Retrotransposon Activity and DNA Methylation Control in Bovine Preimplantation Embryos

Retrotransposon Activiteit en DNA-methylatie Controle in Bovine Preimplantatie Embryo's

Wenwen Li

Thesis submitted in fulfillment of the requirements for the degree of Doctor (PhD) in Veterinary Medicine, Faculty of Veterinary Medicine
Ghent University
2016

Promoter

Prof. Dr. Luc Peelman

Co-promoter

Prof. Dr. Ann van Soom

Department of Nutrition, Genetics and Ethology

Faculty of Veterinary Medicine

Ghent University

Members of the Jury

Prof. Dr. Geert Opsomer (Chairman)
Faculty of Veterinary Medicine
Ghent University

Prof. Dr. Peter Bols
Department of Veterinary Science
University of Antwerp

Prof. Dr. Filip Van Nieuwerburgh
Faculty of Pharmaceutical Sciences
Ghent University

Prof. Dr. Ward De Spiegelaere
Faculty of Veterinary Medicine
Ghent University

Dr. Karen Goossens
Institute for Agricultural and Fisheries Research (ILVO)

Dr. Bart Leemans (Secretary)
Faculty of Veterinary Medicine
Ghent University

List of tables

List of figures

List of abbreviations

5-caC	5-carboxylcytosine
5-fC	5-formylcytosine
5-hmC	5-hydroxymethylcytosine
5-mC	5-methylcytosine
AGO	Argonaute
AI	Artificial insemination
APOBEC	Apolipoprotein B mRNA editing enzyme
AUB	Aubergine
BER	Base excision repair machinery
BLAST	Basic Local Alignment Search Tool
bp	Base pair
BSA	Bovine serum albumin
BSA-ITS	Serum-free supplementation
BSP	Bisulfite sequencing PCR
BWS	Beckwith-Wiedemann syndrome
C	Cytosine
CC	Cumulus cell
cDNA	Complementary DNA
CGI	CpG island
CHM	Choroideremia
COBRA	Combined bisulfite restriction analysis
COCs	Cumulus oocyte complexes
CpG	Cytosine-phosphorous-guanine
DAPI	4’,6-diamidino-2-phenylindole
DMD	Duchenne muscular dystrophy
DNA	Deoxyribonucleic acid
DNMTs	DNA methyltransferases
dpi	Days post insemination

DTT	Dithiothreitol
EDMA	EmbryoGENE DNA Methylation Analysis
EGA	Embryonic genome activation
EGF	Epidermal growth factor
EGFP	Enhanced green fluorescent protein
EN	Endonuclease
ERVs	Endogenous retroviruses
EthD-2	Ethidium Homodimer-2
FBS	Fetal bovine serum
FITC	Fluorescein isothiocyanate
FSH	Follicle-stimulating hormone
gDNA	Genomic deoxyribonucleic acid
HCl	Hydrogen chloride
HELP	High-performance liquid chromatography
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
hpi	Hours post insemination
ICM	Inner cell mass
ICSI	Intracytoplasmic sperm injection
ITS	Insulin-transferrin-selenium
IVC	*In vitro* embryo culture
IVF	*In vitro* fertilization
IVM	*In vitro* oocyte maturation
IVP	*In vitro* embryo production
L1/LINE1	Long interspersed nuclear element 1
LINEs	Long interspersed nuclear elements
lncRNAs	Long non-coding RNAs
LOS	Large offspring syndrome
LTR	Long terminal repeat
MBD-seq	Methylated DNA binding domain protein sequencing

MeDIP-seq	Methylated DNA immunoprecipitation sequencing
MII	Metaphase of the second meiotic division
MIQE	Minimum information for publication of quantitative real-time PCR experiments
miRNA	microRNA
MOET	Multiple ovulation and embryo transfer
NCBI	National Center for Biotechnology Information
ncRNA	Non-coding RNA
nt	Nucleotide
NF	Normalization factor
NGS	Next-generation sequencing
ORF	Open reading frame
oxBS-seq	Oxidative bisulfite sequencing
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PFA	Paraformaldehyde
PI	Propidium Iodide
piRNA	Piwi-interacting RNA
PN	Pronucleus
Pol II	RNA polymerase II
Pol III	RNA polymerase III
qPCR	Quantitative real-time polymerase chain reaction
redBS-seq	Reduced representation bisulfite sequencing
RIFTs	Retrotransposon-initiated fusion transcripts
RISC	RNA-induced silencing complex
RNA	Ribonucleic acid
RNAi	RNA interference
ROS	Reactive oxygen species
RRBS	Reduced representation bisulfite sequencing

rRNA	Ribosomal RNA
RT	Reverse transcriptase
RT-PCR	Reverse transcription polymerase chain reaction
RT-qPCR	Reverse transcription quantitative real-time polymerase chain reaction
SBS	Shotgun bisulfite sequencing
SCNT	Somatic cell nuclear transfer
SINEs	Short interspersed nuclear elements
siRNA	Short interfering RNA
SOF	Synthetic oviduct fluid
ssDNA	Single-stranded DNA
TA	annealing temperature
TALP	Tyrode’s albumin lactate pyruvate
TCM-199	Tissue culture medium 199
TE	Trophectoderm
TEs	Transposable elements
TETs	Ten-eleven translocation enzymes
TPRT	Target-primed reverse transcription
TSD	Target site duplication
TSG	Tumor suppressor gene
UTR	Untranslated region
XCI	X chromosome inactivation
XIC	X-inactivation center
ZP	Zona pellucida

Chapter 1

General introduction

1.1 Introduction of retrotransposons

Transposable elements (TEs), the largest class of mammalian genome sequences, are repetitive DNA sequences, that can move directly or by replication from one locus to another within the genome when they are activated. They were first discovered in maize by Barbara McClintock in the 1940s, were long considered as 'junk' or 'selfish' DNA and dismissed as uninteresting. In 1988, a report of insertions of L1 into exon 14 of the factor VIII gene in two unrelated patients with haemophilia A indicated the activity of TEs (Kazazian *et al.* 1988). Since then it has been found that most eukaryotic genomes allow the expression and accommodate transposition of a few transposon families (Beato *et al.* 1996). Transposons turned out to be important players in genome evolution and gene regulation (Gifford *et al.* 2013). 66% - 69% of the human genome was suggested to be repetitive or repeat-derived (de Koning *et al.*, 2011). This implies that repetitive DNA may have played a larger part in (human) evolution than was previously assumed. The potential relationship between these elements and their hosts should not been underestimated. Recent data suggest that in healthy mammalian tissue TEs are mostly active in undifferentiated cells like germ line cells and early stage embryos (Peaston *et al.* 2004; van den Hurk *et al.* 2007). Therefore, this review will be focused on the activities and functions of TEs during early development.

1.1.1 Classification of retrotransposons

A unified classification system of TEs was proposed by Wicker *et al.* (2007). In this system TEs are classified by class, subclass, order, superfamily, family and subfamily (**Figure 1.1**). Transposons are broadly grouped in two main classes according to their mechanism of transposition, which can be described as either *copy and paste* (class I or retrotransposons) or *cut and paste* (class II or DNA transposons) (Finnegan 1989). Class I retrotransposons, also called RNA transposons, copy themselves to RNA and then back to DNA by reverse transcriptase. Subsequently the DNA copies may integrate in to the genome. This RNA intermediate *copy and paste* mechanism can rapidly increase the copy number of the elements

and therefore increase genome size. Retrotransposons are estimated to comprise about 30 to 40% of the mammalian genome (Varmus 1988; Adelson *et al.* 2009), and functionally they are the most important TEs in the genome. Class II DNA transposons do not involve an RNA intermediate, do not enlarge the genome and the transposition is catalyzed by transposase enzymes. DNA transposons are estimated to occupy very little of the genomes: 4% in human, 2% in cow and only 1% in mouse and rat (**Table 1.1**). They are remnants or fossils of ancient elements, and unlikely to remain transpositionally active (Lander *et al.* 2001). They were not studied in this work and will not be described further in detail.

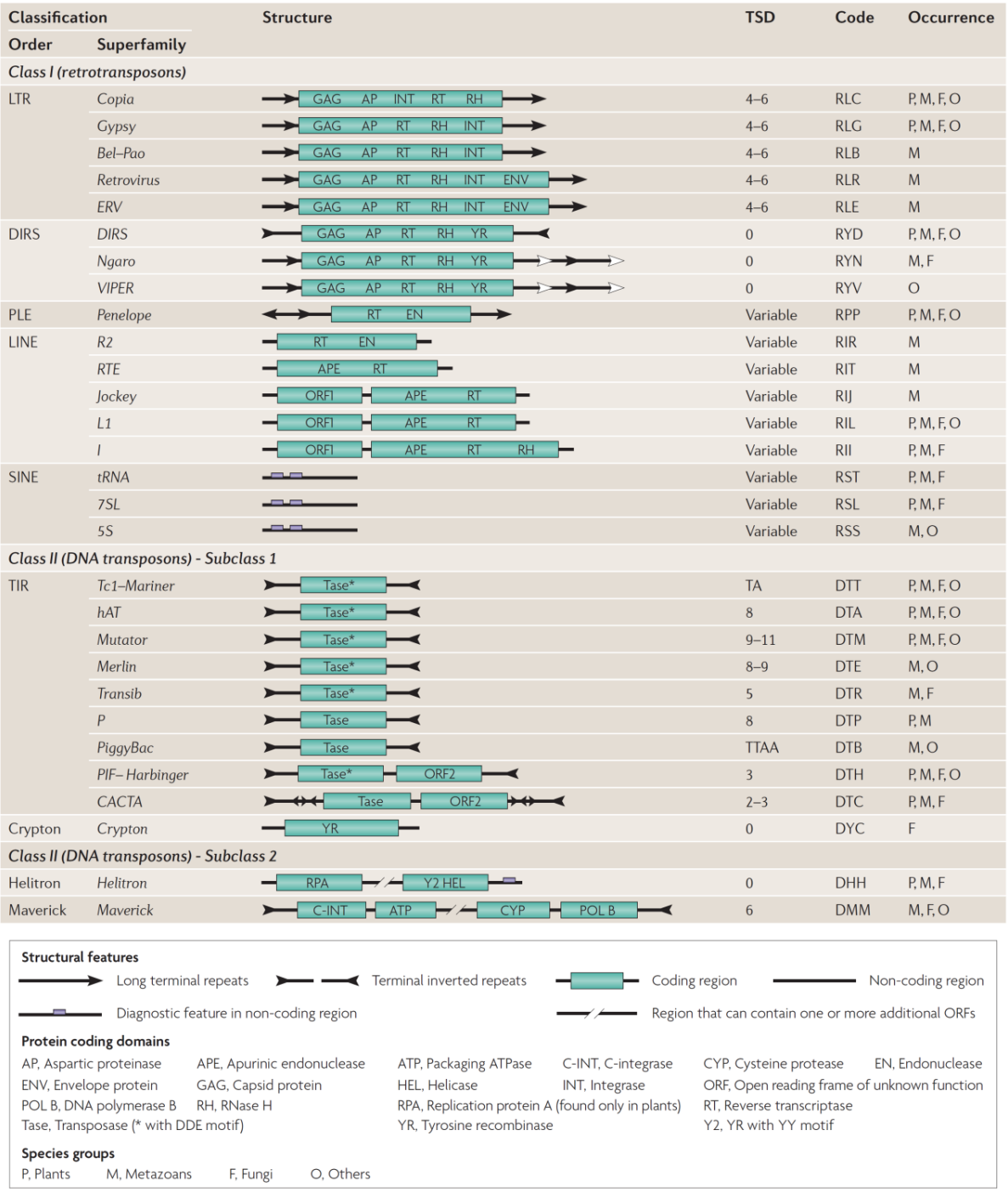


Figure 1.1 Classification of transposable elements hierarchized in classes, subclasses, orders and superfamilies. TSD: target site duplication. Figure from Wicker *et al.* (2007).

Both Class I and Class II TEs can be classified as either 'autonomous transposons' or 'non-autonomous transposons' by their self-sufficiency. Autonomous TEs have an open reading frame (ORF) which can encode transcriptase or reverse transcriptase, so they have the ability to 'jump' in the genome by themselves; while non-autonomous TEs encode defective

polypeptides or cannot encode any, and accordingly require transcriptase or reverse transcriptase from another source (Dewannieux *et al.* 2003; Raiz *et al.* 2012).

Different from older classification systems of retrotransposons based on the presence of a long terminal repeat (LTR) or not (non-LTR), Wicker *et al*. (2007) suggested to divide retrotransposons into five orders based on their mechanistic features, organization and reverse transcriptase phylogeny: LTR retrotransposons, *Dirs*-like elements, *Penelope*-like elements, long interspersed nuclear elements (LINEs) and short interspersed nuclear elements (SINEs). A schematic representation of the TE composition of some species is presented in **Table 1.1**.

LTR retrotransposons

LTR retrotransposons are more abundant in plants compared to animals. They have direct sequence repeats flanking the internal coding region, that range from 100 bp to over 5 kb in size, hence the name long terminal repeats. The size of LTR retrotransposons can range from a few hundred base pairs up to, exceptionally, 25 kb (Ogre) (Neumann *et al.* 2003). All autonomous LTR retrotransposons include ORFs encoding both structural and enzymatic proteins (**Figure 1.1**). The typical ORFs they contain are *gag* - coding for a structural protein for virus-like particles, and *pol* which encodes several enzymes, including aspartic proteinase (AP), reverse transcriptase (RT), RNase H (RH) and DDE integrase (INT) (Wicker *et al.* 2007). The long terminal repeats of LTR retrotransposons do not encode any known proteins, but they contain the promoters and terminators required for transcription.

Table 1.1 Distribution of TEs in the genomes of different species. Calculated by data from *RepeatMasker Genomic Datasets* (http://www.repeatmasker.org/genomicDatasets/RMGenomicDatasets.html). Version of repeat libraries is mentioned after each species.

	Cow (bosTau7)		Pig (susScr3)		Horse (equCab2)		Human (hg38)		Mouse (mm10)		Rat (rn5)	
	Total bp	% genome	Total bp	% genome	Total bp	% genome	Total bp	% genome	Total bp	% genome	Total bp	% genome
LINEs												
L1	373,710,968	13.3	441,128,027	17.5	437,448,496	18	532,540,789	17.5	527,594,044	19.9	478,615,369	18.6
RTE	361,232,407	12.9	3,619,621	0.1	6,510,152	0.3	5,011,001	0.2				
L2	74,482,125	2.7	81,478,182	3.2	127,000,323	5.2	114,260,531	3.7	11,471,810	0.4	10,871,137	0.4
CR1	8,683,417	0.3			13,810,588	0.6	12,390,130	0.4				
SINEs												
tRNA	200,038,987	7.1	287,041,640	11.4	85,229,078	3.5			64,535,745	2.4	77,557,163	3
MIR	64,265,848	2.3	69,142,383	2.7	92,152,925	3.8	87,303,369	2.9	14,704,295	0.6	14,210,833	0.6
Alu							318,958,938	10.5	65,242,138	2.5	44,270,277	1.7
tRNA-Alu									59,746,674	2.3	59,933,252	2.3
RTE	63,577,393	2.3										
LTR retrotranspons												
ERV1	40,529,357	1.4	32,785,059	1.3	46,070,351	1.9	88,857,839	2.9	31,193,184	1.2	26,543,515	1
ERVL	75,560,684	2.7	82,906,424	3.3	122,397,532	5	176,282,644	5.8	156,090,728	5.9	139,962,175	5.4
ERVK	13,088,075	0.5							131,046,475	4.9	97,662,161	3.8
Gypsy					6,831,172	0.3						
DNA transposons												
TcMar	16,264,799	0.6	16,300,429	0.6	20,506,448	0.8	44,227,402	1.5	6,195,274	0.2	5,907,021	0.2
hAT	45,672,733	1.6	48,767,267	1.9	66,773,980	3	66,109,392	2.2	22,679,934	1	22,615,139	1
Other and unknown	11,227,017	0.4	27,755,736	1.1	14,575,909	0.6	24,350,895	0.8	18,662,626	0.7	36,012,143	1.4
All TEs	1,348,333,810	48	1,090,924,768	43	1,039,306,954	43	1,470,292,930	48	1,109,162,927	42	1,014,160,185	39

A common LTR superfamily in the mammalian genome is that of the endogenous retroviruses (ERVs). ERVs supposedly originate from retroviruses that infected the germline of vertebrates and stably integrated their genome in that of the host genomes (Ribet *et al.* 2008). Full ERVs have an *env* sequence (**Figure 1.1**), but most of them do not encode a functional envelope protein anymore and are incapable of horizontal transmission. Although ERVs lost their infectivity, some, such as ERV-K-type family in rats, may still keep their mobility and can significantly multiply in their host (Wang *et al.* 2010).

LINE retrotransposons

LINEs are the biggest mammalian TE component, comprising on average 20% of the mammalian genome (**Table 1.1**). Based on structural features and the phylogeny of the key enzyme - the reverse transcriptase, LINEs are grouped into five main groups, called R2, RTE, *Jockey*, L1 and I. The L1, or LINE-1 element is found in all mammals and has been found active in many species (Warren *et al.* 2008). A typical integral L1 element is approximately 6000 bp, and contains an open reading frame for an RNA binding protein (ORF1) and one coding for an endonuclease and reverse transcriptase (ORF2). The LINE encoded reverse transcriptase recognizes the RNA from which it was translated (Wei *et al.* 2001), and initiates the mobility reactions termed target-primed reverse transcription (TPRT) (Luan *et al.* 1993). This reverse transcription mechanism will be discussed in more detail later in this review.

SINE retrotransposons

SINEs are short DNA sequences, usually 90 to 300 bp in length, and do not have reverse transcriptase protein coding capacity. They are non-autonomous retrotransposons dependent on LINEs for their amplification. The most prominent SINEs are the human *Alu* elements and the rodent B1 elements. The 5’ end of SINEs are derived from tRNA, 5S rRNA, and 7SL RNA with promoter activity for RNA polymerase III (A and B boxes) (Deininger & Batzer 2002). The 3’ end of SINEs was suggested to originate from a corresponding LINE (Ohshima

et al. 1996; Okada & Hamada 1997). However, more recently it was described that the 3' ends of mammalian L1s do not exhibit any similarity to SINEs, except for the presence of 3'-poly(A) repeats. Since the 3'-poly(A) repeats are critical for their retrotransposition, L1s may recognize the poly(A) repeats (Ohshima 2013). This enables the SINE RNA to be reverse transcribed by the partner LINE reverse transcriptase (Ichiyanagi 2013).

Dirs-like elements and Penelope-like elements

Dirs-like elements and *Penelope*-like elements are the least studied retrotransposons, although they were found quite widespread in plants, metazoans, fungi and animals (Goodwin & Poulter 2001; Schostak *et al.* 2008). *Dirs*-like elements differ from other retrotransposons as they lack integrases or aspartic proteases. They encode a tyrosine recombinase, suggesting that they use this enzyme to insert copies into the host genome by recombination (Goodwin & Poulter 2004). The other unusual type, the *Penelope*-like elements, is delimited by two direct repeats flanking a single ORF coding for two protein domains: a reverse transcriptase (RT) similar to telomerase RTs, and an endonuclease (EN) (Cervera & De la Peña 2014). Penelope-like elements were suggested to use the same TPRT mechanism as LINEs, and exhibit a similar evolution pattern as LINEs, originating from deep branching clades dating back to the Precambrian era. However, they have experienced a much higher degree of lineage losses than LINEs (Arkhipova 2006). *Penelope* isolated from *Drosophila virilis* was the first known transpositionally active representative of this class (Evgen'ev *et al.* 1997). Since then similar elements were found in amoebae, fungi, cnidarians, rotifers, flatworms, roundworms, fish, amphibia, and reptilia, but less so in mammals.

1.1.2 Retrotransposons in the bovine genome

The bovine genome, as other mammalian genomes, contains all the main classes of TEs, with the bulk being LINEs and SINEs. As shown in **Table 1.1**, L1s are the dominant retrotransposon type in extant species, such as in cow, pig, horse, human, mouse and rat. With the exception of L1 and LINE RTE in ruminants and marsupials, few retrotransposons

are active in mammalian genomes (Gentles *et al.* 2007). RTE (*BovB*) and L1 (*L1_BT*) comprise 12.9% respectively 13.3% of the bovine genome (data collected from bosTau7 repeat library on *RepeatMasker* Genomic Datasets http://www.repeatmasker.org/species/bosTau.html). There are 1248 full-length *BovB* elements in the bovine genome, but only 9 of them contain a large ORF meeting the domain criteria, and are likely to retrotranspose in cattle. On the other hand, 811 intact *L1_BT* that are able to transpose were identified (Adelson *et al.* 2009).

It is notable that a high number of RTEs in the bovine genome were found compared to other species (**Table 1.1**). RTEs have been found in all ruminantia (mouse deer, deers, giraffes, sheeps, and cows), and are named BovB, after their initial discovery in the bovine genome. However, their relatives, such as Suina (pigs and peccaries), Tylopoda (camels), and Cetacea (whales and dolphins) lack RTE elements. based on the high level of nucleotide identity, and by the phylogenetic relationships, it has been suggested that BovB is introduced into the genome of ancestor to ruminants by horizontal transfer from Squamata 40-50 million years ago (Kordis et al., 1998). After the introduction into the ancestor genomes, new BovB-related SINEs were generated due to rearrangement between older SINEs and BovBs. Based on the proportion of BovB in bovine genome, it was suggested an average increase in genome size by BovB retrotransposition of 0.4% per million years (Nilsson *et al.* 2012).

It was suggested that every species/clade has one or more unique kind of SINE, contributing heavily to species-specific genome sequences (Jurka *et al.* 2007). In the bovine genome, ruminant-specific SINES, such as *BovA* (*Bov-A2, Bov-tA*) and ART2A constitute about 14% of the genomic sequences (Adelson *et al.* 2009). These SINEs are transposed by the *BovB* encoded machinery (Ohshima & Okada 2005). On the other hand, LTR retrotransposons (mainly ERVs) constitute only a very low percentage of the bovine genome as compared to other species listed in **Table 1.1**.

1.1.3 Mechanisms of retrotransposition

Although the mechanism of retrotransposition is not completely understood, a general understanding can be deduced from the mechanisms of retrovirus integration (Varmus *et al.* 1989). L1 is the mostly studied retrotransposon on mechanistic aspects (Ostertag & Kazazian Jr 2001; Ostertag *et al.* 2002). All retrotransposons transpose through the formation of an RNA transcript that is then reverse transcribed and inserted into a new location in the genome. It was found that both LTR elements and LINEs use at least two enzymes: reverse transcriptase, which is a RNA-dependent DNA polymerase used to generate complementary DNA from an RNA template, and an integrase, which is an endonuclease that cleaves at the site of integration to generate a staggered break. These two enzymes are encoded by genes in autonomous retrotransposons, that are not only utilized by the carriers but can also be used by non-autonomous retrotransposons.

In the following paragraphs, a summary of the best known mechanisms will be presented, with indication of the similarities and differences among different types of retrotransposons, to illustrate how these mechanisms have influenced the colonized genomes. The life cycles of both LINE and LTR retrotransposons are presented as reference model in **Figure 1.2**.

Transcription

The insights in the mechanism of LTR retrotransposition are first extensively studied from work on yeast retrotransposons (Ty elements), but it is generally assumed that the mechanism is very similar among LTR retrotransposons from divergent hosts (Havecker *et al.* 2004). LTR retrotransposon RNA is transcribed by RNA polymerase II from a promoter located within the 5’ LTR sequence, and initiates downstream from the promoter region in the middle of the LTR. In contrast, LINEs have a strong, internal sense-stranded promoter harbored within the 5’ UTR for RNA polymerase II (Alexandrova et al. 2012). A less strong anti-sense

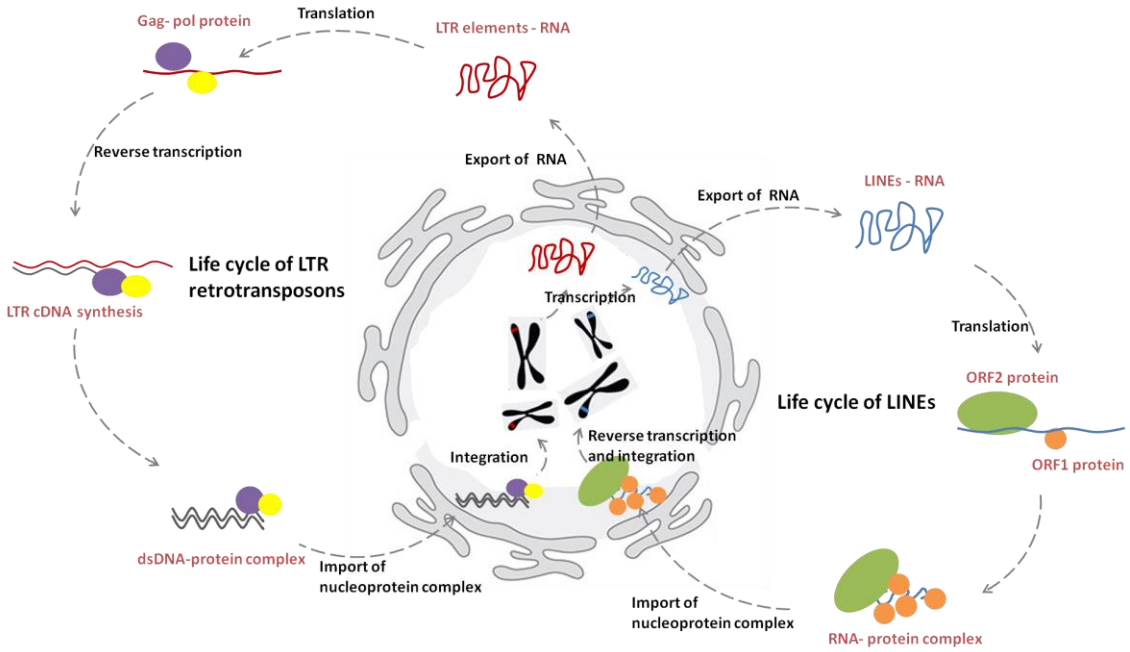


Figure 1.2 Life cycle of LINE and LTR retrotransposons, including transcription, translation, reverse transcription and integration.

promoter was found in human and mouse, initiating expression of anti-sense L1 retrotransposon-initiated fusion transcripts (RIFTs) and modifying the expression of numerous neighboring genes (Mätlik *et al.* 2006). An *in vitro* transcription study revealed that transcription from the human L1 promoter is highly sensitive to tagetitoxin - an inhibitor of RNA polymerase III - indicating that the human L1 promoter is also polymerase III dependent (Kurose *et al.* 1995). Therefore, it is possible that L1 promoters use a hybrid transcription system that takes advantage of factors from both the RNA polymerase II and polymerase III transcription apparatus (Deininger & Batzer 2002).

Transportation

Retrotransposon nucleotides export and entry in the nucleus are important steps in the retrotransposition mechanism. RNAs transcribed by polymerase II are typically modified by

addition of a 7-methylguanosine cap and a poly(A) tail, cleavage and splicing of introns, and then form an export-competent mRNP to be exported to the cytoplasm through nuclear pores (Tollervey & Caceres 2000). Different pathways have been reported for retrotransposons. It was first shown in retroviral replication studies that intron-containing RNA is actively transported from the nucleus to the cytoplasm. Human immunodeficiency virus type 1 (HIV-1) encodes a protein called REV that binds to the *cis*-acting Rev-response element (RRE) present in unspliced viral RNA, thereby mediating their nucleocytoplasmic export (Cullen 1998). LTR elements, like human endogenous retroviruses (HERVs) were found to contain an ORF coding for REV homologs with the same function in RNA export (Magin *et al.* 1999). Also, IAP, a mouse LTR retrotransposon was reported to contain a conserved element termed RTE (RNA transport element) with the same function as RRE, promoting RNA export from nuclei (Nappi *et al.* 2001). Thus, those elements were suggested to be essential for the nucleocytoplasmic transport of LTR retrotransposon full-length RNA (Zolotukhin *et al.* 2008). On the other hand, L1s do not contain any introns, and do not face the problem of exporting unspliced RNA. It was suggested that the mechanism of L1 RNA transport may contain *cis*-elements similar to those of intronless mRNA (Huang & Steitz 2001). Interestingly, in a cultured cell assay, assembled L1 transcripts containing an intron are spliced and retrotranspose appropriately (Moran *et al.* 1996).

For further steps of reverse transcription and integration, both retrotransposon nucleotides and the proteins must re-enter the nucleus and get access to genomic DNA. Due to different reverse transcription mechanisms, LTR retrotransposons and LINEs enter the nucleus in different forms of nucleoprotein complexes (**Figure 1.2**). For LTR retrotransposons, reverse transcription of mRNA happens in the cytoplasm, therefore, they entry the nucleus as nucleoprotein complexes composed by DNA and integrase. For LINEs, both reverse transcription and integration take place in the nucleus, and they enter the nucleus as bigger nucleoprotein complexes composed with mRNA and integrase/reverse transcriptase. These two enzymes are encoded from L1 ORF2, and the protein weight can be up to 150 kDa (Ergün *et al.* 2004). It was suggested the nuclear pore allows the passive diffusion of proteins

with a size up to 110 kDa and even larger ones, depending on the protein structure (Wang & Brattain 2007). If the integrases can not enter the nucleus by passive diffusion, then nuclear localization/export signals will be needed for them to diffuse through the nuclear pore. Ostertag and Kazazian suggested that L1 ORF2 may encode a functional nuclear localization signal by itself or bind an additional protein that contains a nuclear localization signal (Ostertag & Kazazian Jr 2001).

Reverse transcription and integration

The priming of reverse transcription and integration in the genome is different between LTR retrotransposons and LINEs. For LTR retrotransposons, the entire process is catalyzed by reverse transcriptase which has both DNA polymerase and RNase H activities. The most common mechanism of LTRs for priming reverse transcription is by the annealing of the 3' end of a specific cellular tRNA to a complementary region called the primer-binding site (PBS) adjacent to the upstream LTR. This tRNA 3' end serves as a primer for the reverse transcriptase that copies the RNA in a complex series of events into a double-stranded DNA. This process occurs in the cytoplasm, and then the double-stranded DNA molecule is transported to the nucleus and integrated back into the host DNA, adding another copy of the retrotransposon to the genome. This integration process is similar to the integration of DNA transposons, with an element-encoded nuclease making specific nicks in both the element DNA and the integration site to catalyze the integration process (Deininger & Batzer 2002).

Unlike the LTR retrotransposons that transport double-stranded DNA back to the nucleus for integration, LINEs use their RNA in the integration process directly. This mechanism is named target-primed reverse transcription (TPRT). The initial stages of the mechanism have been demonstrated most convincingly with *in vitro* studies of the *Bombys mori* R2 element (Luan *et al.* 1993), and were then extensively studied in the mammalian L1 elements. L1 ORF1 encodes an RNA-binding protein that shows some specificity for binding their RNA (Hohjoh & Singer 1997), and the L1 ORF2 encodes the reverse transcriptase and an

endonuclease that appears to nick the insertion site (Mathias *et al.* 1991; Cost & Boeke 1998). During TPRT, the retrotransposon's endonuclease cleaves one strand of genomic DNA at the target site, producing a 3' OH at the nick, and then the retrotransposon RNA inserts at the nick and the reverse transcriptase uses the free 3' OH to prime the reverse transcription, which is followed by a cleavage of the other DNA strand and integration. Since SINEs and pseudogenes share the same consensus integration site as L1 (Jurka 1997), they are suggested to take advantage of this TPRT mechanism (Boeke 1997).

1.1.4 Expression and mobilization of retrotranposons in mammalian genomes

Transcription

In the mouse, 13% of oocytes' and 7.5% of 2-cell embryos' protein-coding cDNAs contain retrotransposon-derived sequences (Evsikov *et al.* 2004; Peaston *et al.* 2004). These retrotransposon families were suggested to act as alternative promoters and first exons for a subset of host genes, regulating their expression in oocytes and early stage embryos (Peaston *et al.* 2004). The retrotransposon RNA was not only found in early development, but also in many adult tissues by Northern blot (Belancio *et al.* 2010). Full-length L1 RNA was found widespread in human somatic tissues, including ovary, placenta, esophagus, heart muscle and stomach. Translatable spliced transcript (SpORF2), which has the potential to express L1 ORF2, was also detected in lung, thymus, testis and mesenchymal stem cells. Using the same method, L1 RNA was found in rat liver, kidney, and neural cell lines (Witney & Furano 1984). The RNA of more types of retrotransposons was more recently found in different human and mouse tissues using the Cap Analysis Gene Expression (CAGE) technique (Faulkner *et al.* 2009).

Many of the retrotransposon RNAs are not coding for proteins and therefore are classified as long noncoding RNAs (lncRNA). They potentially bind to chromatin-modifying proteins and recruit their catalytic activity to specific sites in the genome to further modulate chromatin states and impact gene expression (Mercer & Mattick 2013). This gene expression regulation

is not limited locally to neighboring genes. These lncRNAs can interact with chromatin at several thousand different locations crossing multiple chromosomes and can modulate large-scale gene expression (Vance & Ponting 2014).

Translation

Using immunohistochemistry and western blot analysis on cell lines and tissues it was found that L1 proteins are present in tumor cell lines, such as human teratocarcinoma, choriocarcinoma, bladder carcinoma, colon carcinoma, melanoma, fibrosarcoma cells and breast carcinomas among others (Leibold *et al.* 1990; Asch *et al.* 1995; Nangia-Makker *et al.* 1998). Besides in the human cancer cell lines, L1 ORF1 protein was detected by immunohistochemistry in mouse germ lines and gonads, including fetal oocytes, leydig cells, prepuberal and adult testes, myoid cells, and cyncytiotrophoblast cells of the placenta (Branciforte & Martin 1994; Trelogan & Martin 1995; Malki *et al.* 2014). The L1 protein expression was also found in neurons (Richardson *et al.* 2014). For LTR retrotransposons, *PEG 10* is a paternally expressed imprinted gene that is thought to have been driven from the Ty3/Gypsy LTR elements. It contains two overlapping ORFs and expresses two proteins: a shorter, *gag*-like protein and a longer, *gag*/*pol* like fusion protein. The protein expression was found in human placenta, testis, and adrenal gland (http://www.proteinatlas.org/ENSG00000242265-PEG10/tissue, Uhlén *el al*. 2015).

Mobilization

A full retrotransposition process includes the final insertion of the repeats. Till today, active retrotransposition in mammalian genomes was found mainly for L1. The insertion activity was first studied using a retrotransposition cassette (Heidmann *et al.* 1988) containing a marker to detect new retrotransposition events. This neomycin phosphotransferase (neo) retrotransposition cassette was used to demonstrate the mobility of mouse IAP elements (Heldmann & Heidmann 1991), pseudogenes (Maestre *et al.* 1995), and mammalian L1

(Moran *et al.* 1996; Naas *et al.* 1998). An enhanced green fluorescent protein (EGFP) gene was later added to the cassette making retrotransposition detection in living cells possible (Ostertag *et al.* 2000). By using this cassette, L1 retrotransposition was revealed both *in vitro* and *in vivo*, mainly in germ cells and early embryos. Human embryonic stem cells express endogenous L1 with retrotransposition, which provided *in vitro* evidence (Garcia-Perez *et al.* 2007). Later *in vivo* evidence was shown in a study of a patient with choroideremia (CHM), in which a full-length L1 insertion was found in the coding region of the CHM gene (van den Hurk *et al.* 2003). Analysis of genomic DNA from the CHM patient's mother indicated L1 retrotransposition can occur very early in human embryonic development and cause choroideremia by insertion in the CHM gene (van den Hurk *et al.* 2007). L1 retrotransposition was also found in male germ cells (Ostertag *et al.* 2002), and L1 and other retrotransposons like HERV-K10 and VL30 were found retrotransposed in human oocytes (Noutsopoulos *et al.* 2006; Georgiou *et al.* 2009). It was suggested that a controlled network of retrotransposon transcripts might serve important roles during gamete development and fertilization, while the uncontrolled ones might lead to genetic disorders (Georgiou *et al*. 2009). Although retrotransposon insertions were reported in germ cells, and a similar level of retrotransposon RNA expression was found in both germ cells and early embryos, a retrotransposon insertion study of transgenic mouse and rat models containing human or mouse L1s showed that L1 insertions predominantly occur during early embryogenesis and less so in germ cells. This was thought to be due to the combined control by DNA methylation and piRNAs in mammalian germ cells (Heras *et al*. 2014).

Retrotransposon insertions were also found in adult tissues/cells, but are often malicious. L1 mobilization was found during both embryonic and adult neurogenesis in rodents and humans, and was suggested to cause extensive somatic mosaicism in the brain (Muotri *et al.* 2005; Coufal *et al.* 2009). Moreover, L1 mediated retrotransposition was found in a variety of tumor types, including colorectal (Miki *et al.* 1992; Solyom *et al.* 2012), lung (Iskow *et al.* 2010), liver (Shukla *et al.* 2013), prostate and ovarian cancers (Lee *et al.* 2012).

Estimating Transposition Frequency

An important question that can be asked considering the potential effects of transposition on the genome is: how often do retrotransposons jump? *Alu* and L1 elements, comprising about 30% of the whole human genome, are the two major active families currently expanding in the human genome (Lander *et al.* 2001; Ostertag & Kazazian Jr 2001; Batzer & Deininger 2002), and thus are the most studied as to the natural retrotransposition rate.

For *Alu's*, two methods for retrotransposition rate estimation were reported (Cordaux *et al.* 2006). The first method utilizes an evolutionary framework, estimating the average *Alu* retrotransposition rate over the past 6 million years. The average rate obtained with this method is one *Alu* insertion in every 18 to 26 individuals over the past 6 million years. The other method estimates the current *Alu* retrotransposition rate by comparing the frequency of the *de novo Alu* insertions involved in genetic diseases in humans to that of the *de novo* nucleotide mutations causing disease in the same set of genes. This method gives an estimation of one new insertion in every 15 individuals (Cordaux *et al.* 2006).

There is more variance in estimation of L1 insertion frequency. By comparing the diploid genome of individual humans with the Human Genome Project reference assembly, it was estimated that the L1 retrotransposition rate is one new insert in 212 births, and one in 21 births and 916 births for *Alu* and SVA respectively (Xing *et al.* 2009). The insertion rate is almost doubled in other groups, by one in 140 and 108 births separately (Ewing & Kazazian 2010; Huang *et al.* 2010). So far, natural retrotransposition rate estimations in other species are not available yet.

1.1.5 Impact of retrotransposon activity on the mammalian genome

1.1.5.1 Formation of new genetic elements

Reverse transcription is one of the key processes that shape eukaryotic genomes. The huge amount of retrotransposons in the genomes is resulting from such events, but it can also produce many new sequence elements in the genome. Here an overview is given of the most important genetic elements formed by the RT activity of retrotransposons.

Formation of processed pseudogenes

Pseudogenes are DNA sequences homologous to a functional gene. They can be divided in two broad groups based on the way they are generated: Duplicate pseudogenes arise by genomic duplication of the functional gene, and become 'genomic fossils' due to mutations that prevent the transcription of the gene, such as within the gene promoter region, or disturb the translation of the gene, for example as consequence of the introduction of a premature stop codon or frameshift. The other type is called processed pseudogenes. They arise from reverse transcription of mRNAs through the activity of the reverse transcriptase encoded by retrotransposons such as LINEs and LTR elements (Esnault *et al.* 2000). DNA copies of the mRNAs are produced and inserted into the genome. The insertion contains coding sequences of the gene, but does normally not contain promoters and introns. They end in a poly(A) tail, and are flanked by short direct repeats (Zheng *et al.* 2007).

Pseudogenes were first defined as non-functional DNA sequences, but recently they were suggested 'not so pseudo anymore'. A genome-wide analysis of EST databases and transcriptional analysis of individual pseudogenes has revealed that up to one third of processed pseudogenes are transcribed, mostly in testes (Babushok *et al.* 2007). More than 1000 pseudogene transcripts were detected in human, 20 of which were suggested to be functionally active (Vinckenbosch *et al.* 2006). Some pseudogenes function as microRNA sponges, which mean they are competitive inhibitors of small RNAs in mammalian cells

(Ebert *et al.* 2007). For example, a pseudogene of *OCT4* (*OCT4-pg4*) functions as a natural microRNA sponge to regulate *OCT4* expression by competing for miR-145 in hepatocellular carcinoma (Wang *et al*. 2013).

Formation of new retrotransposons

Retrotransposon integrations into the genome can lead to the formation of new retrotransposons, such as SVA elements. SVAs are hominoid specific composite elements named after their main components: SINE, VNTR and *Alu* (Wang *et al.* 2005). The first SVA probably appeared in the genome due to the integration of these elements into the same genomic locus (Shen *et al.* 1994). They originated <25 million years ago and have increased to about 3000 copies in the human genome (Wang *et al.* 2005). They were suggested to represent non-autonomous retrotransposons that are mobilized by L1 encoded proteins and are presumed to be still active in the human genome. A *de novo* SVA insertion was found in the *α-spectrin* gene, leading to a variety of hereditary red blood cell disorders (Ostertag *et al.* 2003). Moreover, it was suggested that LTR retrotransposons might have been formed as a fusion product of DNA transposons and non-LTR retrotransposons (Malik & Eickbush 2001).

Formation of chimeric retrogenes

Chimeric retrogenes, found in mammalian and fungal genomes, are bipartite elements composed of DNA copies of cellular transcripts either directly fused to each other or fused to the 3′ part of retrotransposons (Buzdin *et al.* 2007). New chimeric retrogenes, that are often expressed, can be produced by L1 retrotransposition (Kazazian 2004). These genes were suggested to be generated through RNA template switching from retrotransposons like L1 and *Alu* to other small nuclear RNAs during reverse transcription. A total of 82, 116, 66, and 31 retrogenes potentially encoding TE derived proteins and to have evolved new cellular functions, were found in human, mouse, rat and rice respectively (Gogvadze & Buzdin 2009). The chimeric retrogenes in the human genome were suggested to consist of 5′ regions

originating from small nuclear RNAs, such as U6, U3, U5, and 5*S* RNA, and 3′ regions derived from the 3′ ends of L1 or *Alu* elements (Kazazian 2004).

Recombination of retrotransposons

Recombination is another powerful factor of evolution that produces genetic variability by using already existing blocks of biological information (Makalowski 2000). The human glycophorin gene family evolved through several duplication steps that involved recombination between *Alu element*s (Makalowski 2000). Furthermore, *Alu*-derived ectopic recombination generated 492 human-specific deletions, of which 60% were shown to be located in genes, and in at least three cases, exons have been deleted in human genes relative to their chimpanzee orthologs (Sen *et al.* 2006).

Recombination between retrotransposons may cause various diseases. Almost 50 diseases were suggested to be related to *Alu* recombination (Belancio *et al.* 2008; Xing *et al.* 2009). Disorders such as glycogen storage disease and alport syndrome are due to the recombination between L1 elements (Burwinkel & Kilimann 1998; Segal *et al.* 1999) and complete germ cell aplasia is due to recombination between *HERV-I* elements (Kamp *et al.* 2000).

1.1.5.2 Retrotransposons as regulators in the genome

Promoters

A whole-genome analysis revealed that about 24% of the human promoter sequences analyzed contain retrotransposon-derived sequences, making up more than 7% of the total nucleotides in all of the promoters. Those sequences are derived from all types of common human retrotransposons, such as LINEs (1.6%), SINEs (5.3%), and LTRs (0.4%) (Jordan *et al.* 2003). These retrotransposon-derived sequences can act as the main promotor of a gene, for example the antisense L1 and *Alu* sequences of the *HYAL-4* gene, which is necessary for hyaluronan catabolism (van de Lagemaat *et al.* 2003), but mostly they act as alternative

promoters and thus can influence the RNA transcription level and/or change the tissue specificity of the gene it is residing in. There are numerous examples of retrotransposons acting as *cis*-promoters, such as LTR retrotransposons that cause placental-specific expression of *CYP19* (van de Lagemaat *et al.* 2003) and regulate transcription of *NAIP* (Romanish *et al.* 2007).

Surprisingly, retrotransposons were found with a preference in antisense direction when located in gene introns, which was explained by the fact that sense oriented insertions are more likely to be harmful to regulatory motifs, so they are less likely to be fixed (van de Lagemaat *et al.* 2003). Thus, antisense transcription of genes is also promoted (Conley *et al.* 2008). Using cap analysis of gene expression (CAGE) 48718 human gene antisense transcriptional start sites with TE elements were identified (Conley *et al.* 2008). For example, human specific *HERV-K* can generate antisense transcripts of *SLC4A8* and *IFT172* mRNAs (Gogvadze *et al.* 2009). Although they may drive transcription of RNAs that are complementary to the genes, the full function of the antisense transcripts remains unclear.

Transcriptional enhancers

Enhancers are short (50 ~ 1500 bp) regions of DNA that can be bound by transcriptional factors to activate transcription of a gene (Blackwood & Kadonaga 1998). Many retrotransposons have been reported to play an enhancer role in gene regulation. SINEs were reported to act as distal transcriptional enhancers for *Isl1, Fgf8* and *Satb2* genes in mouse (Bejerano *et al.* 2006; Sasaki *et al.* 2008; Tashiro *et al.* 2011) and human SINE *B2* was reported to serve as a domain boundary blocking the influence of repressive chromatin modifications during organogenesis (Lunyak *et al.* 2007). In human, it was shown that LINE sequences reside in an enhancer of apolipoprotein a (*APOA*) and *Alu* sequences are part of the *MAL* gene (also known as CD8A) enhancer (Hambor *et al.* 1993; Yang *et al.* 1998). A study comparing biochemically predicted enhancers in mouse and rat trophoblast stem cells found that species-specific enhancers were highly enriched for ERVs on a genome-wide level

(Chuong *et al.* 2013), and one ERV family (RLTR13D5) contributes hundreds of mouse-specific histone H3 lysine 4 monomethylation (H3K4me1) and H3 lysine 27 acetylation (H3K27ac) defined enhancers that functionally bind *Cdx2*, *Eomes* and *Elf5* core factors k in trophoblast stem cells. The ERV-9 element in the human *β*-globin cluster was reported to be responsible for controlling expression of this cluster in erythroid cells (Long *et al.* 1998). In transgenic zebrafish harboring the *β*-globin ERV-9 coupled to the GFP gene the ERV-9 was shown to possess enhancer activity enabling GFP expression in oocytes (Pi *et al.* 2004).

Environment response elements

Barbara McClintock's 'genomic shock' hypothesis proposed that TE activity could be a response to environmental pressures (McClintock 1983). This hypothesis has been proven by the finding that numerous TEs contain motifs such as hormone response elements and therefore are targets of chemical and molecular environmental changes (Ono *et al.* 1987; Babich *et al.* 1999; Laperriere *et al.* 2007). These hormone response elements influence the transcription of a large number of genes by interaction with intracellular receptors (Beato *et al.* 1996). For example, DR2-type hormone response elements are heavily present in *Alu*s (90%) (Laperriere *et al.* 2007). These *Alu*-DR2 elements were suggested to drive retinoic acid regulated genes, and bind retinoic acid receptors *in vivo* (Laperriere *et al.* 2007).

Insulators

Insulators are genetic boundary elements that block the interaction between distal enhancers and inappropriate target promoters (Kellum & Schedl 1991). Some evidence has been provided that retrotransposons can function as insulators *in vivo* to distinguish blocks of active and transcriptionally silent chromatin. *Gypsy*, a LTR retrotransposon has been widely reported to possess insulator activity in *Drosophila* (Cai & Levine 1995; Gdula *et al.* 1996; Nabirochkin *et al.* 1998; Pai *et al.* 2004). This *gypsy* insulator insertion was suggested blocking enhancers and down regulating the genes *yellow* and *miniwhite* (Dorsett 1993;

Kostyuchenko *et al.* 2008). It might establish higher-order domains of chromatin structure and regulate nuclear organization by tethering the DNA to the nuclear matrix and creating chromatin loops (Byrd & Corces 2003).

Mammalian LINEs and SINEs functional as insulators were found within matrix attachment regions (Purbowasito *et al.* 2004; Akopov *et al.* 2006; Román *et al.* 2011). For example, a B2 retrotransposon in the murine growth hormone locus serves as a boundary to block the influence of repressive chromatin modifications by generating short, overlapping Pol II- and Pol III- driven transcripts (Lunyak *et al.* 2007).

New poly(A) sites

The polyadenylation or poly(A) sites are cleavage recognition sites at which poly(A) tails are added during the processing of the 3'end of pre-mRNA. Many retrotransposons were reported to generate new poly(A) sites. It was estimated that about 8% of all poly(A) sites are associated with TEs (Lee *et al.* 2008). Chen *et al.* (2009) found that 10,000 of the 1.1 million *Alu*s in the human genome are inserted in the 3'UTRs of protein coding genes. Of these about 1% (107 events) is active as poly(A) site. They also found that *Alu* insertions in 3'UTRs are equally in the forward and reverse direction, but 99% of polyadenylation-active *Alu*s are forward oriented. These *Alu*-borne poly(A) sites are intronic and produce truncated transcripts that may impact gene function and / or contribute to gene remodeling (Chen *et al.* 2009).

L1 elements also created premature poly(A) sites with insertions into endogenous genes and produced new mRNA isoforms (Han *et al.* 2004). Many LTR retrotransposons were also suggested as a source of poly(A) sites. For example, HERV were reported as an alternative or major poly(A) site for the genes *PLT* (Goodchild *et al.* 1992), *ZNF195* (Kjellman *et al.* 1999), *HHLA2* and *HHLA3* (Mager *et al.* 1999).

Novel splice sites

Alternative splicing is a mechanism by which different mRNAs are generated from a single gene. Some retrotransposons contain splice donor/acceptor sites that compete with the original ones of exons. *Alu*s were suggested to play a crucial role in the birth of new exons (Lin *et al.* 2008). Complete or partial *Alu*s are found present in coding regions of mature mRNAs, and they were suggested to be recruited in the coding region, possibly because *Alu*s contain motifs that resemble consensus splice sites (Makałowski *et al.* 1994; Sorek *et al.* 2002). Mutation within intronic *Alu*s may yield active splice sites, which makes partial or complete *Alu*s part of the newly generated exon (exonization). It was reported that *Alu* exons with high splicing activities were strongly enriched in the 5'-UTR, and two-thirds of the *Alu* exons significantly altered mRNA transcriptional efficiency (Shen *et al.* 2011). Alternative splicing does not only cause exonization, but also exon skipping. About 1.5% to 4% of the alternatively skipped exons in human were reported to be due to intronic retrotransposons flanking the exons (Lev-Maor *et al.* 2008).

A study of a L1 insertion into the dystrophin gene, which when defect causes Duchenne muscular dystrophy (DMD), disclosed that the L1 insertion altered splicing, skipping exon 44 of the gene (Narita *et al.* 1993). Although LINEs were reported to utilize the same retrotranspositional mechanism as SINEs, the proportion of L1s in gene introns is significantly lower than that of *Alu*s (Buzdin 2004). This difference was suggested to be due to purifying selection acting against accumulation of L1s in genes (Pavlíček *et al.* 2001). The direction of retrotransposon insertion also influences the splicing mRNA events. For example, when *HERV9* and *HERV-K* are present in sense direction in transcribed regions, they show considerable bias for use of strong splice sites, but not when they are in antisense direction (van de Lagemaat *et al.* 2006).

Heterochromatin

Heterochromatin is tightly packed DNA that is inaccessible to polymerases and is therefore not transcribed. One type of heterochromatin called constitutive heterochromatin can affect the genes near itself. Constitutive heterochromatin is composed of predominantly repetitive DNA sequences, including simple repeats, DNA transposons, and retrotransposons (LINEs, SINEs and ERVs), and is mostly found near centromeres, telomeres, but also scattered throughout the chromosomes (Saksouk *et al.* 2015). The other type of heterochromatin called facultative heterochromatin is the result of genes that are silenced through a mechanism such as histone deacetylation or RNAi. Both LINEs (L1) and SINEs (B2) were reported to help in facultative heterochromatin formation during X chromosome inactivation (XCI), (Chow *et al.* 2010).

Transcriptional silencer

Retrotransposons are also known to function as transcriptional silencers of the enclosing genes (Gogvadze & Buzdin 2009). For example, a transcriptional silencer at the distal end of the human tumor-suppressor (*BRCA2*) gene promoter was identified as a full-length *Alu*, and this *Alu* was shown to be involved in the negative regulation of the gene expression in breast cell lines (Sharan *et al.* 1999). An endogenous retroviral sequence, *RTVL-la*, was also reported with a silencer activity. The insertion of *RTVL-la* in *HPR* (92% identical to human haptoglobin gene) reduced HPR expression with 50% (Hatada *et al.* 2003).

1.1.6 Regulation of retrotransposition

In somatic tissues and mature germ cells, retrotransposition and even transcriptional activity of retrotransposons is largely suppressed by several different cellular defense mechanisms. These defense mechanisms will be described in the following part.

DNA methylation

DNA methylation is an important epigenetic regulatory mechanism characterized by the addition of methyl groups to certain nucleotides, mostly at CpG sites (5'-cytosine-phosphate-guanine-3' sites). This methylation results in the conversion of the cytosine to 5-methylcytosine. When located in a gene promoter, DNA methylation typically acts to repress gene transcription. The vast majority of retrotransposons, including LINEs, SINEs and ERVs are densely methylated in normal somatic cells and contained in inactive chromatin. Regulation of this DNA methylation ensures a stable equilibrium between retrotransposons and their host (Schulz *et al.* 2006). CpG sites in retrotransposon promoters are largely methylated in normal cells. In *Alu*s, crucial CpG sites are localized in the essential A and B boxes and their methylation prevents binding of the Pol III co-factors (Liu & Schmid 1993). In L1 elements, the repression by methylation appears to be mediated by the methylcytosine-binding proteins MeCP2 and MBD2 (Yu *et al.* 2001; Steinhoff & Schulz 2004). The effects of methylation on HERV regulation have not been studied in detail. HERV expression at relatively low levels has been reported in some carcinomas with hypomethylated genomes (Wang-Johanning *et al.* 2003; 2007).

A correlation between DNA methylation level and retrotransposon activity has been found: retrotransposons are most activated in hypomethylated cells, like cancer cells, and the retrotransposons themselves are found hypomethylated as well. In various human cancer cell lines, from germ cell cancers to ovarian carcinoma, hypomethylation and strong expression were found for *Alu*, L1 and HERV (Schulz *et al.* 2006).

Besides in cancer cells, retrotransposon promoters undergo two waves of partial demethylation: one occurs during the specification and migration of primordial germ cells and the other during the preimplantation stage (Lee *et al.* 2014). It was reported that mouse embryos deficient for the maintenance DNA methyltransferase DNMT1 lose methylation of several types of TEs (Walsh *et al.* 1998), and a similar phenotype has been observed for the double-mutant mouse for DNMT3a and DNMT3b that encode *de novo* DNA

methyltransferases (Okano *et al.* 1999).

Furthermore, DNA methylation is not the only epigenetic repressor in the control retrotransposons' transcription and movement. Overlapping epigenetic mechanisms evolved to control the expression of TEs in human and mouse embryonic stem cells (Slotkin & Martienssen 2007; Castro-Diaz *et al.* 2014). Another important epigentic regulationary mechanism is by histone modification and different marks appear to be important for different elements and/or developmental stages. Histone H3 trimethylated at lysine9 (H3K9me3) and DNA methylation work together to repress most RT through a constitutive hererochromatin signature. Mutations in the histone H3K9 methyltransferase gene *Suv39* result in modest upregulation of transcripts in mouse embryonic stem cells (Martens *et al.* 2005). In the mouse histone lysine methyation study, it was found different histone modifications have a selective enrichment region over the TEs: The most prominent enrichment for repressive histone modifications was observed over tandem satellite repeats and DNA transposons (combinations of H3K9me3, H3K27me1 and H4K20me3 or me2); only one mark was found for IAP LTRs (H4K20me3); and for LINEs and SINEs, no infomative signals were found (Martens *et al.* 2005). However, Fadloun *et al.* (2013) found that in mouse embryos, L1 and IAP are enriched for both the activating mark H3K4me3 and the repressive mark H3K9me3, when they are expressed at the 2-cell stage, but are only enriched for H3K9me3 at the 8-cell stage as expression begins to decrease. It was also shown that the H3K9me3 exhibited the same tendency as DNA methylation in early mouse embryo development.

Small RNA

Small RNAs are <200 nucleotide in length and are usually non-coding RNAs (Storz 2002). The three major classes of endogenous small RNAs identified in mammals are microRNAs (miRNAs), small interacting RNAs (siRNAs) and PIWI-interacting RNAs (piRNAs) (McManus & Sharp 2002; Siomi *et al.* 2011). miRNA and siRNA are well studied as they are important in post-transcriptional gene silencing. They can bind to mRNAs to regulate their activity, typically causing the degradation of the target mRNAs or inhibiting their translation.

miRNAs and siRNAs are generated from double-stranded precursor RNAs by a process of cleavage (by Dicer or related endonuclease) and modification after which the small RNAs are loaded into the RNA-induced silencing complex (RISC). One of the proteins in RISC, called Argonaute (AGO), cleaves the bound target mRNA (Valencia-Sanchez *et al.* 2006). miRNA induced silencing RISC is guided by miR-128 to bind directly to a target site residing in the L1 ORF2 RNA, which results in destabilization of L1 transcript and subsequently, repression of the L1 protein (Hamdorf *et al.* 2015). siRNAs processed from sense and antisense promoters in the 5'UTR of full-length human L1 were found to suppress L1 retrotransposition by RNAi (Yang & Kazazian 2006), and miRNA from the maternally inherited chromosome 12 was found to inhibit a retrotransposon-like gene (*Rtl1*) expressed from the paternal allele in mouse (Seitz *et al.* 2003). The regulation of retrotransposons by small RNAs was further confirmed by the evidence that ERVs and LINEs serve as substrates for the production of siRNAs (Soifer *et al.* 2005), and a nuclear protein complex involved in miRNA biogenesis binds L1, *Alu*, and SVA-derived small RNAs in human cells (Heras *et al.* 2013).

Another class of small RNAs, PIWI interacting RNA (piRNA, originally P-element induced wimpy testis in *Drosophila*) has been especially linked to transcriptional and post-transcriptional gene silencing of retrotransposons and therefore are also called repeat-associated small interfering RNAs (rasiRNAs) (Saito *et al.* 2006). piRNAs are normally 26-31 nt long. piRNAs are found mainly in germ cells (in fruit fly, mouse, rat and human) derived from TEs and long non-coding RNA, and interact only with PIWI proteins (Siomi *et al.* 2011). Based on work with *Drosophila*, the piRNA pathway is proposed to work with a 'Ping-Pong' mechanism to amplify the inhibitory signal (Grimson *et al.* 2008), as outlined in **Figure 1.3**. In this system, primary antisense piRNAs loaded onto the protein Aubergine (AUB) bind to complementary TE RNAs. The targets are then cleaved by AUB generating new sense piRNAs that are loaded into an Argonaute-3 (AGO3) complex. The AGO3 bound sense piRNAs lack the ability to target transposable element transcripts directly, but are proposed to guide the production of piRNAs that are loaded into AUB by targeting newly exported piRNA cluster transcripts. This AGO3-directed cleavage generates additional

antisense piRNAs that actively silence their target elements, thus reinforcing the cycle through the creation of additional sense piRNAs (Brennecke *et al.* 2007). This mechanism was also observed with piRNAs associated with PIWI family members MILI and MIWI2 isolated from mouse testes (Brennecke *et al.* 2007). It was also found that PIWI proteins associating with piRNAs originating from repetitive targets function in silencing

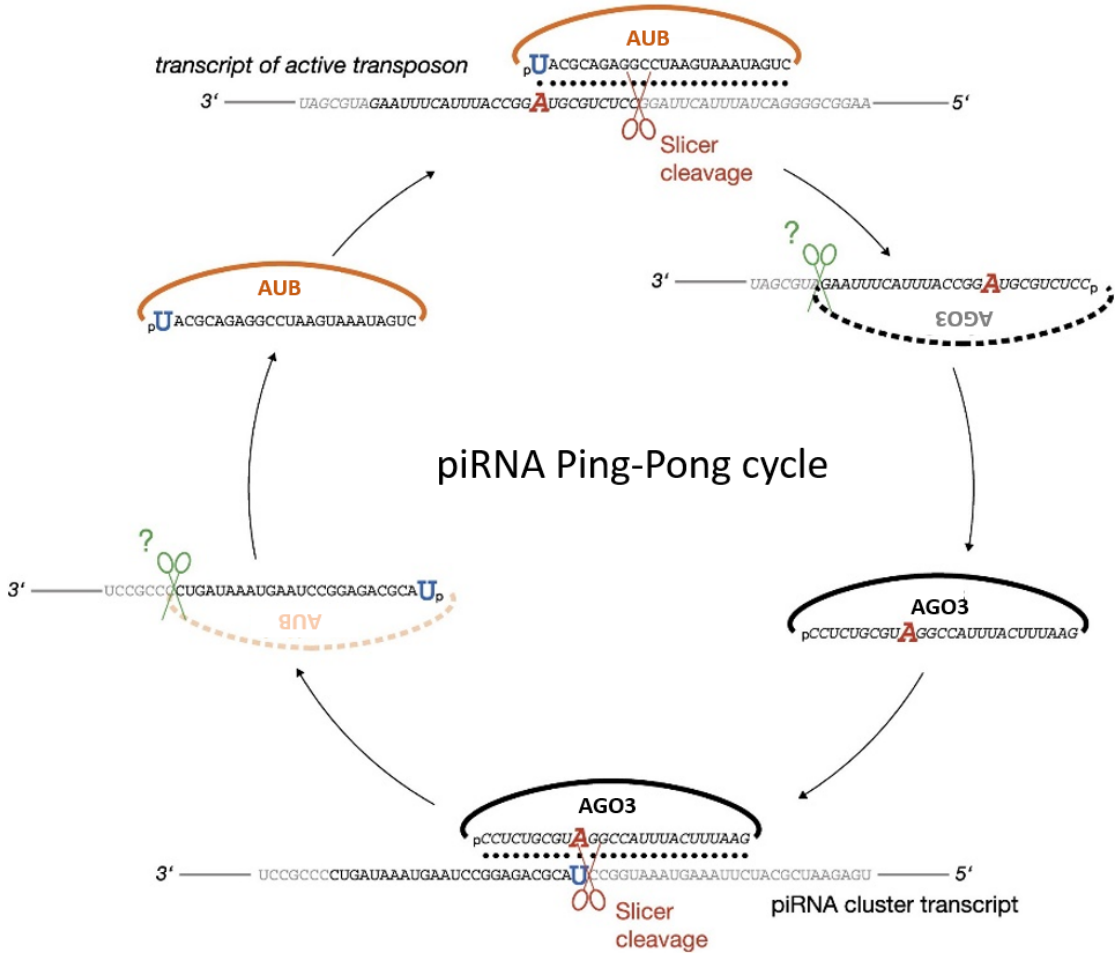


Figure 1.3 The piRNA Ping-Pong amplification cycle consisting of AUB complexes, AGO3 complexes, piPNA cluster transcripts and transcripts of active transposons. Nucleotide cleavage events are shown as scissors. Adapted from Brennecke *et al.* (2007).

retrotransposons through the RNAi mechanism (Saito *et al.* 2006). Recent evidence indicates that the PIWI pathway is also functional in human early embryonic cells and ape pluripotent stem cells controlling expression of L1s (Marchetto *et al.* 2013).

In addition to their role in post-transcriptional regulation, piRNAs also seem active on the transcriptional level. MILI and MIWI2 were reported to play essential roles in establishing *de novo* DNA methylation of retrotransposons (L1 and IAP) in murine fetal testes (Kuramochi-Miyagawa *et al.* 2008). Mice with inactivated MILI and MIWI2 show not only enhanced L1 and IAP expression, but also hypomethylation of the L1 5'URT. This relation between piRNA and *de novo* DNA methylation is also supported by the observation that piRNAs recruit *de novo* DNA methyltransferases at repeated elements (Aravin *et al.* 2008; Aravin & Bourc'his 2008).

Other transcription and post-transcription factors

Some transcription and post-transcription factors interacting with retrotransposon promoters have been identified and shown to act as activator, repressor, or initiator.

Retrotransposons are mostly expressed in early embryonic cells and cancer as discussed above, and this tissue-specific expression is determined in part by transcription factors. Some retrotransposon related transcription factors are listed in **Table 1.2**. The majority of transcription factor binding sites occur in the LTRs of LTR retrotransposons or in UTRs of LINEs and SINEs. Therefore, LTRs and UTRs serve as platforms for recruitment of the transcription factor and co-factor's to regulate expression of the TEs (Robbez-Masson & Rowe 2015). Certain classes of retrotransposons are enriched for particular transcription factors. For example, OCT4/ SOX2 binding is enriched on ERV-K in mouse and ERV1 in human, while TP53 is predominantly found on mer61-type ERV1 in human (Bourque *et al.* 2008).

On the post-transcriptional level 3' repair exonuclease 1 (TREX1) and some cytidine deaminases are reported as potential host factors regulating retrotransposon activity. APOBEC3 (apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3) is a family of cytidine deaminases, converting cytidine to uridine (or deoxycytidine to deoxyuridine). APOBEC3 proteins (3A, 3B, 3F, 3G) have been reported to inhibit retrotransposition of *Alu*s,

ERVs and L1 (Turelli *et al.* 2004; Bogerd *et al.* 2006; Chen *et al.* 2006; Chiu *et al.* 2006; Hulme *et al.* 2007; Khatua *et al.* 2010). A retrotransposition reporter assay in HeLa cells revealed that presence of APOBEC3A reduced the L1 retrotransposition frequency by up to 85%. A combination of APOBEC3B, 3C and 3F inhibited retrotransposition about 75% and down-regulation of APOBEC3C could enhance the movement by about 78% (Muckenfuss *et al.* 2006). APOBEC3G inhibits retrotransposition of IAP and MusD, and induces G to A

Table 1.2 Transcription factors interacting with retrotransposon promoters

Gene name	Function	Regulation of retrotransposons
RUNX3 Runt-Related Transcription Factor 3	It functions as a tumor suppressor, and is frequently deleted or transcriptionally silenced in cancer.	Exogenous expression of *RUNX3*, increases human L1 transcription, and influences L1 5'UTRs in both sense and antisense directions (Yang et al. 2003).
SOX2 SRY (Sex Determining Region Y)-Box 2	It is involved in the regulation of embryonic development and in the determination of cell fate.	L1 promoter recruits *NANOG*, *OCT4* and *SOX2*, and silences expression of *OCT4*, *NANOG*, and *SOX2* resulting in decreased expression of L1 (Närvä et al. 2012).
KRAB-ZFP Kruppel-associated box-zinc finger protein	It is involved in transcriptional repression of RNA polymerase I, II, and III promoters and binding and splicing of RNA	It represses together with its cofactor *KAP1* through histone and DNA methylation, a process resulting in irreversible silencing of TEs (Ecco *et al.*, 2016).
TP53 Tumor Protein P53	It acts as a tumor suppressor in many tumor types.	*TP53* is predominantly found on mer61-type ERV1 in human (Bourque et al. 2008).
NANOG Nanog Homeobox	It is involved in embryonic stem (ES) cell proliferation, renewal, and pluripotency.	LINEs and LTR retrotransposons embedded human-specific transcription factor binding sites consititute 64% of all human-specific binding events for *NANOG* (Glinsky 2015).
OCT4 (***POU5F1***) octamer-binding transcription factor 4	It plays a key role in embryonic development and stem cell pluripotency.	Binds the LTR of HERV-H, which is highly expressed in human ES cells (Santoni et al. 2012). During reprogramming to iPSCs, TEs are co-expressing with *OCT4, SOX2* and *KLF4* (Friedli et al. 2014).
YY1 Yin yang 1	It is involved in repressing and activating a diverse number of promoters.	A *YY1*-binding site is required for accurate human LINE-1 transcription initiation (Athanikar et al. 2004)
CTCF CCCTC-Binding Factor (Zinc Finger Protein)	It plays an essential role in oocyte and preimplantation embryo development by activating or repressing transcription.	Over 10,000 and 5,000 *CTCF* binding events were observed on SINE B2 elements in mouse and rat respectively (Schmidt et al. 2012).
MITF Microphthalmia-Associated Transcription Factor	It regulates the differentiation and development of melanocytes retinal pigment epithelium and is involved in pathways in cancer and gioma.	It serves as transcription activator of HERV-K (Katoh et al. 2011).

hypermutations in their DNA copies (Esnault *et al.* 2005). Hulme *et al.* found that APOBEC3G inhibits *Alu* retrotransposition selectively in an ORF1p-independent manner and suggested that the *Alu* ribonucleoprotein complex may be targeted by APOBEC3G (Hulme *et al.* 2007). *In vitro* experiments showed that overexpression of TREX1 can cause a dramatic reduction in IAP and L1 retrotransposition efficiency (40% and 80% respectively) by mediating the degradation of their reverse transcripts (Stetson *et al.* 2008).

1.1.7 DNA methylation of retrotransposons as marker for global cell methylation status

As the main regulator of the activity of repetitive sequences DNA methylation within retrotransposons has been suggested to be a marker for estimating the global DNA methylation status, mostly in cancer cells (Yang *et al.* 2004; Weisenberger *et al.* 2005; Sunami *et al.* 2011), or as a biomarker for cell growth in stress situations, such as under lead exposure (Wright *et al.* 2010).

As outlined in previous sections DNA methylation is essential for normal development and is associated with a number of key processes within embryonic development, such as genomic imprinting, X-chromosome inactivation, aging and carcinogenesis. Aberrant DNA methylation, both hypermethylation and hypomethylation, has been associated with embryonic development failure, aging, cancer and other diseases (Toyota & Issa 1999; Ehrlich 2002; Shi & Haaf 2002).

Table 1.3 Advantages and disadvantages of selected methods for analysis of global DNA methylation

Methods	Purpose	Pros	Cons
*Sss*I methyltransferase	Non-next-generation sequencing based Global DNA methylation analysis	Inexpensive; Complete coverage of CpG sites	Radioactivity based; unstable SAM and the *SssI* enzymes
Reverse-phase high-performance liquid chromatography (RP-HPLC)		Non-radioactive; Measures all CpG sites	Costs and advanced knowledge of instrumentation.
Anti-5-methylcytosine (5-mC) immunofluorescence staining		Non-radioactive; Measures all CpG sites; DNA methylation loci is visualized	Costs associated with specific antibody and instrumentation; Interference of cell membrane autofluorescence or noise background under observation
Methylated DNA immunoprecipitation followed by sequencing (MeDIP-Seq)	Genome-wide next-generation sequencing based analysis	Coverage is independent of restriction site limitations; Costs of sequencing are not so high due to the reduced representation of the genome; Good coverage of methylome	Costs for specific antibody and sequencing; intensive bioinformatics expertise needed for data analysis; less powerful than equivalent methods to identify differentially methylated regions
Reduced representation bisulfite sequencing (RRBS)		Coverage is independent of restriction site limitations.	High cost associated with sequencing; intensive bioinformatics expertise needed for data analysis; Limited coverage of the methylome in CpG poor region.
Capture of methylated DNA using the methyl DNA binding domain MeCP2 followed by sequencing (MBD-seq)		Coverage is independent of restriction site limitations; good coverage of the methylome.	High cost associated with sequencing; intensive bioinformatics expertise needed for data analysis.

In the last several decades, the technology to determine DNA methylation went through a steady progress. Methods to detect global DNA methylation in the genome have been discussed in detailed in several reviews (Oakeley 1999; Zilberman & Henikoff 2007; Laird 2010; Guerrero-Bosagna 2014). DNA can be digested into single nucleotides and total 5-methylcytosine (5-mC) can be quantified by either high-performance liquid chromatography (HPLC), thin-layer chromatography, or liquid chromatography/mass spectroscopy (Oakeley 1999). Sodium bisulfite conversion based analysis combined with PCR, has also been widely used. These methods include direct bisulfite sequencing PCR (BSP), pyrosequencing, combined bisulfite restriction analysis (COBRA) and so on. Although these PCR based methods were used mostly for gene-specific methylation analysis, they paved the way for more advanced analysis of global DNA methylation. More recently, due to the boom in sequencing technologies, DNA methylation analysis based on next-generation sequencing has made the mapping of DNA methylation feasible on a genome-wide scale. However, each method has its limitations. Non-PCR-based methods require large amounts of good quality DNA, while genome-wide sequencing is still expensive and requires intensive bioinformatics expertise (Boyle *et al.* 2012). The most common analysis methods with their advantages and disadvantages are listed in **Table 1.3**. Among these methods, immunofluorescence staining, which allows visualization of 5-mC by using fluorescent secondary antibody to target anti 5-mC antibody, is the most widely used in embryonic methylation analysis (Dean *et al.* 2001; Beaujean *et al.* 2004; Dobbs *et al.* 2013; Heras *et al.* 2014).

Due to the large amount of retrotransposons in the genome, measuring DNA methylation status of retrotransposons has been suggested as an alternative method for global methylation quantification (Yang *et al.* 2004). It has been used intensively in cancer research, mostly with L1, due to its high proportion in the genome. Recently, other retrotransposons such as IAP (Bakshi & Kim 2014) and *Rlt1* (retrotransposon-like 1) were also suggested for use as epigenetic biomarkers. The advantage of using retrotransposons as global methylation marker is a big coverage of the genome (different coverage depending on retrotransposon families),

low cost and no need of intensive bioinformatics expertise for data analysis. In this thesis, we tested and used a combination of the most widespread retrotransposon families in the bovine genome to represent the global methylation, which provides a more accurate result than using a single retrotransposon.

1.2 Introduction of early embryo development

Preimplantation embryo development comprises the initial stages of mammalian development, before the embryo implants into the mother's uterus. This complex developmental transition begins with the fertilization of the oocyte and runs until the formation of a blastocyst, including first cleavage division, genetic and epigenetic reprogramming, activation of the embryonic genome and differentiation to the inner cell mass (ICM) and trophectoderm (TE) cells. Besides the natural preimplantation development *in vivo*, assisted reproductive technologies (ARTs) have been widely used in many mammalian species for clinical, commercial and research purposes. In cattle, ARTs intend to be used routinely to shorten generational intervals, obtain desired gender and propagate genotype (mainly on milk and meat) in production. In the following pages, the events happening during bovine preimplantation and the influence of ARTs used in bovine embryo production will be described.

1.2.1 *In vivo* bovine preimplantation embryo development

The development of an embryo starts with the fusion of a single capacitated spermatozoon with a mature oocyte. Before this happens, one or more mature oocytes are arrested in metaphase of the second meiotic division (MII), and released in the oviduct, surrounded by a glycoprotein matrix, called zona pellucida (ZP) and several layers of cumulus cells. At mating, more than 10^9 spermatozoa are delivered in the female genital tract, and only several thousands of viable spermatozoa will reach the caudal isthmus of the oviduct to form a sperm reservoir (Van Soom & Kruif 1998). Afterwards, a small number of sperm from this reservoir becomes capacitated and released to ampulla activated by imminent ovulation, and finally

reach the ovulated oocyte (Hunter 1993). The capacitated sperm binds to and penetrates the ZP, fuses with the oolemma and triggers oocyte activation. This activation changes oocyte cell membrane to prevent fusion with other sperm. The penetration of the sperm also activates the oocytes, finishes the second meiotic division, with extrusion of the second polar body, leaving a haploid (n) maternal nucleus and forming a one-cell embryo (zygote). The haploid sperm and oocyte then form paternal and maternal pronuclei (Hyttel *et al.* 1988). The membranes dissolve, leaving no barriers between the male and female chromosomes, and then their chromosomes can combine and become part of a single nucleus in the resulting embryo with full set of chromosomes.

The timing of first cleavage division varies among species. For *in vivo* bovine embryos, the time required for progression to the first cleavage ranges from 23 to 31 hours post fertilization, but this period includes sperm capacitation, oocyte penetration, activation and completion of meiosis II (Barnes & Eyestone 1990), and the first cell cycle is 20-24 h (Eyestone & First 1988). Cleavage to the 4- and 8-cell stage occurs at approximately 36-50 and 56-64 hpi, respectively; and cleavage to 16-cell stage occurs at approximately 80-86 hpi with a fourth cell cycle duration of 21-30 h (Eyestone & First 1988). During the first several divisions, the size of the embryo remains constant, but the nuclei number increases and smaller blastomeres are produced. Around 5 dpi, *in vivo* bovine embryos reach the 32-cell stage and starts compaction (Betteridge & Flechon 1988). Compaction is the first morphogenetic event in which the embryos undergo changes in cell morphology and cell-cell adhesion between the blastomeres until the individual outlines disappear and the formation of a uniform cellular mass, called compacted morula. In this stage, cell polarity is established and initiates the first differentiation event - blastocyst formation.

The first step of blastocyst formation is the differentiation of the trophectoderm (TE) cells and the inner cell mass (ICM). The TE cells acquire the characteristics of epithelial cells in being flattened and joined together by tight junctional complexes, forming one or several cavities, called blastocoel (Watson & Barcroft 2001). ICM, on the other hand, attached and enclosed by TE, will further specify into the epiblast and primitive endoderm. At this time of development, the *in vivo* embryo resides in the uterus and starts expanding. The volume of the

embryo increases by the inflation of the blastocoel cavity, and this expansion thins the ZP until the embryo frees itself by series of expansion and contraction cycles. The ZP is dissolved at the abembryonic pole by enzymes and the embryo bulges out from the ZP envelop. This process is called hatching. In human and rodent, embryos implant soon after hatching. In ruminants, on the other hand, the implantation in the uterus is delayed. In cattle, trophoblast cells start to elongate at 14 dpi and the embryonic membrane can extend the entire length of both uterine horns by 24 dpi. By that time, the bovine conceptus size increases more than 1000-fold during elongation (Maddox-Hyttel *et al.* 2003). This *in vivo* development is considered as the ideal and be used as the gold standard to validate embryo quality under *in vitro* condition.

1.2.2 Transcriptional dynamics in preimplantation embryo

Fertilization brings the two haploid genomes of the gametes into the oocyte cytoplasm, giving rise to a totipotent embryo. The gametic genomes are silenced at this time and the mRNA and proteins are supplied by the oocyte cytoplasm. Then the existing messages are degraded and few new transcripts are produced, which is followed by the embryonic genome activation (EGA). The embryo stage of EGA is species-specific, with 2-cell in mice, 4-cell in human rats and pigs, and 8- to 16-cell in cattle and sheep (Whitworth *et al.* 2004). In cattle, a lot of IVP bovine embryos fail to develop around the 8-16 stage, which is concurrent with the genome activation and therefore also called critical or block-stage (De Sousa *et al.* 1998). Instead of one single global transcriptional switch, the genome activation is suggested to consist of multi-steps. In cattle, embryo development starts the first transcriptional activity already in 1-cell stage. Genes related to different functions have different transcriptional activation times. Genes activated at or before the 4-cell stage are functionally related to RNA processing, translation, and transport, preparing the embryo for the major EGA at 8- to 16-cell, when genes from a broad range of functional categories start to be activated, including transcriptional and translational, as well as protein ubiquitination (Graf *et al.* 2014). Genes expressed around compaction or in the blastocyst stage but not early stages, were suggested

functional in the regulative processes at the onset of differentiation (Ponsuksili *et al.* 2002; Hamatani *et al.* 2004).

At the first differentiation event, blastomeres form either ICM that gives rise to the embryo, or TE that develops into extraembryonic membranes. Lineage commitment towards ICM and TE in under the control of specific transcription factors. For example, in mouse, the ICM is regulated by *Sall4*, *Pou5f1*, *Sox2* and *Nanog*, while TE formation results from a cascade of events involving *Tead4*, *Gata3*, *Cdx2*, *Eomes* and *Elf5* (Chen *et al.* 2010). A global gene expression study showed a total of 870 genes were differentially expressed between bovine ICM and TE (Ozawa *et al.* 2012). Several genes characteristic of ICM (e.g. *NANOG*, *SOX2* and *STAT3*) and TE (e.g. *ELF5*, *GATA3*, and *KRT18*) in mouse and human showed similar patterns in bovine.

1.2.3 Epigenetic dynamics in preimplantation embryo

The term ‘epigenetics’ was introduced and defined as ‘heritable changes in gene expression that occur without any changes in gene sequence’ by Waddington in 1942. During development, cells start in a pluripotent state, then they can differentiate into many cell types, which is controlled by epigenetics activating some genes while inhibiting the expression of others (Reik 2007). There are many types of epigenetic modifications that are known to affect gene expression, including DNA methylation, histone modifications (such as acetylation, phosphorylation, methylation and ubiquitination) and non-histone proteins that bind to chromatin (Berger 2007). Briefly, transcriptionally inactive heterochromatin is characterized by deacetylated histones, methylation of histone H3 lysine 9, and DNA methylation, whereas acetylation of H3 and H4 histones, methylation of histone H3 lysine 4, and low level of DNA methylation are associated with active euchromatin regions (Duranthon *et al.* 2008).

DNA methylation is the most extensively studied epigenetic modification, which plays an important role in the regulation of gene expression and chromatin architecture, in association which histone modification and other chromatin-associated proteins. DNA methylation in mammals is almost exclusively restricted to CpG dinucleotides. The methylation of CpG

dinucleotides can repress transcription either by blocking the binding of transcription factors to the promoters or recruiting histone-modifying protein complexes that repress transcription through the formation of a more condensed chromatin structure.

In mammalian development, there are two critical periods of epigenetic reprogramming: gametogenesis and early preimplantation development (Reik *et al.* 2001). During gametogenesis, genome-wide demethylation occurs, which is followed by remethylation before fertilization. Early embryogenesis is characterized by a second genome-wide demethylation event, a process that is sensitive to environmental factors (Santos et al. 2002). Disturbances in epigenetic reprogramming may lead to developmental problems and early mortality.

In many mammalian species, including human, mouse, cattle, pig and sheep, a global DNA demethylation is observed in preimplantation embryos after fertilization (Dean *et al.* 2001; Beaujean *et al.* 2004; Deshmukh *et al.* 2011). However, there is no consensus on how the demethylation proceeds. It was suggested that the paternal genome is actively demethylated in the zygote before the first cleavage division, while the maternal genome undergoes passive demethylation. During DNA replication preceding each cell division, no new methyl groups are added to the newly synthesized DNA of the maternal genome (Mayer *et al.* 2000; Oswald *et al.* 2000). However, more recent experiments indicate that both maternal and paternal genomes undergo widespread active and passive demethylation in zygotes before the first mitotic division in mouse (Guo *et al.* 2014).

The demethylation is followed by a wave of the *de novo* methylation mediated by *de novo* methyltransferases *DNMT3a* and *DNMT3b* (Okano *et al.* 1999). Different DNA methylation dynamics have been observed in different species. In bovine embryos, demethylation occurs from the 2-cell stage until the 8-cell stage, and increasing methylation level in 16-cell (Dean *et al.* 2001; Dobbs *et al.* 2013). In mouse embryos, DNA methylation is gradually lost during embryo development and reaches its lowest level at the morula stage, and *de novo* methylation starts from the blastocyst stage (Santos *et al.* 2002). In human embryos, DNA

remains intensively methylated until the 4-cell stage and then a quick decrease of the methylation level occurs, followed by the *de novo* methylation at the early blastocyst stage (Fulka *et al.* 2004).

1.2.4 Activation of retrotransposons in early development

Retrotransposons can cause deleterious genomic rearrangements and somatic mosaicism and hence have long been viewed as harmful parasites. However, their relationship with the hosts is much more complicated and more recently it was shown they are necessary in preimplantation development (Sciamanna *et al.* 2011). Two independent approaches were applied to establish whether reverse transcription plays a role in early embryonic development: one is exposing embryos to nevirapine, an RT inhibitor currently employed in AIDS treatment and the other is microinjecting anti-RT antibodies into one blastomere of 2-cell embryo. Both types of RT inhibition stopped embryo development before the blastocyst stage (Pittoggi *et al.* 2003). The injection of antisense oligonucleotides against RT in zygotes also stopped the embryonic development at the 2-4 cell stage (Beraldi et al., 2006). However, when nevirapine was added after the 8-cell stage, the embryo development was not affected, which indicates that the RT activity is required in mouse early embryonenesis, specifically between the late 1-cell and 4-cell stage (Pittoggi *et al.* 2003). Knockout of retrotransposon-derived genes also causes early embryonic lethality owing to defects in the placenta (Ono *et al.* 2005).

Preimplantation development includes stages from fertilization to blastocyst formation, and involves events such as cell proliferation and differentiation, embryonic genome activation, gene imprinting and sex chromosome dosage compensation, which are regulated by epigenetic reprogramming during early embryos (Messerschmidt *et al.* 2014). Here are some hypotheses of retrotransposon function in embryonic activities.

Cell proliferation and differentiation

Not many studies about the involvement of retrotransposons in cell proliferation and differentiation in embryos have been reported. On the other hand, their effect in transformed cells and cancer cells has been broadly studied and this may provide a reference for embryonic research, considering the analogies between early embryogenesis and tumorigenesis (Bailey & Cushing 1925). Typical embryonic genes including *OCT4*, *homeobox* and *twist* family members are re-expressed in cancer cells (Sciamanna *et al.* 2011). Chromosome instability in cancer cells promoted by DNA hypomethylation also occurs commonly in early embryos (Vanneste *et al.* 2009). In line with this background, retrotransposon activity was also found both in embryogenesis and tumorigenesis (Kano *et al.* 2009; Lee *et al.* 2012). In human cell lines, L1 encoded RT was found essential to maintain the cancer cells in a highly proliferating, poorly differentiated or de-differentiated state (Sciamanna *et al.* 2011). In contrast, RT inhibition of L1, either pharmacological or by RNAi, reduced proliferation, induced differentiation and reprogrammed gene expression in human tumorigenic cell lines (Mangiacasale *et al.* 2003; Landriscina *et al.* 2005). In mouse embryonic kidney cells, $L1^{RP}$ (an active human retrotransposon 99% identical to the consensus L1 sequence (Kimberland *et al.* 1999)) was also found to increase proliferation rates and markedly down regulated differentiation programming (Ramos *et al.* 2011). However, RT inhibition of HERV-K by RNAi did not change cell proliferation and differentiation (Oricchio *et al.* 2007).

Genomic imprinting

Genomic imprinting is a gene regulatory mechanism by which certain genes are expressed in a parent-of-origin-specific manner. It is hypothesized that genomic imprinting arose as a by-product of a DNA methylation mechanism that silences foreign DNA, such as retrotransposons. An observation in favor of this 'host defense' hypothesis was made during a comparative analysis of the origin of genomic imprinting in mammals (Suzuki *et al.* 2007). *Peg10* (paternally expressed 10) is a retrotransposon-derived imprinted gene that has an

essential role in proper placental formation in mice (Ono *et al.* 2005), and interestingly, it is also found in marsupials but not in monotremes. *Peg10* is the first example of a differentially methylated region associated with genomic imprinting in marsupials, and the marsupials differentially methylated region was strictly limited to the 5' region of *Peg10*, unlike the eutherian differentially methylated region, which covers the promoter region of *Peg10* and also the adjacent imprinted gene (Suzuki *et al.* 2007). The authors suggested that the genomic imprinting in eutherians can originate from the repression of exogenous DNA sequences and /or retrotransposons by DNA methylation (Suzuki *et al.* 2007). *Peg11*, also known as *Rtl1* (retrotransposon-like 1) is also a retrotransposon-derived imprinted gene (Sekita *et al.* 2008). Additionally, five murine imprinted genes were reported to have arisen through retrotransposition: *Mcts2*, *Nap1l5*, *U2af1-rs1*, *Inpp5f_v2* and *Peg12* (Cowley & Oakey 2010). Unlike *Peg10*, these five retrogenes are derived from cellular mRNA molecules that did not possess autonomous retrotransposon activity.

However, more recently imprinted genes were also found in chickens. This seems to undermine the 'host defense' hypothesis since chickens possess a low level of repeats and active TEs compared to mammals (Frésard *et al.* 2014).

X chromosome inactivation

X chromosome inactivation (XCI), also called Lyonization, is a well-established phenomenon among placental mammals. By this process one of the copies of the X chromosome in female mammals is inactivated in order to maintain the correct dosage relationship of genes between females (XX) and males (XY). One of the two X chromosomes in females is silenced by packaging into transcriptionally inactive heterochromatin in every cell in the early embryo and remains so in all somatic cells throughout life (Lyon 1972). Unlike other gene-silencing mechanisms, XCI is initiated from an X-inactivation center (XIC) on the X chromosome and spreads throughout the chromosome. Since this type of spreading and inactivation are less efficient in autosomes, it was suggested the X chromosome contains 'way stations' or 'boosters' that act as promoters for spreading (Riggs 1990). Lyon further showed that XCI is

mediated by Xist mRNA, which coats the inactive X chromosome and he proposed, based on the evidence that the X chromosomes of human and mouse are rich in L1s, the hypothesis (Lyon 1998) that L1 elements act as booster to promote the spread of Xist mRNA (shown in

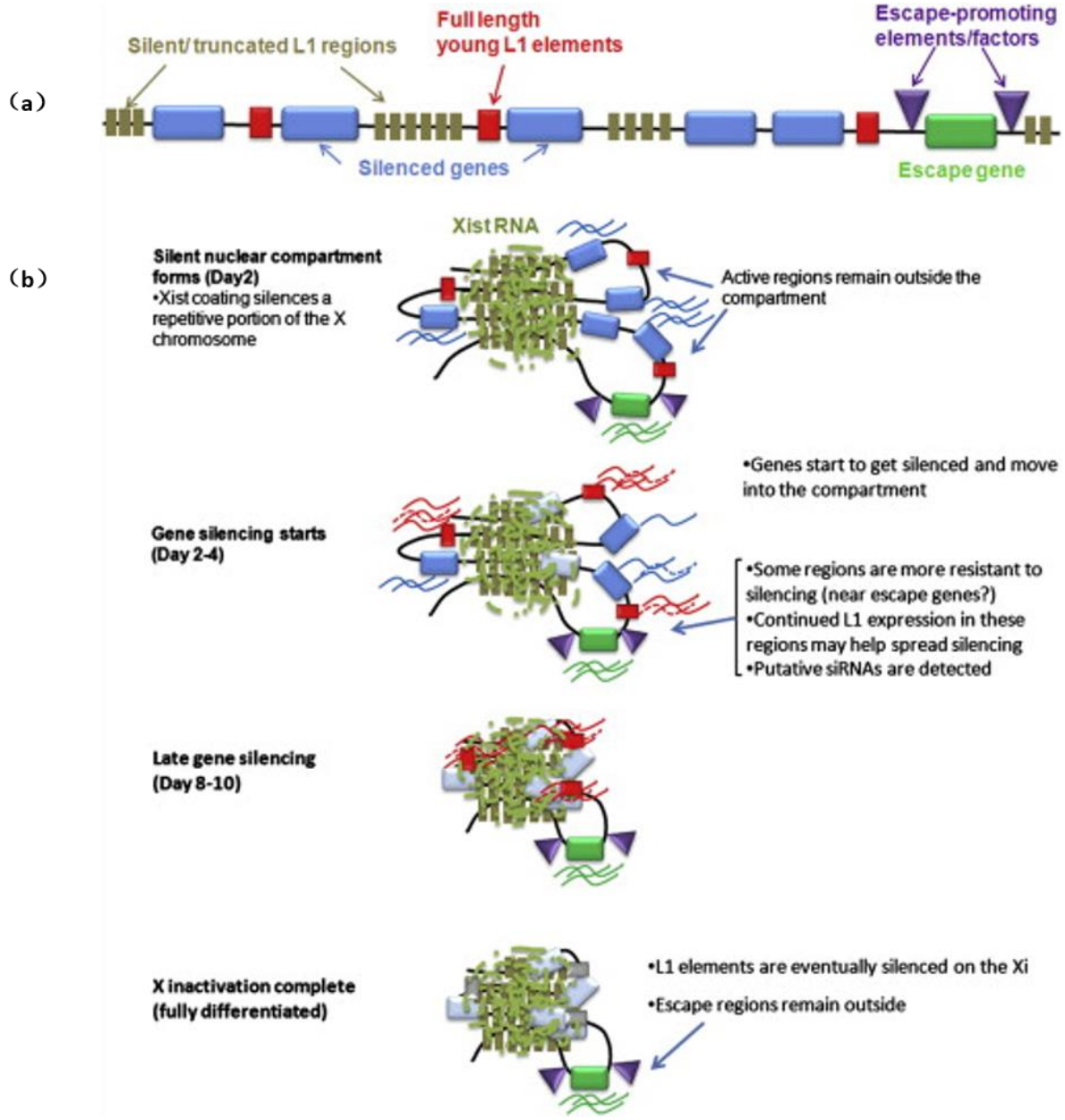


Figure 1.4 Model of LINE participation during XCI. (a) Silent L1s may facilitate assembly of a heterochromatic nuclear compartment early in XCI; (b) transcriptionally active L1 elements may facilitate the spread of silencing into certain regions on the X. Figures from Chow *et al*. (2010).

Figure 1.4). This L1 enrichment was later confirmed by sequence analysis of the human genome. The X chromosome is enriched 2-fold for L1 elements compared to human autosomes (Bailey *et al.* 2000). It was also shown that L1 and L2 are significantly enriched in regions surrounding transcription start sites of genes that are subject to inactivation, while *Alu* and short motifs containing ACG/CGT are significantly enriched in those that escape inactivation (Wang *et al.* 2006). The enrichment of L1 in the mouse X chromosome was confirmed by some other groups (Chow *et al.* 2010; Deng *et al.* 2015), but different results were found by Chureau *et al*. (2002). Using comparative sequence analysis of mouse, human and bovine, they found that human and bovine XIC regions have a high density of L1 with 39% and 46% of sequence length respectively, while the L1 density in mouse XIC region is only 14.5% (Chureau *et al.* 2002). Although more verification and precise elucidation of the relationship between the retrotransposons and XCI is needed, the correlation of repetitive sequences with XCI status cannot be denied (Bailey *et al.* 2000).

1.2.5 ARTs used in bovine embryo production

ARTs including artificial insemination (AI), multiple ovulation and embryo transfer (MOET), *in vitro* embryo production (IVP) and somatic cell nuclear transfer (SCNT) have been developed over decades in bovine production.

AI is the deliberate introduction of sperm into a female's reproductive tract for the purpose of achieving a conception through *in vivo* fertilization rather than sexual intercourse. In this way, it uses the superior genetic merit (the best bull), and prevent exposure to infectious genital diseases. In 1899, the first artificial inseminated cattle were successfully produced by Russian pioneer AI researcher Ilya Ivanovich Ivanov (Webb 1992). Nowadays, this technique is still widely used. MOET uses follicle stimulating hormones (FSH) to induce superovulation, followed by AI to fertilize inside the cow, and then the embryos are transferred into surrogate recipients that are at the same stage of their cycle but not mated. In this way, the merit of genetically superior females can pass to the offspring.

IVP is a newer approach that has seen greatly expanded use for commercial purposes in the last years. Cattle is one the most successful species with the IVP technique, with around

30-40% of inseminated oocytes developing into blastocysts and the birth of innumerable calves worldwide (Rodriguez-Martinez 2012). This technique includes three main steps: *in vitro* oocyte maturation (IVM), i*n vitro* fertilization (IVF) and *in vitro* embryo culture (IVC) (Ward *et al.* 2002). It requires specific laboratory expertise and equipment, and needs to be performed following a strict timing (Van Soom *et al.* 1997). This is the main technique used for research purpose and some problems and the culture condition will be described later.

SCNT is a laboratory strategy for creating a viable embryo from a body cell and an oocyte. Dolly the famous sheep was the first successful case of SCNT, created by transferring a mammary gland cell from an adult sheep to an enucleated oocyte. Cloning by SCNT allows to make a copy of any given animal whose genotype and phenotype are well known. However, SCNT can be inefficient due to, among others, stress placed on both oocyte and the introduced nucleus, resulting in a low percentage of successfully reprogrammed cells. Moreover, with the production of cloned animals becoming more popular, ethical and societal concerns are rising as well.

1.2.6 Influence of IVC on embryo development

IVP of ruminant embryos has become routine and is increasingly available as a commercial service to dairy, meat and wool production. However, the efficiency of producing viable embryos and the development of such embryos are perceived to be inferior to their *in vivo* counterparts. IVP has been reported to alter many embryonic characteristics, such as a darker cytoplasm and a lower buoyant density due to their high lipid content (Pollard & Leibo 1994), a more fragile ZP (Duby *et al.* 1997), differences in intracellular communication and metabolism (Boni *et al.* 1999; Khurana & Niemann 2000), a higher chance of chromosomal abnormalities (mixoploid) (Lonergan *et al.* 2004), a higher apoptotic rate (Gjørret *et al.* 2003), a slower growth rate, a higher thermal sensitivity, and a lower ICM/TE ratio (Van Soom *et al.* 1997) and so on. Furthermore, different gene expression patterns and epigenetic regulation were found in IVP embryos compared to their *in vivo* counterparts.

Influence on gene expression

Many studies have been conducted on the influence of IVP on the gene expression pattern of preimplantation embryos. It was reported that approximately 85% of differentially expressed genes was down-regulated in IVP bovine blastocysts compared to *in vivo* counterparts, most of which are involved in transcriptional and translational events, suggesting that a deficient machinery associated with transcription and translation is behind the inferior quality of IVP embryos (Corcoran *et al.* 2006). Comparing to *in vivo* embryos, genes related to metabolism, growth and differentiation were down-regulated while genes related to stress were up-regulated in IVP bovine embryos. This significant change supports the hypothesis that the IVC system is associated with a considerable amount of (oxidative) stress. Different culture media have a different impact on genes involved in cell communication, differentiation, apoptosis and oxidative stress (Boni *et al.* 1999; Rizos *et al.* 2003). Other culture conditions such as oxygen concentrations were shown to have an impact on gene expression as well (Harvey *et al.* 2004).

Influence on embryo epigenetic change

IVC affects DNA epigenetic patterns of embryos, as was shown in many species. An increased global DNA methylation has been found in IVP rat and mice embryos (Zaitseva *et al.* 2007). In bovine blastocysts, longer exposure to IVC increased DNA methylation profile alterations (Salilew-Wondim *et al.* 2015). IVC was also found to alter gene imprinting. In mouse, loss of methylation in *H19* was reported in embryos cultured in different media (Doherty *et al.* 2000). Serum supplementation in culture medium induced alterations in the DNA methylation pattern of several imprinted genes (*H19*, *Igf2*, *Grb7*, *Grb10* and *Peg1*) in mouse embryos (Velker *et al.* 2012). Large offspring syndrome (LOS), which is associated with loss of methylation at an imprinting region in the *IGF2R* gene, is a well-known phenotype of cattle born after *in vitro* culture of embryos (Young *et al.* 1998). It is characterized by large size at birth, breathing difficulties, reluctance to suckle and sudden perinatal death. A major cause was found in the use of serum in the culture media (Sinclair *et*

al. 1999). The same imprinting disorder has been found in human as well. Beckwith-Wiedemann syndrome (BWS) is an overgrowth disorder usually present at birth, and it shares many features with LOS, such as large tongue, and ear creases and predisposition to childhood cancer. A case control study showed children conceived through IVF have a significant higher chance of BWS (Halliday et al., 2004). Recent studies provide evidence for epigenetic similarities between BWS and LOS, as these syndromes share misregulation of the imprinted genes such as *IGF2R*, *KCNQ1OT1* or *CDKN1C* (Chen *et al.* 2013; Chen *et al.* 2015), which makes LOS a good bovine model of human BWS.

1.2.7 Oxidative stress in early embryo development

IVC and Oxidative stress

As mentioned previously, the suboptimal environment encountered by early embryos during IVC is one big cause of poor blastocyst quality. Under IVC conditions, embryos face a variety of pressures, including heat stress, oxygen tension, light, supplement of serum in the culture media and so on, that impair their further development. All of these have been reported to potentially lead to oxidative stress (Thompson *et al.* 1990; Goto *et al.* 1993; Rizos *et al.* 2003). Oxidative stress is an imbalanced situation of enhanced concentration of oxidants or a lack of antioxidants in cells leading to the presence of reactive oxygen species (ROS) above its biological values. ROS includes superoxide anion (O_2^-) and the hydroxy radical (OH^-). Hydrogen peroxide (H_2O_2) is not a free radical, but usually a product of O_2^- dismutation, catalyzed by the superoxide dismutase. Physiologically, a certain amount of ROS is required for the regulation of several transmembrane signal transduction pathways in cells, and control of cellular fate, like differentiation into specific cell types (Ji *et al.* 2010). ROS act as second messengers by regulating key transcription factors that alter gene expression in embryos. Many of the redox sensitive transcription factors, such as HIF-1, Nf-kB, AP-1, Ref-1 and Nrf-1, are vital to cell signaling pathways that dictate proliferation, differentiation and apoptosis and thus can have a significant role in embryonic development (Dennery 2007). Studies on redox metabolism in murine and bovine oocytes and preimplantation embryos

confirmed that ROS have a key role during IVM, fertilization and embryo development (Tanghe *et al.* 2003; Dumollard *et al.* 2007). Moreover, short-term exposure to H_2O_2 at the end of maturation showed a beneficial effect on subsequent embryo development (Vandaele *et al.* 2010).

On the other hand, high concentration of ROS can induce oxidative modifications of the cell components, thus indirectly causing DNA fragmentation, protein oxidation, lipid peroxidation and mitochondrial damage (Dennery 2007). *In vivo*, the oviduct provides the optimal red/ox environment for an embryo to maintain its oxidative homeostasis by equilibrium between ROS and anti-oxidant defense, which is critical for normal cell division and differentiation. In IVC models, however, an over production of ROS during early embryo development breaks the oxidative homeostasis and affects metabolism, resulting in lower embryo quality (Guerin et al. 2001). For instance, variations in oxygen concentration and temperature have been shown to modulate the rate of ROS production.

Oxygen tension

Oxygen tension is an important regulator of oxidative metabolism. *In vivo* developing embryos are exposed to decreasing oxygen concentration when they travel from oviduct to uterus. An atmosphere with 5% O_2 which is more comparable to the *in vivo* condition, is recommended for *in vitro* embryo culture. Embryos cultured under normal atmospheric oxygen concentration (20% O_2) show more deleterious effects in development and quality in many species, including cattle, horse, pig and human (Yuan *et al.* 2003; Bontekoe *et al.* 2012; Appeltant *et al.* 2015). It was suggested that high oxygen tension increases hydrogen peroxide production, DNA fragmentation, cell apoptosis, and decreases the chance of IVC embryo developing to blastocyst stage (Van Soom *et al.* 2002). The use of 20% O_2 is correlated with increased generation of ROS, and further lead to oxidative stress (Takahashi 2012), and embryos cultured under 20% O_2 have been associated with a great perturbation in global expression profiles (Rinaudo *et al.* 2006).

Influence of oxidative stress on epigenetics

Oxidative stress and DNA methylation are characteristics of various cancer types, therefore the cause-consequence relationship between ROS and DNA methylation was mostly studied in cancer research. ROS-induced oxidative stress is associated with both aberrant hypermethylation of tumor suppressor gene (*TSG*) and global hypomethylation. ROS may also induce site specific hypermethylation by regulation of expression of DNA methyltransferases (DNMTs) or the formation of a new DNMT containing complex (Wu & Ni 2015). ROS may also affect DNA methylation by DNA oxidation or TET mediated hydroxymethylation in cancer (Vanden Berghe 2012). The oxygen level was reported to epigenetically regulate fate switching of neural precursor cells (Mutoh *et al.* 2012). In *in vitro* embryo culture, oxidative stress by means of high O_2 tension alters the expression of epigenome modifying genes such as the ten-eleven translocation (*TET*) gene, which is associated with conversion of 5-methylcytosine to 5-hydroxymethylcytosine at the 16-cell and blastocyst stages of bovine IVF embryos (Burroughs *et al*. 2013). Besides the influence on DNA methylation, ROS was also reported to increase histone acetylation in early embryo development (Chason *et al.* 2011).

Oxidative stress activates retrotransposons

McClintock's (1984) 'genomic shock' hypothesis suggests that environmental stimuli (including oxidative stress) may mobilize transposable elements. This hypothesis is supported by evidence that retrotransposons from various organisms can be transcriptional induced by a variety of environmental stresses, including oxidative stress (Morales *et a*l. 2003). Using H_2O_2 treatment to induce oxidative stress in neuroblastoma cells, L1 expression was found twice as high when compared to control cells (Giorgi et al. 2011). Endogenous L1 expression was also found to be induced by oxidative stress in many other cell types, including vascular smooth muscle cells, ovarian, breast and some somatic carcinomas (Teneng *et al.* 2006). The transcriptional rise of the autonomous retrotransposon L1, potential increases the translation chance of reverse transcriptase, which in turn may facilitate the mobilization of the

transposable elements. SINE elements, such as the human *Alu* and mouse SINE *B2* RNAs were also found transcriptionally increased by cellular stresses (Ichiyanagi 2013).

1.3 References

Adelson D.L., Raison J.M. & Edgar R.C. (2009) Characterization and distribution of retrotransposons and simple sequence repeats in the bovine genome. *Proceedings of the National Academy of Sciences* **106**, 12855-60.

Alexandrova E.A., Olovnikov I.A., Malakhova G.V., Zabolotneva A.A., Suntsova M.V., Dmitriev S.E. & Buzdin A.A. (2012) Sense transcripts originated from an internal part of the human retrotransposon LINE-1 5' UTR. *Gene* **511**, 46-53.

Akopov S.B., Ruda V.M., Batrak V.V., Vetchinova A.S., Chernov I.P., Nikolaev L.G., Bode J. & Sverdlov E.D. (2006) Identification, genome mapping, and CTCF binding of potential insulators within the FXYD5-COX7A1 locus of human Chromosome 19q13. 12. *Mammalian genome* **17**, 1042-9.

Appeltant R., Beek J., Vandenberghe L., Maes D. & Van Soom A. (2015) Increasing the cAMP concentration during in vitro maturation of pig oocytes improves cumulus maturation and subsequent fertilization *in vitro*. *Theriogenology* **83**, 344-52.

Aravin A.A. & Bourc'his D. (2008) Small RNA guides for de novo DNA methylation in mammalian germ cells. *Genes & development* **22**, 970-5.

Aravin A.A., Sachidanandam R., Bourc'his D., Schaefer C., Pezic D., Toth K.F., Bestor T. & Hannon G.J. (2008) A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. *Molecular cell* **31**, 785-99.

Arkhipova I.R. (2006) Distribution and phylogeny of Penelope-like elements in eukaryotes. *Systematic biology* **55**, 875-85.

Asch H.L., Eliacin E., Fanning T.G., Connolly J.L., Bratthauer G. & Asch B.B. (1995) Comparative expression of the LINE-1 p40 protein in human breast carcinomas and normal breast tissues. *Oncology research* **8**, 239-47.

Athanikar J.N., Badge R.M. & Moran J.V. (2004) A YY1-binding site is required for accurate

human LINE-1 transcription initiation. *Nucleic acids research* **32**, 3846-55.

Babich V., Aksenov N., Alexeenko V., Oei S., Buchlow G. & Tomilin N. (1999) Association of some potential hormone response elements in human genes with the Alu family repeats. *Gene* **239**, 341-9.

Babushok D., Ostertag E. & Kazazian Jr H. (2007) Current topics in genome evolution: molecular mechanisms of new gene formation. *Cellular and Molecular Life Sciences* **64**, 542-54.

Bailey J.A., Carrel L., Chakravarti A. & Eichler E.E. (2000) Molecular evidence for a relationship between LINE-1 elements and X chromosome inactivation: the Lyon repeat hypothesis. *Proceedings of the National Academy of Sciences* **97**, 6634-9.

Bailey P. & Cushing H. (1925) Medulloblastoma cerebelli: a common type of midcerebellar glioma of childhood. *Archives of Neurology & Psychiatry* **14**, 192-224.

Bakshi A. & Kim J. (2014) Retrotransposon-based profiling of mammalian epigenomes: DNA methylation of IAP LTRs in embryonic stem, somatic and cancer cells. *Genomics* **104**, 538-44.

Barnes F. & Eyestone W. (1990) Early cleavage and the maternal zygotic transition in bovine embryos. *Theriogenology* **33**, 141-52.

Batzer M.A. & Deininger P.L. (2002) Alu repeats and human genomic diversity. *Nature Reviews Genetics* **3**, 370-9.

Beato M., Chávez S. & Truss M. (1996) Transcriptional regulation by steroid hormones. *Steroids* **61**, 240-51.

Beaujean N., Taylor J., Gardner J., Wilmut I., Meehan R. & Young L. (2004) Effect of limited DNA methylation reprogramming in the normal sheep embryo on somatic cell nuclear transfer. *Biology of reproduction* **71**, 185-93.

Bejerano G., Lowe C.B., Ahituv N., King B., Siepel A., Salama S.R., Rubin E.M., Kent W.J. & Haussler D. (2006) A distal enhancer and an ultraconserved exon are derived from a novel retroposon. *Nature* **441**, 87-90.

Belancio V.P., Hedges D.J. & Deininger P. (2008) Mammalian non-LTR retrotransposons: for better or worse, in sickness and in health. *Genome research* **18**, 343-58.

Belancio V.P., Roy-Engel A.M., Pochampally R.R. & Deininger P. (2010) Somatic expression of LINE-1 elements in human tissues. *Nucleic acids research* **38**, 3909-22.

Beraldi, R., Pittoggi, C., Sciamanna, I., Mattei, E. and Spadafora, C. (2006) Expression of LINE-1 retroposons is essential for murine preimplantation development. *Molecular reproduction and development* **73**, 279-87.

Berger S.L. (2007) The complex language of chromatin regulation during transcription. *Nature* **447**, 407-12.

Bestor T. & Bourc'his D. (2004) Transposon silencing and imprint establishment in mammalian germ cells. In: *Cold Spring Harbor symposia on quantitative biology*, pp. 381-8. Cold Spring Harbor Laboratory Press.

Betteridge K. & Flechon J.-E. (1988) The anatomy and physiology of pre-attachment bovine embryos. *Theriogenology* **29**, 155-87.

Blackwood E.M. & Kadonaga J.T. (1998) Going the distance: a current view of enhancer action. *Science* **281**, 60-3.

Boeke J.D. (1997) LINEs and Alus-the polyA connection. *Nature genetics* **16**, 6-7.

Bogerd H.P., Wiegand H.L., Doehle B.P., Lueders K.K. & Cullen B.R. (2006) APOBEC3A and APOBEC3B are potent inhibitors of LTR-retrotransposon function in human cells. *Nucleic acids research* **34**, 89-95.

Boni R., Tosti E., Roviello S. & Dale B. (1999) Intercellular communication in in vivo-and in vitro-produced bovine embryos. *Biology of reproduction* **61**, 1050-5.

Bontekoe S., Mantikou E., van Wely M., Seshadri S., Repping S. & Mastenbroek S. (2012) Low oxygen concentrations for embryo culture in assisted reproductive technologies. *Cochrane Database of Systematic Reviews* **11,** 7 (CD008950).

Borchert G.M., Lanier W. & Davidson B.L. (2006) RNA polymerase III transcribes human microRNAs. *Nat Struct Mol Biol* **13**, 1097-101.

Bourque G., Leong B., Vega V.B., Chen X., Lee Y.L., Srinivasan K.G., Chew J.-L., Ruan Y., Wei C.-L. & Ng H.H. (2008) Evolution of the mammalian transcription factor binding repertoire via transposable elements. *Genome research* **18**, 1752-62.

Boyle P., Clement K., Gu H., Smith Z.D., Ziller M., Fostel J.L., Holmes L., Meldrim J.,

Kelley F. & Gnirke A. (2012) Gel-free multiplexed reduced representation bisulfite sequencing for large-scale DNA methylation profiling. *Genome Biol* **13**, R92.

Branciforte D. & Martin S.L. (1994) Developmental and cell type specificity of LINE-1 expression in mouse testis: implications for transposition. *Molecular and cellular biology* **14**, 2584-92.

Brennecke J., Aravin A.A., Stark A., Dus M., Kellis M., Sachidanandam R. & Hannon G.J. (2007) Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. *Cell* **128**, 1089-103.

Burwinkel B. & Kilimann M.W. (1998) Unequal homologous recombination between LINE-1 elements as a mutational mechanism in human genetic disease. *Journal of molecular biology* **277**, 513-7.

Buzdin A. (2004) Retroelements and formation of chimeric retrogenes. *Cellular and Molecular Life Sciences CMLS* **61**, 2046-59.

Buzdin A., Gogvadze E. & Lebrun M.-H. (2007) Chimeric retrogenes suggest a role for the nucleolus in LINE amplification. *FEBS letters* **581**, 2877-82.

Byrd K. & Corces V.G. (2003) Visualization of chromatin domains created by the gypsy insulator of Drosophila. *The Journal of cell biology* **162**, 565-74.

Cai H. & Levine M. (1995) Modulation of enhancerá¤-promoter interactions by insulators in the Drosophilaembryo.

Castro-Diaz N., Ecco G., Coluccio A., Kapopoulou A., Yazdanpanah B., Friedli M., Duc J., Jang S.M., Turelli P. & Trono D. (2014) Evolutionally dynamic L1 regulation in embryonic stem cells. *Genes & development* **28**, 1397-409.

Cervera A. & De la Peña M. (2014) Eukaryotic Penelope-like retroelements encode hammerhead ribozyme motifs. *Molecular biology and evolution*, msu232.

Chason R.J., Csokmay J., Segars J.H., DeCherney A.H. & Armant D.R. (2011) Environmental and epigenetic effects upon preimplantation embryo metabolism and development. *Trends in Endocrinology & Metabolism* **22**, 412-20.

Chen C., Ara T. & Gautheret D. (2009) Using Alu elements as polyadenylation sites: A case of retroposon exaptation. *Molecular biology and evolution* **26**, 327-34.

Chen H., Lilley C.E., Yu Q., Lee D.V., Chou J., Narvaiza I., Landau N.R. & Weitzman M.D. (2006) APOBEC3A is a potent inhibitor of adeno-associated virus and retrotransposons. *Current Biology* **16**, 480-5.

Chen L., Wang D., Wu Z., Ma L. & Daley G.Q. (2010) Molecular basis of the first cell fate determination in mouse embryogenesis. *Cell research* **20**, 982-93.

Chen Z., Hagen D.E., Elsik C.G., Ji T., Morris C.J., Moon L.E. & Rivera R.M. (2015) Characterization of global loss of imprinting in fetal overgrowth syndrome induced by assisted reproduction. *Proceedings of the National Academy of Sciences* **112**, 4618-23.

Chen Z., Robbins K.M., Wells K.D. & Rivera R.M. (2013) Large offspring syndrome: a bovine model for the human loss-of-imprinting overgrowth syndrome Beckwith-Wiedemann. *Epigenetics* **8**, 591-601.

Chiu Y.-L., Witkowska H.E., Hall S.C., Santiago M., Soros V.B., Esnault C., Heidmann T. & Greene W.C. (2006) High-molecular-mass APOBEC3G complexes restrict Alu retrotransposition. *Proceedings of the National Academy of Sciences* **103**, 15588-93.

Chow J.C., Ciaudo C., Fazzari M.J., Mise N., Servant N., Glass J.L., Attreed M., Avner P., Wutz A. & Barillot E. (2010) LINE-1 activity in facultative heterochromatin formation during X chromosome inactivation. *Cell* **141**, 956-69.

Chuong E.B., Rumi M.K., Soares M.J. & Baker J.C. (2013) Endogenous retroviruses function as species-specific enhancer elements in the placenta. *Nature genetics* **45**, 325-9.

Chureau C., Prissette M., Bourdet A., Barbe V., Cattolico L., Jones L., Eggen A., Avner P. & Duret L. (2002) Comparative sequence analysis of the X-inactivation center region in mouse, human, and bovine. *Genome research* **12**, 894-908.

Conley A.B., Miller W.J. & Jordan I.K. (2008) Human cis natural antisense transcripts initiated by transposable elements. *Trends in genetics* **24**, 53-6.

Corcoran D., Fair T., Park S., Rizos D., Patel O., Smith G., Coussens P., Ireland J., Boland M. & Evans A. (2006) Suppressed expression of genes involved in transcription and translation in in vitro compared with in vivo cultured bovine embryos. *Reproduction* **131**, 651-60.

Cordaux R., Hedges D.J., Herke S.W. & Batzer M.A. (2006) Estimating the retrotransposition

rate of human Alu elements. *Gene* **373**, 134-7.

Cost G.J. & Boeke J.D. (1998) Targeting of human retrotransposon integration is directed by the specificity of the L1 endonuclease for regions of unusual DNA structure. *Biochemistry* **37**, 18081-93.

Coufal N.G., Garcia-Perez J.L., Peng G.E., Yeo G.W., Mu Y., Lovci M.T., Morell M., O'Shea K.S., Moran J.V. & Gage F.H. (2009) L1 retrotransposition in human neural progenitor cells. *Nature* **460**, 1127-31.

Cowley M. & Oakey R.J. (2010) Retrotransposition and genomic imprinting. *Briefings in functional genomics* **9**, 340-6.

Cullen B.R. (1998) Retroviruses as model systems for the study of nuclear RNA export pathways. *Virology* **249**, 203-10.

de Koning, A. J., Gu, W., Castoe, T. A., Batzer, M. A., & Pollock, D. D. (2011). Repetitive elements may comprise over two-thirds of the human genome. *PLoS Genet* **7(12)**, e1002384.

De Sousa P.A., Watson A.J. & Schultz R.M. (1998) Transient Expression of a Translation Initiation Factor Is Conservatively Associated with Embryonic Gene Activation in Murine and Bovine Embryos. *Biology of reproduction* **59**, 969-77.

Dean W., Santos F., Stojkovic M., Zakhartchenko V., Walter J., Wolf E. & Reik W. (2001) Conservation of methylation reprogramming in mammalian development: aberrant reprogramming in cloned embryos. *Proceedings of the National Academy of Sciences* **98**, 13734-8.

Deininger P.L. & Batzer M.A. (2002) Mammalian retroelements. *Genome research* **12**, 1455-65.

Deng X., Ma W., Ramani V., Hill A., Yang F., Ay F., Berletch J.B., Blau C.A., Shendure J. & Duan Z. (2015) Bipartite structure of the inactive mouse X chromosome. *Genome biology* **16**, 1-21.

Dennery P.A. (2007) Effects of oxidative stress on embryonic development. *Birth Defects Research Part C: Embryo Today: Reviews* **81**, 155-62.

Deshmukh R.S., Østrup O., Østrup E., Vejlsted M., Niemann H., Lucas-Hahn A., Petersen B.,

Li J., Callesen H. & Hyttel P. (2011) DNA methylation in porcine preimplantation embryos developed in vivo and produced by in vitro fertilization, parthenogenetic activation and somatic cell nuclear transfer. *Epigenetics* **6**, 177-87.

Dewannieux M., Esnault C. & Heidmann T. (2003) LINE-mediated retrotransposition of marked Alu sequences. *Nature genetics* **35**, 41-8.

Dobbs K.B., Rodriguez M., Sudano M.J., Ortega M.S. & Hansen P.J. (2013) Dynamics of DNA Methylation during Early Development of the Preimplantation Bovine Embryo. *PLOS ONE* **8**, e66230.

Doherty A.S., Mann M.R., Tremblay K.D., Bartolomei M.S. & Schultz R.M. (2000) Differential effects of culture on imprinted H19 expression in the preimplantation mouse embryo. *Biology of reproduction* **62**, 1526-35.

Dorsett D. (1993) Distance-independent inactivation of an enhancer by the suppressor of Hairy-wing DNA-binding protein of Drosophila. *Genetics* **134**, 1135-44.

Duby R., Hill J., O'Callaghan D., Overstrom E. & Boland M. (1997) Changes induced in the bovine zona pellucida by ovine and bovine oviducts. *Theriogenology* **1**, 332.

Dumollard R., Ward Z., Carroll J. & Duchen M.R. (2007) Regulation of redox metabolism in the mouse oocyte and embryo. *Development* **134**, 455-65.

Duranthon V., Watson A.J. & Lonergan P. (2008) Preimplantation embryo programming: transcription, epigenetics, and culture environment. *Reproduction* **135**, 141-50.

Ebert M.S., Neilson J.R. & Sharp P.A. (2007) MicroRNA sponges: competitive inhibitors of small RNAs in mammalian cells. *Nature methods* **4**, 721-6.

Ecco G., Cassano M., Kauzlaric A., Duc J., Coluccio A., Offner S., Imbeault M., Rowe H.M., Turelli P. & Trono D. (2016) Transposable elements and their KRAB-ZFP controllers regulate gene expression in adult tissues. *Developmental cell* **36**, 611-23.

Ehrlich M. (2002) DNA methylation in cancer: too much, but also too little. *Oncogene* **21**, 5400-13.

Ergün S., Buschmann C., Heukeshoven J., Dammann K., Schnieders F., Lauke H., Chalajour F., Kilic N., Strätling W.H. & Schumann G.G. (2004) Cell type-specific expression of LINE-1 open reading frames 1 and 2 in fetal and adult human tissues. *Journal of*

Biological Chemistry **279**, 27753-63.

Esnault C., Heidmann O., Delebecque F., Dewannieux M., Ribet D., Hance A.J., Heidmann T. & Schwartz O. (2005) APOBEC3G cytidine deaminase inhibits retrotransposition of endogenous retroviruses. *Nature* **433**, 430-3.

Esnault C., Maestre J. & Heidmann T. (2000) Human LINE retrotransposons generate processed pseudogenes. *Nature genetics* **24**, 363-7.

Evgen'ev M.B., Zelentsova H., Shostak N., Kozitsina M., Barskyi V., Lankenau D.-H. & Corces V.G. (1997) Penelope, a new family of transposable elements and its possible role in hybrid dysgenesis in Drosophila virilis. *Proceedings of the National Academy of Sciences* **94**, 196-201.

Evsikov A., De Vries W., Peaston A., Radford E., Fancher K., Chen F., Blake J., Bult C., Latham K. & Solter D. (2004) Systems biology of the 2-cell mouse embryo. *Cytogenetic and genome research* **105**, 240-50.

Ewing A.D. & Kazazian H.H. (2010) High-throughput sequencing reveals extensive variation in human-specific L1 content in individual human genomes. *Genome research* **20**, 1262-70.

Eyestone W. & First N. (1988) Cell cycle analysis of early bovine embryos. *Theriogenology* **29**, 243.

Fadloun A., Le Gras S., Jost B., Ziegler-Birling C., Takahashi H., Gorab E., Carninci P. & Torres-Padilla M.E. (2013) Chromatin signatures and retrotransposon profiling in mouse embryos reveal regulation of LINE-1 by RNA. *Nature Structural & Molecular Biology* **20**, 332-8.

Faulkner G.J., Kimura Y., Daub C.O., Wani S., Plessy C., Irvine K.M., Schroder K., Cloonan N., Steptoe A.L. & Lassmann T. (2009) The regulated retrotransposon transcriptome of mammalian cells. *Nature genetics* **41**, 563-71.

Finnegan D.J. (1989) Eukaryotic transposable elements and genome evolution. *Trends in genetics* **5**, 103-7.

Frésard L., Leroux S., Servin B., Gourichon D., Dehais P., San Cristobal M., Marsaud N., Vignoles F., Bed'Hom B. & Coville J.-L. (2014) Transcriptome-wide investigation of

genomic imprinting in chicken. *Nucleic acids research* **42**, 3768-82.

Friedli M., Turelli P., Kapopoulou A., Rauwel B., Castro-Díaz N., Rowe H.M., Ecco G., Unzu C., Planet E. & Lombardo A. (2014) Loss of transcriptional control over endogenous retroelements during reprogramming to pluripotency. *Genome research* **24**, 1251-9.

Fulka H., Mrazek M., Tepla O. & Fulka J. (2004) DNA methylation pattern in human zygotes and developing embryos. *Reproduction* **128**, 703-8.

Garcia-Perez J.L., Marchetto M.C., Muotri A.R., Coufal N.G., Gage F.H., O'shea K.S. & Moran J.V. (2007) LINE-1 retrotransposition in human embryonic stem cells. *Human molecular genetics* **16**, 1569-77.

Gdula D.A., Gerasimova T.I. & Corces V.G. (1996) Genetic and molecular analysis of the gypsy chromatin insulator of Drosophila. *Proceedings of the National Academy of Sciences* **93**, 9378-83.

Gentles A.J., Wakefield M.J., Kohany O., Gu W., Batzer M.A., Pollock D.D. & Jurka J. (2007) Evolutionary dynamics of transposable elements in the short-tailed opossum Monodelphis domestica. *Genome research* **17**, 992-1004.

Georgiou I., Noutsopoulos D., Dimitriadou E., Markopoulos G., Apergi A., Lazaros L., Vaxevanoglou T., Pantos K., Syrrou M. & Tzavaras T. (2009) Retrotransposon RNA expression and evidence for retrotransposition events in human oocytes. *Human molecular genetics* **18**, 1221-8.

Glinsky G.V. (2015) Transposable elements and DNA methylation create in embryonic stem cells human-specific regulatory sequences associated with distal enhancers and noncoding RNAs. *Genome biology and evolution* **7**, 1432-54.

Gifford W.D., Pfaff S.L. & Macfarlan T.S. (2013) Transposable elements as genetic regulatory substrates in early development. *Trends in cell biology* **23**, 218-26.

Giorgi G., Marcantonio P. & Del Re B. (2011) LINE-1 retrotransposition in human neuroblastoma cells is affected by oxidative stress. *Cell and tissue research* **346**, 383-91.

Gjørret J.O., Knijn H.M., Dieleman S.J., Avery B., Larsson L.-I. & Maddox-Hyttel P. (2003) Chronology of apoptosis in bovine embryos produced in vivo and in vitro. *Biology of*

reproduction **69**, 1193-200.

Gogvadze E. & Buzdin A. (2009) Retroelements and their impact on genome evolution and functioning. *Cellular and Molecular Life Sciences* **66**, 3727-42.

Gogvadze E., Stukacheva E., Buzdin A. & Sverdlov E. (2009) Human-specific modulation of transcriptional activity provided by endogenous retroviral insertions. *Journal of virology* **83**, 6098-105.

Goodchild N.L., Wilkinson D.A. & Mager D.L. (1992) A human endogenous long terminal repeat provides a polyadenylation signal to a novel, alternatively spliced transcript in normal placenta. *Gene* **121**, 287-94.

Goodwin T.J. & Poulter R.T. (2001) The DIRS1 group of retrotransposons. *Molecular biology and evolution* **18**, 2067-82.

Goodwin T.J. & Poulter R.T. (2004) A new group of tyrosine recombinase-encoding retrotransposons. *Molecular biology and evolution* **21**, 746-59.

Goto Y., Noda Y., Mori T. & Nakano M. (1993) Increased generation of reactive oxygen species in embryos cultured in vitro. *Free Radical Biology and Medicine* **15**, 69-75.

Graf A., Krebs S., Heininen-Brown M., Zakhartchenko V., Blum H. & Wolf E. (2014) Genome activation in bovine embryos: review of the literature and new insights from RNA sequencing experiments. *Animal reproduction science* **149**, 46-58.

Grimson A., Srivastava M., Fahey B., Woodcroft B.J., Chiang H.R., King N., Degnan B.M., Rokhsar D.S. & Bartel D.P. (2008) Early origins and evolution of microRNAs and Piwi-interacting RNAs in animals. *Nature* **455**, 1193-7.

Guerrero-Bosagna C. (2014) DNA Methylation Research Methods. *Materials and methods* **3**, 206-13.

Guerin P., El Mouatassim S. & Menezo Y. (2001) Oxidative stress and protection against reactive oxygen species in the pre-implantation embryo and its surroundings. *Human Reproduction Update* **7**, 175-89.

Guo F., Li X., Liang D., Li T., Zhu P., Guo H., Wu X., Wen L., Gu T.-P. & Hu B. (2014) Active and passive demethylation of male and female pronuclear DNA in the mammalian zygote. *Cell stem cell* **15**, 447-58.

Halliday J, Oke K, Breheny S, Algar E & Amor D. (2004) Beckwith-Wiedemann syndrome and IVF: a case-control study. *American Journal of Human Genetics* **75**,526-8.

Hamatani T., Carter M.G., Sharov A.A. & Ko M.S. (2004) Dynamics of global gene expression changes during mouse preimplantation development. *Developmental cell* **6**, 117-31.

Hambor J., Mennone J., Coon M., Hanke J. & Kavathas P. (1993) Identification and characterization of an Alu-containing, T-cell-specific enhancer located in the last intron of the human CD8 alpha gene. *Molecular and cellular biology* **13**, 7056-70.

Hamdorf M., Idica A., Zisoulis D.G., Gamelin L., Martin C., Sanders K.J. & Pedersen I.M. (2015) miR-128 represses L1 retrotransposition by binding directly to L1 RNA. *Nat Struct Mol Biol* **22**, 824-31.

Han J.S., Szak S.T. & Boeke J.D. (2004) Transcriptional disruption by the L1 retrotransposon and implications for mammalian transcriptomes. *Nature* **429**, 268-74.

Harvey A., Kind K., Pantaleon M., Armstrong D. & Thompson J. (2004) Oxygen-regulated gene expression in bovine blastocysts. *Biology of reproduction* **71**, 1108-19.

Hatada S., Grant D.J. & Maeda N. (2003) An intronic endogenous retrovirus-like sequence attenuates human haptoglobin-related gene expression in an orientation-dependent manner. *Gene* **319**, 55-63.

Havecker E.R., Gao X. & Voytas D.F. (2004) The diversity of LTR retrotransposons. *Genome biology* **5**, 225-.

Heidmann T., Heidmann O. & Nicolas J.-F. (1988) An indicator gene to demonstrate intracellular transposition of defective retroviruses. *Proceedings of the National Academy of Sciences* **85**, 2219-23.

Heldmann O. & Heidmann T. (1991) Retrotransposition of a mouse IAP sequence tagged with an indicator gene. *Cell* **64**, 159-70.

Heras S., Forier K., Rombouts K., Braeckmans K., & Van Soom A. (2014). DNA counterstaining for methylation and hydroxymethylation immunostaining in bovine zygotes. *Analytical biochemistry* **454**, 14-6.

Heras S.R., Macias S., Cáceres J.F. & Garcia-Perez J.L. (2014) Control of mammalian

retrotransposons by cellular RNA processing activities. *Mobile genetic elements* **4**, 1173-81.

Heras S.R., Macias S., Plass M., Fernandez N., Cano D., Eyras E., Garcia-Perez J.L. & Cáceres J.F. (2013) The Microprocessor controls the activity of mammalian retrotransposons. *Nat Struct Mol Biol* **20**, 1173-81.

Hohjoh H. & Singer M.F. (1997) Sequence-specific single-strand RNA binding protein encoded by the human LINE-1 retrotransposon. *The EMBO journal* **16**, 6034-43.

Huang C.R.L., Schneider A.M., Lu Y., Niranjan T., Shen P., Robinson M.A., Steranka J.P., Valle D., Civin C.I. & Wang T. (2010) Mobile interspersed repeats are major structural variants in the human genome. *Cell* **141**, 1171-82.

Huang Y. & Steitz J.A. (2001) Splicing factors SRp20 and 9G8 promote the nucleocytoplasmic export of mRNA. *Molecular cell* **7**, 899-905.

Hulme A.E., Bogerd H.P., Cullen B.R. & Moran J.V. (2007) Selective inhibition of Alu retrotransposition by APOBEC3G. *Gene* **390**, 199-205.

Hunter R. (1993) Sperm: egg ratios and putative molecular signals to modulate gamete interactions in polytocous mammals. *Molecular reproduction and development* **35**, 324-7.

Hyttel P., Greve T. & Callesen H. (1988) Ultrastructure of in-vivo fertilization in superovulated cattle. *Journal of reproduction and fertility* **82**, 1-13.

Ichiyanagi K. (2013) Epigenetic regulation of transcription and possible functions of mammalian short interspersed elements, SINEs. *Genes & genetic systems* **88**, 19-29.

Iskow R.C., McCabe M.T., Mills R.E., Torene S., Pittard W.S., Neuwald A.F., Van Meir E.G., Vertino P.M. & Devine S.E. (2010) Natural mutagenesis of human genomes by endogenous retrotransposons. *Cell* **141**, 1253-61.

Ji A.-R., Ku S.-Y., Cho M.S., Kim Y.Y., Kim Y.J., Oh S.K., Kim S.H., Moon S.Y. & Choi Y.M. (2010) Reactive oxygen species enhance differentiation of human embryonic stem cells into mesendodermal lineage. *Experimental & molecular medicine* **42**, 175-86.

Johnson E.J., Sizemore R.C. & Gottlieb A.A. (1995) HIV retrotransposon activity and the immunopathogenesis of AIDS. *Trends in microbiology* **3**, 115-7.

Jordan I.K., Rogozin I.B., Glazko G.V. & Koonin E.V. (2003) Origin of a substantial fraction of human regulatory sequences from transposable elements. *Trends in genetics* **19**, 68-72.

Jurka J. (1997) Sequence patterns indicate an enzymatic involvement in integration of mammalian retroposons. *Proceedings of the National Academy of Sciences* **94**, 1872-7.

Jurka J., Kapitonov V.V., Kohany O. & Jurka M.V. (2007) Repetitive sequences in complex genomes: structure and evolution. *Annu. Rev. Genomics Hum. Genet.* **8**, 241-59.

Kamp C., Hirschmann P., Voss H., Huellen K. & Vogt P.H. (2000) Two long homologous retroviral sequence blocks in proximal Yq11 cause AZFa microdeletions as a result of intrachromosomal recombination events. *Human molecular genetics* **9**, 2563-72.

Kano H., Godoy I., Courtney C., Vetter M.R., Gerton G.L., Ostertag E.M. & Kazazian H.H. (2009) L1 retrotransposition occurs mainly in embryogenesis and creates somatic mosaicism. *Genes & development* **23**, 1303-12.

Katoh I., Mírová A., Kurata S.-i., Murakami Y., Horikawa K., Nakakuki N., Sakai T., Hashimoto K., Maruyama A. & Yonaga T. (2011) Activation of the long terminal repeat of human endogenous retrovirus K by melanoma-specific transcription factor MITF-M. *Neoplasia* **13**, 1081-IN42.

Kazazian H.H. (2004) Mobile elements: drivers of genome evolution. *Science* **303**, 1626-32.

Kazazian H.H., Wong C., Youssoufian H., Scott A.F., Phillips D.G. & Antonarakis S.E. (1988) Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man.

Kellum R. & Schedl P. (1991) A position-effect assay for boundaries of higher order chromosomal domains. *Cell* **64**, 941-50.

Khatua A.K., Taylor H.E., Hildreth J.E.K. & Popik W. (2010) Inhibition of LINE-1 and Alu retrotransposition by exosomes encapsidating APOBEC3G and APOBEC3F. *Virology* **400**, 68-75.

Khurana N.K. & Niemann H. (2000) Energy metabolism in preimplantation bovine embryos derived in vitro or in vivo. *Biology of reproduction* **62**, 847-56.

Kimberland M.L., Divoky V., Prchal J., Schwahn U., Berger W. & Kazazian H.H. (1999)

Full-length human L1 insertions retain the capacity for high frequency retrotransposition in cultured cells. *Human molecular genetics* **8**, 1557-60.

Kjellman C., Sjögren H. O., Salford L.G. & Widegren B. (1999) HERV-F (XA34) is a full-length human endogenous retrovirus expressed in placental and fetal tissues. *Gene* **239**, 99-107.

Kordiš D. & Gubenšek F. (1998) The Bov-B lines found in Vipera ammodytes toxic PLA 2 genes are widespread in snake genomes. *Toxicon* **36**, 1585-90.

Kostyuchenko M., Savitskaya E., Volkov I., Golovnin A. & Georgiev P. (2008) Study of functional interaction between three copies of the insulator from the MDG4 transposable element in the model system of the miniwhite gene of Drosophila melanogaster. In: *Doklady Biochemistry and Biophysics*, pp. 239-43. Springer.

Kuramochi-Miyagawa S., Watanabe T., Gotoh K., Totoki Y., Toyoda A., Ikawa M., Asada N., Kojima K., Yamaguchi Y. & Ijiri T.W. (2008) DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. *Genes & development* **22**, 908-17.

Kurose K., Hata K., Hattori M. & Sakaki Y. (1995) RNA polymerase III dependence of the human L1 promoter and possible participation of the RNA polymerase II factor YY1 in the RNA polymerase III transcription system. *Nucleic acids research* **23**, 3704-9.

Laird P.W. (2010) Principles and challenges of genome-wide DNA methylation analysis. *Nature Reviews Genetics* **11**, 191-203.

Lander E.S., Linton L.M., Birren B., Nusbaum C., Zody M.C., Baldwin J., Devon K., Dewar K., Doyle M. & FitzHugh W. (2001) Initial sequencing and analysis of the human genome. *Nature* **409**, 860-921.

Landriscina M., Fabiano A., Altamura S., Bagalà C., Piscazzi A., Cassano A., Spadafora C., Giorgino F., Barone C. & Cignarelli M. (2005) Reverse transcriptase inhibitors down-regulate cell proliferation in vitro and in vivo and restore thyrotropin signaling and iodine uptake in human thyroid anaplastic carcinoma. *The Journal of Clinical Endocrinology & Metabolism* **90**, 5663-71.

Laperriere D., Wang T.-T., White J.H. & Mader S. (2007) Widespread Alu repeat-driven

expansion of consensus DR2 retinoic acid response elements during primate evolution. *BMC genomics* **8**, 23.

Lee E., Iskow R., Yang L., Gokcumen O., Haseley P., Luquette L.J., Lohr J.G., Harris C.C., Ding L. & Wilson R.K. (2012) Landscape of somatic retrotransposition in human cancers. *Science* **337**, 967-71.

Lee H.J., Hore T.A. & Reik W. (2014) Reprogramming the methylome: erasing memory and creating diversity. *Cell stem cell* **14**, 710-9.

Lee J.Y., Ji Z. & Tian B. (2008) Phylogenetic analysis of mRNA polyadenylation sites reveals a role of transposable elements in evolution of the 3′-end of genes. *Nucleic acids research* **36**, 5581-90.

Leibold D.M., Swergold G.D., Singer M.F., Thayer R.E., Dombroski B.A. & Fanning T. (1990) Translation of LINE-1 DNA elements in vitro and in human cells. *Proceedings of the National Academy of Sciences* **87**, 6990-4.

Lev-Maor G., Ram O., Kim E., Sela N., Goren A., Levanon E.Y. & Ast G. (2008) Intronic Alu s influence alternative splicing. *PLOS Genet* **4**, e1000204.

Lim A.K., Tao L. & Kai T. (2009) piRNAs mediate posttranscriptional retroelement silencing and localization to pi-bodies in the Drosophila germline. *The Journal of cell biology* **186**, 333-42.

Lin L., Shen S., Tye A., Cai J.J., Jiang P., Davidson B.L. & Xing Y. (2008) Diverse splicing patterns of exonized Alu elements in human tissues. *PLOS Genet* **4**, e1000225.

Liu W.-M. & Schmid C.W. (1993) Proposed roles for DNA methylation in Alu transcriptional repression and mutational inactivation. *Nucleic acids research* **21**, 1351-9.

Lonergan P., Pedersen H.G., Rizos D., Greve T., Thomsen P.D., Fair T., Evans A. & Boland M.P. (2004) Effect of the post-fertilization culture environment on the incidence of chromosome aberrations in bovine blastocysts. *Biology of reproduction* **71**, 1096-100.

Long Q., Bengra C., Li C., Kutlar F. & Tuan D. (1998) A long terminal repeat of the human endogenous retrovirus ERV-9 is located in the 5′ boundary area of the human β-globin locus control region. *Genomics* **54**, 542-55.

López-Flores I. & Garrido-Ramos M. (2012) The repetitive DNA content of eukaryotic

genomes.

Luan D.D., Korman M.H., Jakubczak J.L. & Eickbush T.H. (1993) Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. *Cell* **72**, 595-605.

Lunyak V.V., Prefontaine G.G., Núñez E., Cramer T., Ju B.-G., Ohgi K.A., Hutt K., Roy R., García-Díaz A. & Zhu X. (2007) Developmentally regulated activation of a SINE B2 repeat as a domain boundary in organogenesis. *Science* **317**, 248-51.

Lyon M.F. (1972) X-chromosome inactivation and developmental patterns in mammals. *Biological Reviews* **47**, 1-35.

Lyon M.F. (1998) X-chromosome inactivation: a repeat hypothesis. *Cytogenetic and genome research* **80**, 133-7.

Maddox-Hyttel P., Alexopoulos N., Vajta G., Lewis I., Rogers P., Cann L., Callesen H., Tveden-Nyborg P. & Trounson A. (2003) Immunohistochemical and ultrastructural characterization of the initial post-hatching development of bovine embryos. *Reproduction* **125**, 607-23.

Maestre J., Tchenio T., Dhellin O. & Heidmann T. (1995) mRNA retroposition in human cells: processed pseudogene formation. *The EMBO journal* **14**, 6333.

Mager D.L., Hunter D.G., Schertzer M. & Freeman J.D. (1999) Endogenous retroviruses provide the primary polyadenylation signal for two new human genes (HHLA2 and HHLA3). *Genomics* **59**, 255-63.

Magin C., Löwer R. & Löwer J. (1999) cORF and RcRE, the Rev/Rex and RRE/RxRE homologues of the human endogenous retrovirus family HTDV/HERV-K. *Journal of virology* **73**, 9496-507.

Makalowski W. (2000) Genomic scrap yard: how genomes utilize all that junk. *Gene* **259**, 61-7.

Makałowski W., Mitchell G.A. & Labuda D. (1994) Alu sequences in the coding regions of mRNA: a source of protein variability. *Trends in genetics* **10**, 188-93.

Malik H.S. & Eickbush T.H. (2001) Phylogenetic analysis of ribonuclease H domains suggests a late, chimeric origin of LTR retrotransposable elements and retroviruses.

Genome research **11**, 1187-97.

Malki S., van der Heijden G.W., O'Donnell K.A., Martin S.L. & Bortvin A. (2014) A role for retrotransposon LINE-1 in fetal oocyte attrition in mice. *Developmental cell* **29**, 521-33.

Mangiacasale R., Pittoggi C., Sciamanna I., Careddu A., Mattei E., Lorenzini R., Travaglini L., Landriscina M., Barone C. & Nervi C. (2003) Exposure of normal and transformed cells to nevirapine, a reverse transcriptase inhibitor, reduces cell growth and promotes differentiation. *Oncogene* **22**, 2750-61.

Marchetto M.C., Narvaiza I., Denli A.M., Benner C., Lazzarini T.A. & Nathanson J.L. (2013) Differential L1 regulation in pluripotent stem cells of humans and apes. *Nature* **503**.

Martens J.H., O'Sullivan R.J., Braunschweig U., Opravil S., Radolf M., Steinlein P. & Jenuwein T. (2005) The profile of repeat-associated histone lysine methylation states in the mouse epigenome. *The EMBO journal* **24**, 800-12.

Mathias S.L., Scott A.F., Kazazian H.H., Boeke J.D. & Gabriel A. (1991) Reverse transcriptase encoded by a human transposable element. *Science* **254**, 1808-10.

Mätlik K., Redik K. & Speek M. (2006) L1 antisense promoter drives tissue-specific transcription of human genes. *BioMed Research International* **2006**.

Mayer W., Niveleau A., Walter J., Fundele R. & Haaf T. (2000) Embryogenesis: demethylation of the zygotic paternal genome. *Nature* **403**, 501-2.

McClintock B. (1983) The significance of responses of the genome to challenge. *Physiology Or Medicine Literature Peace Economic Sciences*, 180.

McManus M.T. & Sharp P.A. (2002) Gene silencing in mammals by small interfering RNAs. *Nature Reviews Genetics* **3**, 737-47.

Mercer T.R. & Mattick J.S. (2013) Structure and function of long noncoding RNAs in epigenetic regulation. *Nat Struct Mol Biol* **20**, 300-7.

Messerschmidt D.M., Knowles B.B. & Solter D. (2014) DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos. *Genes & development* **28**, 812-28.

Miki Y., Nishisho I., Horii A., Miyoshi Y., Utsunomiya J., Kinzler K.W., Vogelstein B. &

Nakamura Y. (1992) Disruption of the APC gene by a retrotransposal insertion of L1 sequence in a colon cancer. *Cancer research* **52**, 643-5.

Moran J.V., Holmes S.E., Naas T.P., DeBerardinis R.J., Boeke J.D. & Kazazian H.H. (1996) High frequency retrotransposition in cultured mammalian cells. *Cell* **87**, 917-27.

Muckenfuss H., Hamdorf M., Held U., Perković M., Löwer J., Cichutek K., Flory E., Schumann G.G. & Münk C. (2006) APOBEC3 proteins inhibit human LINE-1 retrotransposition. *Journal of Biological Chemistry* **281**, 22161-72.

Muotri A.R., Chu V.T., Marchetto M.C., Deng W., Moran J.V. & Gage F.H. (2005) Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. *Nature* **435**, 903-10.

Mutoh T., Sanosaka T., Ito K. & Nakashima K. (2012) Oxygen Levels Epigenetically Regulate Fate Switching of Neural Precursor Cells via Hypoxia-Inducible Factor 1α-Notch Signal Interaction in the Developing Brain. *Stem Cells* **30**, 561-9.

Naas T.P., DeBerardinis R.J., Moran J.V., Ostertag E.M., Kingsmore S.F., Seldin M.F., Hayashizaki Y., Martin S.L. & Kazazian H.H. (1998) An actively retrotransposing, novel subfamily of mouse L1 elements. *The EMBO journal* **17**, 590-7.

Nabirochkin S., Ossokina M. & Heidmann T. (1998) A nuclear matrix/scaffold attachment region co-localizes with the gypsy retrotransposon insulator sequence. *Journal of Biological Chemistry* **273**, 2473-9.

Nangia-Makker P., Sarvis R., Visscher D.W., Bailey-Penrod J., Raz A. & Sarkar F.H. (1998) Galectin-3 and L1 retrotransposons in human breast carcinomas. *Breast cancer research and treatment* **49**, 171-83.

Nappi F., Schneider R., Zolotukhin A., Smulevitch S., Michalowski D., Bear J., Felber B.K. & Pavlakis G.N. (2001) Identification of a novel posttranscriptional regulatory element by using a rev-and RRE-mutated human immunodeficiency virus type 1 DNA proviral clone as a molecular trap. *Journal of virology* **75**, 4558-69.

Narita N., Nishio H., Kitoh Y., Ishikawa Y., Ishikawa Y., Minami R., Nakamura H. & Matsuo M. (1993) Insertion of a 5'truncated L1 element into the 3'end of exon 44 of the dystrophin gene resulted in skipping of the exon during splicing in a case of Duchenne

muscular dystrophy. *Journal of Clinical Investigation* **91**, 1862.

Närvä E., Rahkonen N., Emani M.R., Lund R., Pursiheimo J.P., Nästi J., Autio R., Rasool O., Denessiouk K. & Lähdesmäki H. (2012) RNA-Binding Protein L1TD1 Interacts with LIN28 via RNA and is Required for Human Embryonic Stem Cell Self-Renewal and Cancer Cell Proliferation. *Stem Cells* **30**, 452-60.

Neumann P., Požárková D. & Macas J. (2003) Highly abundant pea LTR retrotransposon Ogre is constitutively transcribed and partially spliced. *Plant molecular biology* **53**, 399-410.

Noutsopoulos D., Vartholomatos G., Kolaitis N. & Tzavaras T. (2006) SV40 large T antigen up-regulates the retrotransposition frequency of viral-like 30 elements. *Journal of molecular biology* **361**, 450-61.

Oakeley E.J. (1999) DNA methylation analysis: a review of current methodologies. *Pharmacology & therapeutics* **84**, 389-400.

Ohshima K., Hamada M., Terai Y. & Okada N. (1996) The 3'ends of tRNA-derived short interspersed repetitive elements are derived from the 3'ends of long interspersed repetitive elements. *Molecular and cellular biology* **16**, 3756-64.

Ohshima K. & Okada N. (2005) SINEs and LINEs: symbionts of eukaryotic genomes with a common tail. *Cytogenetic and genome research* **110**, 475-90.

Ohshima K. (2013) RNA-mediated gene duplication and retroposons: retrogenes, LINEs, SINEs, and sequence specificity. *International journal of evolutionary biology* **2013**.

Okada N. & Hamada M. (1997) The 3′ ends of tRNA-derived SINEs originated from the 3’ends of LINEs: A new example from the bovine genome. *Journal of Molecular Evolution* **44**, S52-S6.

Okano M., Bell D.W., Haber D.A. & Li E. (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**, 247-57.

Okano M., Takebayashi S., Okumura K. & Li E. (1999) Assignment of cytosine-5 DNA methyltransferases Dnmt3a and Dnmt3b to mouse chromosome bands 12A2–A3 and 2H1 by in situ hybridization. *Cytogenetic and genome research* **86**, 333-4.

Ono M., Kawakami M. & Ushikubo H. (1987) Stimulation of expression of the human endogenous retrovirus genome by female steroid hormones in human breast cancer cell line T47D. *Journal of virology* **61**, 2059-62.

Ono R., Nakamura K., Inoue K., Naruse M., Usami T., Wakisaka-Saito N., Hino T., Suzuki-Migishima R., Ogonuki N. & Miki H. (2005) Deletion of Peg10, an imprinted gene acquired from a retrotransposon, causes early embryonic lethality. *Nature genetics* **38**, 101-6.

Oricchio E., Sciamanna I., Beraldi R., Tolstonog G., Schumann G. & Spadafora C. (2007) Distinct roles for LINE-1 and HERV-K retroelements in cell proliferation, differentiation and tumor progression. *Oncogene* **26**, 4226-33.

Ostertag E.M., DeBerardinis R.J., Goodier J.L., Zhang Y., Yang N., Gerton G.L. & Kazazian H.H. (2002) A mouse model of human L1 retrotransposition. *Nature genetics* **32**, 655-60.

Ostertag E.M., Goodier J.L., Zhang Y. & Kazazian H.H. (2003) SVA elements are nonautonomous retrotransposons that cause disease in humans. *The American Journal of Human Genetics* **73**, 1444-51.

Ostertag E.M. & Kazazian Jr H.H. (2001) Biology of mammalian L1 retrotransposons. *Annual review of genetics* **35**, 501-38.

Ostertag E.M., Prak E.T.L., DeBerardinis R.J., Moran J.V. & Kazazian Jr H.H. (2000) Determination of L1 retrotransposition kinetics in cultured cells. *Nucleic acids research* **28**, 1418-23.

Oswald J., Engemann S., Lane N., Mayer W., Olek A., Fundele R., Dean W., Reik W. & Walter J. (2000) Active demethylation of the paternal genome in the mouse zygote. *Current Biology* **10**, 475-8.

Ozawa M., Sakatani M., Yao J., Shanker S., Yu F., Yamashita R., Wakabayashi S., Nakai K., Dobbs K.B. & Sudano M.J. (2012) Global gene expression of the inner cell mass and trophectoderm of the bovine blastocyst. *BMC developmental biology* **12**, 1.

Pai C.-Y., Lei E.P., Ghosh D. & Corces V.G. (2004) The centrosomal protein CP190 is a component of the gypsy chromatin insulator. *Molecular cell* **16**, 737-48.

Pavlíček A., Jabbari K., Pačes J., Pačes V., Hejnar J. & Bernardi G. (2001) Similar integration but different stability of Alus and LINEs in the human genome. *Gene* **276**, 39-45.

Peaston A.E., Evsikov A.V., Graber J.H., de Vries W.N., Holbrook A.E., Solter D. & Knowles B.B. (2004) Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos. *Developmental cell* **7**, 597-606.

Pi W., Yang Z., Wang J., Ruan L., Yu X., Ling J., Krantz S., Isales C., Conway S.J. & Lin S. (2004) The LTR enhancer of ERV-9 human endogenous retrovirus is active in oocytes and progenitor cells in transgenic zebrafish and humans. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 805-10.

Pittoggi C., Sciamanna I., Mattei E., Beraldi R., Lobascio A.M., Mai A., Quaglia M.G., Lorenzini R. & Spadafora C. (2003) Role of endogenous reverse transcriptase in murine early embryo development. *Molecular reproduction and development* **66**, 225-36.

Pollard J. & Leibo S. (1994) Chilling sensitivity of mammalian embryos. *Theriogenology* **41**, 101-6.

Ponsuksili S., Tesfaye D., El-Halawany N., Schellander K. & Wimmers K. (2002) Stage-specific expressed sequence tags obtained during preimplantation bovine development by differential display RT-PCR and suppression subtractive hybridization. *Prenatal diagnosis* **22**, 1135-42.

Purbowasito W., Suda C., Yokomine T., Zubair M., Sado T., Tsutsui K. & Sasaki H. (2004) Large-scale identification and mapping of nuclear matrix-attachment regions in the distal imprinted domain of mouse chromosome 7. *DNA research* **11**, 391-407.

Raiz J., Damert A., Chira S., Held U., Klawitter S., Hamdorf M., Löwer J., Strätling W.H., Löwer R. & Schumann G.G. (2012) The non-autonomous retrotransposon SVA is trans-mobilized by the human LINE-1 protein machinery. *Nucleic acids research* **40**, 1666-83.

Ramos K.S., Montoya–Durango D.E., Teneng I., Nanez A. & Stribinskis V. (2011) Epigenetic control of embryonic renal cell differentiation by L1 retrotransposon. *Birth Defects Research Part A: Clinical and Molecular Teratology* **91**, 693-702.

Reik W. (2007) Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature* **447**, 425-32.

Reik W., Dean W. & Walter J. (2001) Epigenetic reprogramming in mammalian development. *Science* **293**, 1089-93.

Ribet D., Harper F., Dupressoir A., Dewannieux M., Pierron G. & Heidmann T. (2008) An infectious progenitor for the murine IAP retrotransposon: emergence of an intracellular genetic parasite from an ancient retrovirus. *Genome research* **18**, 597-609.

Richardson S.R., Morell S. & Faulkner G.J. (2014) L1 retrotransposons and somatic mosaicism in the brain. *Annual review of genetics* **48**, 1-27.

Riggs A.D. (1990) Marsupials and mechanisms of X-chromosome inactivation. *Australian Journal of Zoology* **37**, 419-41.

Rinaudo P.F., Giritharan G., Talbi S., Dobson A.T. & Schultz R.M. (2006) Effects of oxygen tension on gene expression in preimplantation mouse embryos. *Fertility and Sterility* **86**, 1265. e1-. e36.

Rizos D., Gutierrez-Adan A., Perez-Garnelo S., De La Fuente J., Boland M. & Lonergan P. (2003) Bovine embryo culture in the presence or absence of serum: implications for blastocyst development, cryotolerance, and messenger RNA expression. *Biology of reproduction* **68**, 236-43.

Robbez-Masson L. & Rowe H.M. (2015) Retrotransposons shape species-specific embryonic stem cell gene expression. *Retrovirology* **12**, 1-12.

Rodriguez-Martinez H. (2012) Assisted reproductive techniques for cattle breeding in developing countries: a critical appraisal of their value and limitations. *Reproduction in Domestic Animals* **47**, 21-6.

Román A.C., González-Rico F.J., Moltó E., Hernando H., Neto A., Vicente-Garcia C., Ballestar E., Gómez-Skarmeta J.L., Vavrova-Anderson J. & White R.J. (2011) Dioxin receptor and SLUG transcription factors regulate the insulator activity of B1 SINE retrotransposons via an RNA polymerase switch. *Genome research* **21**, 422-32.

Romanish M.T., Lock W.M., van de Lagemaat L.N., Dunn C.A. & Mager D.L. (2007)

Repeated recruitment of LTR retrotransposons as promoters by the anti-apoptotic locus NAIP during mammalian evolution. *PLOS Genet* **3**, e10.

Rosser J.M. & An W. (2012) L1 expression and regulation in humans and rodents. *Frontiers in bioscience (Elite edition)* **4**, 2203-25.

Saito K., Nishida K.M., Mori T., Kawamura Y., Miyoshi K., Nagami T., Siomi H. & Siomi M.C. (2006) Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the Drosophila genome. *Genes & development* **20**, 2214-22.

Saksouk N., Simboeck E. & Déjardin J. (2015) Constitutive heterochromatin formation and transcription in mammals. *Epigenetics & chromatin* **8**, 1.

Salilew-Wondim D., Fournier E., Hoelker M., Saeed-Zidane M., Tholen E., Looft C., Neuhoff C., Besenfelder U., Havlicek V. & Rings F. (2015) Genome-Wide DNA Methylation Patterns of Bovine Blastocysts Developed In Vivo from Embryos Completed Different Stages of Development In Vitro. *PLOS ONE* **10**, e0140467.

Santoni F.A., Guerra J. & Luban J. (2012) HERV-H RNA is abundant in human embryonic stem cells and a precise marker for pluripotency. *Retrovirology* **9**.

Santos F., Hendrich B., Reik W. & Dean W. (2002) Dynamic reprogramming of DNA methylation in the early mouse embryo. *Developmental biology* **241**, 172-82.

Sasaki T., Nishihara H., Hirakawa M., Fujimura K., Tanaka M., Kokubo N., Kimura-Yoshida C., Matsuo I., Sumiyama K. & Saitou N. (2008) Possible involvement of SINEs in mammalian-specific brain formation. *Proceedings of the National Academy of Sciences* **105**, 4220-5.

Schmidt D., Schwalie P.C., Wilson M.D., Ballester B., Gonçalves Â., Kutter C., Brown G.D., Marshall A., Flicek P. & Odom D.T. (2012) Waves of retrotransposon expansion remodel genome organization and CTCF binding in multiple mammalian lineages. *Cell* **148**, 335-48.

Schostak N., Pyatkov K., Zelentsova E., Arkhipova I., Shagin D., Shagina I., Mudrik E., Blintsov A., Clark I. & Finnegan D.J. (2008) Molecular dissection of Penelope transposable element regulatory machinery. *Nucleic acids research* **36**, 2522-9.

Schulz W., Steinhoff C. & Florl A. (2006) Methylation of endogenous human retroelements in health and disease. In: *DNA Methylation: Development, Genetic Disease and Cancer* (pp. 211-50. Springer.

Sciamanna I., Vitullo P., Curatolo A. & Spadafora C. (2011) A reverse transcriptase-dependent mechanism is essential for murine preimplantation development. *Genes* **2**, 360-73.

Segal Y., Peissel B., Renieri A., de Marchi M., Ballabio A., Pei Y. & Zhou J. (1999) LINE-1 elements at the sites of molecular rearrangements in Alport syndrome–diffuse leiomyomatosis. *The American Journal of Human Genetics* **64**, 62-9.

Seitz H., Youngson N., Lin S.-P., Dalbert S., Paulsen M., Bachellerie J.-P., Ferguson-Smith A.C. & Cavaillé J. (2003) Imprinted microRNA genes transcribed antisense to a reciprocally imprinted retrotransposon-like gene. *Nature genetics* **34**, 261-2.

Sekita Y., Wagatsuma H., Nakamura K., Ono R., Kagami M., Wakisaka N., Hino T., Suzuki-Migishima R., Kohda T. & Ogura A. (2008) Role of retrotransposon-derived imprinted gene, Rtl1, in the feto-maternal interface of mouse placenta. *Nature genetics* **40**, 243-8.

Sen S.K., Han K., Wang J., Lee J., Wang H., Callinan P.A., Dyer M., Cordaux R., Liang P. & Batzer M.A. (2006) Human genomic deletions mediated by recombination between Alu elements. *The American Journal of Human Genetics* **79**, 41-53.

Shankar R., Grover D., Brahmachari S.K. & Mukerji M. (2004) Evolution and distribution of RNA polymerase II regulatory sites from RNA polymerase III dependant mobile Alu elements. *BMC evolutionary biology* **4**, 37.

Sharan C., Hamilton N.M., Parl A.K., Singh P.K. & Chaudhuri G. (1999) Identification and characterization of a transcriptional silencer upstream of the human BRCA2 gene. *Biochemical and biophysical research communications* **265**, 285-90.

Shen L., Wu L.C., Sanlioglu S., Chen R., Mendoza A.R., Dangel A.W., Carroll M.C., Zipf W.B. & Yu C.-Y. (1994) Structure and genetics of the partially duplicated gene RP located immediately upstream of the complement C4A and the C4B genes in the HLA class III region. Molecular cloning, exon-intron structure, composite retroposon, and breakpoint of gene duplication. *Journal of Biological Chemistry* **269**, 8466-76.

Shen S., Lin L., Cai J.J., Jiang P., Kenkel E.J., Stroik M.R., Sato S., Davidson B.L. & Xing Y. (2011) Widespread establishment and regulatory impact of Alu exons in human genes. *Proceedings of the National Academy of Sciences* **108**, 2837-42.

Shi W. & Haaf T. (2002) Aberrant methylation patterns at the two‐cell stage as an indicator of early developmental failure. *Molecular reproduction and development* **63**, 329-34.

Shukla R., Upton K.R., Muñoz-Lopez M., Gerhardt D.J., Fisher M.E., Nguyen T., Brennan P.M., Baillie J.K., Collino A. & Ghisletti S. (2013) Endogenous retrotransposition activates oncogenic pathways in hepatocellular carcinoma. *Cell* **153**, 101-11.

Sinclair K.D., McEvoy T., Maxfield E., Maltin C., Young L., Wilmut I., Broadbent P. & Robinson J. (1999) Aberrant fetal growth and development after in vitro culture of sheep zygotes. *Journal of reproduction and fertility* **116**, 177-86.

Siomi M.C., Sato K., Pezic D. & Aravin A.A. (2011) PIWI-interacting small RNAs: the vanguard of genome defence. *Nature reviews Molecular cell biology* **12**, 246-58.

Slotkin R.K. & Martienssen R. (2007) Transposable elements and the epigenetic regulation of the genome. *Nature Reviews Genetics* **8**, 272-85.

Soifer H.S., Zaragoza A., Peyvan M., Behlke M.A. & Rossi J.J. (2005) A potential role for RNA interference in controlling the activity of the human LINE-1 retrotransposon. *Nucleic acids research* **33**, 846-56.

Solyom S., Ewing A.D., Rahrmann E.P., Doucet T., Nelson H.H., Burns M.B., Harris R.S., Sigmon D.F., Casella A. & Erlanger B. (2012) Extensive somatic L1 retrotransposition in colorectal tumors. *Genome research* **22**, 2328-38.

Sorek R., Ast G. & Graur D. (2002) Alu-containing exons are alternatively spliced. *Genome research* **12**, 1060-7.

Steinhoff C. & Schulz W. (2004) Transcriptional regulation of the human LINE-1 retrotransposon L1. 2B. *Molecular Genetics and Genomics* **270**, 394-402.

Stetson D.B., Ko J.S., Heidmann T. & Medzhitov R. (2008) Trex1 prevents cell-intrinsic initiation of autoimmunity. *Cell* **134**, 587-98.

Storz G. (2002) An expanding universe of noncoding RNAs. *Science* **296**, 1260-3.

Sunami E., De Maat M., Vu A., Turner R.R. & Hoon D.S. (2011) LINE-1 hypomethylation

during primary colon cancer progression. *PLOS ONE* **6**, e18884.

Suzuki S., Ono R., Narita T., Pask A.J., Shaw G., Wang C., Kohda T., Alsop A.E., Graves J.A.M. & Kohara Y. (2007) Retrotransposon silencing by DNA methylation can drive mammalian genomic imprinting. *PLOS genetics* **3**, e55.

Takahashi M. (2012) Oxidative stress and redox regulation on in vitro development of mammalian embryos. *Journal of Reproduction and Development* **58**, 1-9.

Tanghe S., Van Soom A., Mehrzad J., Maes D., Duchateau L. & de Kruif A. (2003) Cumulus contributions during bovine fertilization in vitro. *Theriogenology* **60**, 135-49.

Tashiro K., Teissier A., Kobayashi N., Nakanishi A., Sasaki T., Yan K., Tarabykin V., Vigier L., Sumiyama K. & Hirakawa M. (2011) A mammalian conserved element derived from SINE displays enhancer properties recapitulating Satb2 expression in early-born callosal projection neurons. *PLOS ONE* **6**, e28497.

Teneng I., Stribinskis V. & Ramos K.S. (2006) Stress-Mediated Activation of LINE-1 in Mammalian Cells. *The FASEB Journal* **20**, A1359.

Thompson J., Simpson A., Pugh P., Donnelly P. & Tervit H. (1990) Effect of oxygen concentration on in-vitro development of preimplantation sheep and cattle embryos. *Journal of reproduction and fertility* **89**, 573-8.

Tollervey D. & Caceres J.F. (2000) RNA processing marches on. *Cell* **103**, 703-9.

Toyota M. & Issa J.-P.J. (1999) CpG island methylator phenotypes in aging and cancer. In: *Seminars in cancer biology*, pp. 349-57. Elsevier.

Trelogan S.A. & Martin S.L. (1995) Tightly regulated, developmentally specific expression of the first open reading frame from LINE-1 during mouse embryogenesis. *Proceedings of the National Academy of Sciences* **92**, 1520-4.

Turelli P., Vianin S. & Trono D. (2004) The innate antiretroviral factor APOBEC3G does not affect human LINE-1 retrotransposition in a cell culture assay. *Journal of Biological Chemistry* **279**, 43371-3.

Uhlén, M., Fagerberg, L., Hallström, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., ... & Olsson, I. (2015). Tissue-based map of the human proteome. *Science* ***347***(6220), 1260419.

Valencia-Sanchez M.A., Liu J., Hannon G.J. & Parker R. (2006) Control of translation and mRNA degradation by miRNAs and siRNAs. *Genes & development* **20**, 515-24.

van de Lagemaat L.N., Landry J.-R., Mager D.L. & Medstrand P. (2003) Transposable elements in mammals promote regulatory variation and diversification of genes with specialized functions. *Trends in genetics* **19**, 530-6.

van de Lagemaat L.N., Medstrand P. & Mager D.L. (2006) Multiple effects govern endogenous retrovirus survival patterns in human gene introns. *Genome biology* **7**, R86.

van den Hurk J.A., Meij I.C., del Carmen Seleme M., Kano H., Nikopoulos K., Hoefsloot L.H., Sistermans E.A., de Wijs I.J., Mukhopadhyay A. & Plomp A.S. (2007) L1 retrotransposition can occur early in human embryonic development. *Human molecular genetics* **16**, 1587-92.

van den Hurk J.A., van de Pol D.J., Wissinger B., van Driel M.A., Hoefsloot L.H., de Wijs I.J., van den Born L.I., Heckenlively J.R., Brunner H.G. & Zrenner E. (2003) Novel types of mutation in the choroideremia (CHM) gene: a full-length L1 insertion and an intronic mutation activating a cryptic exon. *Human genetics* **113**, 268-75.

Van Soom A. & Kruif A.d. (1998) Bovine embryonic development after in vivo and in vitro fertilization. *Reproduction in Domestic Animals* **33**, 261-5.

Van Soom A., Ysebaert M.T. & de Kruif A. (1997) Relationship between timing of development, morula morphology, and cell allocation to inner cell mass and trophectoderm in in vitro-produced bovine embryos. *Molecular reproduction and development* **47**, 47-56.

Van Soom A., Yuan Y.Q., Peelman L.J., de Matos D.G., Dewulf J., Laevens H. & de Kruif A. (2002) Prevalence of apoptosis and inner cell allocation in bovine embryos cultured under different oxygen tensions with or without cysteine addition. *Theriogenology* **57**, 1453-65.

Vance K.W. & Ponting C.P. (2014) Transcriptional regulatory functions of nuclear long noncoding RNAs. *Trends in genetics* **30**, 348-55.

Vandaele L., Thys M., Bijttebier J., Van Langendonckt A., Donnay I., Maes D., Meyer E. &

Van Soom A. (2010) Short-term exposure to hydrogen peroxide during oocyte maturation improves bovine embryo development. *Reproduction* **139**, 505-11.

Vanden Berghe W. (2012) Epigenetic impact of dietary polyphenols in cancer chemoprevention: Lifelong remodeling of our epigenomes. *Pharmacological Research* **65**, 565-76.

Vanneste E., Voet T., Le Caignec C., Ampe M., Konings P., Melotte C., Debrock S., Amyere M., Vikkula M. & Schuit F. (2009) Chromosome instability is common in human cleavage-stage embryos. *Nature medicine* **15**, 577-83.

Varmus H. (1988) Retroviruses. *Science* **240**, 1427-35.

Varmus H., Brown P. & Retroviruses I. (1989) Mobile DNA. *Retroviruses. American Society for Microbiology, Washington, DC*, 53-108.

Velker B.A.M., Denomme M.M. & Mann M.R. (2012) Embryo culture and epigenetics. *Embryo culture: Methods and protocols*, 399-421.

Vinckenbosch N., Dupanloup I. & Kaessmann H. (2006) Evolutionary fate of retroposed gene copies in the human genome. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 3220-5.

Xing J., Zhang Y., Han K., Salem A.H., Sen S.K., Huff C.D., Zhou Q., Kirkness E.F., Levy S. & Batzer M.A. (2009) Mobile elements create structural variation: analysis of a complete human genome. *Genome research* **19**, 1516-26.

Walsh C.P., Chaillet J.R. & Bestor T.H. (1998) Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. *Nature genetics* **20**, 116-7.

Wang-Johanning F., Frost A.R., Jian B., Epp L., Lu D.W. & Johanning G.L. (2003) Quantitation of HERV-K env gene expression and splicing in human breast cancer. *Oncogene* **22**, 1528-35.

Wang-Johanning F., Liu J., Rycaj K., Huang M., Tsai K., Rosen D.G., Chen D.T., Lu D.W., Barnhart K.F. & Johanning G.L. (2007) Expression of multiple human endogenous retrovirus surface envelope proteins in ovarian cancer. *International journal of cancer* **120**, 81-90.

Wang H., Xing J., Grover D., Hedges D.J., Han K., Walker J.A. & Batzer M.A. (2005) SVA

elements: a hominid-specific retroposon family. *Journal of molecular biology* **354**, 994-1007.

Wang L., Guo Z.-Y., Zhang R., Xin B., Chen R., Zhao J., Wang T., Wen W.-H., Jia L.-T. & Yao L.-B. (2013) Pseudogene OCT4-pg4 functions as a natural micro RNA sponge to regulate OCT4 expression by competing for miR-145 in hepatocellular carcinoma. *Carcinogenesis* **34**, 1773-81.

Wang R. & Brattain M.G. (2007) The maximal size of protein to diffuse through the nuclear pore is larger than 60 kDa. *FEBS letters* **581**, 3164-70.

Wang Y., Liška F., Gosele C., Šedová L., Křen V., Křenová D., Ivics Z., Hubner N. & Izsvák Z. (2010) A novel active endogenous retrovirus family contributes to genome variability in rat inbred strains. *Genome research* **20**, 19-27.

Wang Z., Willard H.F., Mukherjee S. & Furey T.S. (2006) Evidence of influence of genomic DNA sequence on human X chromosome inactivation. *PLOS Comput Biol* **2**, e113.

Ward F., Enright B., Rizos D., Boland M. & Lonergan P. (2002) Optimization of in vitro bovine embryo production: effect of duration of maturation, length of gamete co-incubation, sperm concentration and sire. *Theriogenology* **57**, 2105-17.

Warren W.C., Hillier L.W., Graves J.A.M., Birney E., Ponting C.P., Grützner F., Belov K., Miller W., Clarke L. & Chinwalla A.T. (2008) Genome analysis of the platypus reveals unique signatures of evolution. *Nature* **453**, 175-83.

Watson A.J. & Barcroft L.C. (2001) Regulation of blastocyst formation. *Frontiers in Bioscience* **6**, D708-30.

Webb D.W. (1992) Artificial insemination in dairy cattle. University of Florida Cooperative Extension Service, Institute of Food and Agriculture Sciences, EDIS.

Wei W., Gilbert N., Ooi S.L., Lawler J.F., Ostertag E.M., Kazazian H.H., Boeke J.D. & Moran J.V. (2001) Human L1 retrotransposition: cispreference versus trans complementation. *Molecular and cellular biology* **21**, 1429-39.

Weis L. & Reinberg D. (1997) Accurate positioning of RNA polymerase II on a natural TATA-less promoter is independent of TATA-binding-protein-associated factors and initiator-binding proteins. *Molecular and cellular biology* **17**, 2973-84.

Weisenberger D.J., Campan M., Long T.I., Kim M., Woods C., Fiala E., Ehrlich M. & Laird P.W. (2005) Analysis of repetitive element DNA methylation by MethyLight. *Nucleic acids research* **33**, 6823-36.

Whitworth K., Springer G.K., Forrester L.J., Spollen W.G., Ries J., Lamberson W.R., Bivens N., Murphy C.N., Mathialigan N. & Green J.A. (2004) Developmental expression of 2489 gene clusters during pig embryogenesis: an expressed sequence tag project. *Biology of reproduction* **71**, 1230-43.

Wicker T., Sabot F., Hua-Van A., Bennetzen J.L., Capy P., Chalhoub B., Flavell A., Leroy P., Morgante M. & Panaud O. (2007) A unified classification system for eukaryotic transposable elements. *Nature Reviews Genetics* **8**, 973-82.

Witney F. & Furano A. (1984) Highly repeated DNA families in the rat. *Journal of Biological Chemistry* **259**, 10481-92.

Wright R.O., Schwartz J., Wright R.J., Bollati V., Tarantini L., Park S.K., Hu H., Sparrow D., Vokonas P. & Baccarelli A. (2010) Biomarkers of lead exposure and DNA methylation within retrotransposons. *Environmental health perspectives* **118**, 790.

Wu Q. & Ni X. (2015) ROS-mediated DNA methylation pattern alterations in carcinogenesis. *Current drug targets* **16**, 13-9.

Xing J., Zhang Y., Han K., Salem A.H., Sen S.K., Huff C.D., Zhou Q., Kirkness E.F., Levy S. & Batzer M.A. (2009) Mobile elements create structural variation: analysis of a complete human genome. *Genome research* **19**, 1516-26.

Yakovchuk P., Goodrich J.A. & Kugel J.F. (2009) B2 RNA and Alu RNA repress transcription by disrupting contacts between RNA polymerase II and promoter DNA within assembled complexes. *Proceedings of the National Academy of Sciences* **106**, 5569-74.

Yang A.S., Estécio M.R., Doshi K., Kondo Y., Tajara E.H. & Issa J.P.J. (2004) A simple method for estimating global DNA methylation using bisulfite PCR of repetitive DNA elements. *Nucleic acids research* **32**, e38-e.

Yang N. & Kazazian H.H. (2006) L1 retrotransposition is suppressed by endogenously encoded small interfering RNAs in human cultured cells. *Nat Struct Mol Biol* **13**,

763-71.

Yang N., Zhang L., Zhang Y. & Kazazian Jr H.H. (2003) An important role for RUNX3 in human L1 transcription and retrotransposition. *Nucleic acids research* **31**, 4929-40.

Yang Z., Boffelli D., Boonmark N., Schwartz K. & Lawn R. (1998) Apolipoprotein (a) gene enhancer resides within a LINE element. *Journal of Biological Chemistry* **273**, 891-7.

Yoder J.A., Walsh C.P. & Bestor T.H. (1997) Cytosine methylation and the ecology of intragenomic parasites. *Trends in genetics* **13**, 335-40.

Young L.E., Sinclair K.D. & Wilmut I. (1998) Large offspring syndrome in cattle and sheep. *Reviews of reproduction* **3**, 155-63.

Yu F., Zingler N., Schumann G. & Strätling W.H. (2001) Methyl-CpG-binding protein 2 represses LINE-1 expression and retrotransposition but not Alu transcription. *Nucleic acids research* **29**, 4493-501.

Yuan Y., Van Soom A., Coopman F., Mintiens K., Boerjan M., Van Zeveren A., de Kruif A. & Peelman L. (2003) Influence of oxygen tension on apoptosis and hatching in bovine embryos cultured in vitro. *Theriogenology* **59**, 1585-96.

Zaitseva I., Zaitsev S., Alenina N., Bader M. & Krivokharchenko A. (2007) Dynamics of DNA-demethylation in early mouse and rat embryos developed in vivo and in vitro. *Molecular reproduction and development* **74**, 1255-61.

Zheng D., Frankish A., Baertsch R., Kapranov P., Reymond A., Choo S.W., Lu Y., Denoeud F., Antonarakis S.E. & Snyder M. (2007) Pseudogenes in the encode regions: consensus annotation, analysis of transcription, and evolution. *Genome research* **17**, 839-51.

Zilberman D. & Henikoff S. (2007) Genome-wide analysis of DNA methylation patterns. *Development* **134**, 3959-65.

Zolotukhin A.S., Schneider R., Uranishi H., Bear J., Tretyakova I., Michalowski D., Smulevitch S., O'Keeffe S., Pavlakis G.N. & Felber B.K. (2008) The RNA transport element RTE is essential for IAP LTR-retrotransposon mobility. *Virology* **377**, 88-99

Chapter 2

Aims of the study

Autonomous retrotransposons are found to be active and functional in preimplantation embryos. The impact of retrotransposon activity on the genome is unneglectable. They are suggested to participate in many activities during early embryonic development, including genomic imprinting, X chromosome inactivation, cell proliferation and differentiation. The role of retrotransposons in these processes is strongly correlated to the DNA methylation status and thus epigenetic regulation. Based on the 'genomic shock' hypothesis that environmental stimuli (including oxidative stress) may mobilize transposable elements, as first proposed by Barbara McClintock, it can be hypothesized that oxidative stress induces DNA methylation changes and this, in turn, influences retrotransposon expression and eventually embryo health. To test that hypothesis, we applied sustained oxidative stress (20% O_2) during *in vitro* bovine preimplantation embryo development and measured global DNA methylation and expression of bovine autonomous retrotransposons.

The specific aims of the study were:

1. To profile autonomous retrotransposon expression in bovine preimplantation embryos in order to select the expressed ones for further study (**Chapter 3.1**).

2. To select reliable reference genes for evaluating retrotransposon expression in bovine embryos cultured under standard condition and high oxygen stress (**Chapter 3.2**).

3. To study the relationship between DNA methylation and retrotransposon expression under high oxygen tension (**Chapter 4**).

4. To use repetitive sequences as global DNA methylation marker in bovine preimplantation embryos (**Chapter 5**).

Chapter 3

Retrotransposon Expression Profiling and Reference Gene Selection for qPCR Analysis in Bovine Preimplantation Embryos

Chapter 3.1 Retrotransposon Expression Profiling in Bovine Preimplantation Embryos

Li W[1], Goossens K[1], Van Poucke M[1], Vandaele L[2], Van Soom A[2], Peelman L[1]

[1] *Department of Nutrition, Genetics and Ethology, Faculty of Veterinary Medicine, Ghent University, B-9820 Merelbeke, Belgium*

[2] *Department of Obstetrics, Reproduction and Herd Health, Faculty of Veterinary Medicine, Ghent University, B-9820 Merelbeke, Belgium*

Adapted from a poster presentation at the 3rd Mammalian Embryo Genomics meeting, 2011, Bonn Germany.

3.1.1 Abstract

Retrotransposons in general were previously found to be active in preimplantation embryos. The purpose of this study was to make an expression profile of the different autonomous retrotransposons in bovine oocytes and all stages during preimplantation development. *L1_BT*, *BovB*, and *ERV1-1-I_BT* were found to be expressed in all stages whereas the other retrotransposon elements tested showed low or inconsistent expression. These were excluded from further research.

3.1.2 Introduction

Retrotransposons are genetic elements that insert extra copies of themselves throughout the genome via a 'copy and paste' mechanism, through which they can rapidly increase the copy number and therefore increase genome size. Autonomous retrotransposons have an ORF which can encode reverse transcriptase necessary for their reverse transcription and 'jumping ability'. They can not only copy and transport themselves back to the genome, but also facilitate the reverse transcription of non-autonomous retrotransposons, such as SINEs (Raiz *et al.* 2012). The expression of these elements determines when and where the retrotransposition can potentially occur. LINEs and ERVs are two typical autonomous retrotransposons. LINEs are the largest mammalian transposable element components, comprising on average about 20% of mammalian genome - 21.9% in human, 19% in rat, 20.3% in mouse, and even 29.2% in cow. A typical integral L1 element is approximately 6000 bp, and codes for an RNA binding protein (ORF1), and an ORF2 for endonuclease and reverse transcriptase. ERVs are another superfamily in the mammalian genome. A typical structure of ERVs comprises a central part with three major genes (*gag*, *pol* and *env*) flanked by two long terminal repeats (LTRs).

Both LINEs and ERVs are known to play biologically significant roles in mutation, disease (i.e. breast cancer in human) and genome evolution, and to have a strong expression in mouse oocytes and early embryos, possibly due to demethylation during these stages (Peaston *et al.* 2004). However, little is known about the ERV families' expression in different bovine embryonic stages. The aim of the present study was to analyse the specific autonomous retrotransposon expression during cattle preimplantation embryo development.

3.1.3 Material and methods

Materials

Unless stated otherwise, all chemicals, reagents and media were obtained from Sigma (USA) and Life Technologies (Belgium).

Autonomous retrotransposon selection and primer design

Repbase (Jurka *et al.* 2005) was used to select autonomous retrotransposon classes originating from *Bos taurus* with a potential complete open reading frame (ORF) of reverse transcriptase. These include LTR retrotransposons, Endogenous Retroviruses (ERVs) and non-LTR retrotransposons. In this way, seven bovine autonomous retrotransposon families with complete reverse transcriptase ORFs were found: *L1_BT*, *BovB*, *ERV1-1-I_BT*, *BtERVF2-I*, *ERV2-1-I_BT*, *ERV2-2-I_BT*, and *ERV2-3-I_BT*.

As the aim was to amplify as many as possible individual repeats of each retrotransposon family present in the genome, the repeat sequences for each retrotransposon were collected from the bovine genome (UMD_3.1 and Btau_4.2) using NCBI BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi). The selected sequences were further checked by NCBI ORF Finder (Wheeler *et al.* 2000) to determine the ORFs for each of these

retrotransposon repeat sequences. Then, MultAlin (Corpet 1988) and ClustalW2 (Larkin *et al.* 2007) were used to identify conserved regions in the reverse transcriptase ORFs for each of the seven autonomous retrotransposon families. For some retrotransposons, such as *L1_BT*, which was suggested to have 811 complete copies in the genome (Adelson *et al.* 2009), 80 representative copies of those were used in alignment in the present study.

Based on the conserved sequences within ORFs, specific primers were designed for each of the seven families using Primer3 software (Rozen & Skaletsky 1999). All amplicons are located in the ORFs of the retrotranposons (**Figure 3.1**). The number of amplicons in the genome was checked using BiSearch software (Arányi *et al.* 2006) with default settings and listed in **Table 3.1**. Mfold (Zuker 2003) was used to account for any secondary structure in the PCR amplicons.

RT-PCR amplified products were analyzed in 2% agarose gels containing ethidium bromide. For cloning, RT-PCR bands of interest were excised from agarose gels, purified using the GENECLEAN® II Kit (MP Biomedicals, Belgium), cloned into pCR 2.1 and the re-amplified products were sequenced for verification. All the amplicons are listed in **Table 3.2**.

There are many studies about repetitive sequence PCR amplification (Bratthauer & Fanning 1992; Georgiou *et al.* 2009; Macia *et al.* 2011), but in the present study, our purpose was to amplify as many copies as possible. Therefore, extra steps such as genome-wide BLAST to find the copies, and alignment for conserved sequences to design primers were required. With the analysis of BiSearch, amplicon numbers of each pair of primers were found as well, and were listed in **Table 3.1**.

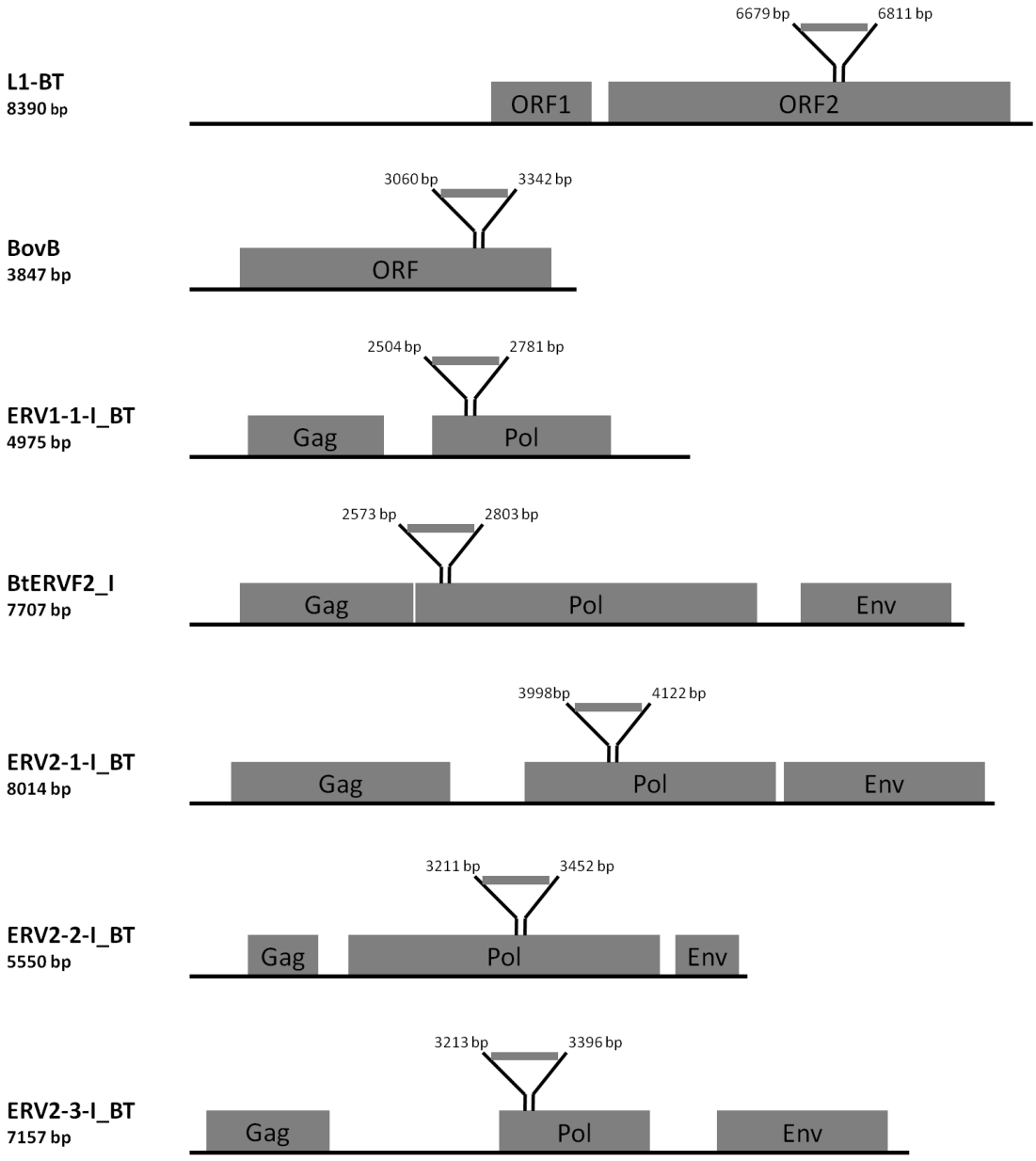


Figure 3.1 Structures of the retrotransposons in the study, and locations of PCR amplicons within retrotransposon ORFs.

Table 3.1 Information of the primers used in the study. The amplicon number for *GAPDH* is based on bovine cDNA database (UMD 3.1.86); while the potential amplicon numbers for retrotransposons are based on bovine genome (UMD_3.1.86).

Retrotransposon name	Potential amplicon numbers (BiSearch)	Sequence 5' – 3'	Amplicon length cDNA	Ta (°C)
GAPDH	1	TTCAACGGCACAGTCAAGG	119	62
		ACATACTCAGCACCAGCATCAC		
L1_BT	6606	GCAATCCCTATCAAGCTACCA	133	60
		TGATTCCTCCAGTTCCATTCTT		
BovB	25639	CCTCAGATATGCAGATGACACC	283	60
		GTTGGTCATAACTTTCCTTCCA		
ERV1-1-I_BT	103	TGTTAAGCTCAAAGACCCACAC	277	60
		CCGTTCAGGAATTTCAGACAA		
BtERVF2_I	17	ACCTTCCCTTGGATCTACCC	231	63
		TAGGGGTGTTGTGTGGTGAA		
ERV2-1-I_BT	15	GGCTTAAAAATTGCCCCAGA	125	60
		AAAGTAACAAGGTGATCTTTTCTCAA		
ERV2-2-I_BT	13	CCTGACCCTAATAGTCCTAGACAG	243	60
		TGCCCTATTTTCTCAATGTATGA		
ERV2-3-I_BT	5	GCTTTTAGTGTCCCTTCTACAAAT	186	60
		TGCCAAAAGAATATCATCCAT		

Table 3.2 Sequences of retrotransposon PCR amplicons

Retrotransposon name	**Sequence**
L1_BT (133 bp)	TGATTCCTCCAGTTCCATTCTTCTTTCTCAAGATCGCTTTGGCTATTCGAGGTTTTTTGTATTTCCA TACAAATTGTGAAATTATTTGTTCTAGCTCTGTGAAAAATACTGTTGGTAGCTTGATAGGGATT GC
BovB (283 bp)	CCTCAGATATGCAGATGACACCACCCTTATGGTAGAAAGTAAAGAGGAACTAAAAAGCCTCTT GATGAAAGTGAAAGAGGAGAGTGAAAAAGCTGGCTTAAAGCTCAACATTCAGAAAATGAAGA TCATGGCATCCGGTCCCATCACTTCATGGGAAATAGATGGGGAAACAGTGGAAACAGTGTCAG ACTTTATTTTTCTGGGCTCCAAAATCACGACAGATGGTGACAGCAGCCATGAAATTAAAAGACG CTTACTCCTTGGAAGGAAAGTTATGACCAAC
ERV1-1-I_BT (277 bp)	TGTTAAGCTCAAAGACCCACACGTATTTCCGCATAAGAAGCAGTATCCACTGAAACCTGAAGTT AAGGAAGAGTTAAAACCCATCATCAAAAATTTAAAGGAGCAGGGACTATTAATTCCCTGTAAC AGTCCTTGCAACACTCCTATTTTGGGTATAAAGAAATCGAATGGTAAATGAAGACTAGTTCAAG ATTTATGAATAATAAATGAGGCTGTAGTTCCTTTACACCCCGTGGTGCCTAATCCTTATACTCTA TTGTCTGAAATTCCTGAACGGAA
BtERVF2_I (231 bp)	ACCTTCCCTTGGATCTACCCACCTTANACCCCCAAGTCTGGGACACTGATCACCCATCCATAGCC AAACATCATCCCCCAGTCCACATTACCCTAAAAGACCCCTCGACTATAATCTCCCAACAGTACTC GCCACCCCGAAGACCCACAAGGGACTTAAGCCTATCATAGATCGTCTTCTTCAAGCCCCTATCCT AATTCCTAACCATTCACCACACAACACCCCTA
ERV2-1-I_BT (125 bp)	GGCTTAAAAATTGCCCCAGAGAAAATTCAAGTCAATCCGCCCATAACTTACTTAGGGCGGGTTA TCAATTCAGAAACTGTGACTCATGCCCCATTAAAATTGAGAAAAGATCACCTTGTTACTTT
ERV2-2-I_BT (243 bp)	CCTGACCCTAATAGTCCTAGACAGCTAACAAAAGAAGTGGAAGAAGAATTAAAATTTGTTGAA AAATGCATTCAACAAGCTTTCACAACTCAGCTGGATCATACCCAACCGGTCTGTTTATATATATA TATTCCCCACCAAACATTCACCTACTGCAATCATAGCTCAATACAGCCCGATAGAGTGGGTATAT CTACAGACTAAACAGTTAAA-ATTTTCTTATCATACATTGAGAAAATAGGGCA
ERV2-3-I_BT (186 bp)	GCTTTTAGTGTCCCTTCTACAAATTATAAAGAACCTATGAAAAGATATCAGTGGCAAGTTTTACC TCAGGGAATGGCTAATAGTCCTACTCTCTGTCAGAAATTTGTTGCTCAAGCTTTAAAAACCACTA GGTCCTTGTACTCCCAAGTATATATTATTCATTATATGGATGATATTCTTTTGGCA

Oocytes and in vitro embryo collection

In vitro bovine embryos were produced in serum-free media according to our standard procedures (Wydooghe *et al.* 2013). Briefly, bovine ovaries were collected at the local slaughterhouse. Immature cumulus-oocyte complexes were aspirated from follicles between 2 and 8 mm diameter and matured in groups of 60 in 500 µl modified bicarbonate-buffered TCM-199 supplemented with 50 mg/ml gentamycin and 20 ng/ml EGF for 22 hrs at 38.5°C in a 5% CO_2 incubator. After maturation, the cumulus-oocyte complexes were inseminated with frozen-thawed bovine spermatozoa at a final concentration of 10^6 spermatozoa/ml. The cumulus cells and excess spermatozoa were removed from presumed zygotes by vortex after 21 hrs of incubation. The presumed zygotes were incubated in groups of 25 in 50 µl synthetic oviductal fluid supplemented with essential and non-essential amino acids (SOFaa), 0.4% BSA and ITS (5 µg/ml insulin + 5 µg/ml transferrin + 5 ng/ml selenium) covered with mineral oil, and incubated at 38.5°C at 5% CO_2, 5% O_2 and 90% N_2 (as control group). Embryos with good morphology (some examples are shown in **Figure 3.2**) were collected at specific time points: 2 to 4- cells at 36-40 hpi, 5 to 8- cells at 64 hpi, 9 to 16-cell at 5 dpi, morula at 6 dpi, normal blastocyst at 7 dpi, expanded blastocyst at 8 dpi and hatched blastocyst at 9 dpi, together with immature and mature oocytes.

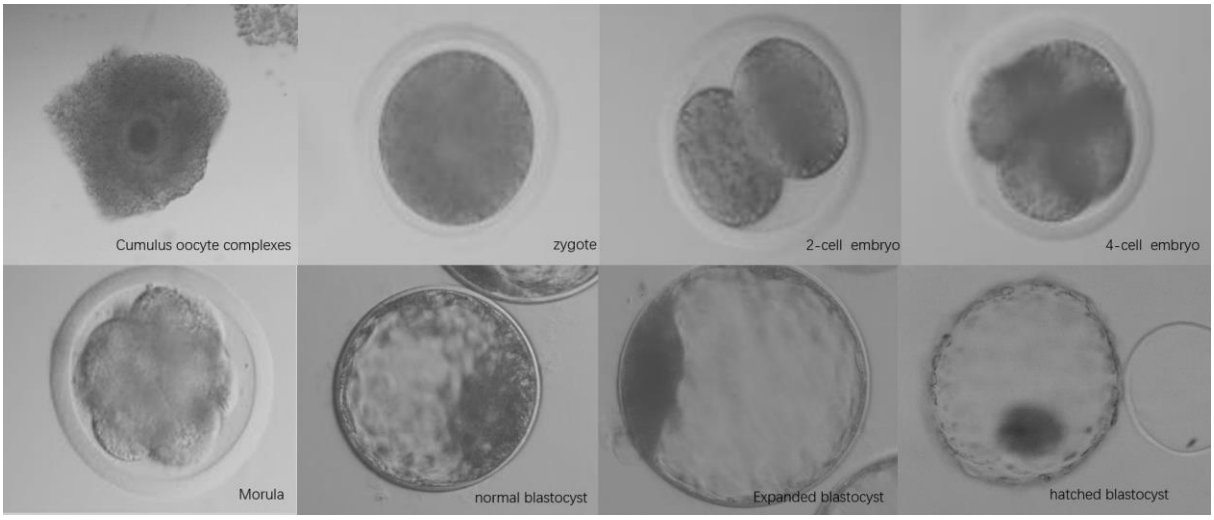


Figure 3.2 Embryos produced by IVF at different stages of development

Oocytes and embryos were further washed three times with RNase-free PBS, frozen in pools of five in lysis buffer (10% RNAsin Plus RNase inhibitor (Promega, Netherlands), 5% DTT (Promega, Netherlands), and 0.8% Igepal CA-630 in RNase free water) and stored at -80 °C, until use. For each stage, 3 pools of 5 oocytes or embryos were collected.

RNA extraction and cDNA synthesis

Total RNA was isolated from pooled oocytes or embryos from each stage, using the RNeasy Micro kit (Qiagen, Belgium) according to the manufacturer's instructions, including a genomic DNA removal step in the procedure. The extracted RNA was dissolved in 14 µl of RNase-free water. A minus RT control was performed with primers based on the genomic sequence of *GAPDH* to check for contaminating genomic DNA, and only the RNA samples without DNA contamination were used for reverse transcription (Goossens *et al.* 2005). First strand cDNA was generated from the total amount of RNA using the iScript cDNA synthesis kit (BioRad, Belgium) according to the product manual. After reverse transcription, the cDNA was 2.5 times diluted and used for downstream PCR.

Retrotransposon expression profiling by polymerase chain reaction (PCR)

All PCR reactions were performed in a 10 µl volume on the Eppendorf Mastercycler PCR System (Eppendorf, Belgium), with 0.5 U FastStart Taq DNA Polymerase and 1 µl 10x reaction buffer (Roche, Belgium), 200 µM dNTPs (Bioline Reagents Ltd., UK), 500 nM of each primer (LDT, Belgium) and 2.5 µl of diluted cDNA of pooled oocytes or embryos. The PCR program consisted of initial denaturation at 95°C for 5 min, followed by 40 cycles of 15 sec at 95°C, 15 sec at the specific annealing temperature for each retrotransposon (**Table 3.1**) and 30 sec at 72°C, and a final 10 min elongation at 72°C. The PCR products were verified by electrophoresis for confirmation of expression.

3.1.4 Results and discussion

The PCR results of the seven autonomous retrotransposon families in bovine oocytes and preimplantation embryos are summarized in **Table 3.3**. *L1_BT*, *BovB*, *ERV1-1-I_BT*, and *ERV2-1-1_BT* show strong RNA expression crossing the development from oocyte to blastocyst stage, while other ERVs have specific, developmentally regulated expression patterns. Similar differences in LTR retrotransposon expression pattern were found in mouse and it was suggested that normal repressive chromatin structure including these loci is established sequentially during the oocyte to embryo transition and preimplantation stages (Peaston *et al.* 2004). Also, the potential copies for *BtERVF2-I*, *ERV2-2-I_BT* and *ERV2-3-I_BT* in the bovine genome are 17, 13 and 5, which may lack functional promoters and/or relulatory elements to guarantee their transcription. This also explains the variable expression in the pooled oocytes or embryos of the same stage. Although *ERV2-1-1_BT* is expressed in all the developmental stages, the fact that there is no expression in some replicates makes it unsuitable for further quantitive study.

It is notable that expression of *BtERVF2-I* starts from the 9-16 cell stage and *ERV2-2-I_BT* is more consistently expressed from this stage on compared to earlier stages. The specific function of these two retrotransposons is still not clear, but the increased expression at this stage may be due to the embryonic genome activation (EGA). EGA is an important event for maternal-to-embryonic transcriptome transition. The major EGA in bovine embryos was identified to start from the 8-cell onwards, with a marked increase in number of differentially expressed genes, compared with the 4-cell stage (Graf *et al.* 2014). Expression quantification using the ratio of ERV copies / luciferase reporter transcripts showed a 1000-fold increase in post-EGA stage (morula) ERV expression compared to the pre-EGA stage (4-cell) in bovine embryos (Bui *et al*. 2009). This massive transcriptional activation of ERV retrotransposons is also suggested as a defining event of genome reprogramming in bovine embryos.

Table 3.3 RNA expression of the retrotransposons in 3 pools of 5 bovine oocytes and preimplantation embryos determined by RT-PCR

Results are shown for each of three replicates at each stage. (●), strong expression; (○), weak expression; (□), no expression

	Immature oocyte	Mature oocyte	Zygote	2–4 cell	5–8 cell	9–16 cell	Morula	Early blastocyst	Normal blastocyst	Expanded blastocyst	Hatched blastocyst
L1-BT	●●●	●●●	●●●	●●●	●●●	●●●	●●●	●●●	●●●	●●●	●●●
BovB	●●●	●●●	●●●	●●●	●●●	●●●	●●●	●●●	●●●	●●●	●●●
ERV1–1-I_BT	●●●	●●●	●●●	●●●	●●●	●●●	●●●	●●●	●●●	●●●	●●●
BtERVF2-I	□□□	□□□	□□□	□□□	□□□	○□□	○□□	○○○	○□□	●●○	●●□
ERV2–1-I_BT	●●●	●●□	●●□	●●●	●●□	●●●	●●□	●●●	●●●	●●●	●●●
ERV2–2-I_BT	○□□	○□□	□□□	○□□	□□□	○○○	○○○	○○○	●○○	□□□	○□□
ERV2–3-I_BT	○□□	○□□	□□□	○□□	□□□	○○□	○□□	□□□	□□□	○□□	□□□

3.1.5 Conclusions

Three autonomous retrotransposons, *L1_BT*, *BovB*, and *ERV1-1-I_BT*, with complete ORFs were found with consistent expression in bovine oocytes and crossing all stages during preimplantation embryo development. These retrotransposons were analysed during further research in this thesis.

3.1.6 Acknowledgements

The authors like to thank Isabel Lemahieu and Petra Van Damme for their excellent technical assistance with *in vitro* production of bovine embryos. We also thank Ruben Van Gansbeke for his assistance in this study.

Chapter 3.2 Reference Gene Selection for Normalization of RT-qPCR Data in Sexed Bovine Preimplantation Embryo under Low and High Oxygen Tension

Wenwen Li[1], Karen Goossens[1], Mario Van Poucke[1], Ann Van Soom[2], Luc Peelman[1]

[1] *Department of Nutrition, Genetics and Ethology, Faculty of Veterinary Medicine, Ghent University, B-9820 Merelbeke, Belgium*

[2] *Department of Obstetrics, Reproduction and Herd Health, Faculty of Veterinary Medicine, Ghent University, B-9820 Merelbeke, Belgium*

Adapted from a poster presentation at the 4th Mammalian Embryo Genomics meeting, 2013, Quebec city, Canada.

3.2.1 Abstract

Real-time quantitative polymerase chain reaction (RT-qPCR) is an efficient and accurate tool for gene expression analysis in preimplantation embryos. To normalize gene expression data, stable reference genes are needed. The stability of reference genes can vary under different developing stages or culture conditions, hence it is essential to determine reference gene stability before initiating a gene expression analysis. In the present study, we have analyzed and selected reference genes for normalization of RT-qPCR data of bovine preimplantation embryos, considering gender, developmental stages and culture conditions. *GAPDH, YWHAZ, 18S rRNA* and *SDHA* were identified as the most stable genes overall (all embryos stages, both male and female, and cultured under 5% and 20% O_2). Oxygen tension shows no effect on the gene expression stability ranking order, while in female embryos, *HPRT1* and *H2A* are the most stably expressed genes.

3.2.2 Introduction

Preimplantation embryo development is a dynamic developmental process from fertilization until formation of a blastocyst, including first cleavage division, activation of the embryonic genome and differentiation to the inner cell mass (ICM) and trophectoderm (TE) cells. These developmental activities are directed by gene expression regulation dynamics leading to complex expression patterns that not only vary from embryonic stage to stage, but also differ between *in vivo* and *in vitro* produced embryos (Niemann & Wrenzycki 2000).

To detect and analyze gene expression levels, RT-qPCR (reverse transcription quantitative polymerase chain reaction) is one of the most widely used methods. It is a highly specific and sensitive tool for quantifying RNA levels in preimplantation embryos. The reliability of RT-qPCR results, besides a good experimental and primer design, also depends on accurate applications of all the procedures, such as quality RNA extraction, cDNA synthesis, dilution series, pipetting, use of appropriate controls, and analysis methods (Gál *et al.* 2006). Unlike cell lines and single tissue samples, embryonic cells have a heterogeneous nature, which leads

to a greater variation in the endogenous biological processes, and a greater variation in the sensitivity of the cells to administered compounds (Zhang *et al.* 2003). Therefore, a reliable normalization procedure is of utmost importance.

Constitutively expressed genes minimally influenced by the experimental treatments, are used in most experiments to normalize the gene expression levels. Reference genes have been previously validated in bovine preimplantation embryos (Goossens *et al.* 2005). However, it is essential to determine reference gene stability before initiating a gene expression study in new conditions. The stability of reference genes can vary under different culture conditions, for example high oxygen tension. Another factor that can have an influence on the gene expression is the sex of the embryo. It was reported that male and female preimplantation embryos display a sex-specific transcriptional regulation. Using DNA microarrays almost one third of the expressed genes were reported as having different expression levels in male and female bovine blastocysts (Bermejo-Alvarez *et al*. 2010). To incorporate this potential gender specific expression, an embryo sexing method is needed before proceeding with the RT-qPCR analysis. Normally, the sexing technique is based on chromosomal difference between the sexes (presence or absence of Y chromosome). By PCR amplification of embryo genomic DNA (from whole embryos or some blastomeres), Y-specific sequences can be detected. However, in the present study, an RNA-based embryo sexing method was evaluated before proceeding with the RT-qPCR analysis in order to avoid consumption of cells for DNA extraction and reduce the already limited material for RNA expression.

3.2.3 Material and methods

Material

Unless stated otherwise, all chemicals, reagents and media were obtained from Sigma (USA) and Life Technologies (Belgium).

Experimental design

The scheme of the study is show in **Figure 3.3**. For the reference gene selection, embryos were collected from all stages cultured under 5% O_2 and 20% O_2. RT-qPCR was performed with 20 experimental groups of 6 single embryos, including 2-cell, 4-cell, 8-cell male, 8-cell female, 16-cell male, 16-cell female, morula male, morula female, blastocyst male and female.

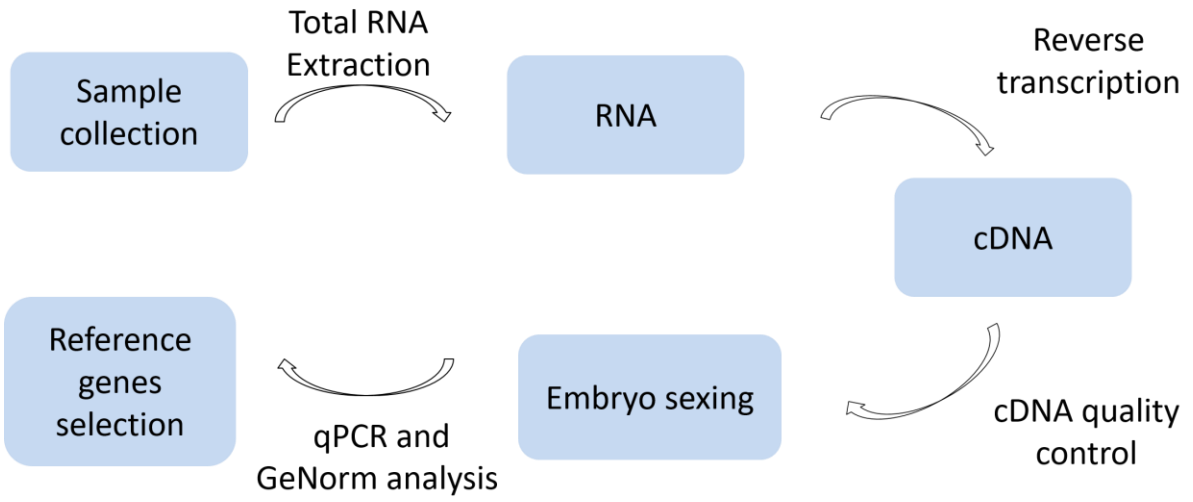


Figure 3.3 Scheme of reference gene selection in bovine embryos

In vitro embryo production

In vitro bovine embryos were produced in serum-free media according to our standard procedures, as mentioned in **Chapter 3.1** *Method*. For the high oxygen tension group, embryos were cultured under 20% O_2, instead of 5% O_2. Embryos with good morphology were collected at specific time points: 2-cell and 4- cell at 36-40 hpi, 8- cells at 64 hpi, 16-cell at 5 dpi, morula at 6 dpi, and expanded blastocyst at 8 dpi. Embryo developmental competence was evaluated at 8 dpi by the percentage of blastocysts obtained from presumptive zygotes.

For evaluating the RNA-based sex determination test, 8-cell embryos and blastocysts were cut into two halves. One half of the embryo was used to test the RNA-based sexing method and the other half was used in a classic DNA-based sexing method. Embryos used for the DNA test were washed and frozen with PBS, and stored at -80 °C until use. All embryos for RNA extraction were further washed three times with RNase-free PBS, frozen individually in lysis buffer (10% RNAsin Plus RNase inhibitor (Promega, Netherlands), 5% DTT (Promega, Netherlands), and 0.8% Igepal CA-630 in RNase free water) and stored at -80 °C, until use.

Embryo RNA extraction, cDNA synthesis and quality control

Total RNA was isolated from the individual embryos of each stage, using the RNeasy Micro kit (Qiagen, Belgium) according to the manufacturer's instructions, including a genomic DNA removal step in the procedure. The extracted RNA was dissolved in 14 µl of RNase-free water. A minus RT control was performed with *GAPDH* primers to check for contaminating genomic DNA. First strand cDNA was generated from the total amount of RNA using the iScript cDNA synthesis kit (BioRad, Belgium) according to the product manual. After reverse transcription, the cDNA was 2.5 times diluted and used for downstream PCR.

An *YWHAZ* (tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, zeta) primer assay was used to check embryo cDNA integrity for the RT-qPCR experiment. The assay has one forward primer and three reverse primers, which can amplify amplicons of 109, 497, 909 bp. Samples with three amplicons are considered as sufficient quality for RT-qPCR (primers shown in **Table 3.4**).

RNA-based embryo sex determination

An RNA-based method for embryo sexing was performed using a Y-linked gene and an X-linked gene as an internal control. The Y-linked gene DEAD (Asp-Glu-Ala-Asp) box polypeptide 3 (*DDX3Y*) primers were taken from Hamilton *et al*. (2012). According to our preliminary results from expression profiling of *DDX3Y* among embryo stages, *DDX3Y* was

found expressed from the 8-cell stage on, maybe due to embryo genome activation (EGA). Consequently, RNA expression of this gene can only be used for gender determination from the 8-cell stage onwards. As X-linked internal control hypoxanthine phosphoribosyltransferase 1 (*HPRT1*) was used, and the primers were taken from previous work of the lab (Goossens *et al.* 2005). Primer sequences are given in **Table 3.4**. The PCR protocol used was as described above for the retrotransposon expression profiling experiment with an annealing temperature of 60°C.

Table 3.4 Information of the primers used for embryo sexing and RNA integrity check

Gene name	Sequence 5' – 3'	Amplicon length cDNA	Ta (°C)
DDX3Y	GGACGTGTAGGAAACCTTGG	225	60
	GCCAGAACTGCTACTTTGTCG		
HPRT1	TGCTGAGGATTTGGAGAAGG	154	60
	CAACAGGTCGGCAAAGAACT		
AMXY	CAGCCAAACCTCCCTCTGC	280 on X chr	58
	CCCGCTTGGTCTTGTCTGTTGC	217 on Y chr	
YWHAZ assay +1	GAGCAAAAGACGGAAGGTGCT		60
YWHAZ assay -1	TCCCCACCAGGACATACCAA	909	
YWHAZ assay -2	TCCGATGTCCACAATGTCAAGT	497	
YWHAZ assay -3	CCAAAAGAGACAGTACATCATTGCA	109	

Embryo DNA preparation and DNA sexing method

Genomic DNA of half 8-cell embryos and half blastocysts was isolated according to our routine lab protocol, including Proteinase K treatment, with minor changes (Van Poucke *et al.* 2005). In brief, each sample was added to 20 µl Lysis Buffer K (10 mM Tris-HCl pH 8; 50

mM KCl; 0.5% Tween 20) supplemented with 500 µg/ml proteinase K (Roche Diagnostics, Belgium) and incubated for 45 min at 56°C to release the DNA. The lysate was incubated for 10 min at 95°C to inactivate the proteinase K and centrifuged at 16100 g for 1 min to pellet down the cell debris.

The amelogenin gene present on both the X and the Y chromosome in the bovine genome was used for the embryo sexing. The amelogenin gene on the X chromosome has a different size compared to the gene on the Y (also called AMXY) chromosome and therefore can be used for sexing. The primers used are listed in **Table 3.4**.

RT-qPCR analysis

Seven candidate gene primer sequences for RT-qPCR experiments were taken from previous work in the laboratory (Goossens *et al.* 2005). Six single embryos were used for each developmental stage and sex (2-cell, 4-cell, 8-cell male and female, 16-cell male and female, morula male and female, expanded blastocyst male and female) from the high oxygen (20% O_2) and control (5% O_2) groups separately.

RT-qPCR reactions were performed in 10 µl reaction volume on a BioRad CFX 96 PCR Detection system, including 5 µl Sso Advanced SYBR Green Supermix (BioRad, Belgium), 300 nM each primer and 2.5 µl diluted embryo cDNA. The PCR program consisted of an initial denaturation step at 95°C for 3 min to activate the Taq DNA polymerase, followed by 40 cycles of denaturation for 5 s at 95°C and a combined primer annealing-extension step for 30 s at specific primer annealing temperatures, during which fluorescence was measured. A melting curve was produced afterwards by heating the samples from 70°C to 95°C in 0.5°C increments for 5 s and fluorescence was monitored at the same time to confirm a single specific peak for each pair of primers. Each reaction was run in duplicate. PCR efficiencies were calculated by a relative standard curve derived from a pooled bovine cDNA mixture (a 10-fold dilution series with five measuring points).

Determination of reference genes by geNorm

Raw RT-qPCR data were calculated taking into account exact PCR efficiencies and were analysed using the geNorm program (Vandesompele *et al.* 2002). This program ranks the stability of gene expression from a set of tested candidate reference genes by M value in a given sample panel. The M value for each reference gene is the average pairwise variation for that gene with all the other tested control genes. A higher M value indicates a greater variation of gene expression. Stepwise exclusion of the gene with the highest M value and recalculation of the M value allows the ranking of the tested genes according to their expression stability. Finally, a normalization factor (NF) was calculated based on the geometric mean of the given housekeeping genes.

3.2.4 Results and discussion

In vitro embryo production under low and high oxygen tension

In six replicate IVF experiments (both a control group at 5% O_2 and an oxidative stress group at 20% O_2), 20% O_2 had an impact on embryo development. The blastocyst rate at 8 d.p.i. was significantly ($P < 0.01$) reduced under conditions of oxidative stress (20% O_2) compared to control (23.34 ± 5.47% vs 45.32 ± 7.73%, respectively; **Figure 3.4**).

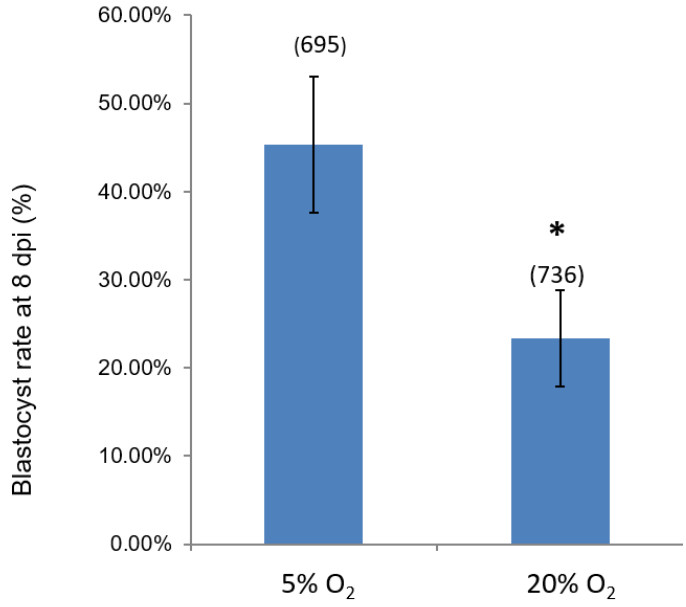


Figure 3.4 Blastocyst rate at 8 dpi from presumed zygotes under normal *in vitro* embryo culture or oxidative stress (n= presumed zygotes number, in 6 replicate IVF experiments). Data are shown as Mean±SEM, and an asterisk indicates a significant difference between two culture conditions (P<0.01) by paired student t-test.

RNA quality of the embryo samples

RNA samples without genomic DNA contamination were further tested with the *YWHAZ* primer assay. For each cDNA sample, three amplicons were found, with the largest of 909 bp (**Figure 3.5**). Since the primers designed for qPCR only amplify fragments < 300 bp in length, we suppose that samples with a 497 bp-amplicon have sufficient quality for further qPCR analysis.

(a)

(b)

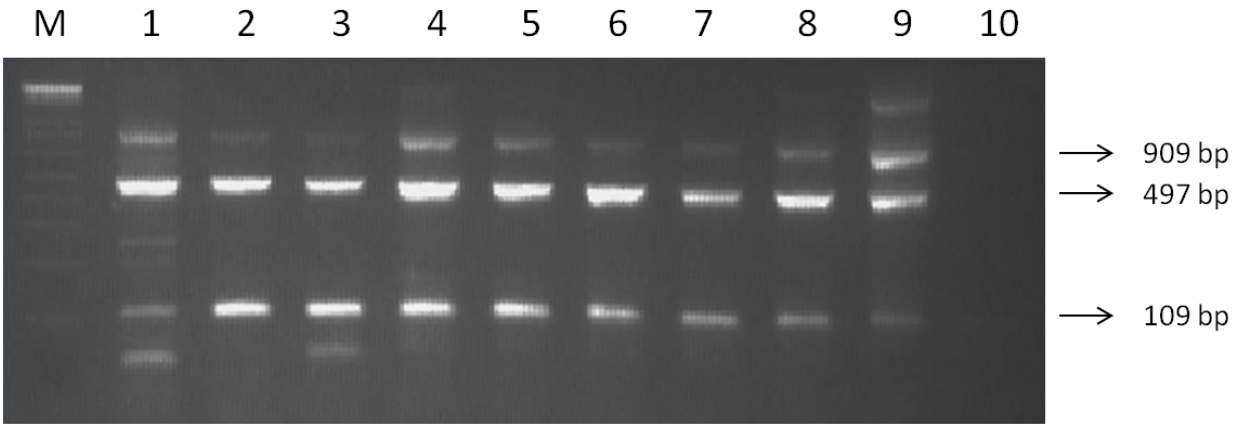


Figure 3.5 RNA quality control primer assay. (a) YWHAZ primers design (b) embryo samples checked by the primer assay; samples are used when 909 bp can be amplified. M: DNA marker; 1-8: single preimplantation embryo cDNA (1&2: 2-cell; 3&4: 4-cell; 5-6: 8-cell; 7: morula; 8: blastocyst); 9: bovine tissue cDNA; 10: negative control.

RNA-based embryo sex determination

Comparison of the results from the *DDX3Y-HPRT1* RNA-based method and the AMXY DNA-based gender marker showed the *DDX3Y-HPRT1* RNA-based method is reliable for bovine embryo sexing from the 8-cell stage onwards (**Figure 3.6**). Moreover, four blastomeres from an 8-cell embryo can provide enough RNA for this test. Therefore, embryos used in the present study were all sexed from the 8-cell stage onwards.

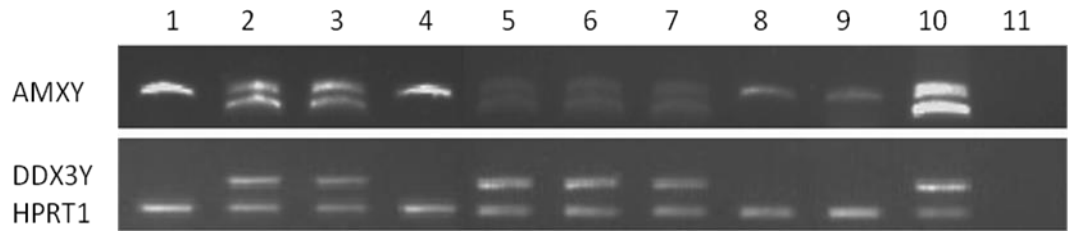


Figure 3.6 Accordance of DNA-based and RNA-based embryo sexing results. Genomic AMXY was used as a gender marker (upper penal), and Y-specific gene *DDX3Y* expression was used for sexing while *HPRT1* was added as an internal control. 1-4: 8-cell; 5-8 blastocyst; 9: female tissue DNA and RNA; 10: male tissue DNA and RNA; 11: negative control. 1, 4, 8 are female embryos and 2, 3, 5, 6, 7 are male embryos.

Reference gene selection

A good reference gene should express stably over all experimental groups. The selection of reference genes for qPCR in bovine preimplantation embryos has been reported before (Goossens *et al.* 2005). Different with that study, we selected reference genes for individual embryo qPCR instead of embryo groups, sexed the embryos and tested the influence of high oxygen tension (20% O_2 culture).

Table 3.5 Ranking of expression stability (M) for each embryo developmental stage by geNorm

2-cell	**M**	**4-cell**	**M**	**8-cell**	**M**	**16-cell**	**M**	**Morula**	**M**	**Blastocyst**	**M**
YWHAZ/ HPRT1	0.426	*GAPDH/ YWHAZ*	0.461	*YWHAZ/ ACTB*	0.811	*YWHAZ/ SDHA*	0.927	*ACTB/ SDHA*	0.980	*YWHAZ/ ACTB*	0.546
GAPDH	0.733	*ACTB*	0.638	*GAPDH*	0.913	*GAPDH*	1.046	*YWHAZ*	1.020	*SDHA*	0.578
SDHA	0.909	*HPRT1*	0.870	*H2A*	0.975	*ACTB*	1.131	*GAPDH*	1.054	*GAPDH*	0.637
H2A	0.969	*H2A*	0.949	*18S rRNA*	1.070	*18S rRNA*	1.200	*H2A*	1.098	*H2A*	0.708
ACTB	1.024	*SDHA*	1.013	*HPRT1*	1.165	*HPRT1*	1.344	*18S rRNA*	1.174	*18S rRNA*	0.761
18S rRNA	1.123	*18S rRNA*	1.212	*SDHA*	1.270	*H2A*	1.578	*HPRT1*	1.337	*HPRT1*	0.905

Reference genes were analysed by geNorm and ranked according to the gene stability measure (M) (Vandesompele *et al.* 2002). The ranking of gene expression stability is shown

by stages (**Table 3.5**) and by other sample combinations (**Table 3.6**). As expected, the expression stability ranking of the reference genes varied among stages and varied according to gender (from the 8-cell to blastocyst stage). Compared to the results of Goossens *et al.* (2005), we found no difference in the stability ranking of the reference genes between pooled and individual embryos. For both, the most stable genes are *GAPDH*, *YWHAZ* and *18S rRNA*. When gene expression is calculated only within one developmental stage, the gene expression stability (M) is relatively low, and the stability ranking order is slightly different. It is notable that in female embryos the stability rank is different from male embryos. *HPRT1* is the most stable reference gene in female embryos, maybe due to the X-linked expression, with a double expression before X-inactivation compared to males (**Table 3.6**). In male embryos, *GAPDH* and *YWHAZ* are the most stable genes. High oxygen tension does not change the ranking significantly. *GAPDH* and *YWHAZ* are the most stable reference genes, with low M-value, when all stages, both sexes and oxygen tension are considered (**Table 3.6**).

Table 3.6 Ranking of expression stability (M) for different sample combinations by geNorm

All female (8-cell to Blastocyst)	M	All male (8-cell to Blastocyst)	M	All 5% O_2 (2-cell to Blastocyst)	M	All 20% O_2 (2-cell to Blastocyst)	M	**All**	**M**
HPRT1/ H2A	1.339	*GAPDH/ YWHAZ*	1.025	*GAPDH/ YWHAZ*	1.066	*GAPDH/ YWHAZ*	0.972	*GAPDH/ YWHAZ*	1.016
YWHAZ	1.436	*18S rRNA*	1.231	*HPRT1*	1.407	*18S rRNA*	1.395	*18S rRNA*	1.458
GAPDH	1.503	*SDHA*	1.477	*H2A*	1.468	*SDHA*	1.546	*SDHA*	1.566
18S rRNA	1.637	*ACTB*	1.632	*18S rRNA*	1.648	*ACTB*	1.696	*HPRT1*	1.742
SDHA	1.722	*HPRT1*	1.907	*SDHA*	1.746	*HPRT1*	1.918	*H2A*	1.849
ACTB	1.896	H2A	2.053	*ACTB*	1.861	*H2A*	2.076	*ACTB*	1.976

Surprisingly, we found that *18S rRNA* is always among the least stable genes in **Table 3.5**, while one of the most stable genes in **Table 3.6**. In this reference gene selection experiment, we found that *18s rRNA* is one of the least stable genes in single embryonic stage ranking, but one of the most stable across the embryo stages, which indicates *18s rRNA* is relatively stable expressed in different stages compared to other candidate genes.

After gene expression stability (M) calculation and ranking by geNorm, the number of reference genes chosen for RT-qPCR normalization needs to be determined. It was suggested that in order to measure gene expression levels accurately, a normalization factor (NF) calculated by the geometric mean of multiple control genes instead of one is required (Vandesompele *et al.* 2002). These authors recommend three reference genes as the minimum number for RT-qPCR normalization. The optimal number of reference genes can be determined by a pairwise variation ($V_{n/n+1}$) between the sequential normalization factors (NF_n and NF_{n+1}) after successive inclusion of less stable reference genes. In other words, stepwise inclusion of more control genes until the (n+1)th gene has no significant contribution to the recalculated NF. **Figure 3.7** shows the pair wise variation V_n/V_{n+1} between 2 sequential normalization factors NF_n and NF_{n+1} for all samples (all stages, both genders, 5% and 20% O_2). In this case, $V_{3/4}$ (0.360) is much lower than $V_{2/3}$ (0.534), while adding the 5th gene does not change the pairwise variation value (0.356). Adding the 6th gene reduces the value to 0.299. However, due to the limited RNA amount that can be extracted from single embryos, and the relative small gain of adding more genes, four reference genes is considered a good compromise in this case. The four most stable genes are *GAPDH*, *YWHAZ*, *18S rRNA* and *SDHA*.

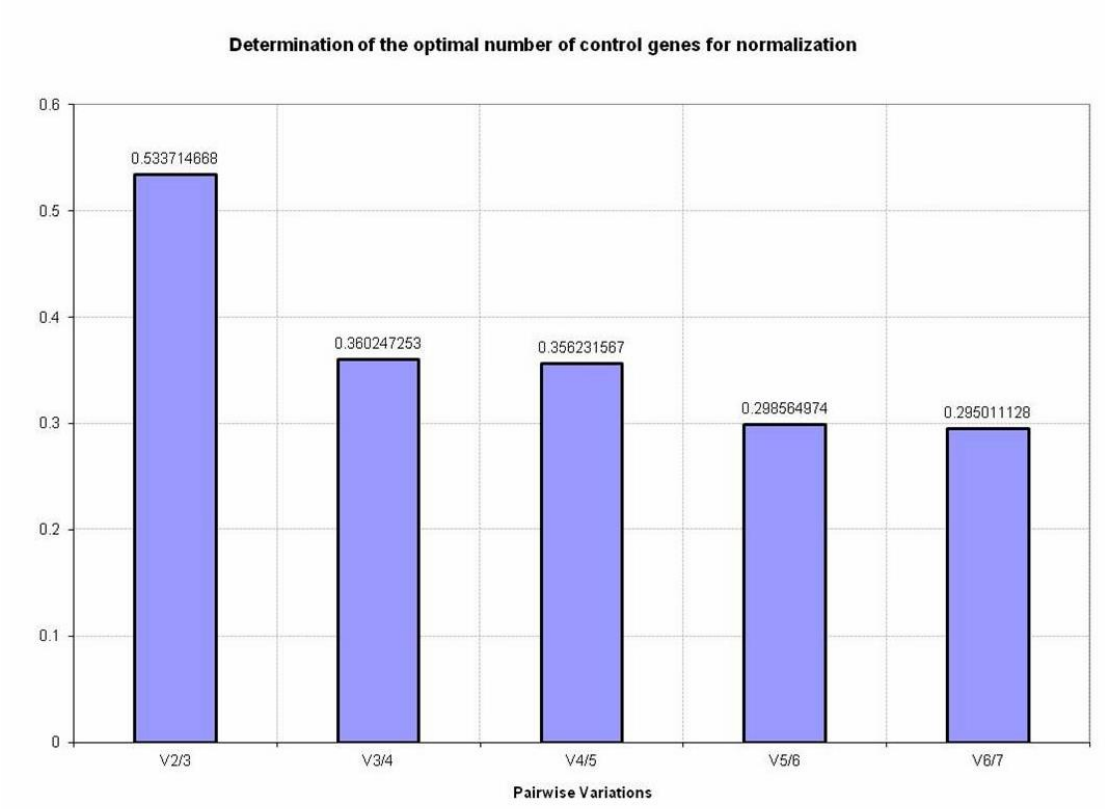


Figure 3.7 Determination of optimal number of control genes for normalization of all samples. The optimal number of control genes was calculated by geNorm. The value of the pair wise variation significantly reduces to 0.360 for $V_{3/4}$, and inclusion of the 5th gene shows no difference in calculation.

3.2.5 Conclusions

In conclusion, methods of quality controlled RNA extraction from embryos and RNA-based embryo sexing were optimized, and a selection of reference genes for RT-qPCR analysis for individual (male or/and female) embryos under 5% and 20% O_2 was made. *GAPDH*, *YWHAZ*, *18S rRNA* and *SDHA* are the most stable genes over all the embryonic stages (both male and female, and cultured under 5% and 20% O_2). Oxygen tension shows no effect on the gene expression stability ranking order.

3.2.6 Acknowledgements

The authors like to thank Isabel Lemahieu and Petra Van Damme for their excellent technical assistance.

3.3 References

Adelson D.L., Raison J.M. & Edgar R.C. (2009) Characterization and distribution of retrotransposons and simple sequence repeats in the bovine genome. *Proceedings of the National Academy of Sciences* **106**, 12855-60.

Arányi T., Váradi A., Simon I. & Tusnády G.E. (2006) The BiSearch web server. *BMC bioinformatics* **7**, 431.

Bermejo-Alvarez, P., Rizos, D., Rath, D., Lonergan, P. and Gutierrez-Adan, A. (2010) Sex determines the expression level of one third of the actively expressed genes in bovine blastocysts. *Proceedings of the National Academy of Sciences* **107**, 3394-9.

Bratthauer G. & Fanning T. (1992) Active LINE-1 retrotransposons in human testicular cancer. *Oncogene* **7**, 507-10.

Bui, L. C., Evsikov, A. V., Khan, D. R., Archilla, C., Peynot, N., Henaut, A., ... & Duranthon, V. (2009). Retrotransposon expression as a defining event of genome reprograming in fertilized and cloned bovine embryos.*Reproduction*, **138(**2), 289-299.

Corpet F. (1988) Multiple sequence alignment with hierarchical clustering. *Nucleic acids research* **16**, 10881-90.

Gál Á.B., Carnwath J.W., Dinnyes A., Herrmann D., Niemann H. & Wrenzycki C. (2006) Comparison of real-time polymerase chain reaction and end-point polymerase chain reaction for the analysis of gene expression in preimplantation embryos. *Reproduction, Fertility and Development* **18**, 365-71.

Georgiou I., Noutsopoulos D., Dimitriadou E., Markopoulos G., Apergi A., Lazaros L., Vaxevanoglou T., Pantos K., Syrrou M. & Tzavaras T. (2009) Retrotransposon RNA

expression and evidence for retrotransposition events in human oocytes. *Human molecular genetics* **18**, 1221-8.

Goossens K., Van Poucke M., Van Soom A., Vandesompele J., Van Zeveren A. & Peelman L.J. (2005) Selection of reference genes for quantitative real-time PCR in bovine preimplantation embryos. *BMC developmental biology* **5**, 27.

Graf, A., Krebs, S., Heininen-Brown, M., Zakhartchenko, V., Blum, H., & Wolf, E. (2014). Genome activation in bovine embryos: review of the literature and new insights from RNA sequencing experiments. *Animal reproduction science*, **149**(1), 46-58.

Hamilton C.K., Combe A., Caudle J., Ashkar F.A., Macaulay A.D., Blondin P. & King W.A. (2012) A novel approach to sexing bovine blastocysts using male-specific gene expression. *Theriogenology* **77**, 1587-96.

Jurka J., Kapitonov V.V., Pavlicek A., Klonowski P., Kohany O. & Walichiewicz J. (2005) Repbase Update, a database of eukaryotic repetitive elements. *Cytogenetic and genome research* **110**, 462-7.

Larkin M., Blackshields G., Brown N., Chenna R., McGettigan P.A., McWilliam H., Valentin F., Wallace I.M., Wilm A. & Lopez R. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947-8.

Macia A., Muñoz-Lopez M., Cortes J.L., Hastings R.K., Morell S., Lucena-Aguilar G., Marchal J.A., Badge R.M. & Garcia-Perez J.L. (2011) Epigenetic control of retrotransposon expression in human embryonic stem cells. *Molecular and cellular biology* **31**, 300-16.

Niemann H. & Wrenzycki C. (2000) Alterations of expression of developmentally important genes in preimplantation bovine embryos by in vitro culture conditions: implications for subsequent development. *Theriogenology* **53**, 21-34.

Peaston, A.E., Evsikov, A.V., Graber, J.H., de Vries, W.N., Holbrook, A.E., Solter, D. and Knowles, B.B. (2004) Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos. *Developmental cell* **7**, 597-606.

Pérez-Novo C.A., Claeys C., Speleman F., Van Cauwenberge P., Bachert C. & Vandesompele J. (2005) Impact of RNA quality on reference gene expression stability. *Biotechniques*

39, 52-6.

Raiz J., Damert A., Chira S., Held U., Klawitter S., Hamdorf M., Löwer J., Strätling W.H., Löwer R. & Schumann G.G. (2012) The non-autonomous retrotransposon SVA is trans-mobilized by the human LINE-1 protein machinery. *Nucleic acids research* **40**, 1666-83.

Rozen S. & Skaletsky H. (1999) Primer3 on the WWW for general users and for biologist programmers. In: *Bioinformatics methods and protocols* (pp. 365-86. Springer.

Van Poucke M., Vandesompele J., Mattheeuws M., Van Zeveren A. & Peelman L.J. (2005) A dual fluorescent multiprobe assay for prion protein genotyping in sheep. *BMC infectious diseases* **5**, 13.

Vandesompele J., De Preter K., Pattyn F., Poppe B., Van Roy N., De Paepe A. & Speleman F. (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome biology* **3**, research0034.

Wheeler D.L., Chappey C., Lash A.E., Leipe D.D., Madden T.L., Schuler G.D., Tatusova T.A. & Rapp B.A. (2000) Database resources of the national center for biotechnology information. *Nucleic acids research* **28**, 10-4.

Wydooghe E., Heras S., Dewulf J., Piepers S., Van Den Abbeel E., De Sutter P., Vandaele L. & Van Soom A. (2013) Replacing serum in culture medium with albumin and insulin, transferrin and selenium is the key to successful bovine embryo development in individual culture. *Reproduction, Fertility and Development*.

Zhang Q.J., Chadderton A., Clark R.L. & Augustine-Rauch K.A. (2003) Selection of normalizer genes in conducting relative gene expression analysis of embryos. *Birth Defects Research Part A: Clinical and Molecular Teratology* **67**, 533-44.

Zuker M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic acids research* **31**, 3406-15.

Chapter 4

High Oxygen Tension Increases Global Methylation in Bovine 4-cell Embryos and Blastocysts but does not Affect General Retrotransposon Expression

Wenwen Li[1], Karen Goossens[1], Mario Van Poucke[1], Katrien Forier[2,3], Kevin Braeckmans[2,3], Ann Van Soom[4] and Luc Peelman[1]

[1]*Department of Nutrition, Genetics and Ethology, Faculty of Veterinary Medicine, Ghent University, Heidestraat 19, 9820 Merelbeke, Belgium.*

[2]*Laboratory of General Biochemistry and Physical Pharmacy, Ghent University, Harelbekestraat 72, 9000 Ghent, Belgium.*

[3]*Center for Nano- and Biophotonics, Harelbekestraat 72, 9000 Ghent, Belgium.*

[4] *Department of Obstetrics, Reproduction and Herd Health, Faculty of Veterinary Medicine, Ghent University, Salisburylaan 133, 9820 Merelbeke, Belgium.*

Adapted from Reproduction, Fertility and Development 28(7): 948-959, 2014

http://dx.doi.org/10.1071/RD14133.

4.1 Abstract

Retrotransposons are transposable elements that insert extra copies of themselves throughout the genome via an RNA intermediate using a '*copy and paste*' mechanism. They account for more than 44% of the bovine genome and have been reported to be functional, especially during preimplantation embryo development. In the present study, we tested whether high oxygen tension (20% O_2) influences global DNA methylation analysed by immunofluorescence staining of developing bovine embryos and whether this has an effect on the expression of some selected retrotransposon families. High oxygen tension significantly increased global DNA methylation in 4-cell embryos and blastocysts. A significant expression difference was observed for *ERV1-1-I_BT* in female blastocysts, but no significant changes were observed for the other retrotransposon families tested. Therefore, the study indicates that global DNA methylation is not necessarily correlated with retrotransposon expression in bovine preimplantation embryos.

Keywords: bovine preimplantation embryo, DNA methylation, oxidative stress, reference genes.

4.2 Introduction

Transposable elements, the largest class of mammalian genome sequences, were long considered as 'junk' or 'selfish' DNA. The major part of transposable elements in mammals is made up of retrotransposons also called RNA transposons. They insert extra copies of themselves throughout the genome via an RNA intermediate using a '*copy and paste*' mechanism. Retrotransposons account for more than 44% of the bovine genome (Adelson *et al.* 2009). Most retrotransposons have accumulated mutations and are evolutionarily inactivated (Horie *et al.* 2007). Those still potentially capable of jumping within the genome play an important role in gene regulation in developing embryos, especially autonomous retrotransposons with an open reading frame (ORF) for reverse transcriptase, because not only can they transport themselves back to the genome, but they also help the insertion of short interspersed elements (SINEs; Dewannieux *et al*. 2003; Raiz *et al.* 2012). Endogenous reverse transcriptase encoded from autonomous retrotransposons is also a mediator of cell proliferation and differentiation (Spadafora 2004). It was proposed that differential retrotransposon expression triggers sequential reprogramming of the embryonic genome in preimplantation embryos (Peaston *et al.* 2004). It was also shown that expression of autonomous long interspersed element (LINE)-1 retrotransposons is essential for preimplantation embryo development in the mouse (Beraldi *et al*. 2006) and that elimination of these autonomous retrotransposons resulted in developmental arrest at the 2-cell stage and deviant expression of approximately 40% of genes sampled. Most retrotransposition events are demonstrated to take place in early embryos, gametes and cancer cells (Ostertag *et al.* 2002; Kigami *et al.* 2003; Schulz 2006). Transposition is markedly downregulated in non-pathological differentiated tissues with a highly methylated genome (Reik *et al*. 2001; Shi *et al*. 2007; Howard *et al*. 2008). DNA methylation is a major epigenetic mechanism in silencing retrotransposons (Bestor and Bourc'his 2004), influencing DNA transcription by preventing or enhancing the binding of regulatory transcription factors to the promoter region, and thus determining which genes are eventually transcribed. Because of the enormous percentage of retrotransposons in the whole genome, it was suggested that most methylated

cytosines are within transposons (Yoder *et al*. 1997), and more than 90% of all 5-methylcytosines lie within cytosine-phosphorous-guanine (CpG) islands of transposons, including LINEs and SINEs (Asada *et al*. 2006). An *in vitro* study suggested that DNA methylation of the LINE-1 retrotransposon promoter is essential for LINE-1 silencing (Hata and Sakaki 1997) and activation of retrotransposons during preimplantation embryo development correlates with the loss of DNA methylation in mice (Peaston *et al*. 2004).

Oxidative stress is generated by excessive production of reactive oxygen species and/or reduction of antioxidant defenses, and has been reported to induce epigenetic changes in different cell types (Ehrlich 2002; Burroughs *et al*. 2013). Oxidative stress by means of high O_2 tension during embryo culture alters the expression of epigenome modifying genes such as the ten-eleven translocation (*TET*) gene, which is associated with conversion of 5-methylcytosine to 5-hydroxymethylcytosine at the 16-cell and blastocyst stages of bovine IVF embryos (Burroughs *et al*. 2013).

McClintock's (1984) 'genomic shock' hypothesis suggests that environmental stimuli may mobilise transposable elements. This hypothesis is supported by evidence that retrotransposons from various organisms can be mobilised by a variety of environmental stresses, including oxidative stress (Morales *et al*. 2003). Based on these findings, it can be hypothesised that excess oxidative stress induces DNA methylation changes and this, in turn, influences retrotransposon expression and eventually embryo health. To test this hypothesis, we applied sustained oxidative stress (20% O_2) during *in vitro* bovine preimplantation embryo development and measured global DNA methylation and expression of three bovine retrotransposon families in different preimplantation stages up to the blastocyst, also taking into account embryo gender, because this is known to affect one-third of the actively expressed genes in bovine blastocysts (Bermejo-Alvarez *et al*. 2010b).

4.3 Materials and methods

Materials

Unless stated otherwise, all chemicals, reagents and media were obtained from Sigma (USA) and Life Technologies (Belgium).

Experimental design

To study retrotransposon expression in bovine embryos, preliminary expression profiling of potentially autonomous retrotransposon families was described in **Chapter 3.1** by reverse transcription-polymerase chain reaction (RT-PCR). Because no consistent results could be obtained with four of the seven families tested due to low expression and/or high variation, these four families were left out, and *L1_BT*, *BovB* and *ERV1-1-I_BT* were used for quantification by reverse transcription quantitative polymerase chain reaction (RT-qPCR) in further experiments.

Global DNA methylation was measured by immunofluorescence staining of embryos cultured at 5% and 20% O_2. Because DNA methylation differences between the two groups were observed only in 4-cell embryos and blastocysts, these specific stages were used for the RT-qPCR experiment. Before performing the RT-qPCR experiment with the three selected retrotransposon families, a set of reference genes was evaluated for normalisation of the final RT-qPCR data of these two embryo stages.

In vitro embryo production

In vitro bovine embryos were produced in serum-free media according to standard procedures (Wydooghe *et al.* 2014). Briefly, bovine ovaries were collected at the local slaughterhouse. Immature cumulus-oocyte complexes (COCs) were aspirated from follicles between 2 and 8 mm in diameter and matured in groups of 60 in 500 µl modified bicarbonate-buffered TCM-199 supplemented with 50 mg/ml gentamicin and 20 ng/ml epidermal growth factor

(EGF) for 22 h at 38.5°C in a 5% CO_2 incubator. After maturation, the COCs were inseminated with frozen-thawed bovine spermatozoa at a final concentration of 10^6 spermatozoa/ml. The cumulus cells and excess spermatozoa were removed from presumptive zygotes by vortexing after 21 h incubation. The presumptive zygotes were incubated in groups of 25 in 50 µl synthetic oviductal fluid supplemented with essential and non-essential amino acids (SOFaa), 0.4% bovine serum albumin (BSA) and ITS (5 µg/ml insulin + 5 µg/ ml transferrin + 5 ng/ml selenium), covered with mineral oil and incubated at 38.5°C under 5% CO_2, 5% O_2 and 90% N_2 (as control group) or 5% CO_2 in air (20% O_2; as the high oxygen tension group). Embryos with good morphology were collected for immunofluorescence staining and gene expression at specific time points: 2-cell and 4- cell at 36-40 hpi, 8- cell at 64 hpi, 16-cell at 5 dpi, morula at 6 dpi, and expanded blastocyst at 8 dpi.

Embryos destined for gene expression analysis were washed three times with RNase-free phosphate-buffered saline (PBS), frozen individually or in pools of five in lysis buffer (10% RNAsin Plus RNase inhibitor (Promega, Netherlands), 5% dithiothreitol (Promega, Netherlands) and 0.8% Igepal CA-630 in RNase free water) and stored at -80°C, whereas embryos for immunofluorescence staining were washed in warm PBS supplemented with 0.5% BSA (PBS-BSA), fixed for 20 min in 4% paraformaldehyde at room temperature and then stored in PBS-BSA at 4°C until use.

DNA methylation immunofluorescence staining

Staining was performed as described by Beaujean *et al.* (2004). In all cases, at least 10 embryos from each stage (2-cell to blastocyst) and both culture conditions (5% O_2 and 20% O_2) were washed extensively in PBS-BSA before further processing, and all steps were performed at room temperature if not detailed otherwise. Briefly, fixed embryos were permeabilised with 0.5% Triton X-100 and 0.05% Tween overnight at 4°C, and treated with 4 M HCl for 1 h at 37°C before being neutralised with 100 mM Tris-HCl (pH 8.0) for 10 min. After three washes in PBS-BSA, embryos were blocked with 10% goat serum for 1 h to block non-specific binding sites. Next, samples were incubated overnight at 4°C with a mouse

monoclonal antibody against 5-methylcytosine (1: 100 dilution in blocking solution; Gentaur, Belgium), followed by a 30 min wash with PBS-BSA and 1 h incubation with a secondary antibody, namely goat anti-mouse fluorescein isothiocyanate (FITC; 1 : 100 dilution in blocking solution; Invitrogen, Belgium). Nuclei were stained with 25 µg/mlpropidium iodide (Invitrogen, Belgium) for 30 min in the dark. Embryos were finally deposited on slides and mounted in DABCO under a coverslip.

Images were taken with the Nikon (Japan) C1si confocal laser scanning microscope system using an oil-free objective (×40) and excitation wavelengths of 488 and 637 nm. The depth of the confocal slide was 4 µm and Z-stacks were merged by ImageJ software (Abràmoff *et al*. 2004) to produce a two-dimensional image under each channel. The fluorescence quantification was corrected for background by subtracting the mean intensity of the cytoplasmic area surrounding each nucleus (Beaujean *et al*. 2004). All nuclei from 2-cell to morula embryos were analysed and 25 nuclei were randomly selected per blastocyst to analyse. Due to the DNA methylation dynamics, one microscopy setting will induce over exposure or underexposure, which leads to wrong quantification. Therefore, we applied one setting for each embryonic stage, and the DNA methylation level between high oxygen tension and the control group are only compared within each stage. Statistical analysis was performed using an unpaired Student's *t*-test comparing the methylation level (corrected fluorescence) of individual nuclei from embryos cultured at 20% O_2 or 5% O_2 of each developing stage.

RNA extraction and cDNA synthesis

For RNA extraction, total RNA was isolated from the individual embryos of each stage, using the RNeasy Micro kit (Qiagen, Belgium) according to the manufacturer's instructions, including a genomic DNA removal step in the procedure. The extracted RNA was dissolved in 14 µl of RNase-free water. A minus RT control was performed with *GAPDH* primers to check for contaminating genomic DNA. First strand cDNA was generated from the total amount of RNA using the iScript cDNA synthesis kit (BioRad, Belgium) according to the

product manual. After reverse transcription, the cDNA was 2.5 times diluted and used for downstream PCR.

RNA quality control and RNA-based sexing for embryos

Because the cDNA amount is too low to be evaluated by traditional methods, an *YWHAZ* primer assay was applied to check embryo cDNA integrity for the RT-qPCR experiment, as described in **Chapter 3.2**. Samples with three amplicons of 109, 497 and 909 bp are considered as sufficient quality for RT-qPCR. As described in **Chapter 3.2**, embryos from 8-cell stage can be sexed with the *DDX3Y-HPRT1* RNA-based method, and the expression of retrotransposons was analysized for both genders.

Reference gene selection by geNorm

We have described reference gene selection for normalization of qPCR data in bovine preimplantation embryos before. However, a re-selection of the reference genes based on the new experimental parameters improves RT-qPCR normalization, as described in **Chapter 3.2**. Based on the results of the DNA methylation immunofluorescence staining only the relevant embryo stages were further studied. The stability of housekeeping genes and the optimal number of reference genes were recalculated with the geNorm program including only the stages showing differential methylation staining.

DNA methyltransferase and retrotransposon expression analysis by RT-qPCR

Retrotransposon families with quantifiable expression (*L1_BT*, *BovB* and *ERV1-1-I_BT*) were selected for RT-qPCR. Of the DNA methyltransferases (*DNMT*s), three members have been shown to be active (Okano et al. 1999): *DNMT1* is the most abundantly expressed methyltransferase and is considered the key maintenance methyltransferase, whereas *DNMT3a* and *DNMT3b* can methylate hemimethylated and unmethylated DNA, as required for *de novo* methylation. Therefore, expression of these three members was determined.

RT-qPCR was only performed on the stages with differential methylation staining results, resulting in six experimental groups (4-cell, male and female expanded blastocyst from both 5% O_2 and 20% O_2 culture). Six single embryos were used per experimental group.

All RT-qPCR experiments were performed according to the MIQE (Minimum Information for Publication of Quantitative Real- Time PCR Experiments) guidelines (Bustin *et al*. 2009). qPCR reactions were performed in 10 µL reaction volume on a BioRad CFX 96 PCR Detection system, including 5 µL Sso Advanced SYBR Green Supermix (BioRad), 300 nM each primer and 2.5 µL diluted embryo cDNA. The PCR program consisted of an initial denaturation step at 95°C for 3 min to activate the Taq DNA polymerase, followed by 40 cycles of denaturation for 5 s at 95°C and a combined primer annealing-extension step for 30 s at specific primer annealing temperatures as given in **Table 4.1**, during which fluorescence

Table 4.1 Information on the primers used in the study, and those with qPCR efficiency were used for expression quantification. The amplicon number for *DNMTs* is based on bovine cDNA databases (UMD 3.1.86); while the potential amplicon numbers for retrotransposons are based on bovine genome (UMD_3.1.86).

Retrotransposon name	Amplicon numbers (by BiSearch)	Sequence 5′ – 3′	Amplicon length cDNA	Ta (°C)	qPCR efficiency
DNMT1	1	AGCGCCTCAGCTAAAATCAA ACAAACACCGCATACGACAC	157	58	105.85%
DNMT3a	1	GCATTGTGTCTTGGTGGATG CTTGTTGTAGGTGGCCTGGT	158	60	104.35%
DNMT3b	1	AAGACCGGCCTTTCTTCTGGATGT TGTGAGCAGCAGACACTTTGATGG	129	57	97.23%
L1_BT	6606	GCAATCCCTATCAAGCTACCA TGATTCCTCCAGTTCCATTCTT	133	60	100.30%
BovB	25639	CCTCAGATATGCAGATGACACC GTTGGTCATAACTTTCCTTCCA	283	60	101.10%
ERV1-1-I_BT	103	TGTTAAGCTCAAAGACCCACAC CCGTTCAGGAATTTCAGACAA	277	60	107.00%

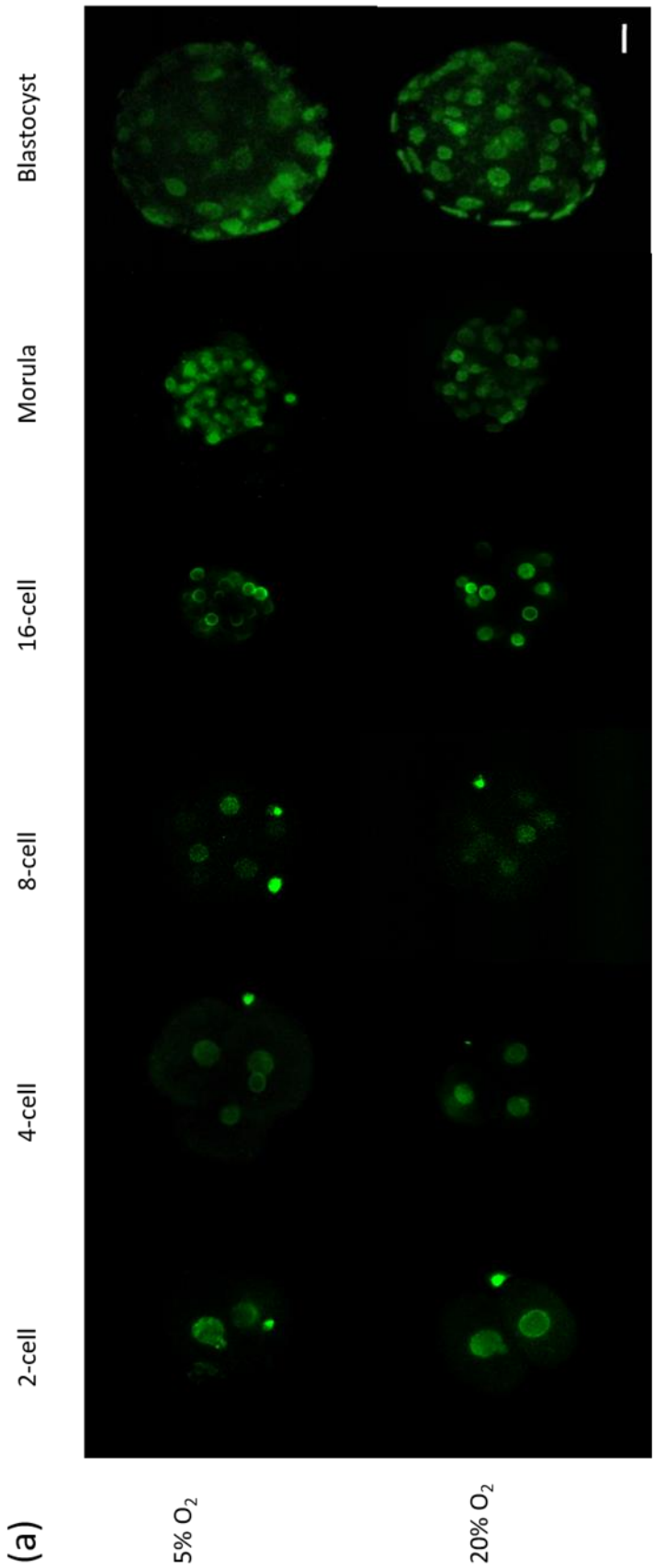
was measured. A melting curve was produced afterwards by heating the samples from 70°C to 95°C in 0.5°C increments for 5 s and fluorescence was monitored at the same time to confirm a single specific peak for each pair of primers. Each reaction was run in duplicate. PCR efficiencies were calculated by a relative standard curve derived from a pooled bovine cDNA mixture (a 10-fold dilution series with five measuring points). All PCR efficiencies were between 90% and 110%.

The geometric mean of the reference genes was used to calculate an accurate normalisation factor. The mean quantity of each transcript was divided by the respective normalisation factor to obtain a normalised value according to the method described by Hellemans *et al.* (2007). The highest value was assigned 100%, and the other normalised values were divided by the calibrator value to generate relative expression levels.

4.4 Results

DNA methylation immunofluorescence staining

Representative images of global DNA methylation staining for embryos from the 2-cell to blastocyst stage are shown in **Figure 4.1a**, with quantification of normalised DNA methylation levels summarised in **Figure 4.1b** (control group at all stages set to 100%). There was a significant ($P < 0.05$) increase in normalised DNA methylation levels at 20% O_2 in embryos at the 4-cell and blastocyst stages based on mean fluorescence per nucleus compared with the control group (151.5 ± 11.9% vs 100 ± 9.6%, respectively, for 4-cell embryos; 220.1 ± 8.6% vs 100 ± 4.3%, respectively, for blastocysts). No differences were observed at the other developmental stages. Consequently, we used only 4-cell and blastocyst embryos to test the hypothesis put forward in the Introduction.


(a)
5% O_2
20% O_2
2-cell
4-cell
8-cell
16-cell
Morula
Blastocyst

(b)

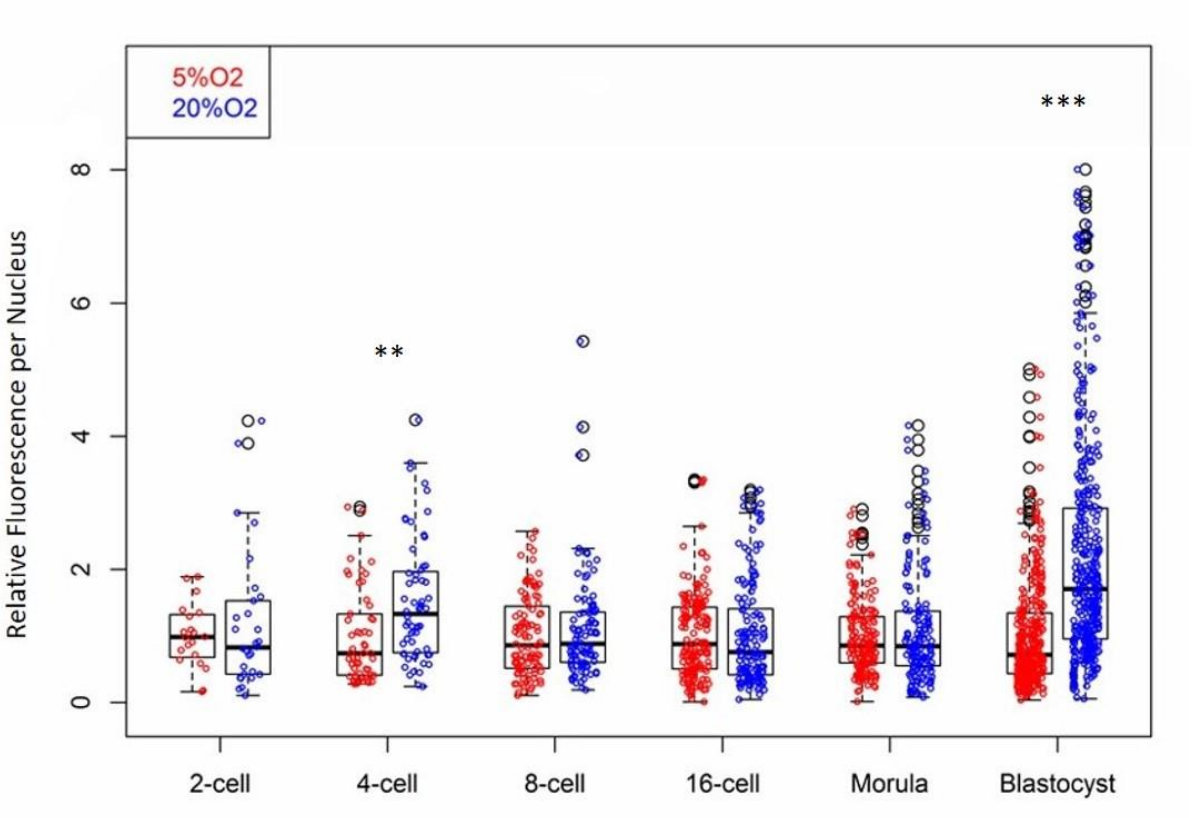


Figure 4.1 Global DNA methylation analysis in bovine *in vitro* embryos cultured at 5% O_2 or 20% O_2 from 2-cell stage to blastocyst. (a) Confocal Z-stack images of 5-methylcytosine immunodetection in bovine preimplantation embryos at 5% O_2 (upper panel) and 20% O_2 (lower panel). The scale bar represents 20 μm. (b) Statistical analysis by corrected fluorescence quantification per nucleus. Corrected fluorescence of all stages from the 5% O_2 culture is set as 1 (red), and 20% O_2 cultured embryos are presented relatively to the corresponding stage. The non-parametric test Kruskal-Wallis was used to compare corrected fluorescence between two groups since the variances were not equal. Significances were presented as *<0.05; ** < 0.01; ***<0.001.

Reference gene selection

A good reference gene is stably expressed over all experimental groups. In the present study, because we found significant changes in DNA methylation between 20% O_2 and 5% O_2 treatments in 4-cell and blastocyst embryos, reference gene selection and retrotransposon expression were further focused on these two stages. With geNorm analysis, three reference genes, namely *GAPDH*, *YWHAZ* and *SDHA* were chosen for normalisation of retrotransposon expression data in 4-cell embryos and blastocysts (taking into account gender and oxygen tension).

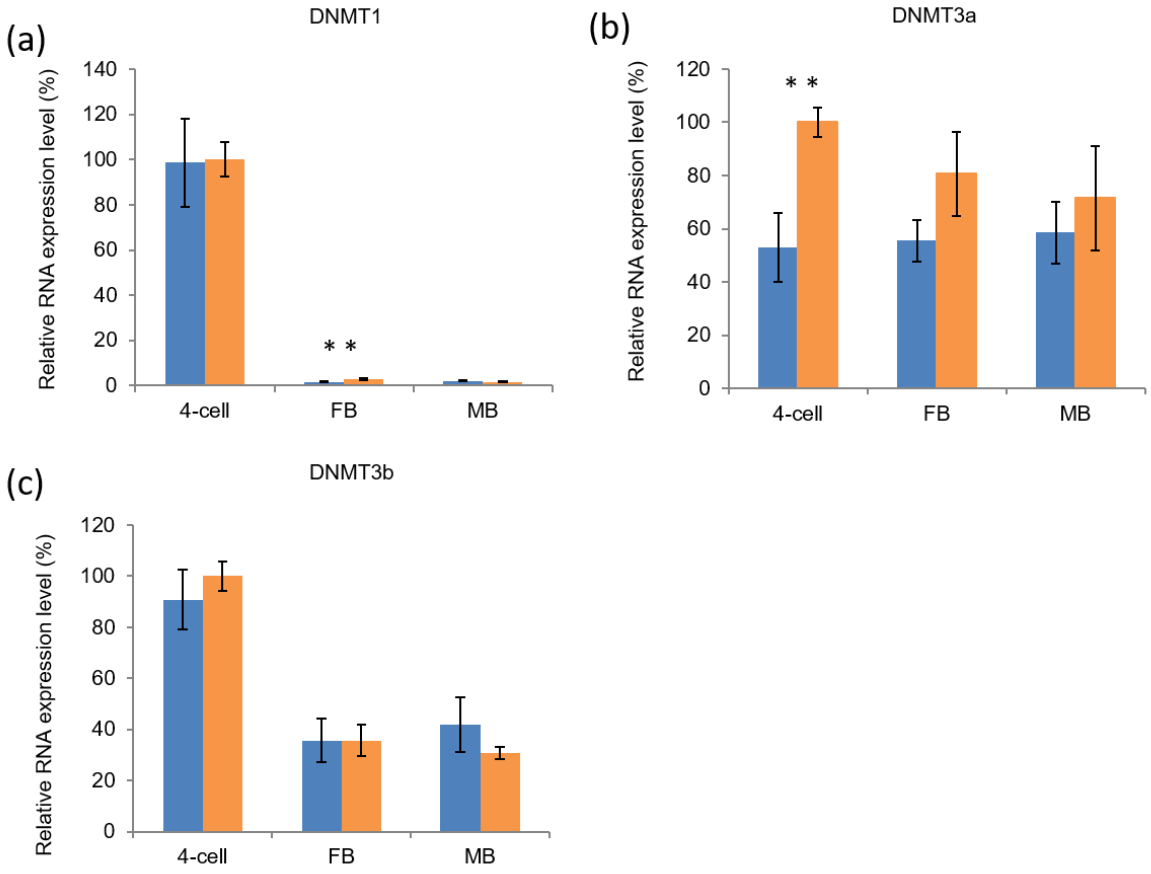


Figure 4.2 Relative DNMTs mRNA expression levels in 4-cell stage and blastocyst cultured at 5% O_2 (blue) or 20% O_2 (orange) by RT-qPCR. For each gene, expression levels were compared to the highest expression, set as 100%. Data shown are Mean±SEM. The unpaired student-t test was used to compare gene expression of each stage between two groups. Significant differences between two culture conditions were presented as *<0.05; ** < 0.01; ***<0.001. MB: male blastocyst; FB: female blastocyst.

Expression of DNMTs

Expression of *DNMT1* was higher in female, but not male, blastocysts cultured under 20% O_2 compared with 5% O_2. *DNMT3a* expression was higher at the 4-cell stage in embryos cultured under 20% O_2 compared with 5% O_2. No significant differences were found in *DNMT3b* expression for embryos cultured under 20% O_2 versus 5% O_2, but there was a decrease in expression between 4-cell embryos and blastocysts. In addition, *DNMT1* expression dropped markedly in blastocysts compared with 4-cell embryos. There was no significant change in *DNMT3a* expression between 4-cell embryos and blastocysts (**Figure 4.2**).

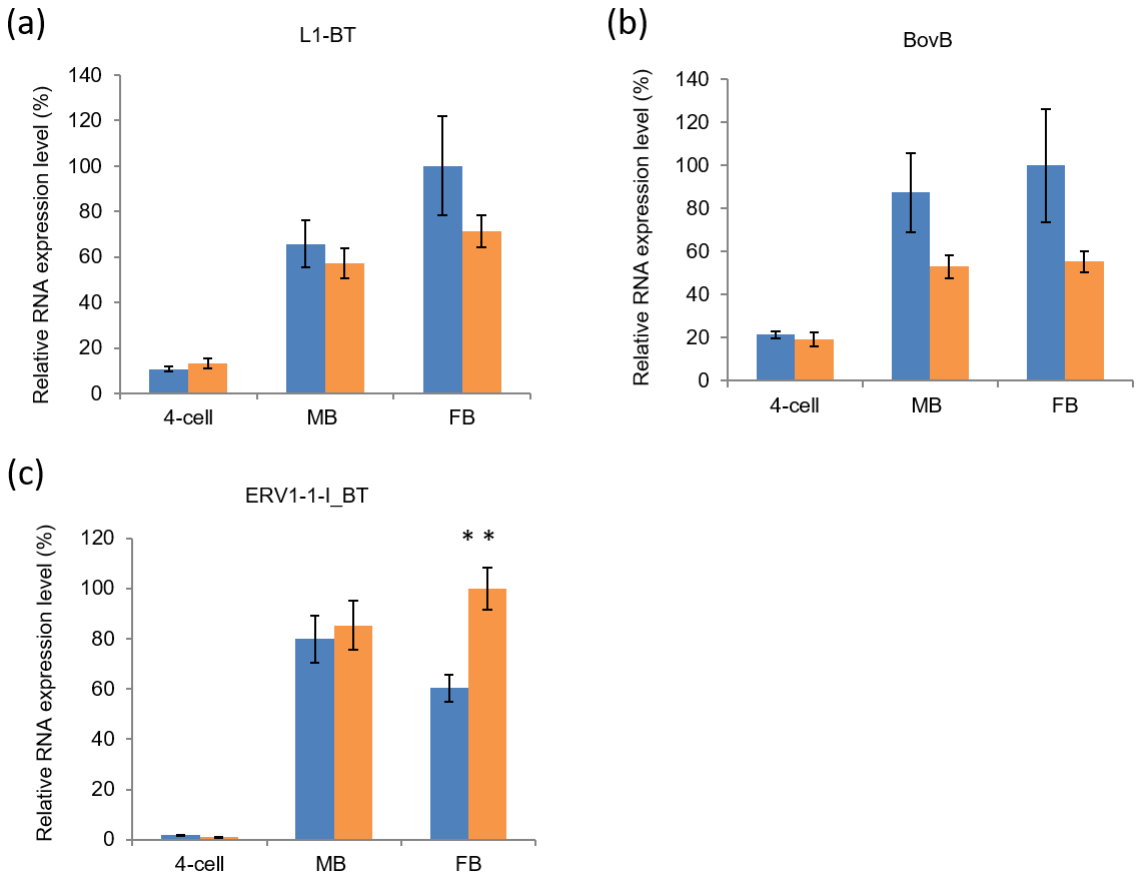


Figure 4.3 Relative retrotransposon mRNA expression levels in 4-cell stage and blastocyst cultured at 5% O_2 (blue) or 20% O_2 (orange) by RT-qPCR. For each gene, expression levels were compared to the highest expression, set as 100%. Data shown are Mean±SEM. The unpaired student-t test was used to compare gene expression of each stage between two groups. Significant differences between two culture conditions were presented as *<0.05; ** < 0.01; ***<0.001. MB: male blastocyst; FB: female blastocyst.

Retrotransposon expression

RT-qPCR was performed for *L1_BT*, *BovB* and *ERV1-1-I_BT* on single embryo samples from the 4-cell stage and sexed blastocysts cultured at 20% O_2 and 5% O_2 tension (stages chosen according to the results of the DNA methylation study). In general, for the three retrotransposon families tested (**Figure 4.3**), the transcription level was significantly lower 4-cell embryos than blastocyst regardless of oxygen tension or gender. For *L1_BT*, *BovB* and *ERV1-1-I_BT*, there were no differences between 5% and 20% O_2 culture conditions in 4-cell embryos.

In blastocysts, there was a significant increase in *ERV1-1-I_BT* expression at 20% O_2 in female blastocysts ($P < 0.01$; **Figure 4.3c**), but no significant difference was observed for other retrotransposons.

Correlation between the DNMTs and retrotransposon expression

As can be seen in Figure 4.2 and Figure 4.3 the expression of DNMTs dropped from 4-cell to blastocyst, while the expression of retrotransposons was increased in blastocysts. Based on these results, a correlation of the expression pattern between the two embryonic stages was checked with SPSS (**Figure 4.4**). A negative correlation between *DNMT1*/*DNMT3b* and the retrotransposon expression was found. No linear relationship has been found between *DNMT3a* and the retrotransposon expression.

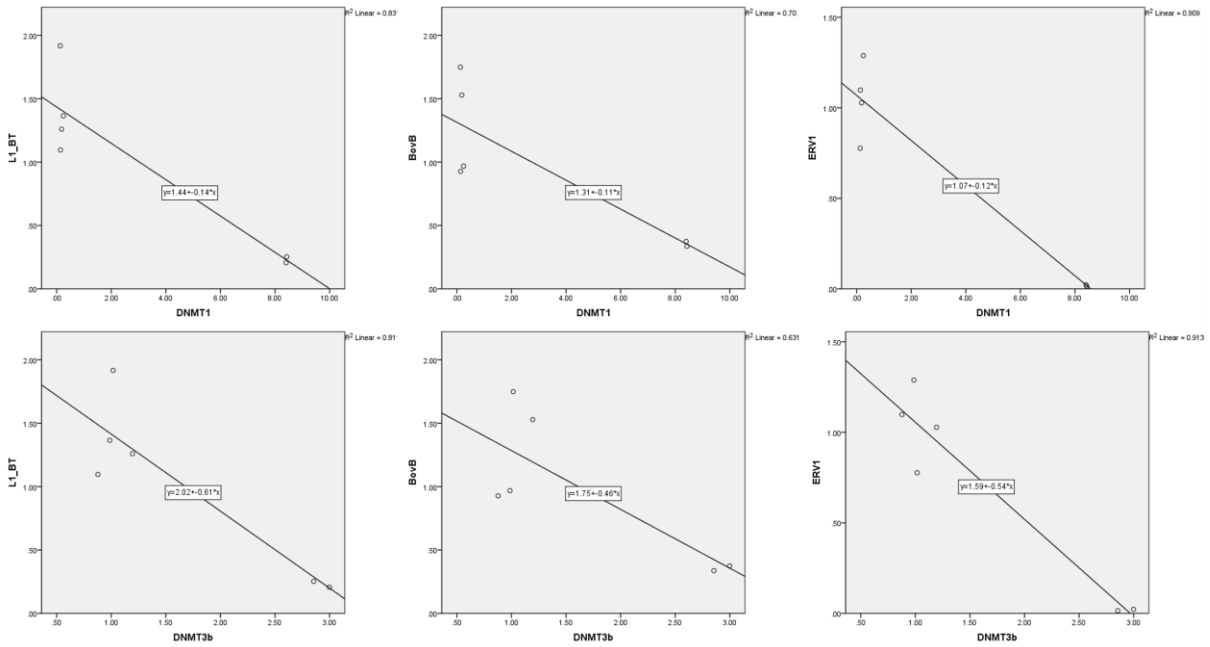


Figure 4.4 Linear relationship between the expression of *DNMT*s and retrotransposons.

4.4 Discussion

Following Barbara McClintock's (1984) 'genomic shock' hypothesis, we hypothesised that oxidative stress induces changes in DNA methylation and that this, in turn, influences retrotransposon silencing and eventually embryo health. In the present study, we confirmed that 20% O_2 leads to a significantly lower 8 dpi blastocyst rate compared with culture at 5% O_2. This is in agreement with other studies (Van Soom *et al*. 2002; Yuan *et al*. 2003; Burroughs *et al*. 2013). In those studies, it was found that culture at 20% O_2 leads to a lower blastocyst rate and a higher apoptotic cell ratio, and that 20% O_2 has an effect on epigenetic-related gene expression. It has been proposed recently that oxidative stress may affect DNA methylation by DNA oxidation or TET protein-mediated hydroxymethylation (Perillo *et al.* 2008; Vanden Berghe 2012).

An indirect immunofluorescence staining method with antibodies against 5-methylcytosine has been widely used in the past decade to detect dynamic changes in preimplantation embryo

DNA methylation (Santos and Dean 2006). The paternal and maternal genomes are both highly methylated at the moment of fertilization, and previous studies with immunofluorescence staining showed that in bovine preimplantation embryos the paternal genome undergoes sharp active demethylation within only a few hours of fertilization, whereas the maternal genome undergoes passive demethylation from cleavage division to the 8-cell stage. *De novo* methylation starts at the 8-cell stage (Dean *et al.* 2001; Reik *et al*. 2001; Dobbs *et al*. 2013). In the present study, immunofluorescence staining revealed a significant increase in methylation at the 4-cell stage under high oxygen tension, but not at the 2- or 8-cell stages. This points to a more moderate rate or a delay in demethylation induced by oxidative stress. A similar trend was found in cloned bovine embryos (Bourc'his *et al*. 2001; Dean *et al.* 2001). Conversely, higher expression of *DNMT3a* was found at the 4-cell stage under high oxygen tension. It has been reported that after the first cell division and another round of DNA replication, the maternal chromosome consists of a hemimethylated and a fully demethylated sister chromatid (Haaf 2006), and *DNMT3a* shows methylation activity on both unmethylated and hemimethylated DNA (Hsieh 2005); thus, we suggest that oxidative stress may accelerate *de novo* methylation at the 4-cell stage. Normal IVF bovine embryos reach the lowest methylation level at the 8-cell stage (Reik *et al*. 2001; Dobbs *et al*. 2013). This is a critical survival period, as is indicated by the observation that approximately half the embryos stop growing at this stage (Bourc'his *et al*. 2001). Thus, it is possible that embryos with an altered methylation status may stop developing at the 4- to 8-cell stage, explaining the drop in methylation status observed after the 4-cell stage because only correctly demethylated embryos develop further. More intensive 5-methylcytosine staining was found in blastocysts under high oxygen tension, whereas increased *DNMT1* expression was found only in female blastocysts. It has been reported that overexpression of *DNMT1* causes genomic hypermethylation (Biniszkiewicz *et al*. 2002).

We did not find different methylation levels between genders in the present study; however, there are some conflicting reports about methylation in both sexes. It was reported that *DNMT3a* and *DNMT3b* were significantly upregulated in male compared with female bovine

blastocysts and a higher methylation level for a repeated sequence was found in male blastocysts (Bermejo-Alvarez *et al*. 2008; Bermejo-Álvarez *et al*. 2010a). However, others have reported more intense labelling for 5-methylcytosine in female embryos at the 6- to 8-cell stage and lower staining in male embryos at the blastocyst stage (Dobbs *et al.* 2013). Given these results, DNA methylation in both sexes may be best approached separately in future studies.

A recent study using a new staining method challenges the widely accepted idea that preimplantation development goes through a global demethylation phase (Li and O'Neill 2012). In that study, the authors found different results of paternal demethylation during mouse zygote maturation when a trypsin digestion step was included. They claimed that this is due to the onset of a progressive acid-resistant antigenic masking of 5-methylcytosine and that trypsin digestion of zygotes can remove this masking (Li and O'Neill 2012). However, it was not mentioned whether trypsin digestion changed methylation detection at later embryo stages or in other species, and this step should be taken into account in future staining also. In addition to anti-5-methylcytosine immunofluorescence staining, DNA methylation is being studied more and more by sequencing-based methods, including bisulfite sequencing (BS-seq; Nakanishi *et al*. 2012; Smith *et al*. 2012), methylated DNA binding domain protein sequencing (MBD-seq), methylated DNA immunoprecipitation sequencing (MeDIP-seq; Li *et al*. 2010) and whole-genome shotgun bisulfite sequencing (SBS; Kobayashi *et al*. 2012). However, these methods differ in CpG coverage, resolution, quantitative accuracy and cost (Harris *et al*. 2010), so one should be cautious when interpreting the results obtained with only one of these techniques.

In blastocysts, we observed weaker methylation in the inner cell mass (ICM) than in trophectoderm (TE) cells under both low and high oxygen tension. Other studies showed a species-specific methylation pattern between ICM and TE at the blastocyst stage. For instance, higher methylation levels in ICM than TE cells have been found in mouse and sheep (Dean *et al*. 2001; Beaujean *et al*. 2004), whereas more methylation in TE cells than the ICM was

found in humans and rabbits (Fulka *et al.* 2004; Shi *et al.* 2004). However, there are conflicting data as to whether DNA methylation is more extensive in ICM or TE cells in bovine blastocysts (Dean *et al.* 2001; Hou *et al.* 2007; Dobbs *et al.* 2013). A possible explanation for the conflicting results may be differences in the staining methods used. For this reason, we tested some modifications to the immunofluorescence staining method. Interventions such as prolonging the time of DNA denaturation or antibody incubation, or cutting the edge of the blastocysts to make it easier for the antibodies to reach ICM cells, did not change the methylation levels we observed. It has also been suggested that different embryo production methods generate different methylation patterns in pigs and cattle (Santos *et al.* 2003; Deshmukh *et al.* 2011), so another explanation for the differences observed may be the influence of the serum-free embryo culture medium that we used. Compared with *in vivo* embryo and serum-plus *in vitro* embryo culture, blastocysts produced under serum-free condition have slower development (Rizos *et al*. 2003) and the slower development may also cause a delay in the *de novo* methylation of the ICM.

We further tested whether the differences in methylation caused by the high O_2 tension had an effect on the expression of the three selected retrotransposon families. Only *ERV1-1-I_BT* was significantly differentially expressed with higher expression ($P = 0.007$) in female blastocysts in the 20% O_2 group. This is surprising because a higher methylation level, as observed in blastocysts cultured at 20% O_2 compared with 5% O_2, would be expected to lead to lower gene expression. For the other two retrotransposon families tested, there was a tendency for lower expression, but the differences did not reach statistical significance (**Figure 4.3**). This unexpected result may be because the global methylation status as determined here is not representative for specific loci containing certain genes. The fact that the difference was observed in female and not in male blastocysts may be a consequence of the complex interactions between chromatin modifiers encoded by the X-chromosome induced by oxidative stress. Because X-inactivation happens between 8 and 14 dpi in bovine embryos (De La Fuente *et al*. 1999), female embryos carrying two active X-chromosomes can potentially produce twice the amount of X-linked transcripts relative to the male embryos

with only one X-chromosome (Epstein *et al*. 1978); the damage to X-chromosomes is also double in female than male embryos. This may also explain the reports of one group that not only do male embryos develop faster than female embryos, but that more male than female embryos reach the expanded blastocyst stage under *in vitro* culture conditions (Gutiérrez-Adán *et al*. 1996, 2001). ERV elements were found to be significantly enriched in bovine X-chromosomes using BLAST and Retrotector v1.0 (Sperber *et al*. 2007; Garcia-Etxebarria and Jugo 2010). Together, these results indicate that regulation of retrotransposon expression is not a simple consequence of DNA methylation.

As found here, except for ERV1-1-I_BT in female blastocysts, high O_2 tension-induced DNA methylation does not have a significant effect on the expression of other retrotransposon families in this study. The same phenomenon has been found recently in mouse zygotes (Inoue *et al*. 2012), as well as in MORC (microrchidia) ATPases mutants in *Arabidopsis* and *Caenorhabditis elegans* (Moissiard *et al*. 2012). Inoue *et al*. (2012) used short interference (si) RNA-mediated depletion of *Tet3*, which is responsible for oxidation of 5-methylcytosine to 5-hydroxymethylcytosine (Gu *et al.* 2011), to inhibit 5-methylcytosine oxidation in mouse zygotes and found that the transcriptional levels of LINE-1 and ERV1 were not significantly correlated with changes in DNA methylation. It was also found that DNA methylated genes and transposable elements were derepressed, with no loss of DNA methylation or histone modification, which was associated with decondensation of pericentromeric heterochromatin and disruption of the chromosome superstucture (Moissiard *et al*. 2012). In addition to DNA methylation and chromosome structure, endogenous siRNAs have been reported to play an important role in retrotransposon silencing (Yang and Kazazian 2006; Tam *et al.* 2008; Watanabe *et al.* 2008).

In conclusion, in the present study we found that oxidative stress (20% O_2) induced higher DNA methylation at the 4-cell and blastocyst stages in bovine embryo development. However, a link between global DNA methylation and general retrotransposon expression was not found under oxidative stress conditions, but in bovine implantation embryos stages. The

regulation of retrotransposon expression is probably, like that of other genes, a combination of various mechanisms with an interplay of different types of regulatory factors and not a simple consequence of methylation-demethylation.

4.5 Acknowledgements

The authors thank Petra Van Damme and Isabel Lemahieu for their excellent technical assistance with the *in vitro* culture of embryos.

4.6 References

Abràmoff, M.D., Magalhães, P.J. and Ram, S.J. (2004) Image processing with ImageJ. *Biophotonics international* **11**, 36-42.

Adelson, D.L., Raison, J.M. and Edgar, R.C. (2009) Characterization and distribution of retrotransposons and simple sequence repeats in the bovine genome. *Proceedings of the National Academy of Sciences* **106**, 12855-60.

Asada, K., Kotake, Y., Asada, R., Saunders, D., Broyles, R.H., Towner, R.A., Fukui, H. and Floyd, R.A. (2006) LINE-1 hypomethylation in a choline-deficiency-induced liver cancer in rats: dependence on feeding period. *BioMed Research International* **2006(1)**, 17142.

Beaujean, N., Taylor, J., Gardner, J., Wilmut, I., Meehan, R. and Young, L. (2004) Effect of limited DNA methylation reprogramming in the normal sheep embryo on somatic cell nuclear transfer. *Biology of reproduction* **71**, 185-93.

Beraldi, R., Pittoggi, C., Sciamanna, I., Mattei, E. and Spadafora, C. (2006) Expression of LINE-1 retroposons is essential for murine preimplantation development. *Molecular reproduction and development* **73**, 279-87.

Bermejo-Álvarez, P., Lonergan, P., Rath, D., Gutiérrez-Adan, A. and Rizos, D. (2010) Developmental kinetics and gene expression in male and female bovine embryos produced in vitro with sex-sorted spermatozoa. *Reproduction, Fertility and Development* **22**, 426-36.

Bermejo-Alvarez, P., Rizos, D., Rath, D., Lonergan, P. and Gutierrez-Adan, A. (2008) Epigenetic differences between male and female bovine blastocysts produced in vitro. *Physiological genomics* **32**, 264-72.

Bermejo-Alvarez, P., Rizos, D., Rath, D., Lonergan, P. and Gutierrez-Adan, A. (2010) Sex determines the expression level of one third of the actively expressed genes in bovine blastocysts. *Proceedings of the National Academy of Sciences* **107**, 3394-9.

Bestor, T. and Bourc'his, D. (2004) Transposon silencing and imprint establishment in mammalian germ cells. In: *Cold Spring Harbor symposia on quantitative biology*, pp.

381-8. Cold Spring Harbor Laboratory Press.

Biniszkiewicz, D., Gribnau, J., Ramsahoye, B., Gaudet, F., Eggan, K., Humpherys, D., Mastrangelo, M. A., Jun, Z., Walter, J. and Jaenisch, R. (2002) Dnmt1 overexpression causes genomic hypermethylation, loss of imprinting, and embryonic lethality. *Molecular and cellular biology* **22**, 2124-35.

Bird, A.P. and Wolffe, A.P. (1999) Methylation-Induced Repression- Belts, Braces, and Chromatin. *Cell***99**, 451-4.

Bourc'his, D., Le Bourhis, D., Patin, D., Niveleau, A., Comizzoli, P., Renard, J. P., and Viegas-Pequignot, E. (2001). Delayed and incomplete reprogramming of chromosome methylation patterns in bovine cloned embryos. *Current Biology* **11(19)**, 1542-46.

Burroughs, C.A., Williamson, G.L., Golding, M.C. and Long, C.R. (2012) Oxidative stress induced changes in epigenetic modifying gene mRNA in pre-implantation *in vitro* bovine embryos. *Reproduction, Fertility and Development* **25**, 149-.

Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W. and Shipley, G.L. (2009) The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clinical chemistry* **55**, 611-22.

De La Fuente, R., Hahnel, A., Basrur, P.K. and King, W.A. (1999) X inactive-specific transcript (Xist) expression and X chromosome inactivation in the preattachment bovine embryo. *Biology of reproduction* **60**, 769-75.

Dean, W., Santos, F., Stojkovic, M., Zakhartchenko, V., Walter, J., Wolf, E. and Reik, W. (2001) Conservation of methylation reprogramming in mammalian development: aberrant reprogramming in cloned embryos. *Proceedings of the National Academy of Sciences* **98**, 13734-8.

Deshmukh, R.S., Østrup, O., Østrup, E., Vejlsted, M., Niemann, H., Lucas-Hahn, A., Petersen, B., Li, J., Callesen, H. and Hyttel, P. (2011) DNA methylation in porcine preimplantation embryos developed in vivo and produced by in vitro fertilization, parthenogenetic activation and somatic cell nuclear transfer. *Epigenetics* **6**, 177-87.

Dewannieux, M., Esnault, C. and Heidmann, T. (2003) LINE-mediated retrotransposition of

marked Alu sequences. *Nature genetics* **35**, 41-8.

Dobbs, K.B., Rodriguez, M., Sudano, M.J., Ortega, M.S. and Hansen, P.J. (2013) Dynamics of DNA Methylation during Early Development of the Preimplantation Bovine Embryo. *PLOS ONE* **8**, e66230.

Ehrlich, M. (2002) DNA methylation in cancer: too much, but also too little. *Oncogene* **21**, 5400-13.

Epstein, C.J., Smith, S., Travis, B. and Tucker, G. (1978) Both X chromosomes function before visible X-chromosome inactivation in female mouse embryos. *Nature* **274**, 500-3

Fulka, H., Mrazek, M., Tepla, O. and Fulka, J. (2004) DNA methylation pattern in human zygotes and developing embryos. *Reproduction* **128**, 703-8.

Garcia-Etxebarria, K. and Jugo, B.M. (2010) Genome-wide detection and characterization of endogenous retroviruses in Bos taurus. *Journal of virology* **84**, 10852-62.

Gu, T.-P., Guo, F., Yang, H., Wu, H.-P., Xu, G.-F., Liu, W., Xie, Z.-G., Shi, L., He, X. and Jin, S.-g. (2011) The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature* **477**, 606-10.

Gutiérrez-Adán, A., Behboodi, E., Andersen, G., Medrano, J. and Murray, J. (1996) Relationship between stage of development and sex of bovine IVM-IVF embryos cultured in vitro versus in the sheep oviduct. *Theriogenology* **46**, 515-25.

Gutierrez-Adan, A., Granados, J., Pintado, B. and De La Fuente, J. (2001) Influence of glucose on the sex ratio of bovine IVM/IVF embryos cultured in vitro. *Reproduction, Fertility and Development* **13**, 361-5.

Haaf, T. (2006). Methylation dynamics in the early mammalian embryo: implications of genome reprogramming defects for development. In *DNA Methylation: Development, Genetic Disease and Cancer*, 13-22.

Harris, R.A., Wang, T., Coarfa, C., Nagarajan, R.P., Hong, C., Downey, S.L., Johnson, B.E., Fouse, S.D., Delaney, A., Zhao, Y., Olshen, A., Ballinger, T., Zhou, X., Forsberg, K.J., Gu, J., Echipare, L., O'Geen, H., Lister, R., Pelizzola, M., Xi, Y., Epstein, C.B., Bernstein, B.E., Hawkins, R.D., Ren, B., Chung, W.-Y., Gu, H., Bock, C., Gnirke, A.,

Zhang, M.Q., Haussler, D., Ecker, J.R., Li, W., Farnham, P.J., Waterland, R.A., Meissner, A., Marra, M.A., Hirst, M., Milosavljevic, A. and Costello, J.F. (2010) Comparison of sequencing-based methods to profile DNA methylation and identification of monoallelic epigenetic modifications. *Nat Biotech* **28**, 1097-105.

Hata, K. and Sakaki, Y. (1997) Identification of critical CpG sites for repression of L1 transcription by DNA methylation. *Gene* **189**, 227-34.

Hellemans, J., Mortier, G., De Paepe, A., Speleman, F. and Vandesompele, J. (2007) qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome biology* **8**, R19.

Hemberger, M. (2007) Epigenetic landscape required for placental development. *Cellular and Molecular Life Sciences* **64**, 2422-36.

Horie, K., Saito, E. S., Keng, V.W., Ikeda, R., Ishihara, H. and Takeda, J. (2007) Retrotransposons influence the mouse transcriptome: implication for the divergence of genetic traits. *Genetics* **176**, 815-27.

Hou, J., Liu, L., Lei, T., Cui, X., An, X. and Chen, Y. (2007) Genomic DNA methylation patterns in bovine preimplantation embryos derived from *in vitro* fertilization. *Science in China Series C: Life Sciences* **50**, 56-61.

Howard, G., Eiges, R., Gaudet, F., Jaenisch, R. and Eden, A. (2007) Activation and transposition of endogenous retroviral elements in hypomethylation induced tumors in mice. *Oncogene* **27**, 404-8.

Hsieh, C. L. (2005). The de novo methylation activity of Dnmt3a is distinctly different than that of Dnmt1. *BMC biochemistry* **6(1)**, 6.

Inoue, A., Matoba, S. and Zhang, Y. (2012) Transcriptional activation of transposable elements in mouse zygotes is independent of Tet3-mediated 5-methylcytosine oxidation. *Cell research* **22(12)**, 1640-6.

Kigami, D., Minami, N., Takayama, H. and Imai, H. (2003) MuERV-L is one of the earliest transcribed genes in mouse one-cell embryos. *Biology of reproduction* **68**, 651-4.

Kobayashi H., Sakurai T., Imai M., Takahashi N., Fukuda A.,Yayoi O. and Kono T. (2012) Contribution of Intragenic DNA Methylation in Mouse Gametic DNA Methylomes to

Establish Oocyte-Specific Heritable Marks. *PLOS Genet* **8(1):** e1002440.

Li, N., Ye, M., Li, Y., Yan, Z., Butcher, L.M., Sun, J., Han, X., Chen, Q., zhang, X. and Wang, J. (2010) Whole genome DNA methylation analysis based on high throughput sequencing technology. *Methods* **52**, 203-12.

Li, Y. and O'Neill, C. (2012) Persistence of Cytosine Methylation of DNA following Fertilisation in the Mouse. *PLOS ONE* **7**, e30687.

McClintock, B. (1984) The significance of responses of the genome to challenge. *Science* **226(4676)**, 792-801.

Moissiard, G., Cokus, S.J., Cary, J., Feng, S., Billi, A.C., Stroud, H., Husmann, D., Zhan, Y., Lajoie, B.R. and McCord, R.P. (2012) MORC family ATPases required for heterochromatin condensation and gene silencing. *Science* **336**, 1448-51.

Morales, J.F., Snow, E.T. and Murnane, J.P. (2003) Environmental factors affecting transcription of the human L1 retrotransposon. II. Stressors. *Mutagenesis* **18**, 151-8.

Nakanishi, M. O., Hayakawa, K., Nakabayashi, K., Hata, K., Shiota, K., and Tanaka, S. (2012). Trophoblast-specific DNA methylation occurs after the segregation of the trophectoderm and inner cell mass in the mouse periimplantation embryo. *Epigenetics* ***7***(2), 173-182.

Okano M., Bell D.W., Haber D.A. & Li E. (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**, 247-57.

Ostertag, E.M., DeBerardinis, R.J., Goodier, J.L., Zhang, Y., Yang, N., Gerton, G.L. and Kazazian, H.H. (2002) A mouse model of human L1 retrotransposition. *Nature genetics* **32**, 655-60.

Peaston, A.E., Evsikov, A.V., Graber, J.H., de Vries, W.N., Holbrook, A.E., Solter, D. and Knowles, B.B. (2004) Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos. *Developmental cell* **7**, 597-606.

Perillo, B., Ombra, M.N., Bertoni, A., Cuozzo, C., Sacchetti, S., Sasso, A., Chiariotti, L., Malorni, A., Abbondanza, C. and Avvedimento, E.V. (2008) DNA oxidation as triggered by H3K9me2 demethylation drives estrogen-induced gene expression.

Science **319**, 202-6.

Raiz, J., Damert, A., Chira, S., Held, U., Klawitter, S., Hamdorf, M., Löwer, J., Strätling, W.H., Löwer, R. and Schumann, G.G. (2012) The non-autonomous retrotransposon SVA is trans-mobilized by the human LINE-1 protein machinery. *Nucleic acids research* **40**, 1666-83.

Reik, W., Dean, W. and Walter, J. (2001) Epigenetic reprogramming in mammalian development. *Science* **293**, 1089-93.

Rizos, D., Gutierrez-Adan, A., Perez-Garnelo, S., De La Fuente, J., Boland, M. and Lonergan, P. (2003) Bovine embryo culture in the presence or absence of serum: implications for blastocyst development, cryotolerance, and messenger RNA expression. *Biology of reproduction* **68**, 236-43.

Santos, F. and Dean, W. (2006) Using immunofluorescence to observe methylation changes in mammalian preimplantation embryos. In: *Nuclear Reprogramming.* 129-38. Springer.

Santos, F., Zakhartchenko, V., Stojkovic, M., Peters, A., Jenuwein, T., Wolf, E., Reik, W. and Dean, W. (2003) Epigenetic marking correlates with developmental potential in cloned bovine preimplantation embryos. *Current Biology* **13**, 1116-21.

Schulz, W.A. (2006) L1 retrotransposons in human cancers. *BioMed Research International* **2006**.

Shi, W., Dirim, F., Wolf, E., Zakhartchenko, V. and Haaf, T. (2004) Methylation reprogramming and chromosomal aneuploidy in in vivo fertilized and cloned rabbit preimplantation embryos. *Biology of reproduction* **71**, 340-7.

Shi, W. and Haaf, T. (2002) Aberrant methylation patterns at the two-cell stage as an indicator of early developmental failure. *Molecular reproduction and development* **63**, 329-34.

Shi, X., Seluanov, A. and Gorbunova, V. (2007) Cell divisions are required for L1 retrotransposition. *Molecular and cellular biology* **27**, 1264-70.

Smith, Z. D., Chan, M. M., Mikkelsen, T. S., Gu, H., Gnirke, A., Regev, A. and Meissner, A. (2012). A unique regulatory phase of DNA methylation in the early mammalian embryo. *Nature **484*****(7394)**, 339-344.

Spadafora, C. (2004) Endogenous reverse transcriptase: a mediator of cell proliferation and

differentiation. *Cytogenetic and genome research* **105**, 346-50.

Sperber, G.O., Airola, T., Jern, P. and Blomberg, J. (2007) Automated recognition of retroviral sequences in genomic data - RetroTector©. *Nucleic acids research* **35**, 4964-76.

Tam, O.H., Aravin, A.A., Stein, P., Girard, A., Murchison, E.P., Cheloufi, S., Hodges, E., Anger, M., Sachidanandam, R. and Schultz, R.M. (2008) Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. *Nature* **453**, 534-8.

Van Soom, A., Yuan, Y.Q., Peelman, L.J., de Matos, D.G., Dewulf, J., Laevens, H. and de Kruif, A. (2002) Prevalence of apoptosis and inner cell allocation in bovine embryos cultured under different oxygen tensions with or without cysteine addition. *Theriogenology* **57**, 1453-65.

Vanden Berghe, W. (2012) Epigenetic impact of dietary polyphenols in cancer chemoprevention: Lifelong remodeling of our epigenomes. *Pharmacological Research* **65**, 565-76.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A. and Speleman, F. (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome biology* **3**, research0034.

Watanabe, T., Totoki, Y., Toyoda, A., Kaneda, M., Kuramochi-Miyagawa, S., Obata, Y., Chiba, H., Kohara, Y., Kono, T. and Nakano, T. (2008) Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature* **453**, 539-43.

Wydooghe, E., Heras, S., Dewulf, J., Piepers, S., Van Den Abbeel, E., De Sutter, P., Vandaele, L. and Van Soom, A. (2013) Replacing serum in culture medium with albumin and insulin, transferrin and selenium is the key to successful bovine embryo development in individual culture. *Reproduction, Fertility and Development* **26**, 717-24.

Yang, N. and Kazazian, H.H. (2006) L1 retrotransposition is suppressed by endogenously encoded small interfering RNAs in human cultured cells. *Nat Struct Mol Biol* **13**, 763-71.

Yoder, J.A., Walsh, C.P. and Bestor, T.H. (1997) Cytosine methylation and the ecology of intragenomic parasites. *Trends in genetics* **13**, 335-40.

Yuan, Y., Van Soom, A., Coopman, F., Mintiens, K., Boerjan, M., Van Zeveren, A., de Kruif, A. and Peelman, L. (2003) Influence of oxygen tension on apoptosis and hatching in bovine embryos cultured in vitro. *Theriogenology* **59**, 1585-96.

Chapter 5

Repeats as Global DNA Methylation Marker in Bovine Preimplantation Embryos

Wenwen Li[1], Ann Van Soom[2] and Luc Peelman[1]

[1] *Department of Nutrition, Genetics and Ethology, Faculty of Veterinary Medicine, Ghent University, Heidestraat 19, 9820 Merelbeke, Belgium.*

[2] *Department of Obstetrics, Reproduction and Herd Health, Faculty of Veterinary Medicine, Ghent University, Salisburylaan 133, 9820 Merelbeke, Belgium.*

Adapted from Czech Journal of Animal Science. In press.

5.1 Abstract

DNA methylation undergoes dynamic changes and is a crucial part of the epigenetic regulation during mammalian early development. To determine the DNA methylation levels in bovine embryos, we applied a bisulfite sequencing based method aimed at repetitive sequences including three retrotransposons (*L1_BT*, *BovB* and *ERV1-1-I_BT*) and *Satellite I*. A more accurate estimate of the global DNA methylation level compared to previous methods using only one repeat sequence, like *Alu*, could be made by calculation of the Weighted Arithmetic Mean of multiple repetitive sequences, considering the copy number of each repetitive sequence. *Satellite I* and *L1_BT* showed significant methylation reduction at the blastocyst stage, while *BovB* and *ERV1-1-I_BT* showed no difference. The mean methylation level of the repetitive sequences during preimplantation development was lowest at the blastocyst stage. No methylation difference was found between embryos cultured in 5% and 20% O_2. Because mutations of CpGs negatively influence the calculation accuracy, we checked the mutation rate of the sequenced CpG sites. *Satellite I* and *L1_BT* showed a relatively low mutation rate (1.92% and 3.72% respectively) while that of *ERV1-1-I_BT* and *BovB* was higher (11.95% and 24% respectively). Therefore, we suggest using a combination of repeats with low mutation rate, taking into account the proportion of each sequence, as a relatively quick marker for the global DNA methylation status of preimplantation stages and eventually also for other cell types.

Keywords

Preimplantation embryo; DNA methylation; retrotransposon; bisulfite sequencing

5.2 Introduction

Preimplantation development from separate parental germ cells to a fertilized zygote, and further to a blastocyst is a complex process including cell proliferation and differentiation. Epigenetic mechanisms play a crucial role here, allowing for activation of some genes and silencing of others. Among all the epigenetic mechanisms, DNA methylation is best known. It is associated with key processes in embryo development, including genomic imprinting, X-chromosome inactivation and repression of repetitive elements to maintain the genome stability (Messerschmidt *et al.* 2014).

Bovine embryos can be successfully produced *in vitro* by various procedures. However, compared to *in vivo* embryos, *in vitro* produced and cloned embryos still remain of inferior quality and are associated with more losses during pregnancy (Farin *et al.* 2006). Furthermore, calves derived from *in vitro* produced embryos can be affected by the large offspring syndrome (LOS), and an altered methylation pattern that disturbs the gene expression during preimplantation is suggested as the primary cause for LOS (Young *et al*. 2001). The DNA methylation pattern of sperm is suggested to be predictive of embryo quality during IVF (Aston *et al.* 2015). It is proposed that incomplete nuclear reprogramming in cloned embryos, which contributes to a low developmental success, is caused by failure of DNA demethylation (Bourc'His *et al.* 2001).

A common method used to study DNA methylation of preimplantation embryos is 5-methylcytosine (5-mC) immunofluorescence staining (Dean *et al.* 2001; Hou *et al.* 2007; Dobbs *et al.* 2013). However, conflicting results were found among these studies, for example as to whether methylation is more intensive in ICM or TE cells in IVF bovine blastocysts, and the staining result can be affected by the protocol used (Li *et al*. 2012). Therefore, another simple and direct method is needed for determining the global DNA methylation in preimplantation embryos.

Bisulfite sequencing is another approach to evaluate DNA methylation at CpG dinucleotides, mostly applied for specific *loci*. The bisulfite (HSO_3^-) treatment converts cytosine residues to

uracil, but leaves 5-methylcytosine residues unaffected (Frommer *et al.* 1992). By comparing the sequences the methylation pattern can be determined. More recently, new methods based on bisulfite conversion and whole genome sequencing were developed (Guo *et al.* 2014; Salilew-Wondim *et al.* 2015). These platforms can determine the methylation status of thousands of genes at CpG site-level, but are still very expensive and require intensive bioinformatics (data analysis) expertise, which makes it not practical when a large number of treatments or samples are studied. Therefore, a genomic marker that is cheap, relatively easy to type and representative for the global DNA methylation status is wanted. The aim of this study was to develop a routine technique that can predict embryo quality based on the DNA methylation status.

Satellite DNA sequences have been used for this purpose, due to the presence of high order repeats that are from several hundred to several thousand bp in length (Sawai *et al.* 2011; Couldrey & Wells 2013). Besides satellite DNA, most of the DNA methylation in mammalian genomes is found in retrotransposons (Schulz *et al.* 2006). Retrotransposons, which account for almost half of the genome (44% of bovine genome (Adelson *et al.* 2009) and 45% of human genome (Lander *et al.* 2001)), were previously suggested as biomarker for global methylation status (Klose & Bird 2006). Recent research on epigenetics of mammalian preimplantation embryos found that transposable elements show a similar dynamic trend towards global methylation (Guo *et al.* 2014). Among all types of retrotransposons, L1 is mostly studied and has been used for indicating methylation changes in cancer cells (Hsiung *et al.* 2007), and a recent study of global methylation found a good correlation between L1 methylation and total amount of 5-methylcytosine measured by liquid chromatography mass spectrometry in murine cells and tissues (Newman *et al.* 2012).

In this paper, we selected three bovine retrotransposons with complete internal promoter sequence (*L1_BT*, *BovB* and *ERV1-1-I_BT*), together with *Satellite I* DNA to test as marker for global DNA methylation estimation in bovine preimplantation embryo development. Since *in vitro* culture conditions alter DNA methylation (Salilew-Wondim *et al.* 2015) and different methylation patterns under oxygen stress (20% O_2) compared to normal (5% O_2) were found

in a previous study by 5-methylcytosine immunofluorescence staining (Li *et al.* 2014), we included this in the study.

5.3 Materials and Methods

Materials

Unless stated otherwise, all chemicals, reagents and media were obtained from Sigma (USA) and Life Technologies (Belgium).

Primer design

L1_BT, *BovB* and *ERV1-1-I_BT*, three autonomous retrotransposons were used for the methylation analysis. First, retrotransposon sequences were retrieved from Repbase (Jurka *et al.* 2005). The amplicons were targeted to the internal promoter regions upstream of the open reading frame (ORF) from each retrotransposon sequence. In order to amplify as many as possible repeats to represent the whole genome, we used NCBI blast bovine genome (Btau_4.2) and MultiAlin (Hemberger 2007) to find the consensus sequence of each promoter, and then designed BSP primers based on the consensus sequences with BiSearch software (Aranyi *et al.* 2006). Primers are listed with predicted amplicon number in **Table 5.1**.

IVF embryo production and sample collection

In vitro bovine embryos were produced in serum-free media as previously reported (Li *et al*. 2014). The control group and high oxygen tension group were incubated at 38.5°C under 5% CO_2, 5% O_2 and 90% N_2, and 5% CO_2 in air (20% O_2) separately. Embryos were collected at specific time points (2-4 cell at 36-40 h post-insemination (h.p.i.), 8-cell at 64 h.p.i., expanded blastocyst at 8.d.p.i).

Embryos were washed three times with phosphate-buffered saline (PBS), and then frozen in 3 pools of embryos for each stage (75 for 2-4 cell, 25 for 8-cell and 5 for blastocyst) at –80°C until use.

Table 5.1 Characteristics of the bisulfite primers used in the methylation analysis

Repetitive sequence name	Amplicon numbers (by BiSearch)	Sequence 5' – 3'	Amplicon length cDNA	CpG numbers	Ta (°C)
Satellite1 (Wroclawska *et al.* 2000)	859	TTGGTTTTTAGGTTATGTAGGAG	379	25	55
		AATACACCAAACCCAATAAAAT			
L1_BT	303	TAATTAAAATTTTTTGGGGGTTTG	254	14	55
		TAACCCTAAACTACATACACCTCCC			
BovB	467	AGGAGGGTTTAGAGGAGTTATTTTA	210	6	58
		TCTATAAATATAATTTCAATATATTTACCC			
ERV1-1-I_BT	44	TTTTTGGTATATTTTTATTTTATTT	214	11	55
		AATTATTAACTCCCATCTATAAAAA			

DNA extraction and bisulfite conversion

Embryo DNA extraction and bisulfite conversion were performed with EZ Methylation Direct kit (Zymo research, USA), according to the manufacturer's instructions with minor changes. Instead of incubating 20 min, we applied 3 h of incubation in M-Digestion buffer and Proteinase K for embryo lysis and genomic DNA release, and the bisulfite converted DNA was eluted in 20 µl water. Blood samples were taken from healthy cows, and genomic DNA was released with Proteinase K from 200 µl blood. Twenty µl of unpurified DNA solution was taken for bisulfite conversion directly with EZ Methylation Gold kit (Zymo research, USA), according to the manufacturer's instructions, and eluted in 20 µl water. All converted DNA samples were checked by PCR with GAPDH primers (F: TTCAACGGCACAGTCAAGG; R: ACATACTCAGCACCAGCATCAC) to check for conversion completion. These primers amplify genomic DNA, but not converted DNA, so samples without a PCR amplicon, which represent samples without unconverted genomic DNA, were used for further methylation study.

Amplifying, Cloning and sequencing

After bisulfite treatment, retrotransposons and satellite DNA were amplified by PCR. To amplify the part of the repeat region, we used the following PCR mixture: 0.5 U FastStart Taq DNA Polymerase and 1 µl 10x reaction buffer (Roche, Belgium), 200 µM dNTPs (Bioline Reagents Ltd., UK), 500 nM of each primer (LDT, Belgium) and 2 µl of bisulfite converted DNA. The PCR program consisted of initial denaturation at 95°C for 5 min, followed by 30 cycles of 15 sec at 95°C, 15 sec at the specific annealing temperature (**Table 5.1**) and 30 sec at 72°C, and a final 10 min elongation at 72°C. The PCR products were verified by electrophoresis (2 µl out of 10 µl).

For each retrotransposon or satellite DNA, PCR products for three replicate samples were purified from gel (GENECLEAN II kit), checked by electrophoresis, pooled and cloned in plasmid pCR2.1, and transformed into DH5α competent cells (Invitrogen™, Belgium). Selected colonies were verified by PCR with universal and reversal primers (U:

CGACGTTGTAAAACGACGGCCAG; R: CACAGGAAACAGCTATGACCATGATTACG), and those colonies with correct amplicon size with a bright band on gel were sequenced.

Methylation analysis

Sequence analyses and statistical comparisons (nonparametric two-tailed Mann- Whitney test) were performed using the QUMA web service (quma.cdb.riken.jp). "Strict CpG site check of bisulfite sequence" was selected for repetitive sequence analysis. In this case, the divergence of retrotransposons is considered and a more corrected result can be obtained (Kumaki *et al.* 2008). Compared to regular genes, retrotransposons have a higher rate of mutation, so sequences that passed 90% bisulfite conversion were included in the analysis (normally 95% for genes). A *P*-value threshold of ≤ 0.05 was chosen to identify significant differences in the variance between groups. The target genomic sequences of each repeat are listed in **Table 5.2**.

Since different repetitive sequences have their own amplified number in the BS-PCR, we propose a method by using more than one retrotransposon as a global methylation marker considering the proportion of each repetitive sequence by calculating the Weighted Arithmetic Mean of the multiple repetitive sequences used, according to the following formula.

$$\bar{x} = \frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i}$$

x- methylation ratio of each repetitive sequence (calculated by QUMA)

n- number of repetitive sequences

w-amplicon number of each repetitive sequence from BiSearch (Listed in **Table 5.1**)

Table 5.2 Unconverted target genomic sequences of each repeat.

Satellite1	TTGGTCTCTAGGCCATGCAGGAGACGAAGGCCtTCATCTCTCGA
	TGACGGGGGAATCTCGGGGTTGTTCTCGAGCGGCGGCCCCAGT
	GTGCGTTTTCTCACGAGGTACGACGGCGAGGTCAGTGAGCCTCt
	CGTGGGGCGCCAGGGAAGTCGGGTCTCCTTGCGAGTGGCGAGG
	GGGAGCGCGTCACTGCTCCCGAGCCATGGTAGGGGAATCTGGC
	CTCGAGACATGTTGAAGAAGGTCTCTCGAGGCCTTTCCCGGGTT
	GAGGCAGGAAACCCTGGGTTCCCTCGACtTGTGCAGGTGACCTC
	AGGGGAATTCTCATGGTGGCTCtGAGAAGCCAGGGAAACTGGA
	GGTGATAGGGGCCTCTCGGGACTCTACTGGGTTTGGTGCATT
L1_BT	TAATTAAAATTTTTTGGGGGTTTGGACGGTTAACATCTGCCTGA
	GAAGGTGCGCCGGTTTTACACCCAGATAACCGAGTGGCGGGGA
	GGCGATAAGTCGCAGCATTGGCGCTCGCCAAACACCTCATCAC
	CTGAGCTGCTCGGACCTGGGAAGAGCACAAAACGCAGGCCCA
	ACTGAGTTTGCGCCTTTGAGGACTACCCGAGTGCCTGAATTTGA
	GCGGCTTGGACCTGGGAGGTGCATGTAGTTTAGGGTTA
BovB	AGGAGGGCCTAGAGGAGCTATCCCACGTTGAAGGTCAGGAAG
	GGCGGCGGTGAGGAGATACCCCTCGTCCAAGGTAAGGAGCAAT
	GGTTGCGCTTTGCTGGAGCAGCTGTGAAGAGATATCCCATGCC
	CAAGGTAAGAGAAACCCAAGTAAGACGGTAGGTGTTGCAAGA
	GGGCATCAGAGGGCAAACACACTGAAATCATACTCACAGA
ERV1-1-I_BT	CCCCTGGTACATCCCCACCCCATTTCGGTGGTAGAACCGGGAG
	GGACGAGGACGGCGCCTGCGTCAGTAAGGGACAGACTAAGTC
	CGACCAGGAAGGAAAAGCTTTTGGTGTAATGTCTGTCTACACC
	CCCATCTAGAGCAGGGAGGGACGCCTCCGGTAGAAAAATGGC
	GTTGGTCGCTTTTTTCTCTCTTACAGATGGGAGCTAACAATT

5.4 Results

Methylation changes during embryo development

The DNA methylation levels of the four repetitive sequences in embryos and blood samples (used as reference) are indicated in **Figure 5.1** (methylation status of each CpG site) and **Figure 5.2** (methylation level of each repetitive sequence). As expected, the promoter regions in all four repetitive sequences were significantly hypermethylated in blood samples (with *Satellite I* 87.1%, *L1_BT* 89.5%, *BovB* 63.4% and *ERV1-1-I_BT* 88.6% of CpGs methylated) compared to the preimplantation embryos. The methylation level of different stages of embryo development in standard culture (5% O_2) varies. In *Satellite I* DNA and *L1_BT*, there is a sharp decreased methylation in blastocyst as compared to the 2-4 and 8-cell stages (*P*-value of 0.0072 and 0.0175 respectively); in *ERV1-1-I_BT*, an increased methylation was found in 8-cell stage (*P*-value of 0.0362); while no significant methylation difference was found in *BovB*.

(a) Satellite I methylation status

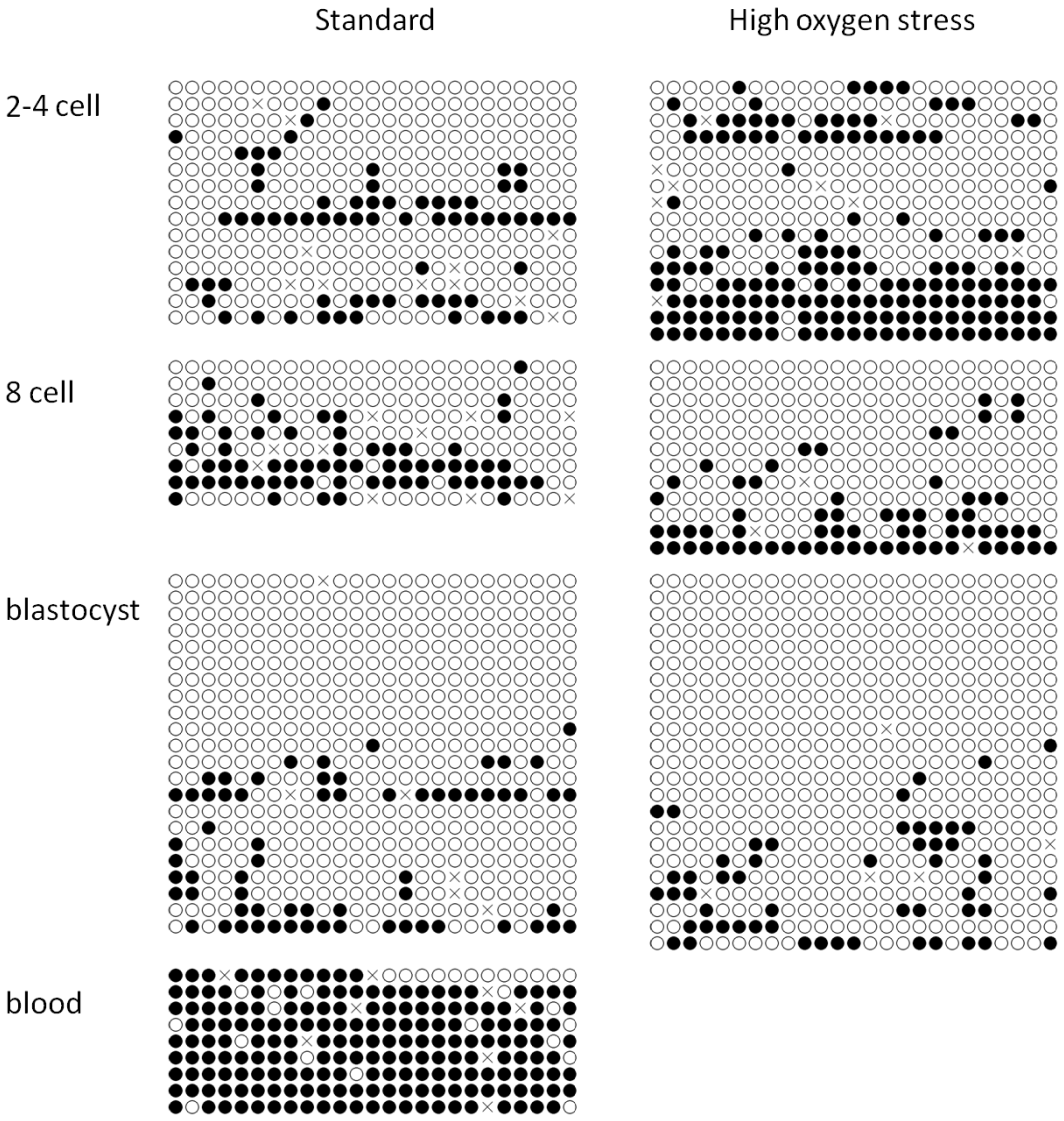
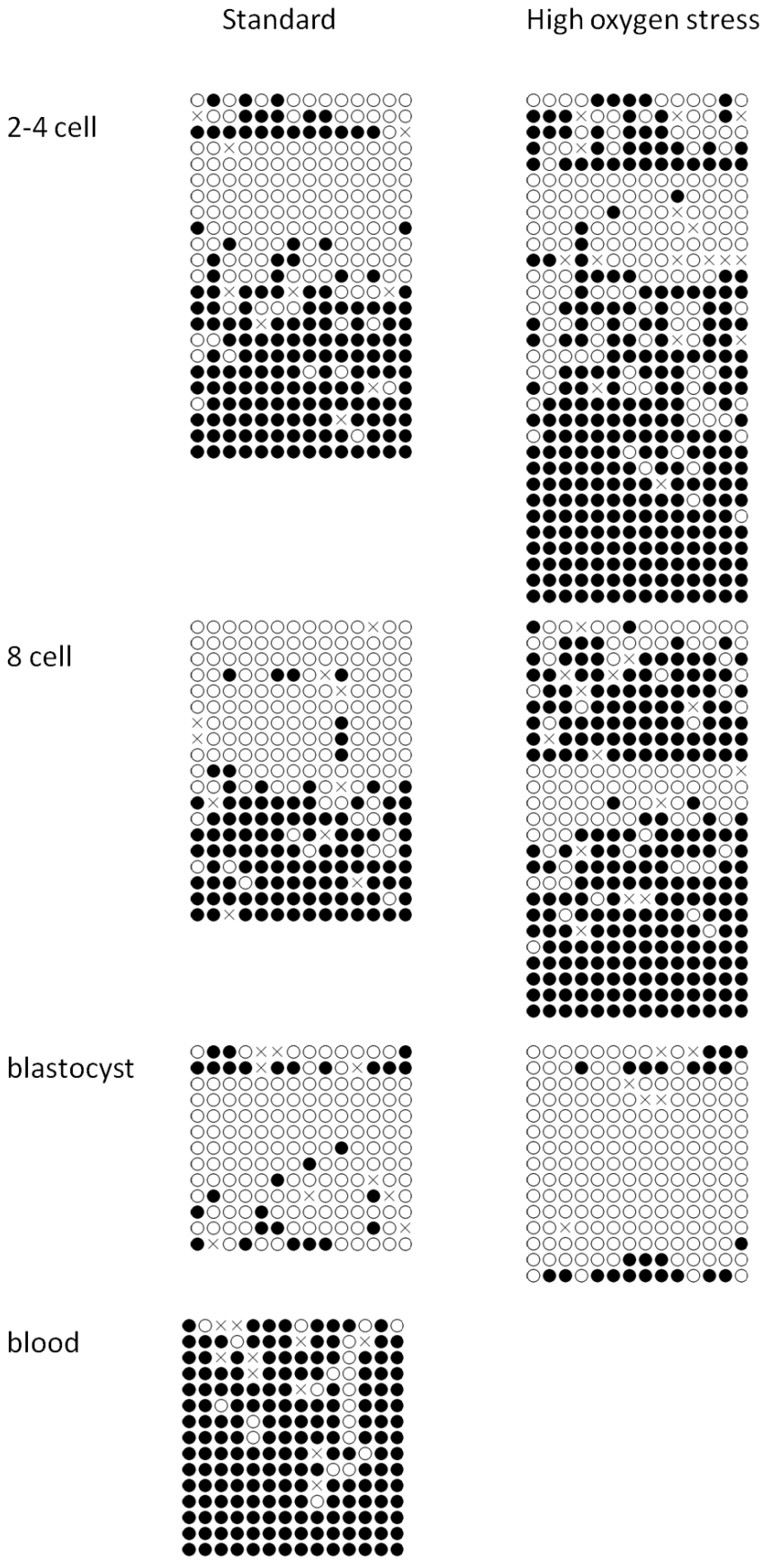
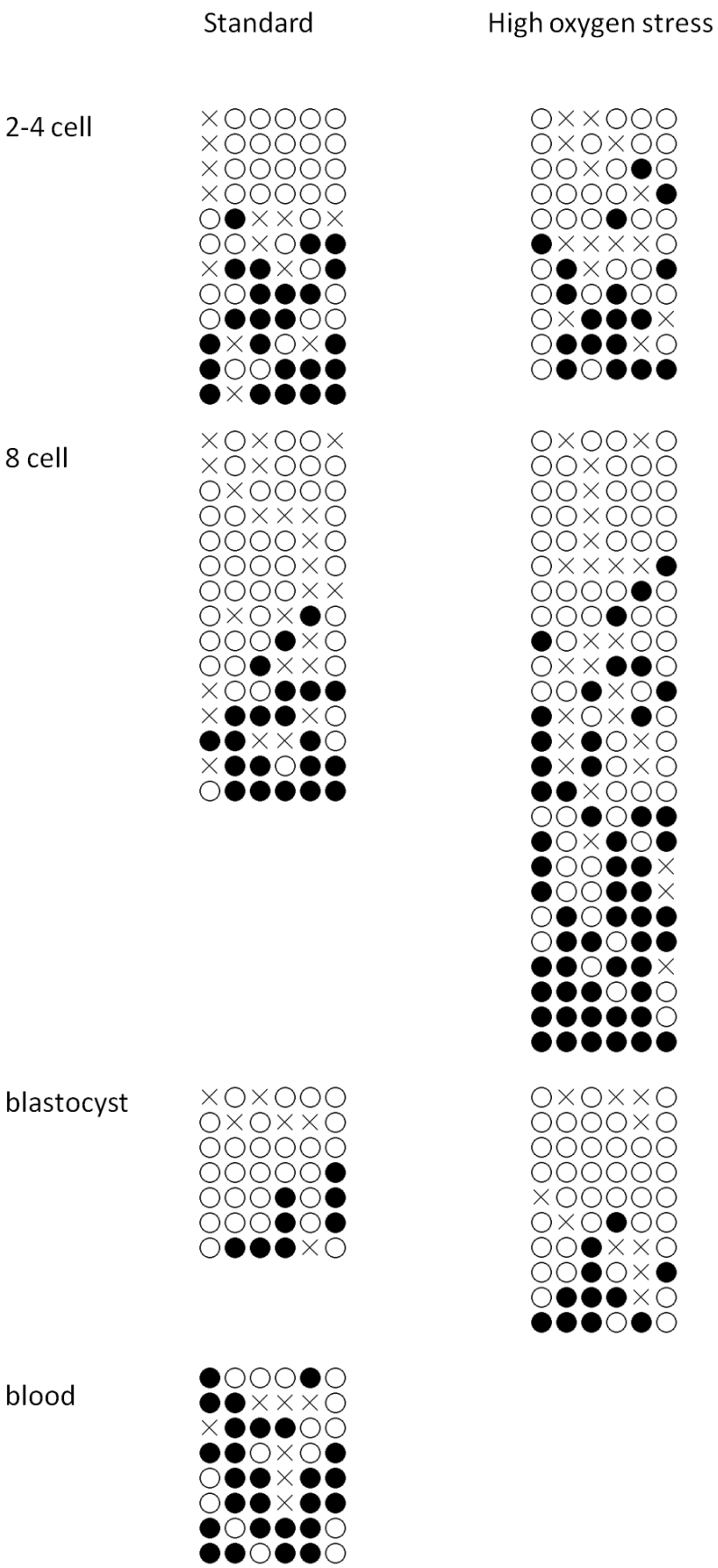


Figure 5.1 Methylation status of each CpG site of the repeats. Filled, methylated; open, unmethylated; cross, mutation. (a) *Satellite I* methylation status (b) *L1_BT* methylation status (c) *BovB* methylation status (d) *ERV1-1-I_BT* methylation status.

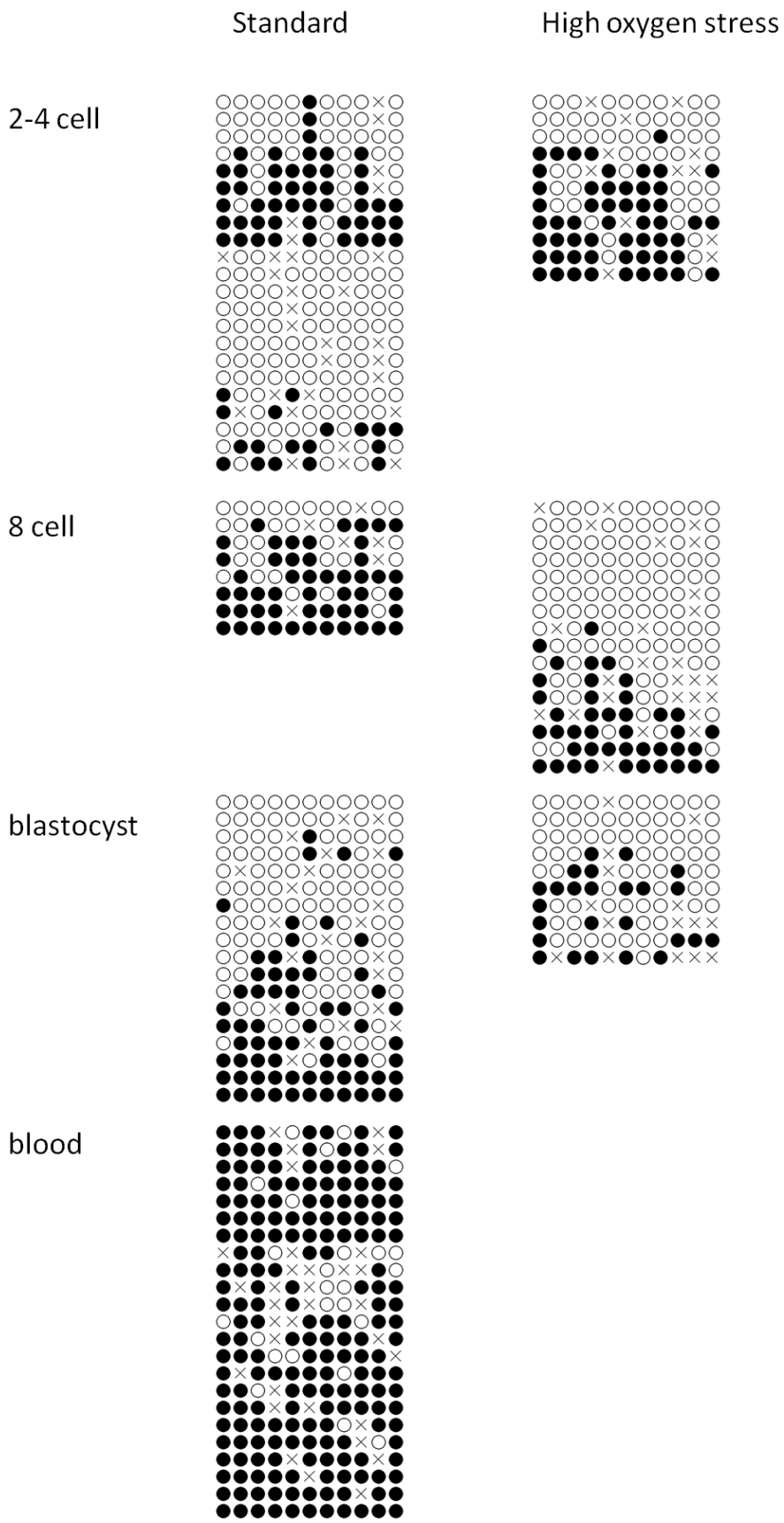
(b) *L1_BT* methylation status



(c) *BovB* methylation status



(d) *ERV1-1-I_BT* methylation status



Methylation changes under high oxygen tension

Oxygen stress (20% O_2) does not seem to have an influence on the methylation level of the four repeat families studied with the exception of *L1_BT* which was more methylated in the 8-cell stage under 20% O_2 (70.8%) compared to 41.8% under 5% O_2, with *P*-value of 0.0203 (**Figure 5.2b)**.

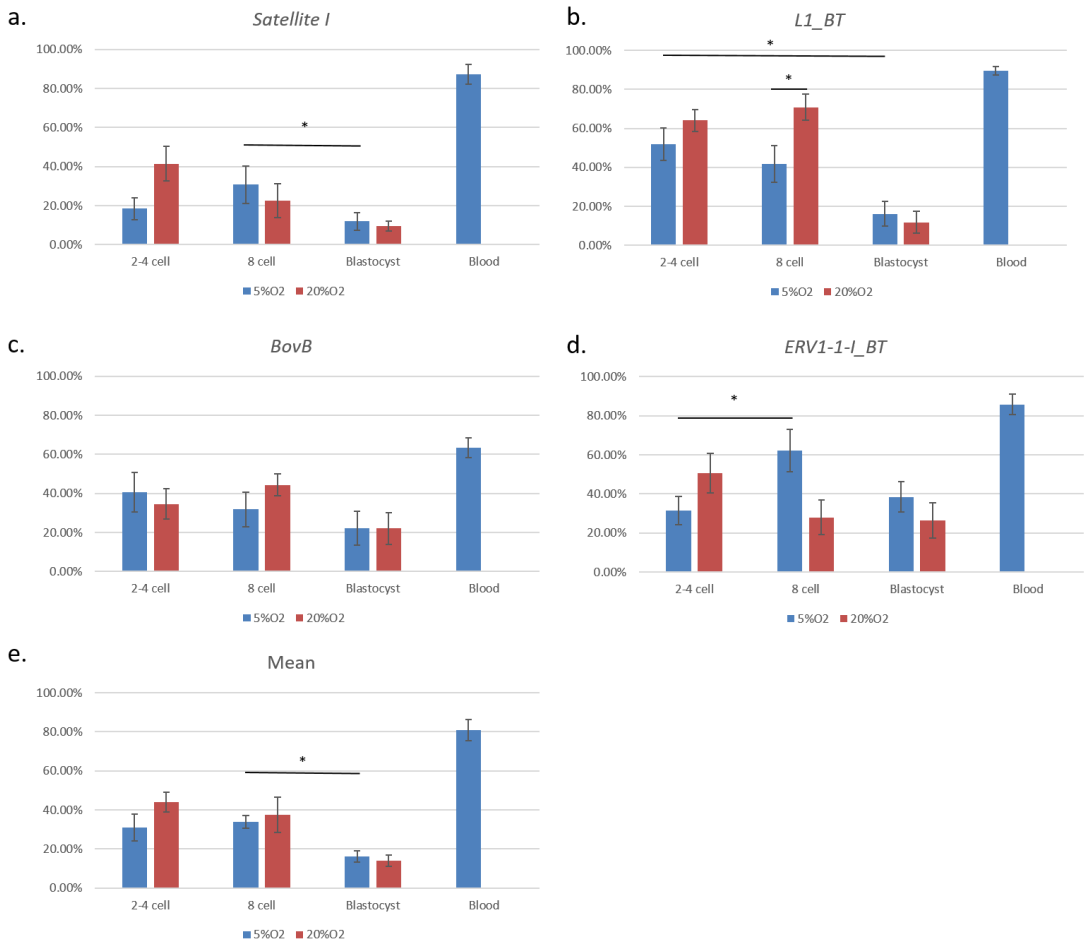


Figure 5.2 Methylation comparison of different embryo stages and treatment. a.–d. methylation comparison of each repetitive sequence. e. Comparison of Weighted Arithmetic Mean methylation. Data are shown as Mean±SEM, and an asterisk indicates significant difference between two groups. (P<0.05).

Retrotransposon mutation frequency

Since repetitive sequences are variable, CpG sites may be present in some copies of the repeat and not in others. To estimate the CpG mutation frequency in the four repeats studied we took the consensus sequence of the repetitive elements as reference sequence and applied the "strict CpG site check" on QUMA. In that case, TG is calculated as unmethylated; CG as methylated and others as site mutation. The mutation rate thus estimated was 1.92% in *Satellite I*, 3.72% in *L1_BT*, 24.03% in *BovB* and 11.95% in *ERV1-1-I_BT*.

Global methylation estimation by combining the four repetitive sequences

The global methylation level was calculated by using the Weighted Arithmetic Mean of the four repetitive sequences studied as shown in **Figure 5.2e**. No significant methylation difference was observed between embryos produced under 5% O_2 and 20% O_2 culture, but methylation at the blastocyst stage was much lower than that in earlier stages.

5.5 Discussion

The alteration of DNA methylation patterns in preimplantation embryo development has been studied widely in different species (Bourc'His *et al.* 2001; Fulka *et al.* 2004). Global DNA demethylation and *de novo* methylation in this period has been shown to guide and restrict differentiation and prevent cell regression into an undifferentiated state, and on the other hand are also crucial to establish pluripotency (Messerschmidt *et al.* 2014). An accurate quantitative assay representative for global DNA methylation may not exist, but repetitive genomic sequences such as satellite DNA and L1 which accounts for almost 20% of the mammalian genome have been widely used as a means for estimating global methylation status in cancer research (Yang *et al.* 2004).

In the present study, we evaluated bisulfite sequencing of four repetitive sequences as global DNA methylation marker during bovine early embryo development and used it to check if oxygen stress (20% O_2) has an influence on the global methylation level of preimplantation embryos. *Satellite I* and *L1_BT* showed the lowest methylation level at the blastocyst stage.

This is consistent with *DNMT1* (DNA cytosine-5-methyltransferase1) RNA expression and the results of a recent report (Salilew-Wondim et al. 2015). We previously found *DNMT1* expression dropped markedly in blastocysts compared to earlier stages (4-cell embryo) (Li *et al.* 2014), implying hypomethylation in the blastocyst. However, the *DNMT1* expression was not corroborated by the DNA methylation level as measured by 5-mC immunofluorescence staining which indicated the highest level in the blastocyst stage (Li *et al.* 2014). In other species, conflicting results were found as well. In human embryos, 5-cytosine immunofluorescence staining showed an increased methylation in blastocysts, compared to morula (Fulka *et al.* 2004); while Guo *et al.* (2014) found, by using a bisulfite sequencing based method - reduced representation bisulphite sequencing (RRBS), the lowest DNA methylation level in blastocysts. The same conflicting results were found also in mouse embryos, with high methylation levels compared to previous stages found in blastocysts by staining (Dean *et al.* 2001); whereas RRBS showed the lowest methylation level in mouse morula/blastocyst stages (Smith *et al.* 2012). These conflicting results were not only found between these two analytical methods, but also between studies using the same staining method. There is disagreement in whether methylation is higher in ICM or TE of blastocysts (Dean *et al.* 2001; Hou *et al.* 2007; Dobbs *et al.* 2013; Li *et al.* 2014), and differences in staining protocols maybe responsible for the conflicting results.

There are several possible explanations for the different results obtained with 5-mC staining. First of all, changes in the staining protocol may lead to different results. Li and O'Neill (2012) used a step of trypsin digestion in methylation staining and a different result of paternal demethylation during mouse zygote maturation was found. Besides the staining method itself, the lack of standard quantification for staining may contribute to different results as well. The ratio of 5-mC to a DNA counter stain is often used for methylation level estimation. The most common dyes for DNA counter staining are Hoechst, DAPI, Propidium Iodide (PI) and Ethidium homodimer-2 (EthD-2) which only bind to double-stranded DNA, or weakly bind to single-stranded DNA and RNA as well. Therefore, this widely used ratio presents actually single-stranded methylated DNA/ total double-stranded DNA, which can be

easily influenced by the denaturation of DNA. Furthermore, the fluorescence intensity observed may be affected by cell size or shape and even the placement orientations under the microscope. The compact composition of morula and ICM in blastocysts also make the staining and methylation quantification more difficult.

Bisulfite sequencing of repetitive sequences, on the other hand, has its limitations as well and we tried to improve on that. First of all, the high mutation rate of repetitive sequences makes the number of potential targets for DNA methylation variable between the individual members of a repeat family. This is the reason we did not include SINEs in the study, since they are super mutated. For example, almost two-thirds of the CpG methylation sites in *Alu* elements (the most widespread SINE family in the human genome) are mutated (Yang *et al.* 2004). In the present study, we calculated the mutation rate of CpG sites of the four repetitive sequences studied. *Satellite I* showed the lowest mutation rate (1.92%), while *BovB* has the highest mutation rate (24.03%) making it less suitable as global methylation marker when used individually.

Another factor that affects the methylation result is primer design, which can influence the number of amplified repetitive sequence directly. For example, the *Alu* bisulfite primers used in a recent study could amplify only 16 different copies, with length differing from 46 bp up to 1017 bp (Lisanti *et al.* 2013), making it dubious as marker for global methylation status. In the present study, we tried to amplify as many as possible copies of the repeat families to represent the whole genome. All primers were designed based on the consensus sequence of the promoters and the amplified number (listed in **Table 5.1**) could be checked by BiSearch. To improve the method even more we calculated the Weighted Arithmetic Mean of the four repetitive sequences studied and use that for calculation of the global methylation status.

5.6 Conclusion

In conclusion, this study is the first report using *L1_BT*, *BovB*, *ERV1-1-I_BT* and *Satellite I* DNA as global methylation marker in bovine early embryo development. The mean methylation of the repeats showed the same tendency as observed in a recent DNA

methylation study (Salilew-Wondim *et al.* 2015), with blastocysts showing the lowest methylation level. We also compared DNA methylation in embryos cultured under normal (5% O_2) with high oxygen tension (stress group, 20% O_2), and found only significant hypermethylation in *L1_BT* at 8-cell under stress. Although the accuracy and sensitivity of the method used needs to be tested in more independent studies, we propose using the Weighted Arithmetic Mean of several repetitive sequences as global methylation marker. Especially the repeats with (relatively) low mutation rate and more copies targeted during PCR, like *Satellite I* DNA and *L1_BT*, are promising.

5.7 Acknowledgements

The authors like to thank Petra Van Damme and Isabel Lemahieu for their excellent technical assistance with *in vitro* culture of embryos. We also thank Linda Impe for the great help with colony PCR and sequencing.

5.8 References

Adelson D.L., Raison J.M. & Edgar R.C. (2009): Characterization and distribution of retrotransposons and simple sequence repeats in the bovine genome. *Proceedings of the National Academy of Sciences* **106**, 12855-60.

Aranyi T., Varadi A., Simon I. & Tusnady G.E. (2006): The BiSearch web server. *BMC bioinformatics* **7**, 431-7.

Aston K.I., Uren P.J., Jenkins T.G., Horsager A., Cairns B.R., Smith A.D. & Carrell D.T. (2015): Aberrant sperm DNA methylation predicts male fertility status and embryo quality. *Fertility and Sterility* **104(6)**: 1388-97.

Bourc'His D., Le Bourhis D., Patin D., Niveleau A., Comizzoli P., Renard J.-P. & Viegas-Pequignot E. (2001): Delayed and incomplete reprogramming of chromosome methylation patterns in bovine cloned embryos. *Current Biology* **11**, 1542-6.

Couldrey C. & Wells D.N. (2013): DNA methylation at a bovine alpha satellite I repeat CpG site during development following fertilization and somatic cell nuclear transfer. *PLOS ONE* **8**, e55153.

Dean W., Santos F., Stojkovic M., Zakhartchenko V., Walter J., Wolf E. & Reik W. (2001): Conservation of methylation reprogramming in mammalian development: aberrant reprogramming in cloned embryos. *Proceedings of the National Academy of Sciences* **98**, 13734-8.

Dobbs K.B., Rodriguez M., Sudano M.J., Ortega M.S. & Hansen P.J. (2013): Dynamics of DNA Methylation during Early Development of the Preimplantation Bovine Embryo. *PLOS ONE* **8**, e66230.

Farin P.W., Piedrahita J.A. & Farin C.E. (2006): Errors in development of fetuses and placentas from in vitro-produced bovine embryos. *Theriogenology* **65**, 178-91.

Frommer M., McDonald L.E., Millar D.S., Collis C.M., Watt F., Grigg G.W., Molloy P.L. & Paul C.L. (1992): A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proceedings of the National Academy of Sciences* **89**, 1827-31.

Fulka H., Mrazek M., Tepla O. & Fulka J. (2004): DNA methylation pattern in human zygotes and developing embryos. *Reproduction* **128**, 703-8.

Guo H., Zhu P., Yan L., Li R., Hu B., Lian Y., Yan J., Ren X., Lin S. & Li J. (2014): The DNA methylation landscape of human early embryos. *Nature* **511**, 606-10.

Hemberger M. (2007): Epigenetic landscape required for placental development. *Cellular and Molecular Life Sciences* **64**, 2422-36.

Hou J., Liu L., Lei T., Cui X., An X. & Chen Y. (2007): Genomic DNA methylation patterns in bovine preimplantation embryos derived from in vitro fertilization. *Science in China Series C: Life Sciences* **50**, 56-61.

Hsiung D.T., Marsit C.J., Houseman E.A., Eddy K., Furniss C.S., McClean M.D. & Kelsey K.T. (2007): Global DNA methylation level in whole blood as a biomarker in head and neck squamous cell carcinoma. *Cancer Epidemiology Biomarkers & Prevention* **16**, 108-14.

Jurka J. (1997): Sequence patterns indicate an enzymatic involvement in integration of mammalian retroposons. *Proceedings of the National Academy of Sciences* **94**, 1872-7.

Klose R.J. & Bird A.P. (2006): Genomic DNA methylation: the mark and its mediators. *Trends in biochemical sciences* **31**, 89-97.

Kumaki Y., Oda M. & Okano M. (2008): QUMA: quantification tool for methylation analysis. *Nucleic acids research* **36**, 170-5.

Lander E.S., Linton L.M., Birren B., Nusbaum C., Zody M.C., Baldwin J., Devon K., Dewar K., Doyle M. & FitzHugh W. (2001): Initial sequencing and analysis of the human genome. *Nature* **409**, 860-921.

Li Y. & O'Neill C. (2012) Persistence of Cytosine Methylation of DNA following Fertilisation in the Mouse. *PLOS ONE* **7**, e30687.

Li W., Goossens K., Van Poucke M., Forier K., Braeckmans K., Van Soom A. & Peelman L.J. (2014): High oxygen tension increases global methylation in bovine 4-cell embryos and blastocysts but does not affect general retrotransposon expression. *Reproduction, Fertility and Development* **28(7)**, 948-59.

Lisanti S., Omar W.A., Tomaszewski B., De Prins S., Jacobs G., Koppen G., Mathers J.C. & Langie S.A. (2013): Comparison of methods for quantification of global DNA methylation in human cells and tissues. *PLOS ONE* **8(11)**, e79044.

Messerschmidt D.M., Knowles B.B. & Solter D. (2014): DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos. *Genes & development* **28**, 812-28.

Newman M., Blyth B.J., Hussey D.J., Jardine D., Sykes P.J. & Ormsby R.J. (2012): Sensitive quantitative analysis of murine LINE1 DNA methylation using high resolution melt analysis. *Epigenetics* **7**, 92-105.

Salilew-Wondim, D., Fournier, E., Hoelker, M., Saeed-Zidane, M., Tholen, E., Looft, C, Neuhoff C., Besenfelder U., Havlicek V., Rings F., Gagne D., Sirard M.A., Robert C., Saadi H.A.S., Gad A., Schellander K. & Tesfaye D. (2015): Genome-Wide DNA Methylation Patterns of Bovine Blastocysts Developed In Vivo from Embryos Completed Different Stages of Development In Vitro. *PLOS ONE* **10(11)**, e0140467.

Sawai K., Takahashi M., Fujii T., Moriyasu S., Hirayama H., Minamihashi A., Hashizume T. & Onoe S. (2011): DNA methylation status of bovine blastocyst embryos obtained from various procedures. *Journal of Reproduction and Development* **57**, 236-41.

Schostak N., Pyatkov K., Zelentsova E., Arkhipova I., Shagin D., Shagina I., Mudrik E., Blintsov A., Clark I. & Finnegan D.J. (2008): Molecular dissection of Penelope transposable element regulatory machinery. *Nucleic Acids Research* **36**, 2522-29.

Schulz W., Steinhoff C., Florl A. (2006): Methylation of endogenous human retroelements in health and disease. In: Doerfler W., Bohm P. (ed.): DNA Methylation: Development, Genetic Disease and Cancer. Springer Berlin Heidelberg, Germany, 211-50.

Smith Z.D., Chan M.M., Mikkelsen T.S., Gu H., Gnirke A. & Regev A., Meissner A. (2012): A unique regulatory phase of DNA methylation in the early mammalian embryo. *Nature*, **484** 339-44.

Wroclawska, E., Brant, J. O., Yang, T. P. & Moore, K. (2010): Improving efficiencies of locus-specific DNA methylation assessment for bovine in vitro produced embryos. *Systems Biology in Reproductive Medicine* **56(1)**, 96-105.

Yang A.S., Estecio M.R., Doshi K., Kondo Y., Tajara E.H. & Issa J.P.J. (2004): A simple method for estimating global DNA methylation using bisulfite PCR of repetitive DNA

elements. *Nucleic Acids Research* **32(3)**, e38.

Young L.E., Fernandes K., McEvoy T.G., Butterwith S.C., Gutierrez C.G., Carolan C., Broadbent P.J., Robinson J.J., Wilmut I. & Sinclair K.D. (2001): Epigenetic change an Igf2r is associated with fetal overgrowth after sheep embryo culture. *Nature genetics* **27**,153-4.

Chapter 6

General discussion and conclusions

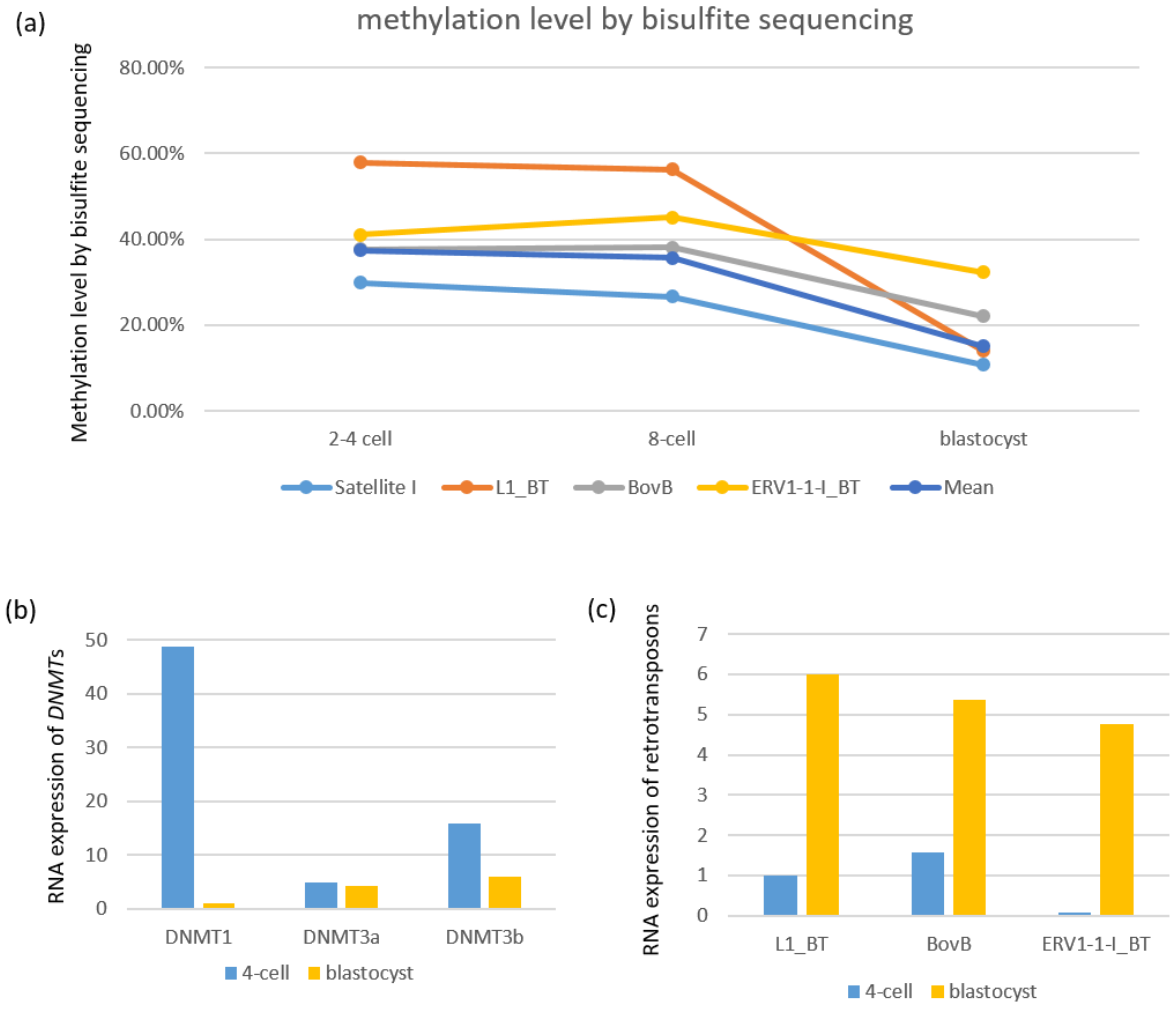
The aim of the present thesis was to perform a detailed investigation of the retrotransposon expression and to analyse its relation with global DNA methylation during bovine preimplantation embryo development. In this final chapter, the results of and the difficulties encountered during the study will be discussed. We will consequently focus on (1) correlation of DNA methylation and retrotransposon expression; (2) influence of oxidative stress on methylation and retrotransposon expression; (3) problems in immunofluorescence staining; (4) comparison of the methylation research methods used in the thesis; (5) DNA methylation dynamics during bovine preimplantation embryo development; and (6) retrotransposons used as global DNA methylation marker.

6.1 Correlation of DNA methylation and expression of retrotransposons during preimplantation development

DNA methylation is an important epigenetic mechanism regulating retrotransposon repression, and a correlation between DNA methylation and retrotransposon expression (especially L1 elements) has been established in many studies. In this thesis, we have studied the relation between DNA methylation and the expression of three autonomous retrotransposons (*L1-BT*, *BovB* and *ERV1-1-I_BT*) in bovine preimplantation embryos.

Immunofluorescence staining and RNA expression of *DNMT*s and three retrotransposons in 4-cell and blastocyst stages was performed (**Chapter 4)**, while in a second set of experiments, bisulfite sequencing (BS) of repetitive sequences was done (**Chapter 5)**. The results of the studies are summarized in **Figure 6.1**. All retrotransposons and *satellite I* DNA show a decrease in methylation level during development reaching the lowest level in blastocysts (**Figure 6.1a)**. The RNA expression of *DNMT1*, dramatically dropping in blastocysts compared to earlier stages, is in line with the decrease in DNA methylation level (**Figure 6.1b**). DNA methylation dynamics in bovine early development is discussed in more detail in **6.5**. No difference was observed in expression of *DNMT3a* and *DNMT3b* between 4-cell and blastocyst. As *DNMT3a* and *DNMT3b* were suggested to be essential for *de novo* methylation

in mammalian early development (Okano *et al.* 1999) their expression pattern together with the BS result indicates global *de novo* methylation is not happening at the blastocyst stage.



(d)

		DNA methyltransferases RNA expression			retrotransposon RNA expression			Methylation on repeat CpGs				
		DNMT1	DNMT3a	DNMT3b	L1_BT	BovB	ERV1-1-I_BT	me_MEAN	me_ satellite I	me_L1_BT	me_BovB	me_ERV1-1-I_BT
DNA methyltransferases RNA expression	DNMT1		0.212	0.004	0.011	0.018	0.01	0.038	0.119	0.011	0.019	0.261
	DNMT3a	0.212		0.489	0.194	0.158	0.25	0.106	0.052	0.158	0.324	0.167
	DNMT3b	0.004	0.489		0.019	0.03	0.05	0.047	0.136	0.015	0.013	0.267
retrotransposon RNA expression	L1_BT	0.011	0.194	0.019		0.019	0.041	0.061	0.144	0.032	0.026	0.366
	BovB	0.018	0.158	0.03	0.019		0.111	0.102	0.17	0.08	0.078	0.388
	ERV1-1-I_BT	0.01	0.25	0.05	0.041	0.111		0.039	0.083	0.013	0.027	0.28
Methylation on repeat CpGs	me_MEAN	0.038	0.106	0.047	0.061	0.102	0.039		0.025	0.009	0.102	0.124
	me_satellite I	0.119	0.052	0.136	0.144	0.17	0.083	0.025		0.063	0.219	0.063
	me_L1_BT	0.011	0.158	0.015	0.032	0.08	0.013	0.009	0.063		0.054	0.178
	me_BovB	0.019	0.324	0.013	0.026	0.078	0.027	0.102	0.219	0.054		0.367
	me_ ERV1-1-I_BT	0.261	0.167	0.267	0.366	0.388	0.28	0.124	0.063	0.178	0.367	

Figure 6.1 DNA methylation and retrotransposon expression in bovine preimplantation embryos. (a) Methylation level determined by bisulfite sequencing of repetitive sequences. (b) RNA expression of methyltransferases. (c) RNA expression of retrotransposons. (d) Correlation between DNA methylation and retrotransposon expression. The correlation was tested using Pearson's Correlation Coefficient. Negative correlations are shown in blue, and positive correlations are shown in orange, numbers represent the p-value, dark blue and orange represent significant linear relationship ($p<0.05$).

The RNA expression of retrotransposons showed a negative correlation with DNA methylation during preimplantation development (**Figure 6.1d**), with low DNA methylation level and high RNA expression in the blastocyst stage compared to the earlier stages. Furthermore, this negative correlation exists also when single retrotransposon classes are considered. In blastocysts, the methylation level is *L1_BT* < *BovB* < *ERV1-1-I_BT*, while the RNA expression is *L1_BT* > *BovB* > *ERV1-1-I_BT*. Guo *et al.* (2014) found that during demethylation in human embryos, evolutionarily younger LINEs and SINEs are demethylated to a milder extent compared to older elements. However, this phenomenon was not observed in our study, because *L1_BT* is a younger repeat than *BovB* (Adelson *et al.* 2009), but is more

demethylated in blastocysts. A recent report showed different DNA methylation mechanisms between LINEs and LTR retrotransposons: the methylation of LINEs is dependent on piRNAs, but the methylation of LTR elements is Mili-dependent, and Miwi2-independent (Nagamori *et al.* 2015).

6.2 Oxidative stress influence on methylation and retrotransposon expression

Oxidative stress is generated by excessive production of reactive oxygen species and/or reduction of antioxidant defences, and has been reported to induce epigenetic changes in different cell types (Franco *et al.* 2008; Giacco *et al.* 2010). Mild oxidative stress is important in embryo development, as it regulates key transcription factors that influence cell-signaling pathways involved in proliferation, differentiation and apoptosis, but high oxidative stress can cause DNA damage, epigenetic change, defective embryo development, and can even be lethal (Guerin *et al.* 2001; Dennery 2007; Donkena *et al.* 2010). For *in vitro* culture of embryos, excess oxidative stress can be induced by many factors, including: heat, oxygen tension, freeze-thaw process, light, serum in the culture media and so on (Thompson *et al.* 1990; Goto *et al.* 1993; Bilodeau *et al.* 2000; Rizos *et al.* 2003).

High concentration of oxygen was reported by our group to influence the efficiency of embryo production and embryo quality (Yuan *et al.* 2003). Oxidative stress by high concentration of oxygen during embryo culture was also reported altering the expression of epigenome modifying genes such as the ten-eleven translocation (*TET*) gene, which is associated with conversion of 5-methylcytosine to 5-hydroxymethylcytosine at the 16-cell and blastocyst stages of bovine IVF embryos (Burroughs *et al.* 2012). It has also been proposed that oxidative stress may affect DNA methylation by DNA oxidation or TET mediated hydroxymethylation in cancer (Vanden Berghe 2012). Marked perturbations were also found in global patterns of gene expression in mouse embryos cultured in 20% O_2 as compared with 5% O_2 (Rinaudo *et al.* 2006). Therefore, in this thesis, we applied 20% O_2 culture to expose bovine embryos to excess oxidative stress, and found a significant decrease in day 8 blastocyst rate, 23.34% compared to 45.32% in the control group. Furthermore, we found an

influence of high oxygen tension on DNA methylation by 5-mC immunofluorescence staining at 4-cell and blastocyst stages, but not when using bisulfite sequencing (BS) based method, this may also due to the different methylation coverage of the two methods, which is discussed in **Chapter 6.4**.

When looking at retrotransposon expression, we found that oxygen tension only increased *ERV1-1-I_BT* in female blastocysts, which may be explained by enrichment of ERVs in the bovine X-chromosome (Sperber *et al.* 2007; Garcia-Etxebarria & Jugo 2010). Moreover, the females carry a double amount of X-linked transcripts before X-chromosome inactivation, so the effect is also more obvious than in male embryos. Giorgi *et al.* (2011) used hydrogen peroxide (H_2O_2) treatment to induce oxidative stress in neuroblastoma cells, and found L1 expression was twice as high when compared to control cells. Two ERV families were found with significantly increased expression in preimplantation rabbit embryos cultered in two different IVC conditions when compared to *in vivo* (Salvaing *et al.* 2016). However, with high oxygen treatment, we did not find the same result in bovine embryos.

A recent study found no significant difference in DNA methylation between mouse embryos cultured under 5% O_2 and 20% O_2, while culturing embryos in both oxygen concentrations resulted in a significant increase of epigenetic defects later on in placental tissues compared to naturally conceived controls (de Waal et al. 2014). This result indicates that even culturing embryos in a 5% O_2 environment is far from optimal. Many other factors during IVP, such as heat, light, culture media can all induce oxidative stress to the embryo. Using IVC embryo culture under reduced oxygen concentration, as a standard control to study global methylation and retrotransposon activity, may already induce oxidative stress and disturb the transcriptional and epigenetic dynamics. A recent study showed that although embryos cultured under 5% O_2 have up-regulation of genes involved in cell morphogenesis, which is relevant for embryo development and blastocyst formation, but no difference in expression of genes involved in functions as oxidative phosphorylation and stress processes when compared with 20% O_2 culture (Mantikou *et al.* 2016).

The influence of oxidative stress on DNA methylation is suggested to work via many pathways, including regulation of methylation related enzymes (TETs) (Burroughs *et al*. 2012) and DNA repair pathways in which DNA lesions function as a substrate for binding DNMTs, resulting in global hypomethylation (Wachsman 1997). Oxidative stress may also induce site-specific hypermethylation by either the up-regulation of DNMTs or formation of new DNMT containing complexes (Wu & Ni 2015). However, it is also involved in metabolic pathways, apoptosis, pregnancy recognition regulation of growth factors and other cellular processes (reviewed by Takahashi 2012), and interaction among these processes may lead to a more complex influence on gene expression and epigenetics than was previously anticipated. Therefore, it can be concluded that high oxygen concentration is not an ideal stressor to study stress effects on DNA methylation and activity of retrotransposons. More specific stimuli for influencing DNA methylation, such as to expose embryos to 5-Aza-2'-deoxycytidine (DNA methyltransferase inhibitor), should be used, to test the 'genomic shock' hypothesis. Besides, *in vivo* produced embryos would be more ideal as a standard for this study to avoid extra stress from IVC procedure, and evaluation or quantification of the stress or stimuli should be taken into account, instead of only using blastocyst rate.

6.3 Difficulties encountered in the immunofluorescence staining

Immunofluoresent staining is often used to detect DNA methylation status in preimplantation embryos, because it can show DNA methylation changes on 5-mC (and other cytosine modifications, such as 5-hmC) within the genome from single embryos, and even single nuclei. 5-mC immunofluorescence staining on preimplantation embryos has been reported widely across different species (Santos *et al.* 2002; Beaujean *et al.* 2004; Fulka *et al.* 2004; Shi *et al.* 2004; Deshmukh *et al.* 2011). However, the results obtained with immunofluorescence staining can be highly influenced by the protocol used. Therefore, careful validation and control are required for reliable results of the cytosine staining. A recent review described determinants of valid measurements of this staining in the early

embryos (Salvaing *et al.* 2015). Some crucial steps we encountered in the staining described in **Chapter 4** are discussed here.

Denaturation of DNA

DNA denaturation is a very important step in 5-mC immunofluorescence staining. It is crucial for detecting and quantifing DNA methylation (5-mC) efficiently and accurately. Double-stranded DNA blocks antibody binding to 5-mC. Therefore successful detection of 5-mC requires DNA denaturation to make the nucleotides accessible for the antibody. HCl is widely used for denaturing DNA (Dean *et al.* 2001; Beaujean *et al.* 2004; Dobbs *et al.* 2013). HCl breaks the hydrogen bonds between complementary DNA strands converting double-stranded DNA into single-stranded DNA. Ideally, complete denaturation of DNA makes all 5-methylcytosine exposed to the antibody, but a high HCl concentration or long time treatment may disturb DNA counter staining. For example, Kennedy *et al*. (2000) observed divergent bromodeoxyuridine (BrdU) staining results under different HCl concentrations. BrdU is a synthetic nucleoside analog of thymidine commonly used in the detection of proliferating cells in living tissues. As 5-mC staining, BrdU antibodies only bind to single-stranded DNA. They showed that BrdU staining is more visable with increasing HCl concentration while the DAPI counter staining disappeared in high HCl concentrations, because it binds only to double-stranded DNA. It was also reported that when 4N HCl is used for denaturation, DAPI and Hoechst are failed to bind DNA (Heras *et al.* 2014). Furthermore, the denaturation protocol required adjustments for different tissues or samples. A denaturation protocol using combined HCl and trypsin treatment was suggested for mouse zygote 5-mC immunofluorescence staining. The extra trypsin treatment showed different 5-mC levels in zygotic PN5 and metaphase (Li & O'Neill 2012). The explanation is that a range of methyl binding domain (MBD) proteins are highly specifically binding to 5-mC and the treatment with trypsin can remove the MBDs exposing the 5-mC to the anti 5-mC antibody, but so far this phenomenon is not studied in later mouse embryo stages.

Image acquisition

Normally, embryo images are acquired and analyzed in a z-stack, which combines multiple images taken at different focal distances to provide a composite image with a greater depth of field. In other words, the object is 'sliced' for single image, and the images are z-stacked to form a composite 3D image of the object. The image slices should go through the sample, and the number of slices depends on the thickness of the sample and the depth between each slice. In that case, the more slices, the less information is lost. However, during the image acquisition, bleaching of the fluorescence can cause the image to become darker over time, and lead to bias between the first and the last image acquired, especially for expanded blastocysts, that can be more than 200 μm in diameter.

Moreover, fluorescent photomicrography suffers from common problems as other forms of optical microscopy and photomicrography, including overexposure and underexposure, which may lead to wrong quantification results. Image acquisition systems should be set so that the fluorescent intensity of the images avoids over- or underexposure among the z-stack images and samples. However, preimplantation embryos experience rapid DNA methylation dynamics, thus one setting for all is not possible. Therefore, in the present study we applied one setting for each embryo stage to compare DNA methylation between two culture conditions (5% and 20% O_2) (**Chapter 4)**, but not among different stages, to avoid wrong results caused by overexposure.

Quantification of the fluorescence

The methylation level is estimated by the fluorescence of a selected region of interest corrected for background by subtracting the mean intensity of the cytoplasmic area surrounding each nucleus. There are different ways for fluorescence quantification. The two most frequently used are: all 5-mC fluorescence of nucleus z-stack (Beaujean *et al.* 2004) and 5-mC/DNA fluorescence ratio (Reis *et al.* 2011). The former's advantage is all fluorescence of the nucleus can be taken into account, but the disadvantage is fluorescence of overlapping

nuclei may be included in the analysis and increase the fluorescence intensity; the latter method can avoid the overlapping, but the result is depending on the DNA staining. As mentioned in **Chapter 5**, The most common dyes for DNA counter staining are Hoechst, DAPI, Propidium Iodide (PI) and Ethidium homodimer-2 (EthD-2) which only bind to double-stranded DNA, or weakly bind to single-stranded DNA and RNA as well. Therefore, this widely used ratio of 5-mC/DNA presents actually single-stranded methylated DNA/total double-stranded DNA, which can be easily influenced by the denaturation status of the DNA. Therefore, anti-single-stranded DNA antibodies are recommended to use for this normalization.

6.4 Comparison of DNA methylation experiments

The major and traditional DNA methylation involves the transfer of a methyl group to the carbon 5-position of the cytosine base, often at CG dinucleotides (CpG), to produce 5-methylcytosine (5-mC). 5-mC is implicated in numerous cellular processes during development by gene regulation. However, during the last several years, new methylation modifications of cytosine have been reported. Ten-eleven translocation (TET) enzymes are cytosine oxygenases that convert 5-mC to 5-hydroxymethylcytosine (5-hmC) and further to 5-formylcytosine (5-fC) and 5-carboxylcytosine (5-caC) (Ito *et al.* 2011). It is suggested that these intermediates could have functions in controlling cell identity and the transition between the cell states, especially 5-fC, which was reported to bind gene regulatory factors (Iurlaro *et al.* 2013). Based on that, new insights of mechanisms that regulate DNA methylation have been suggested by many researchers (Tahiliani *et al.* 2009; Seisenberger *et al.* 2012; Song *et al.* 2012; Ficz 2015), shown in **Figure 6.2**. DNA methylation is established and maintained by DNMTs. Active demethylation can be initiated by TETs by converting 5-mC to 5-hmC, 5-hmC to 5-fC and 5-fC to 5-caC. 5-fC and 5-caC can then be excised by thymine DNA glycosylase (TDG) to generate an abasic site. This abasic site can be repaired to a cytosine by the base excision repair (BER) pathway. 5-hmC may be deaminated by AID and APOBEC1 to 5-hmU, then to an abasic site by TDG or single-strand selective monofunctional uracil

DNA glycosylase (SMUG1), or directly be dehydroxymethylated by DNMT3a/3b (Chen *et al.* 2012). Besides, all methylcytosine variants can be passively demethylated back to cytosine through DNA replication.

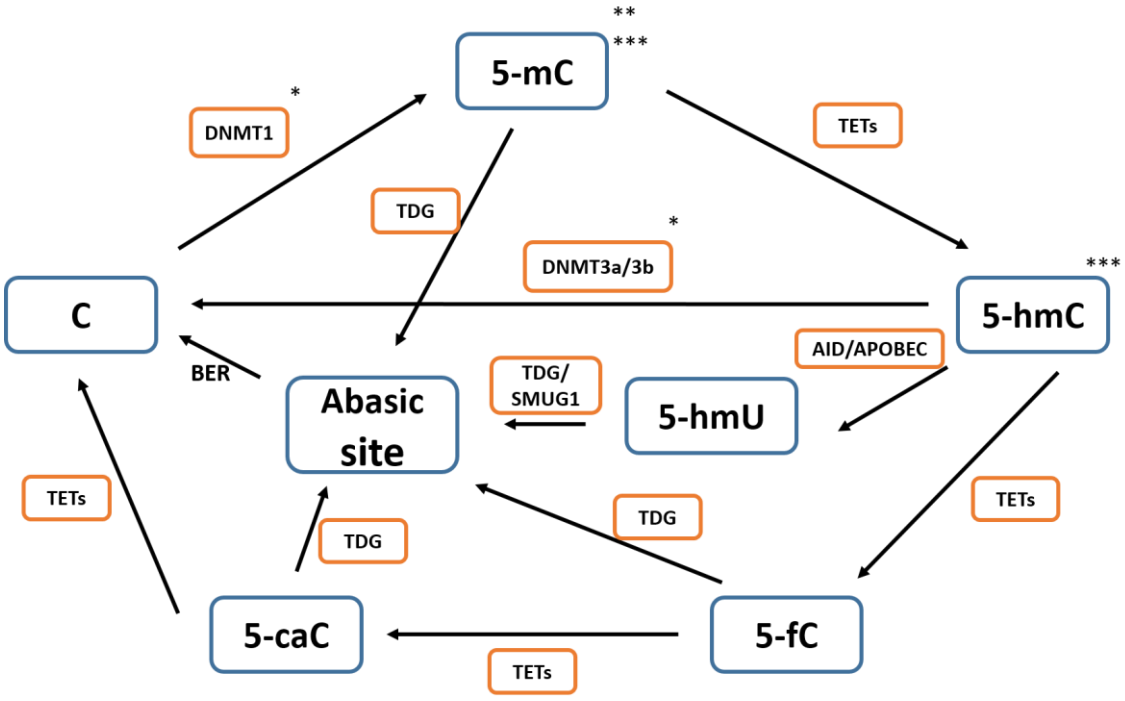


Figure 6.2 Different cytosine modifications and their interactions. * studied by RNA expression experiment; ** studied by immunofluorescence staining; *** studied by bisulfite sequencing.

We applied three methods to study DNA methylation dynamics in bovine preimplantation embryos: anti-5-mC immunofluorescence staining, gene expressing of *DNMT*s, and bisulfite sequencing of repetitive sequences. However, the results of these three methods were not exact in line with each other, which has been mentioned in **6.1**. This difference may be due to the fact that the three methods target different steps in the methylation biochemistry (indicated in **Figure 6.2**) or to a potential bias inherent to the methods used or a combination of both.

Using the immunofluorescence staining method described in **Chapter 4**, all methylcytosines throughout the genome are detected, not only those restricted to the CpG islands. Gene expressing of *DNMT*s was studied also in **Chapter 4**. *DNMT1* is responsible for the

maintenance of DNA methylation during replication, and overexpression of *DNMT1* causes genomic hypermethylation (Biniszkiewicz *et al.* 2002). *DNMT3a* and *DNMT3b* are *de novo* methyltransferases responsible for methylation of DNA in the absence of a template. They function as dehydroxymethylases that convert 5-hmC to C (Chen *et al.* 2012). Therefore, the RNA expression of *DNMT*s is positively correlated to the DNA methylation level in the embryos. Bisulfite sequencing (**Chapter 5**) can not distinguish 5-hmC from 5-mC, so the result of the experiment represents the total amount of 5-mC and 5-hmC (Jin *et al.* 2010). However, recently modifications of bisulfite sequencing, oxidative bisulfite sequencing (oxBS-Seq) and reduced bisulfite sequencing (redBS-Seq), were developed to quantitatively determine 5-mC, 5-hmC and 5-fC (Booth *et al.* 2013; Booth *et al.* 2014). With these modifications, repetitive sequences can potentially be used as global marker for 5-mC, 5-hmC and 5-fC separately. Nevertheless, we found consistently that the methylation level is relatively high in 4-cell embryos and low in blastocysts in both the *DNMT*s expression and bisulfite sequence experiments.

Another factor potentially introducing differences observed is variation in DNA methylation level during progression of the cell cycle. Expression of *DNMT1* and *3b* was shown significantly downregulated in G_0/G_1 while *DNMT3a* mRNA levels were less sensitive to cell cycle alterations (Robertson *et al.* 2000). Brown *et al.* (2007) found, using immunofluorescence staining, that the global levels of DNA methylation decreased in G_1 and increase during the S phase, while the DNA methylation level of repetitive sequences changed little throughout the cell cycle by bisulfite sequencing (Brown *et al.* 2007). In line with this last observation is a report of Vandiver and colleagues. Using whole genome bisulfite sequencing they found no global changes during cell cycle phases and that global DNA methylation is stable during replication and cell cycle arrest (Vandiver *et al*. 2015).

Table 6.1 DNA methylation patterns in mammalian preimplantation embryos. EDMA: EmbryoGENE DNA Methylation Analysis.

Species	Method	Feature	Reference
Human	RRBS	Lowest methylation at blastocyst (ICM).	(Guo *et al.* 2014; Smith *et al.* 2014)
Human	Staining	Lowest methylation at morula, and more methylation in TE than ICM.	(Fulka *et al.* 2004)
Mouse	RRBS	Lowest methylation at blastocyst (ICM).	(Smith *et al.* 2012)
Mouse	Staining	Demethyltion until morula, ICM but not TE undergoes *de novo* methylation.	(Dean *et al.* 2001)
Cow	EDMA	Demethylation until blastocyst.	(Saadi *et al.* 2014)
Cow	Staining	Lowest methylation at 6-8 cell, and followed by *de novo* methylation. More methylation in ICM than TE.	(Dean *et al.* 2001)
Cow	Staining	Lowest methylation at 6-8 cell. More methylation in TE than ICM.	(Dobbs *et al.* 2013)
Cow	Staining	Lowest methylation at morula stage, and more methylation in TE than ICM.	(Hou et al. 2007)
Pig	Staining	No loss of methylation during development, and more methylation in ICM than TE.	(Fulka *et al.* 2006)
Sheep	Staining	Methylation declines until blastosyct, and more methylation in ICM than TE.	(Beaujean *et al.* 2004)
Rabbit	Staining	No loss of methylation during development, and more methylation in TE than ICM.	(Shi *et al.* 2004)

6.5 DNA methylation reprogramming during preimplantation embryo development

An important step in the DNA methylation reprogramming in mammalian preimplantation embryos is demethylation of parental genomes after fertilization, followed by *de novo* methylation of the embryo genome. These methylation changes were first observed in mouse embryos. The detected *de novo* methylation began at the blastocyst stage and was more

intense in ICM than TE (Carlson *et al*. 1992). Later, the methylation patterns in preimplantation embryos were studied in many other species (listed in **Table 6.1**). For bovine embryos, different opinions still exist on the exact timing of *de novo* methylation. There are reports indicating that the global methylation drops during cleavage and reaches the lowest point at 6-8 cells stages, and increases again at the 16-cell stage of development (Dean *et al.* 2001; Dobbs *et al.* 2013), while other reports indicated that demethylation was only observed after the 8-cell stage and persisted throughout the morula stage (Hou *et al.* 2007). On the other hand, 5-mC stained blastocysts showed intense fluorescence, mostly in TE cells (**Figure 4.1**), but we did not compare it with other stages statistically. Whether there is more methylation in TE cells or in ICM cells in bovine blastocysts is still debated. Different methylation patterns in the ICM and TE are shown in **Figure 6.3** (Dean *et al.* 2001; Hou *et al.* 2007). This variability may be due to the differences in staining or embryo culture protocols used.

However, in contrast to the variation in DNA methylation patterns observed in different species or even one species as carried out by immunofluorescent staining, the sequencing based methods showed a similar tendency, which is that the methylation is decreasing during early development and is very low at the blastocyst stage; and less methylation is found in ICM in mouse, human and cow (Smith *et al.* 2012; Guo *et al.* 2014; Saadi *et al.* 2014; Smith *et al.* 2014). Our results from *DNMT*s expression and bisulfite sequencing experiments indicate that in blastocysts the DNA methylation still remains very low and is even the lowest compared to other stages in BS experiment, which is in consonance with these results. There is still lack of data from sequencing based method from other species. Therefore, we would strongly suggest that the 'species-specific' methylation which has been claimed by staining in the past, now needs to be confirmed by novel sequencing based technology.

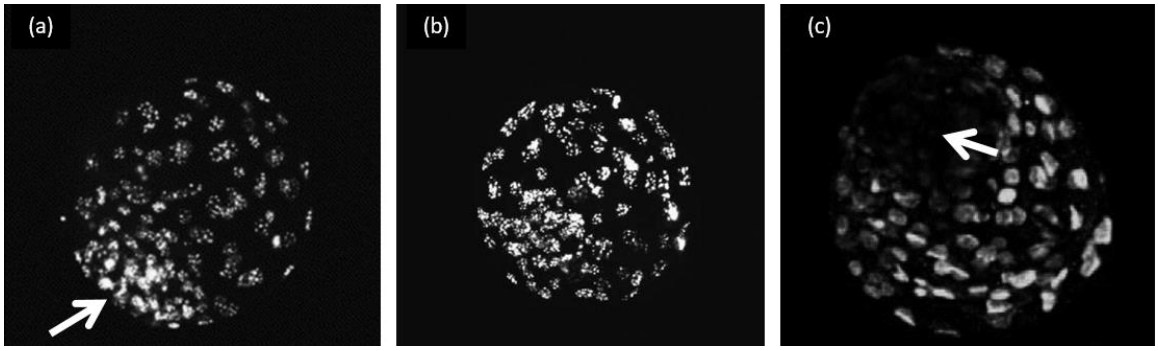


Figure 6.3 5-mC immunofluorescence staining of bovine blastocyst. (a) high methylated in ICM (Dean *et al.* 2001). (b) homogeneously methylated (Dean *et al.* 2001). (c) high methylated in trophectoderm (Hou *et al.* 2007).

6.6 Retrotransposons as DNA methylation marker

Recent research on epigenetics of mammalian preimplantation embryos found that transposable elements show a similar dynamic trend as global methylation (Guo *et al.* 2014). A bovine study also showed that the majority of the repetitive elements, particularly LINEs, SINEs and LTR retrotransposons, are hypermethylated in spermatozoa compared to blastocysts, which was in line with the methylation pattern of all types of promoter, intronic and exonic regions, non-CpG islands and CpG islands studied (Saadi *et al.* 2014). Therefore, those retrotransposons can be considered as global methylation markers in embryo development.

An advantage of using retrotransposons as methylation marker is the great coverage of the genome. In bovine, retrotransposons take almost half of the genome, and *L1_BT* and *BovB* alone take more than 20%. Other sequencing based methods, such as EDMA and RRBS, have a genome coverage of about 5%, respectively 10% of all CpG sites (Wu & Zhang 2012; Saadi *et al.* 2014). In this thesis, we suggested to use a weighted arithmetic mean of the repeat methylation, taking each retrotransposon or repeat genome coverage, which makes the calculation more accurate.

However, when repetitive sequences are used as global methylation marker, there are two factors that need to be considered. First is the mutation rate of repetitive sequences. High mutation rate of retrotransposons makes the number of potential targets for DNA methylation variable between the individual members of a repeat family. When selecting a retrotransposon as a global methylation marker, low mutation rate is the first criterion. For example, almost two-thirds of the CpG methylation sites of *Alu* are mutated, which makes *Alu* not suitable to be used as a methylation marker. Another factor that affect the methylation results is primer design for bisulfite PCR, which can influence the number of amplified repetitive sequences directly. In the present study, we tried to amplify as many as possible copies of the repeat families to represent the whole genome. All primers were designed based on the consensus sequence of each family and the amplified number was checked by BiSearch using the weighted arithmetic mean of methylation level of each repetitive sequence family.

6.7 General conclusions

The general conclusions of the thesis are:

1) Three autonomous retrotransposons, *L1_BT*, *BovB*, and *ERV1-1-I_BT*, with complete ORFs were found with consistent expression in bovine oocytes and crossing all stages during preimplantation embryo development. These retrotransposons were analysed during further research in this thesis.

2) A reliable assay was designed for the normalization of RT-qPCR analysis in bovine preimplantation embryos, considering developmental stages, gender and oxygen tension. Different stability rankings of gene expression were found for different sample group combinations. The four most stable reference genes (*GAPDH*, *YWHAZ*, *18S rRNA* and *SDHA*) were selected when all samples were considered. Specifically, *HPRT1* and *H2A* are the most stable genes for female embryo normalization, and for 4-cell and blastocyst embryos analysis, *GAPDH*, *YWHAZ* and *SDHA* are used for an accurate normalization, allowing small expression differences to be reliably measured.

3) Oxygen tension, as an oxidative stress, reduced the embryonic development, and induced RNA expression of *ERV-1-1-I_BT* in female blastocyst, which suggesting it is more harmful to female embryo than male ones.

4) A global DNA methylation immunofluorescence staining for embryo from the 2-cell to blastosyt stage cultured under 5% O_2 and 20% O_2 was performed. By this analytical method, significant increased methylation was found in 4-cell and blastocyst stages under 20% O_2 comparing to 5% O_2.

5) RT-qPCR analysis of *DNMT*s and retrotransposons showed a negative correlation with high expression of *DNMT1* and low expression of retrotransposon at 4-cell stage, and an opposite of expression at blastocyst stage.

6)A bisulfite sequence analysis of retrotransposon promoter regions showed a DNA methylation pattern in line with the *DNMT1* expression, both high in 4-cell and low in blastocyst stage. Moreover, an assay using Weighted Arithmetic Mean of the multiple repetitive sequences was designed, as a promising quick marker for the global DNA methylation status in preimplantation stages, while attention should be paid to primer design and mutation rates of the repeats selected.

6.8 Perspectives for future research

In this thesis, we induced oxidative stress (high oxygen concentration during embryo development), and studied if the stress triggers activation of retrotransposons and the relationship between global DNA methylation and retrotransposon expression. Unfortunately, we succeeded only partially in obtaining our goals. A correlation between global DNA methylation and retrotransposon expression was shown. However, this stressor does not give a significant influence, which may be due to the involvement of many cellular processes and interaction among the processes may lead to a more complex influence to gene expression and epigenetics, or because even under 5% O_2, the *in vitro* culture condition is still prone to

inducing stress and is far from optimal. Using only the blastocyst rate as an evaluation of the stress is not sufficient, so oxidative stress monitoring or evaluation, such as quantification of ROS during the development should be taken into consideration in the future. On the other hand, more specific stimuli for affecting DNA methylation should be used, such as to expose embryos to 5-Aza-2'-deoxycytidine (DNA methyltransferase inhibitor) to test the 'genomic shock' hypothesis.

Furthermore, methyltransferases RNA expression and bisulfite sequencing of repeats in the thesis showed a decreasing methylation pattern from cleavage to blastocyst in bovine embryos. However, considering the different patterns in previous 5-mC immunofluorescence staining results, we propose that methylation dynamics in preimplantation embryos should be studied by more than one method. Besides, conflicting results remain concerning the different methylation levels detected between ICM and TE in bovine blastocysts using different detection methods. We strongly suggest reliably separating ICM from TE for both epigenetic and genetic studies.

Also, during the last several years new insights in the mechanisms that regulate DNA methylation have been found and new methylation modifications of cytosine of which the biological significance is unclear have been detected. With the development of new technologies: oxidative bisulfite sequencing (oxBS-Seq) and reduced bisulfite sequencing (redBS-Seq), methylation intermediates can be studied separately, and multiple repeats can be used as the marker for to quantitatively determine the global 5-mC, 5-hmC and 5-fC.

In this thesis, we focused on three autonomous retrotransposons: *L1_BT*, *BovB* and *ERV1-1-I_BT*, but their non-autonomous counterparts can play regulatory functions also, thus non-autonomous retrotransposons should be taken into consideration for further studies.

6.9 References

Adelson D.L., Raison J.M. & Edgar R.C. (2009) Characterization and distribution of retrotransposons and simple sequence repeats in the bovine genome. *Proceedings of the National Academy of Sciences* **106**, 12855-60.

Beaujean N., Taylor J., Gardner J., Wilmut I., Meehan R. & Young L. (2004) Effect of limited DNA methylation reprogramming in the normal sheep embryo on somatic cell nuclear transfer. *Biology of reproduction* **71**, 185-93.

Bilodeau J.F., Chatterjee S., Sirard M.A. & Gagnon C. (2000) Levels of antioxidant defenses are decreased in bovine spermatozoa after a cycle of freezing and thawing. *Molecular reproduction and development* **55**, 282-8.

Biniszkiewicz D., Gribnau J., Ramsahoye B., Gaudet F., Eggan K., Humpherys D., Mastrangelo M.-A., Jun Z., Walter J. & Jaenisch R. (2002) Dnmt1 overexpression causes genomic hypermethylation, loss of imprinting, and embryonic lethality. *Molecular and cellular biology* **22**, 2124-35.

Booth M.J., Marsico G., Bachman M., Beraldi D. & Balasubramanian S. (2014) Quantitative sequencing of 5-formylcytosine in DNA at single-base resolution. *Nature chemistry* **6**, 435.

Booth M.J., Ost T.W., Beraldi D., Bell N.M., Branco M.R., Reik W. & Balasubramanian S. (2013) Oxidative bisulfite sequencing of 5-methylcytosine and 5-hydroxymethylcytosine. *Nature protocols* **8**, 1841-51.

Brown S.E., Fraga M.F., Weaver I.C., Berdasco M. & Szyf M. (2007) Variations in DNA methylation patterns during the cell cycle of HeLa cells. *Epigenetics* **2**, 54-65.

Burroughs C.A., Williamson G.L., Golding M.C. & Long C.R. (2012) Oxidative stress induced changes in epigenetic modifying gene mRNA in pre-implantation *in vitro* bovine embryos. *Reproduction, Fertility and Development* **25**, 149-.

Carlson L.L., Page A. & Bestor T. (1992) Properties and localization of DNA methyltransferase in preimplantation mouse embryos: implications for genomic imprinting. *Genes & development* **6**, 2536-41.

Chen C.-C., Wang K.-Y. & Shen C.-K.J. (2012) The mammalian de novo DNA methyltransferases DNMT3A and DNMT3B are also DNA 5-hydroxymethylcytosine dehydroxymethylases. *Journal of Biological Chemistry* **287**, 33116-21.

de Waal E., Mak W., Calhoun S., Stein P., Ord T., Krapp C., Coutifaris C., Schultz R.M. & Bartolomei M.S. (2014) In vitro culture increases the frequency of stochastic epigenetic errors at imprinted genes in placental tissues from mouse concepti produced through assisted reproductive technologies. *Biology of reproduction* **90**, 22.

Dean W., Santos F., Stojkovic M., Zakhartchenko V., Walter J., Wolf E. & Reik W. (2001) Conservation of methylation reprogramming in mammalian development: aberrant reprogramming in cloned embryos. *Proceedings of the National Academy of Sciences* **98**, 13734-8.

Dennery P.A. (2007) Effects of oxidative stress on embryonic development. *Birth Defects Research Part C: Embryo Today: Reviews* **81**, 155-62.

Deshmukh R.S., Østrup O., Østrup E., Vejlsted M., Niemann H., Lucas-Hahn A., Petersen B., Li J., Callesen H. & Hyttel P. (2011) DNA methylation in porcine preimplantation embryos developed in vivo and produced by in vitro fertilization, parthenogenetic activation and somatic cell nuclear transfer. *Epigenetics* **6**, 177-87.

Dobbs K.B., Rodriguez M., Sudano M.J., Ortega M.S. & Hansen P.J. (2013) Dynamics of DNA Methylation during Early Development of the Preimplantation Bovine Embryo. *PLOS ONE* **8**, e66230.

Donkena K.V., Young C.Y. & Tindall D.J. (2010) Oxidative stress and DNA methylation in prostate cancer. *Obstetrics and gynecology international* **2010**.

Ficz G. (2015) New insights into mechanisms that regulate DNA methylation patterning. *Journal of Experimental Biology* **218**, 14-20.

Franco R., Schoneveld O., Georgakilas A. G., & Panayiotidis, M. I. (2008) Oxidative stress, DNA methylation and carcinogenesis. *Cancer letters* **266(1)**, 6-11.

Fulka H., Mrazek M., Tepla O. & Fulka J. (2004) DNA methylation pattern in human zygotes and developing embryos. *Reproduction* **128**, 703-8.

Fulka J., Fulka H., Slavik T., Okada K. & Fulka Jr J. (2006) DNA methylation pattern in pig

in vivo produced embryos. *Histochemistry and cell biology* **126**, 213-7.

Garcia-Etxebarria K. & Jugo B.M. (2010) Genome-wide detection and characterization of endogenous retroviruses in Bos taurus. *Journal of virology* **84**, 10852-62.

Giacco F., & Brownlee M. (2010) Oxidative stress and diabetic complications. *Circulation research* **107(9)**, 1058-1070.

Giorgi G., Marcantonio P. & Del Re B. (2011) LINE-1 retrotransposition in human neuroblastoma cells is affected by oxidative stress. *Cell and tissue research* **346**, 383-91.

Goto Y., Noda Y., Mori T. & Nakano M. (1993) Increased generation of reactive oxygen species in embryos cultured in vitro. *Free Radical Biology and Medicine* **15**, 69-75.

Guerin P., El Mouatassim S. & Menezo Y. (2001) Oxidative stress and protection against reactive oxygen species in the pre-implantation embryo and its surroundings. *Human Reproduction Update* **7**, 175-89.

Guo H., Zhu P., Yan L., Li R., Hu B., Lian Y., Yan J., Ren X., Lin S. & Li J. (2014) The DNA methylation landscape of human early embryos. *Nature* **511**, 606-10.

Heras S., Forier K., Rombouts K., Braeckmans K., & Van Soom A. (2014). DNA counterstaining for methylation and hydroxymethylation immunostaining in bovine zygotes. *Analytical biochemistry* **454**, 14-6.

Hou J., Liu L., Lei T., Cui X., An X. & Chen Y. (2007) Genomic DNA methylation patterns in bovine preimplantation embryos derived from in vitro fertilization. *Science in China Series C: Life Sciences* **50**, 56-61.

Ito S., Shen L., Dai Q., Wu S.C., Collins L.B., Swenberg J.A., He C. & Zhang Y. (2011) Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* **333**, 1300-3.

Iurlaro M., Ficz G., Oxley D., Raiber E.-A., Bachman M., Booth M.J., Andrews S., Balasubramanian S. & Reik W. (2013) A screen for hydroxymethylcytosine and formylcytosine binding proteins suggests functions in transcription and chromatin regulation. *Genome biology* **14**, 1.

Jin S.-G., Kadam S. & Pfeifer G.P. (2010) Examination of the specificity of DNA methylation

profiling techniques towards 5-methylcytosine and 5-hydroxymethylcytosine. *Nucleic acids research* **38**, e125-e.

Kennedy B.K., Barbie D.A., Classon M., Dyson N. & Harlow E. (2000) Nuclear organization of DNA replication in primary mammalian cells. *Genes & development* **14**, 2855-68.

Li Y. & O'Neill C. (2012) Persistence of Cytosine Methylation of DNA following Fertilisation in the Mouse. *PLOS ONE* **7**, e30687.

Mantikou E., Jonker M.J., Wong K.M., van Montfoort A.P., de Jong M., Breit T., Repping S. & Mastenbroek S. (2016) Factors affecting the gene expression of in vitro cultured human preimplantation embryos. *Human reproduction* **31**, 298-311.

Nagamori I., Kobayashi H., Shiromoto Y., Nishimura T., Kuramochi-Miyagawa S., Kono T. & Nakano T. (2015) Comprehensive DNA methylation analysis of retrotransposons in male germ cells. *Cell reports* **12**, 1541-7.

Okano M., Bell D.W., Haber D.A. & Li E. (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**, 247-57.

Reis S.A., Adenot P., Daniel N., Archilla C., Peynot N., Lucci C.M., Beaujean N. & Duranthon V. (2011) Dynamics of DNA methylation levels in maternal and paternal rabbit genomes after fertilization. *Epigenetics* **6**, 987-93.

Rinaudo P.F., Giritharan G., Talbi S., Dobson A.T. & Schultz R.M. (2006) Effects of oxygen tension on gene expression in preimplantation mouse embryos. *Fertility and Sterility* **86**, 1265. e1-36.

Rizos D., Gutierrez-Adan A., Perez-Garnelo S., De La Fuente J., Boland M. & Lonergan P. (2003) Bovine embryo culture in the presence or absence of serum: implications for blastocyst development, cryotolerance, and messenger RNA expression. *Biology of reproduction* **68**, 236-43.

Robertson K.D., Keyomarsi K., Gonzales F.A., Velicescu M. & Jones P.A. (2000) Differential mRNA expression of the human DNA methyltransferases (DNMTs) 1, 3a and 3b during the G0/G1 to S phase transition in normal and tumor cells. *Nucleic acids research* **28**, 2108-13.

Saadi H.A.S., O'Doherty A.M., Gagné D., Fournier É., Grant J.R., Sirard M.-A. & Robert C. (2014) An integrated platform for bovine DNA methylome analysis suitable for small samples. *BMC genomics* **15**, 1.

Salvaing J., Li Y., Beaujean N. & O'Neill C. (2015) Determinants of valid measurements of global changes in 5′-methylcytosine and 5′-hydroxymethylcytosine by immunolocalisation in the early embryo. *Reproduction, Fertility and Development* **27**, 755-64.

Salvaing J., Peynot N., Bedhane M., Veniel S., Pellier E., Boulesteix C., Beaujean N., Daniel N. & Duranthon V. (2016) Assessment of 'one-step'versus 'sequential'embryo culture conditions through embryonic genome methylation and hydroxymethylation changes. *Human reproduction* **31**, 2471-83.

Santos F., Hendrich B., Reik W. & Dean W. (2002) Dynamic reprogramming of DNA methylation in the early mouse embryo. *Developmental biology* **241**, 172-82.

Seisenberger S., Andrews S., Krueger F., Arand J., Walter J., Santos F., Popp C., Thienpont B., Dean W. & Reik W. (2012) The dynamics of genome-wide DNA methylation reprogramming in mouse primordial germ cells. *Molecular cell* **48**, 849-62.

Shi W., Dirim F., Wolf E., Zakhartchenko V. & Haaf T. (2004) Methylation reprogramming and chromosomal aneuploidy in in vivo fertilized and cloned rabbit preimplantation embryos. *Biology of reproduction* **71**, 340-7.

Smith Z.D., Chan M.M., Humm K.C., Karnik R., Mekhoubad S., Regev A., Eggan K. & Meissner A. (2014) DNA methylation dynamics of the human preimplantation embryo. *Nature* **511**, 611-5.

Smith Z.D., Chan M.M., Mikkelsen T.S., Gu H., Gnirke A., Regev A. & Meissner A. (2012) A unique regulatory phase of DNA methylation in the early mammalian embryo. *Nature* **484**, 339-44.

Song C.-X., Yi C. & He C. (2012) Mapping recently identified nucleotide variants in the genome and transcriptome. *Nature biotechnology* **30**, 1107-16.

Sperber G.O., Airola T., Jern P. & Blomberg J. (2007) Automated recognition of retroviral sequences in genomic data-RetroTector©. *Nucleic acids research* **35**, 4964-76.

Tahiliani M., Koh K.P., Shen Y., Pastor W.A., Bandukwala H., Brudno Y., Agarwal S., Iyer L.M., Liu D.R. & Aravind L. (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324**, 930-5.

Takahashi M. (2012) Oxidative stress and redox regulation on in vitro development of mammalian embryos. *Journal of Reproduction and Development* **58**, 1-9.

Thompson J., Simpson A., Pugh P., Donnelly P. & Tervit H. (1990) Effect of oxygen concentration on in-vitro development of preimplantation sheep and cattle embryos. *Journal of reproduction and fertility* **89**, 573-8.

Vanden Berghe W. (2012) Epigenetic impact of dietary polyphenols in cancer chemoprevention: Lifelong remodeling of our epigenomes. *Pharmacological Research* **65**, 565-76.

Vandiver A. R., Idrizi A., Rizzardi L., Feinberg A. P., & Hansen, K. D. (2015). DNA methylation is stable during replication and cell cycle arrest. *Scientific reports* **5**, 17911.

Wachsman J.T. (1997) DNA methylation and the association between genetic and epigenetic changes: relation to carcinogenesis. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **375**, 1-8.

Wu H. & Zhang Y. (2012) Early embryos reprogram DNA methylation in two steps. *Cell stem cell* **10**, 487-9.

Wu Q. & Ni X. (2015) ROS-mediated DNA methylation pattern alterations in carcinogenesis. *Current drug targets* **16**, 13-9.

Yuan Y., Van Soom A., Coopman F., Mintiens K., Boerjan M., Van Zeveren A., de Kruif A. & Peelman L. (2003) Influence of oxygen tension on apoptosis and hatching in bovine embryos cultured in vitro. *Theriogenology* **59**, 1585-96.

Summary

Retrotransposons, or RNA intermediated transposable elements, have been considered as ‘junk’ or ‘selfish’ DNA and dismissed as uninteresting for a long time. However, they were found active and functional in preimplantation embryos in the last decades. The impact of retrotransposon activity on the genome such as formation of new genetic elements and regulators in the genome is unneglectable. They are also suggested to participate in many activities during early embryo development, like genomic imprinting, X chromosome inactivation, cell proliferation and differentiation.

In **Chapter 1**, the current literature on retrotransposons in mammalian genomes is reviewed. An overview of the classification and composition of retrotransposons in the genome is given, with emphasis of the bovine genome. Furthermore, the mechanisms of retrotransposition are described in detail and the activities of retrotransposons on the level of transcription, translation and mobilization are explained. The impact of retrotransposon activity on the mammalian genome is discussed, illustrating that retrotransposons influence the genome by formatting new genetic elements and working as regulators. The defense mechanisms against retrotransposon to prevent their disturbance of the genomes are also given in **Chapter 1**.

In the latter part of **Chapter 1**, activation of retrotransposons in early development is discussed. Retrotransposons participate in many embryonic activities, including cell proliferation and differentiation, genomic imprinting, and X-chromosome inactivation. These activities are suggested to be mainly a consequence of DNA methylation loss during epigenetic reprogramming of preimplantation. Therefore, a description of the dynamics of DNA methylation during bovine preimplantation embryo development is included. Those facts reveal the importance of the study of the relation between DNA methylation and retrotransposon expression in embryo development. At last, a short review of using retrotransposons as marker for global DNA methylation status is given.

The aims of the study are presented in **Chapter 2**. The general aim of the thesis was to study retrotransposon expression and their DNA methylation control in bovine preimplantation embryos. The first aim was to find out autonomous retrotransposons with consistent RNA

expression, and reliable reference genes for gene expression normalization during bovine preimplantation embryo development. The second aim was to study the relation between DNA methylation and retrotransposon expression under high oxygen. The third aim was to use repetitive sequences as global DNA methylation marker in bovine preimplantation embryos.

To achieve these goals, we first profiled autonomous retrotransposon expression in bovine preimplantation embryos. *L1_BT, BovB and ERV1-1-I_BT* were found to be expressed throughout all stages of preimplantation development (**Chapter 3.1**). Furthermore, reference genes were evaluated and selected for estimating retrotransposon expression in bovine embryos of different stages and under different conditions (**Chapter 3.2**).

In **Chapter 4**, the relation between DNA methylation and retrotransposon expression under influence of high oxygen tension as stressor was studied. The global DNA methylation of embryos was estimated by 5-methylcytosine immunofluorescent (5-mC) staining. We found a significant increase in DNA methylation under 20% O_2 at the 4-cell stage and in blastocysts. Gene expression of DNA methytransferases (*DNMT*s) and three retrotransposons (*L1_BT, BovB and ERV1-1-I_BT*) in 4-cell embryos and blastocysts was analyzed by RT-qPCR. Unexpectedly, the retrotransposon expression was not correlated with the global DNA methylation level estimated by 5-mC staining, but was negatively correlated with the expression of *DNMT1*.

We further studied the DNA methylation level of retrotransposons during embryo development by bisulfite sequencing (**Chapter 5**). Using this method, a negative correlation between the methylation level (*L1_BT* < *BovB* < *ERV1-1-I_BT*) and their RNA level (*L1_BT* > *BovB* > *ERV1-1-I_BT*) was found. Additionally, we developed a global DNA methylation marker using a combination of repetitive sequences with low mutation rate, taking in to account the proportion of each sequence.

The general discussion and the conclusions are presented in **Chapter 6**. The discussion focuses on the following aspects: 1) correlation of DNA methylation and retrotransposon expression; 2) influence of oxidative stress on methylation and retrotransposon expression; 3) problems in immunofluorescence staining; 4) comparison of the methylation research methods used in the thesis; 5) DNA methylation dynamics during bovine preimplantation embryo development; and 6) retrotransposons used as global DNA methylation marker.

Samenvatting

Retrotransposons of RNA geïntermedieerde transposons werden lange tijd als 'junk' of 'egoïstisch' DNA beschouwd en van de hand gedaan als oninteressant. In de laaste decennia werd echter gevonden dat ze actief en fuctioneel zijn in pre-implantatie embryo's. De gevolgen van retrotransposon activiteit in het genoom, zoals vorming van nieuwe genetische elementen en regulatoren in het genoom is niet verwaarloosbaar. Er wordt nu verondersteld dat ze betrokken zijn bij vele activiteiten tijdens de vroege embryo ontwikkeling, zoals genomische imprinting, X chromosoom inactivatie, cel proliferatie en differentiatie.

In **hoofdstuk 1** wordt de huidige literatuur over retrotransposons in zoogdiergenomen beoordeeld. Een overzicht van de indeling en de samenstelling van retrotransposons in het genoom wordt gegeven, met de nadruk op het rundergenoom. Daarnaast worden de mechanismen van retrotranspositie beschreven en de activiteiten van retrotransposons op het niveau van transcriptie, translatie en mobilisatie worden toegelicht. De impact van retrotransposonactiviteit op het zoogdiergenoom wordt besproken als illustratie dat retrotransposons het genoom beïnvloeden door vorming van nieuwe genetische elementen en het werken als regulatoren. De beschermingsmechanismen om de verstoring van het genoom door retrotransposons te voorkomen worden ook beschreven in **hoofdstuk 1**.

In het laatste deel van **hoofdstuk 1** wordt de activering van retrotransposons in de vroege ontwikkeling besproken. Retrotransposons nemen deel aan tal van embryonale activiteiten, waaronder cel proliferatie en differentiatie, genomische imprinting en X-chromosoom inactivatie. Het werd gesuggereerd dat deze activiteiten vooral een gevolg zijn van het verlies van DNA-methylatie tijdens epigenetische herprogrammering tijdens de pre-implantatie ontwikkeling. Daarom werd een beschrijving van de DNA methylatie in boviene pre-implantatie embryoontwikkeling inbegrepen. Deze feiten laten het belang zien van het onderzoek naar de relatie tussen DNA methylatie en retrotransposonexpressie in embryo-ontwikkeling. Tot slot wordt een kort overzicht van het gebruik van retrotransposons als marker voor de wereldwijde status van DNA-methylatie gegeven.

De doelstellingen van de studie worden gepresenteerd in **Hoofdstuk 2**. Het algemene doel

van het proefschrift was om retrotransposonexpressie en de DNA-methylatie controle bij pre-implantatie embryo's van runderen te bestuderen. Het eerste doel was om autonome retrotransposons met consistente RNA expressie en betrouwbare referentiegenen voor genexpressie normalisering tijdens pre-implantatie embryo-ontwikkeling van runderen te identificeren. Het tweede doel was om de relatie tussen DNA-methylatie en retrotransposonexpressie onder hoge zuurstof te bestuderen. Het derde doel was om repetitieve sequenties als globale DNA-methylatie merker in pre-implantatie embryo's van runderen te gebruiken.

Om deze doelen te bereiken hebben we eerst de expressie van autonome retrotransposons geprofileerd in pre-implantatie embryo's van runderen. Expressie van *L1_BT*, *BovB* en *ERV1-1-I_BT* werd in alle stadia van de pre-implantatie ontwikkeling aangetroffen **(hoofdstuk 3.1**). Verder werden referentiegenen voor het bepalen van retrotransposonexpressie in runderembryo's van verschillende stadia en onder verschillende omstandigheden geëvalueerd en geselecteerd (**hoofdstuk 3.2**).

In **hoofdstuk 4** werd de relatie tussen DNA methylatie en retrotransposonexpressie onder invloed van hoge zuurstofspanning als stressor bestudeerd. De globale DNA methylatie van embryo's werd geschat via 5-methylcytosine immunofluorescentie (5-mC) kleuring. We vonden een significante toename van DNA methylering onder 20% O_2 in het 4-cel stadium en in blastocysten. Genexpressie van DNA methytransferases (DNMTs) en drie retrotransposons (*L1_BT*, *BovB* en *ERV1-1-I_BT*) in 4-cel embryo's en blastocysten werd verder geanalyseerd door RT-qPCR. Onverwacht bleek de retrotransposonexpressie niet gecorreleerd te zijn met het globale DNA methylatieniveau geschat door 5-mC kleuring, maar was het wel negatief gecorreleerd met de expressie van *DNMT1*.

Verder onderzochten we het DNA methylatieniveau van retrotransposons tijdens de embryonale ontwikkeling door bisulfietsequencing (**hoofdstuk 5**). Met behulp van deze methode werd een negatieve correlatie tussen het methylatieniveau (*L1_BT <BovB <ERV1-1-I_BT*) en hun RNA-niveau (*L1_BT> BovB> ERV1-1-I_BT*) gevonden. Daarnaast

hebben we een globale DNA methylatiemerker ontwikkeld gebruik makende van een combinatie van repetitieve sequenties met een lage mutatiefrequentie en rekening houdend met het aandeel van elke sequentie.

De algemene discussie en de conclusies worden gepresenteerd in **hoofdstuk 6**. De discussie richt zich op de volgende aspecten: 1) correlatie van DNA-methylatie en retrotransposonexpressie; 2) invloed van oxidatieve stress op methylatie en retrotransposonexpressie; 3) problemen bij immunofluorescentiekleuring; 4) vergelijking van de methylatie onderzoeksmethoden gebruikt in het proefschrift; 5) DNA-methylatie dynamiek tijdens runderen pre-implantatie embryo-ontwikkeling en 6) het gebruik van retrotransposons als globale DNA-methylatie merker.

Curriculum Vitae

Wenwen Li

Department of Nutrition, Genetics and Ethology, Ghent University, Heidestraat 19, 9820 Merelbeke, Belgium,

Email: wenwen.li @ugent.be / imwenwenli@gmail.com

Wenwen Li obtained her bachelor degree in Biotechnology from Northwest A&F University (September 2002 to July 2006), and master degree in Animal Science from the same University (September 2007 to July 2010). In October 2010, Wenwen Li started her PhD study at the Department of Nutrition, Genetics and Ethology, Faculty of Veterinary Medicine, Ghent University, Belgium.

Journal articles

Li W., Van Soom A., & Peelman L. (2016). Repeats as Global DNA Methylation Marker in Bovine Preimplantation Embryos. *Czech Journal of Animal Science* (In press). (A1)

Li W., Goossens K., Van Poucke M., Forier K., Braeckmans K., Van Soom A., & Peelman L. (2014). High oxygen tension increases global methylation in bovine 4-cell embryos and blastocysts but does not affect general retrotransposon expression. *Reproduction, Fertility and Development* **28(7)**, 948-56. (A1)

Li W., Zhang J., Lei A., Qian Y. (2010). Fusion expression of cecropin B in escherichia coli with lactose as the inducer. *Acta Agriculturae Boreali-Occidentalis Sinica*. 23(9) (A3)

Chen J., **Li W.**, & Qian Y. (2009). Recovery of vesicular stomatitis virus recombinants expressing E2 protein of CSFV. *Progress in Veterinary* **30(9)**, 7-10. (A3)

Presentation at international scientific congresses

Li W., Goossens K., Van Poucke M., Vandaele L., Van Soom A., & Peelman L. (2011) Expression of Endogenous Retrovirus (ERV) Families in Bovine Preimplantation Embryos. 3rd Mammalian Embryo Genomics meeting, 2011, Bonn Germany. (C3)

Li W., Goossens K., Forier K., Braeckmans K., Van Soom A., & Peelman L. (2013) High oxygen tension alters DNA methylation levels of bovine *in vitro* preimplantation embryos. 1st General Meeting of Epiconcept, Antalya Turkey. (C3)

Li W., Goossens K., Van Poucke M., Van Soom A., Peelman L. (2013). Reference gene election for normalization of qPCR data in sexed bovine preimplantation embryos under low and high oxygen. 4th Mammalian Embryo Genomics meeting, Quebec City Canada. (C3).

Acknowledgements

Finally, I am nearly at the end of this journey. This is the moment I was waiting and looking forward to for many years, and the moment I want to take the chance to express my deepest thank to these people for accompanying me on my way to the achievement of the degree of PhD.

First, I would like to thank my promoter Professor Luc Peelman, for accepting me as a research student, and the support and guidance throughout my PhD. I will never forget the sunny morning that you picked me up at Gent Sint-Pieter station. To my co-promoter Professor Ann Van Soom, thank you for leading me to the fascinating world of embryo research.

I would also like to give my gratitude to the exam committee members: Prof. Geert Opsomer, Prof. Peter Bols, Prof. Filip Van Nieuwerburgh, Prof. Ward De Spiegelaere, Dr. Karen Goossens and Dr. Bart Leemans. Thank you for critically reviewing and improving my thesis with the valuable comments.

I want to extend my warm thank to all my colleagues from LDG: Ruben, thank you for arranging my first accommodation and preparing a lamp with a Chinese character '安', helping me with all the administration, solving all my computer problems, driving me to the slaughterhouses, and so many other things. Karen, thank you for guiding my research in my first three years, and teaching me all the experimental techniques. Other members: Linda, Tim, Caro, Tiphanie, Domi, Mario, Bart, Jozefien, Alice, Sara, Philippe, Evy, Xiaoyuan, Cindy and Christel. Thank you for not only helping me at work, but also giving me so many fun memories after work, the birthday treatments, new year's calendars, quiz night, Christmas parties, chocolate workshop, beer festivals and so on. I would also like to thank everyone from LGT. To those who have left the lab: Oliwia, Marina, Sofie and Laetitia, thank you for being always friendly and always helping me whenever I need a hand. The 'fresh blood': Sean, Bregje, Francis, Hanne and Zifu, thank you for your generous encouragement.

I would also like to thank my colleagues from IVF lab: Petra, Isabel, Eline, Hilde, Sonia, Ruth, Bart, Nerea, Lynn, Maaike, and Feline. Thank you for the help with the experiments, puncturing, fertilizing, and SOFing. Special thanks go to my dear Sonia, Bart, Lynn and Nerea, thank you for your support and encouragement, and always being good friends to me.

I am also very grateful for the friendships I gained during my stay in Belgium. Dear Sara, thank you for not only being a good friend, but also providing a family to me in Belgium. I would like to thank your mom Gilberte as well. Thank you both for having me for Christmas dinners, taking me to the amazing hallerbos, and always think of me when you plan some fun things. My dear Marina, how can I thank you enough? Thank you for always there being supportive when I feel down and when I need help, and thank you for introducing me to work with great people at EcoTreasures. Dear Wei, you are not only working hard but also working smart. Thank you for inspiring me with the way you think and you do, and thank you for helping me develop the interests in cycling. Leilei, thank you for being a good friend that we could talk about anything, and thank you for your company in the last six months. This PhD is not easy, but you are almost there, my dear!

Also to my parents, thank you for giving me a carefree growing environment, and giving me the full freedom to explore the world.

My most warmly thanks are going to my beloved husband. Thank you for your patience, support and love during these years.

The journey of PhD will end here, but all the memories you gave me will be a lifetime treasure.

Wenwen

January 2017