chief
Biology of Reproduction

_____
If you would like your personal information to be removed from the database, please contact the publication office.

Creative Commons License Deed

Attribution-NonCommercial 4.0 International (CC BY-NC 4.0)

This is a human-readable summary of (and not a substitute for) the license.

You are free to:

Share — copy and redistribute the material in any medium or format

Adapt — remix, transform, and build upon the material

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

NonCommercial — You may not use the material for commercial purposes.

No additional restrictions — You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

Notices:

You do not have to comply with the license for elements of the material in the public domain or where your use is permitted by an applicable exception or limitation.

No warranties are given. The license may not give you all of the permissions necessary for your intended use. For example, other rights such as publicity, privacy, or moral rights may limit how you use the material.

License information obtained at https://creativecommons.org/licenses/by-nc/4.0/, access date April 23, 2020.

Creative Commons Legal Code

Attribution-NonCommercial 4.0 International

Official translations of this license are available in other languages.

Creative Commons Corporation ("Creative Commons") is not a law firm and does not provide legal services or legal advice. Distribution of Creative Commons public licenses does not create a lawyer-client or other relationship. Creative Commons makes its licenses and related information available on an "as-is" basis. Creative Commons gives no warranties regarding its licenses, any material licensed under their terms and conditions, or any related information. Creative Commons disclaims all liability for damages resulting from their use to the fullest extent possible.

Using Creative Commons Public Licenses

Creative Commons public licenses provide a standard set of terms and conditions that creators and other rights holders may use to share original works of authorship and other material subject to copyright and certain other rights specified in the public license below. The following considerations are for informational purposes only, are not exhaustive, and do not form part of our licenses.

Considerations for licensors: Our public licenses are intended for use by those authorized to give the public permission to use material in ways otherwise restricted by copyright and certain other rights. Our licenses are irrevocable. Licensors should read and understand the terms and conditions of the license they choose before applying it. Licensors should also secure all rights necessary before applying our licenses so that the public can reuse the material as expected. Licensors should clearly mark any material not subject to the license. This includes other CC-licensed material, or material used under an exception or limitation to copyright.

Considerations for the public: By using one of our public licenses, a licensor grants the public permission to use the licensed material under specified terms and conditions. If the licensor's permission is not necessary for any reason–for example, because of any applicable exception or limitation to copyright–then that use is not regulated by the license. Our licenses grant only permissions under copyright and certain other rights that a licensor has authority to grant. Use of the licensed material may still be restricted for other reasons, including because others have copyright or

other rights in the material. A licensor may make special requests, such as asking that all changes be marked or described. Although not required by our licenses, you are encouraged to respect those requests where reasonable.

Creative Commons Attribution-NonCommercial 4.0 International Public License

By exercising the Licensed Rights (defined below), You accept and agree to be bound by the terms and conditions of this Creative Commons Attribution-NonCommercial 4.0 International Public License ("Public License"). To the extent this Public License may be interpreted as a contract, You are granted the Licensed Rights in consideration of Your acceptance of these terms and conditions, and the Licensor grants You such rights in consideration of benefits the Licensor receives from making the Licensed Material available under these terms and conditions.

Section 1 – Definitions.

a. **Adapted Material** means material subject to Copyright and Similar Rights that is derived from or based upon the Licensed Material and in which the Licensed Material is translated, altered, arranged, transformed, or otherwise modified in a manner requiring permission under the Copyright and Similar Rights held by the Licensor. For purposes of this Public License, where the Licensed Material is a musical work, performance, or sound recording, Adapted Material is always produced where the Licensed Material is synched in timed relation with a moving image.

b. **Adapter's License** means the license You apply to Your Copyright and Similar Rights in Your contributions to Adapted Material in accordance with the terms and conditions of this Public License.

c. **Copyright and Similar Rights** means copyright and/or similar rights closely related to copyright including, without limitation, performance, broadcast, sound recording, and Sui Generis Database Rights, without regard to how the rights are labeled or categorized. For purposes of this Public License, the rights specified in Section 2(b)(1)-(2) are not Copyright and Similar Rights.

d. **Effective Technological Measures** means those measures that, in the absence of proper authority, may not be circumvented under laws fulfilling obligations under Article 11 of the WIPO Copyright Treaty adopted on December 20, 1996, and/or similar international agreements.

e. **Exceptions and Limitations** means fair use, fair dealing, and/or any other exception or limitation to Copyright and Similar Rights that applies to Your use of the Licensed Material.

f. **Licensed Material** means the artistic or literary work, database, or other material to which the Licensor applied this Public License.

g. **Licensed Rights** means the rights granted to You subject to the terms and conditions of this Public License, which are limited to all Copyright and Similar Rights that apply to Your use of the Licensed Material and that the Licensor has authority to license.

h. **Licensor** means the individual(s) or entity(ies) granting rights under this Public License.

i. **NonCommercial** means not primarily intended for or directed towards commercial advantage or monetary compensation. For purposes of this Public License, the exchange of the Licensed Material for other material subject to Copyright and Similar Rights by digital file-sharing or similar means is NonCommercial provided there is no payment of monetary compensation in connection with the exchange.

j. **Share** means to provide material to the public by any means or process that requires permission under the Licensed Rights, such as reproduction, public display, public performance, distribution, dissemination, communication, or importation, and to make material available to the public including in ways that members of the public may access the material from a place and at a time individually chosen by them.

k. **Sui Generis Database Rights** means rights other than copyright resulting from Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases, as amended and/or succeeded, as well as other essentially equivalent rights anywhere in the world.

l. **You** means the individual or entity exercising the Licensed Rights under this Public License. **Your** has a corresponding meaning.

Section 2 – Scope.

a. **License grant**.

1. Subject to the terms and conditions of this Public License, the Licensor hereby grants You a worldwide, royalty-free, non-sublicensable, non-exclusive, irrevocable license to exercise the Licensed Rights in the Licensed Material to:

   A. reproduce and Share the Licensed Material, in whole or in part, for NonCommercial purposes only; and

   B. produce, reproduce, and Share Adapted Material for NonCommercial purposes only.

2. Exceptions and Limitations. For the avoidance of doubt, where Exceptions and Limitations apply to Your use, this Public License does not apply, and You do not need to comply with its terms and conditions.

3. Term. The term of this Public License is specified in Section 6(a).

4. Media and formats; technical modifications allowed. The Licensor authorizes You to exercise the Licensed Rights in all media and formats whether now known or hereafter created, and to make technical modifications necessary to do so. The Licensor waives and/or agrees not to assert any right or authority to forbid You from making technical modifications necessary to exercise the Licensed Rights, including technical modifications necessary to circumvent Effective Technological Measures. For purposes of this Public License, simply making modifications authorized by this Section 2(a)(4) never produces Adapted Material.

5. <u>Downstream recipients</u>.
   A. <u>Offer from the Licensor – Licensed Material</u>. Every recipient of the Licensed Material automatically receives an offer from the Licensor to exercise the Licensed Rights under the terms and conditions of this Public License.
   B. <u>No downstream restrictions</u>. You may not offer or impose any additional or different terms or conditions on, or apply any Effective Technological Measures to, the Licensed Material if doing so restricts exercise of the Licensed Rights by any recipient of the Licensed Material.

6. <u>No endorsement</u>. Nothing in this Public License constitutes or may be construed as permission to assert or imply that You are, or that Your use of the Licensed Material is, connected with, or sponsored, endorsed, or granted official status by, the Licensor or others designated to receive attribution as provided in Section 3(a)(1)(A)(i).

b. **Other rights**.

1. Moral rights, such as the right of integrity, are not licensed under this Public License, nor are publicity, privacy, and/or other similar personality rights; however, to the extent possible, the Licensor waives and/or agrees not to assert any such rights held by the Licensor to the limited extent necessary to allow You to exercise the Licensed Rights, but not otherwise.
2. Patent and trademark rights are not licensed under this Public License.
3. To the extent possible, the Licensor waives any right to collect royalties from You for the exercise of the Licensed Rights, whether directly or through a collecting society under any voluntary or waivable statutory or compulsory licensing scheme. In all other cases the Licensor expressly reserves any right to collect such royalties, including when the Licensed Material is used other than for NonCommercial purposes.

Section 3 – License Conditions.

Your exercise of the Licensed Rights is expressly made subject to the following conditions.

a. **Attribution**.

1. If You Share the Licensed Material (including in modified form), You must:

   A. retain the following if it is supplied by the Licensor with the Licensed Material:
      i. identification of the creator(s) of the Licensed Material and any others designated to receive attribution, in any reasonable manner requested by the Licensor (including by pseudonym if designated);

        ii. a copyright notice;

       iii. a notice that refers to this Public License;

       iv. a notice that refers to the disclaimer of warranties;

       v. a URI or hyperlink to the Licensed Material to the extent reasonably practicable;

    B. indicate if You modified the Licensed Material and retain an indication of any previous modifications; and

    C. indicate the Licensed Material is licensed under this Public License, and include the text of, or the URI or hyperlink to, this Public License.

2. You may satisfy the conditions in Section 3(a)(1) in any reasonable manner based on the medium, means, and context in which You Share the Licensed Material. For example, it may be reasonable to satisfy the conditions by providing a URI or hyperlink to a resource that includes the required information.

3. If requested by the Licensor, You must remove any of the information required by Section 3(a)(1)(A) to the extent reasonably practicable.

4. If You Share Adapted Material You produce, the Adapter's License You apply must not prevent recipients of the Adapted Material from complying with this Public License.

Section 4 – Sui Generis Database Rights.

Where the Licensed Rights include Sui Generis Database Rights that apply to Your use of the Licensed Material:

a. for the avoidance of doubt, Section 2(a)(1) grants You the right to extract, reuse, reproduce, and Share all or a substantial portion of the contents of the database for NonCommercial purposes only;

b. if You include all or a substantial portion of the database contents in a database in which You have Sui Generis Database Rights, then the database in which You have Sui Generis Database Rights (but not its individual contents) is Adapted Material; and

c. You must comply with the conditions in Section 3(a) if You Share all or a substantial portion of the contents of the database.

For the avoidance of doubt, this Section 4 supplements and does not replace Your obligations under this Public License where the Licensed Rights include other Copyright and Similar Rights.

Section 5 – Disclaimer of Warranties and Limitation of Liability.

a. **Unless otherwise separately undertaken by the Licensor, to the extent possible, the Licensor offers the Licensed Material as-is and as-available, and makes no representations**

or warranties of any kind concerning the Licensed Material, whether express, implied, statutory, or other. This includes, without limitation, warranties of title, merchantability, fitness for a particular purpose, non-infringement, absence of latent or other defects, accuracy, or the presence or absence of errors, whether or not known or discoverable. Where disclaimers of warranties are not allowed in full or in part, this disclaimer may not apply to You.

b. **To the extent possible, in no event will the Licensor be liable to You on any legal theory (including, without limitation, negligence) or otherwise for any direct, special, indirect, incidental, consequential, punitive, exemplary, or other losses, costs, expenses, or damages arising out of this Public License or use of the Licensed Material, even if the Licensor has been advised of the possibility of such losses, costs, expenses, or damages. Where a limitation of liability is not allowed in full or in part, this limitation may not apply to You.**

c. The disclaimer of warranties and limitation of liability provided above shall be interpreted in a manner that, to the extent possible, most closely approximates an absolute disclaimer and waiver of all liability.

Section 6 – Term and Termination.

a. This Public License applies for the term of the Copyright and Similar Rights licensed here. However, if You fail to comply with this Public License, then Your rights under this Public License terminate automatically.

b. Where Your right to use the Licensed Material has terminated under Section 6(a), it reinstates:

   1. automatically as of the date the violation is cured, provided it is cured within 30 days of Your discovery of the violation; or
   2. upon express reinstatement by the Licensor.

   For the avoidance of doubt, this Section 6(b) does not affect any right the Licensor may have to seek remedies for Your violations of this Public License.

c. For the avoidance of doubt, the Licensor may also offer the Licensed Material under separate terms or conditions or stop distributing the Licensed Material at any time; however, doing so will not terminate this Public License.

d. Sections 1, 5, 6, 7, and 8 survive termination of this Public License.

Section 7 – Other Terms and Conditions.

a. The Licensor shall not be bound by any additional or different terms or conditions communicated by You unless expressly agreed.

b. Any arrangements, understandings, or agreements regarding the Licensed Material not stated herein are separate from and independent of the terms and conditions of this Public License.

Section 8 – Interpretation.

a. For the avoidance of doubt, this Public License does not, and shall not be interpreted to, reduce, limit, restrict, or impose conditions on any use of the Licensed Material that could lawfully be made without permission under this Public License.

b. To the extent possible, if any provision of this Public License is deemed unenforceable, it shall be automatically reformed to the minimum extent necessary to make it enforceable. If the provision cannot be reformed, it shall be severed from this Public License without affecting the enforceability of the remaining terms and conditions.

c. No term or condition of this Public License will be waived and no failure to comply consented to unless expressly agreed to by the Licensor.

d. Nothing in this Public License constitutes or may be interpreted as a limitation upon, or waiver of, any privileges and immunities that apply to the Licensor or You, including from the legal processes of any jurisdiction or authority.

Creative Commons is not a party to its public licenses. Notwithstanding, Creative Commons may elect to apply one of its public licenses to material it publishes and in those instances will be considered the "Licensor." The text of the Creative Commons public licenses is dedicated to the public domain under the CC0 Public Domain Dedication. Except for the limited purpose of indicating that material is shared under a Creative Commons public license or as otherwise permitted by the Creative Commons policies published at creativecommons.org/policies, Creative Commons does not authorize the use of the trademark "Creative Commons" or any other trademark or logo of Creative Commons without its prior written consent including, without limitation, in connection with any unauthorized modifications to any of its public licenses or any other arrangements, understandings, or agreements concerning use of licensed material. For the avoidance of doubt, this paragraph does not form part of the public licenses.

Creative Commons may be contacted at creativecommons.org.

Appendix C

Supplementary File 2.1 Original and Processed miRNA Expression Data

Microsoft Office Excel document with original and processed data for miRNA

expression in oocytes, 8-cell stage embryos, and blastocyst stage embryos. Spreadsheets include

unmodified output data from miRDeep2, processed expression data (log $(y \pm 1)$ transformation,

calculation of ratio with respect to average oocyte RPM, and $\log_2$ transformation), and results of

ANOVA and pairwise statistical tests with FDR-adjusted P values. File accessible online at

https://doi.org/10.1093/biolre/ioy190.

Appendix D

Supplementary File 2.2 Predicted mRNA Targets
for Differentially Expressed miRNAs.

Microsoft Office Excel document with lists of target mRNAs for over- or under

expressed miRNAs in 8-cell stage embryos compared to mature oocytes predicted using

TargetScan. This document also includes results of Metascape gene ontology, KEGG pathway

and Reactome pathway enrichment analyses, as well as ontology analysis using the DAVID

EASE platform. File accessible online at https://doi.org/10.1093/biolre/ioy190.

Appendix E

Supplementary File 2.3 Differentially Expressed
sncRNAs and mapping of pilRNAs to transposable elements

Microsoft Office Excel document data for sncRNA expression in oocytes, 8-cell stage embryos, and blastocyst stage embryos. Summary tables for sequence reads for snoRNA, snRNA, pilRNA, mitochondrial rRNA, genomic rRNA, mitochondrial tRNA, genomic tRNA, and miscellaneous sncRNA. Mapping of pilRNAs to TEs and analysis of sequence reads for snoRNAs. File accessible online at https://doi.org/10.1093/biolre/ioy190.

Appendix F

**Supplementary File 3.1. Read Lengths and
Annotation Summaries of sncRNA**

Microsoft Excel document detailing the read lengths and annotation results for small

RNA sequencing. This file includes multiple spreadsheets, as follows: "Read lengths" provides

the number of reads for each sample at 17 to 92 nt in length; "Annotation summary 17-32 nt"

provides the number of reads per sample for each classification category for those reads 17-32 nt

in length; and "Annotation summary 32+ nt" provides the number of reads per sample for each

classification category for those reads >32 nt in length.

For each spreadsheet, samples are named by the following convention: for embryo type,

in vitro fertilized (IVF) or somatic cell nuclear transfer (NT); and for developmental stage,

fibroblasts (Fb), oocytes (Oo), 2-cell stage embryo (2c), 8-cell stage embryo (8c), morula stage

embryo (Mo), blastocyst stage embryo (Bl), blastocyst derived cells (BSCs). Numbers indicate

the sample replicate.

This file is accessible online at https://doi.org/10.26078/8jjx-1v90.

Appendix G

Supplementary File 3.2. Normalized and Variance-Stabilized Values for Annotated sncRNAs

Microsoft Excel document detailing normalized and variance-stabilized reads generated by RNA sequencing of small non-coding RNA from cattle embryos. This file includes two spreadsheets, the first providing the normalized reads per sequence per sample and the second providing the variance-stabilized reads per sequence per sample. The first column (starting row 4) for each spreadsheet provides the identifier(s) of the mapped sequence, and the remaining columns provide values for each sample ID.

Supplementary file 2 row definitions.

ID	Assigned sample ID
Sample type	Indicates whether the sample was IVF or scNT embryo, or fibroblast or oocyte samples
Developmental stage	Indicates the developmental stage at which the sample was acquired, as applicable

Abbreviations are: for embryo type, *in vitro* fertilized (IVF) or somatic cell nuclear transfer (NT); and for developmental stage, fibroblasts (Fb), oocytes (Oo), 2-cell stage embryo (2c), 8-cell stage embryo (8c), morula stage embryo (Mo), blastocyst stage embryo (Bl), blastocyst derived cells (BSCs).

This file is accessible online at https://doi.org/10.26078/8jjx-1v90.

Appendix H

Supplementary File 3.3. Results of DESeq2 Analyses
for miRNAs, tRNA Fragments and piRNAs

Microsoft Excel document with results of DESeq2 differential expression analysis

comparing scNT vs IVF embryos at each developmental stage and sequential comparisons of

developmental stage for both IVF and scNT embryos. This file includes three spreadsheets

providing results of DESeq2 analyses for miRNAs, tRNA fragments and piRNAs. For each

spreadsheet, the pairwise comparisons are noted with the two naming conventions. First, for

sequential comparisons by stage: "embryo type: sample A vs. sample B" (e.g., IVF: Oo vs 2c).

Second, for comparisons of scNT vs IVF embryos at a specific stage: "scNT vs IVF: stage" (e.g.,

scNT vs IVF: Mo). Abbreviations are: for embryo type, *in vitro* fertilized (IVF) or somatic cell

nuclear transfer (NT); and for developmental stage, fibroblasts (Fb), oocytes (Oo), 2-cell stage

embryo (2c), 8-cell stage embryo (8c), morula stage embryo (Mo), blastocyst stage embryo (Bl),

blastocyst derived cells (BSCs).

For each comparison, results of the DESeq2 analyses are provided as outlined below:

Supplementary file 3 column definitions.

ID	Official sncRNA name, for sncRNA type (miRNA, tRNA, piRNA) indicated by spreadsheet tab
baseMean	The mean of normalized counts of all samples, normalizing for sequencing depth
Log2FoldChange	Average $\log_2$ fold change calculated for comparison indicated in the top row name
lfcSE	$\log_2$ standard error of the calculated average $\log_2$ fold change
stat	Wald statistic
P-value	Wald test unadjusted P-value
Padj	False discovery rate adjusted P-value using the Benjamini-Hochberg method

For comparisons that did not yield any significant differences, no results are shown.

This file is accessible online at https://doi.org/10.26078/8jjx-1v90.

Appendix I

Supplementary File 3.4. TargetScan Prediction of mRNA
Targets for Differentially Expressed miRNAs

Microsoft Excel document providing results of TargetScan prediction of mRNA targets

for differentially expressed miRNAs. This file includes multiple spreadsheets, each providing the

results of the TargetScan prediction of mRNA targets for differentially expressed miRNAs. The

spreadsheets are labeled according to the specific comparison made. For comparisons of scNT vs

IVF embryos at a specific stage: "scNT vs IVF at stage" (e.g., scNT vs IVF at Mo). For

sequential comparisons by stage: "stage A vs stage B for embryo type" (e.g., Oo vs 2c for IVF).

Abbreviations are: for embryo type, *in vitro* fertilized (IVF) or somatic cell nuclear

transfer (NT); and for developmental stage, fibroblasts (Fb), oocytes (Oo), 2-cell stage embryo

(2c), 8-cell stage embryo (8c), morula stage embryo (Mo), blastocyst stage embryo (Bl),

blastocyst derived cells (BSCs).

Supplementary file 4 column definitions.

Ortholog of target gene	Gene symbol of predicted target gene
Representative transcript	Ensembl ID for representative transcript of predicted target gene
Gene name	Name of predicted target gene
Representative miRNA	ID of the miRNA molecule querried
Cumulative weighted context++ score	Score that quantifies the predicted efficacy of the miRNA binding sites
Total context++ score	Score that is predictive for all types of interactions, including those of miRNAs that are not highly conserved
miRNA family	The ID for the family of miRNA molecules to which this specific miRNA belongs

This file is accessible online at https://doi.org/10.26078/8jjx-1v90.

Appendix J

**Supplementary File 3.5. Results of Metascape Ontology Analyses
for Predicted mRNA targets of differentially Expressed miRNAs**

Microsoft Excel document providing results of Metascape ontology analysis for

biological processes associated with differentially expressed miRNAs. This file includes multiple

spreadsheets, each providing the results of the Metascape ontology enrichment analysis. The

spreadsheets are labeled according to the specific comparison made. For comparisons of scNT vs

IVF embryos at a specific stage: "scNT vs IVF at stage" (e.g., scNT vs IVF at Mo). For

sequential comparisons by stage: "stage A vs stage B for embryo type" (e.g., Oo vs 2c for IVF).

Abbreviations are: for embryo type, *in vitro* fertilized (IVF) or somatic cell nuclear

transfer (NT); and for developmental stage, fibroblasts (Fb), oocytes (Oo), 2-cell stage embryo

(2c), 8-cell stage embryo (8c), morula stage embryo (Mo), blastocyst stage embryo (Bl),

blastocyst derived cells (BSCs).

Supplementary file 5 column definitions.

GroupID	Indicates whether the term is a member of the cluster of terms, or the term that serves as the overall label (summary)
Category	Specifies the type of ontology term (biological process, cellular compartment, or molecular function)
Term	Accession number for the ontology term
Description	Descriptive name of the ontology term
LogP	Log_{10} p value
Log(q-value)	Log_{10} of the false discovery rate-corrected p-value (the q-value)
InTerm_InList	Number of genes within the query set that are annotated by that term/number of genes in the complete list for that term
Genes	Gene ID accession numbers for all genes that were annotated with that term
Symbols	Gene symbols for all genes that were annotated with that term

This file is accessible online at https://doi.org/10.26078/8jjx-1v90.

Appendix K

Supplementary File 3.6. Mapping of piRNAs to Transposable Elements

Microsoft Excel document providing results of ProTrac mapping of putative piRNAs to transposable elements in the bovine genome. This file includes two spreadsheets providing results of mapping of piRNAs to transposable elements in the bovine genome in either the forward, "TE mapping (+)", or reverse, "TE mapping (-)" directions.

Abbreviations are: for embryo type, *in vitro* fertilized (IVF) or somatic cell nuclear transfer (NT); and for developmental stage, fibroblasts (Fb), oocytes (Oo), 2-cell stage embryo (2c), 8-cell stage embryo (8c), morula stage embryo (Mo), blastocyst stage embryo (Bl), blastocyst derived cells (BSCs). Numbers indicate the sample replicate.

This file is accessible online at https://doi.org/10.26078/8jjx-1v90.

Appendix L

Supplementary File 4.1. Normalized Reads for Annotated mRNA Transcripts

Microsoft Excel document providing normalized sequencing reads per sample. This file provides the normalized reads per sample for transcripts >200 nt in length that mapped to the bovine genome. The first column provides the Ensemble transcript accession.

Abbreviations are: for embryo type, *in vitro* fertilized (IVF) or somatic cell nuclear transfer (NT); and for developmental stage, fibroblasts (Fb), oocytes (Oo), 2-cell stage embryo (2c), 8-cell stage embryo (8c), morula stage embryo (Mo), blastocyst stage embryo (Bl), blastocyst derived cells (BSCs). Numbers indicate the sample replicate.

This file is accessible online at https://doi.org/10.26078/svr1-5r53.

Appendix M

Supplementary File 4.2. Results of DESeq2 Analyses for Annotated mRNAs

Microsoft Excel document with results of DESeq2 differential expression analysis comparing scNT vs IVF embryos at each developmental stage and sequential comparisons of developmental stage for both IVF and scNT embryos. This file includes three spreadsheets providing results of DESeq2 analyses for mRNAs. For each spreadsheet, the pairwise comparisons are noted with the two naming conventions. First, for sequential comparisons by stage: "embryo type: sample A vs. sample B" (e.g., IVF: Oo vs 2c). Second, for comparisons of scNT vs IVF embryos at a specific stage: "scNT vs IVF: stage" (e.g., scNT vs IVF: Mo).

Abbreviations are: for embryo type, *in vitro* fertilized (IVF) or somatic cell nuclear transfer (NT); and for developmental stage, fibroblasts (Fb), oocytes (Oo), 2-cell stage embryo (2c), 8-cell stage embryo (8c), morula stage embryo (Mo), blastocyst stage embryo (Bl), blastocyst derived cells (BSCs).

For each comparison, results of the DESeq2 analyses are provided as outlined below:

Supplementary file 2 column definitions.

ID	Ensemble transcript accession ID
baseMean	The mean of normalized counts of all samples, normalizing for sequencing depth
Log2FoldChange	Average $\log_2$ fold change calculated for comparison indicated in the top row name
lfcSE	$\log_2$ standard error of the calculated average $\log_2$ fold change
stat	Wald statistic
P-value	Wald test unadjusted P-value
Padj	False discovery rate adjusted P-value using the Benjamini-Hochberg method

This file is accessible online at https://doi.org/10.26078/svr1-5r53.

Appendix N

**Supplementary File 4.3. Results of Metascape Ontology
Analyses for Differentially Expressed mRNAs**

Microsoft Excel document with results of Metascape ontology analyses for biological

processes associated with predicted mRNA targets of differentially expressed miRNAs. This file

includes multiple spreadsheets, each providing the results of the Metascape ontology enrichment

analysis. The spreadsheets are labeled according to the specific comparison made. For

comparisons of scNT vs IVF embryos at a specific stage: "scNT vs IVF at stage" (e.g., scNT vs

IVF at Mo). For sequential comparisons by stage: "stage A vs stage B for embryo type" (e.g., Oo

vs 2c for IVF).

Abbreviations are: for embryo type, *in vitro* fertilized (IVF) or somatic cell nuclear

transfer (NT); and for developmental stage, fibroblasts (Fb), oocytes (Oo), 2-cell stage embryo

(2c), 8-cell stage embryo (8c), morula stage embryo (Mo), blastocyst stage embryo (Bl),

blastocyst derived cells (BSCs).

Supplementary file 5 column definitions.

GroupID	Indicates whether the term is a member of the cluster of terms, or the term that serves as the overall label (summary)
Category	Specifies the type of ontology term (biological process, cellular compartment, or molecular function)
Term	Accession number for the ontology term
Description	Descriptive name of the ontology term
LogP	Log_{10} p value
Log(q-value)	Log_{10} of the false discovery rate-corrected p-value (the q-value)
InTerm_InList	Number of genes within the query set that are annotated by that term/number of genes in the complete list for that term
Genes	Gene ID accession numbers for all genes that were annotated with that term
Symbols	Gene symbols for all genes that were annotated with that term

This file is accessible online at https://doi.org/10.26078/svr1-5r53.

Appendix O

Supplementary File 4.4. Cytoscape Networks for Differentially Expressed miRNAs and Their Predicted mRNA Targets

PDF document depicting Cytoscape networks for differentially expressed miRNAs and their predicted targets. Networks depicted within this PDF document were created using data for differentially miRNAs and their predicted mRNA targets (TargetScan total context ++ score <−0.35) using Cytoscape with a force-directed layout. Only miRNAs within families annotated by TargetScan as conserved or broadly conserved were included in the network analysis. MiRNAs are represented by squares and target mRNAs as circles and both are shaded according to their log2 fold change for the comparison indicated. Edges between miR and mRNA nodes indicate that the connected mRNA is a predicted target of the miRNA in *Bos taurus*.

Abbreviations are: for embryo type, *in vitro* fertilized (IVF) or somatic cell nuclear transfer (NT); and for developmental stage, fibroblasts (Fb), oocytes (Oo), 2-cell stage embryo (2c), 8-cell stage embryo (8c), morula stage embryo (Mo), blastocyst stage embryo (Bl), blastocyst derived cells (BSCs).

This file is accessible online at https://doi.org/10.26078/svr1-5r53.

VITA

Jocelyn Cuthbert was born in Pleasanton California, and grew up in the foothills of Colorado. She attended Colorado State University for her undergraduate training and earned dual Bachelor of Science degrees in Equine Science and Agricultural Business. During her undergraduate career, Ms. Cuthbert also engaged in scientific research focusing on epigenetics in equine animal models and genetics research in plants. Ms. Cuthbert then matriculated to the graduate doctoral training program in the Department of Animal, Dairy and Veterinary Sciences at Utah State University under the mentorship of Major Advisors Dr. Abby Benninghoff and Dr. Kenneth White. Her doctoral research focused on reproduction, development and epigenetics. Ms. Cuthbert was the recipient of two prestigious fellowships in support of her doctoral training, including the USU Presidential Doctoral Research Fellowship and a U.S. Department of Agriculture predoctoral fellowship. Ms. Cuthbert was the first graduate student at USU earn both the Empowering Teaching Excellence Teaching Scholar Certificate and the Master of Teaching Certificate. While pursuing her own training, she also collaborated with faculty to propose and help create a campus wide graduate student track for instructional training.