

**EXPRESSION OF THE BOVINE DNA (CYTOSINE 5)  
METHYLTRANSFERASE FAMILY DURING PREIMPLANTATION  
DEVELOPMENT AND ABERRATIONS INDUCED BY SOMATIC CELL  
NUCLEAR TRANSFER**

A Dissertation

by

MICHAEL CAMERON GOLDING

Submitted to the Office of Graduate Studies of  
Texas A&M University  
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

December 2003

Major Subject: Veterinary Physiology

**EXPRESSION OF THE BOVINE DNA (CYTOSINE 5)  
METHYLTRANSFERASE FAMILY DURING PREIMPLANTATION  
DEVELOPMENT AND ABERRATIONS INDUCED BY SOMATIC CELL  
NUCLEAR TRANSFER**

A Dissertation

by

MICHAEL CAMERON GOLDING

Submitted to Texas A&M University  
in partial fulfillment of the requirements  
for the degree of

DOCTOR OF PHILOSOPHY

Approved as to style and content by:

---

Mark E. Westhusin  
Chair of Committee

---

Duane C. Kraemer  
(Member)

---

James N. Derr  
(Member)

---

Terry L. Thomas  
(Member)

---

Glen A. Laine  
(Head of Department)

December 2003

Major Subject: Veterinary Physiology

## **ABSTRACT**

Expression of the Bovine DNA (cytosine 5) Methyltransferase Family During  
Preimplantation Development and Aberrations Induced by Somatic Cell Nuclear  
Transfer. (December 2003)

Michael Cameron Golding, H.B.Sc., University of Western Ontario

Chair of Advisory Committee: Dr. Mark Westhusin

Bovine preimplantation embryos derived from nuclear transfer experiments exhibit a global state of genomic hypermethylation that likely account for the large number of developmental abnormalities observed to date. The central hypotheses of this work is that the genomic hypermethylation and improper epigenetic reprogramming reported in studies of bovine nuclear transfer, are in large part due to abnormal expression and regulation of the DNA methyltransferase proteins.

Bovine Dnmt mRNAs display strong sequence homology to those of human and mouse and similar to other species, exist as multiple isoforms. Two of these splice variants, which have been termed Dnmt2 $\gamma$  and Dnmt3a4 represent previously unreported sequence combinations. Investigation of bovine DNA methyltransferase expression in the bovine oocyte and early preimplantation development has revealed an intricate system divergent from observations previously reported in the mouse. Specifically, the somatic version of Dnmt1 along with Dnmt2, 3a and 3b are all expressed during these initial stages of bovine development. Further, real time analyses of the Dnmt transcripts

in cloned and *in vitro* produced embryos reveal significant differences in the mRNA expression levels of Dnmt1 and 2 but not Dnmt3a and 3b suggesting that the *de novo* methyltransferases may be functioning normally while Dnmt1 and Dnmt2 are aberrantly methylating the genome during a critical time when methylation levels should be receding. Real time PCR analysis of the Dnmt transcripts in fetal and adult tissues has revealed a developmental and tissue specific expression pattern suggesting that proper expression and function of these enzymes is a key element in the process of differentiation. These results are further supported by studies of Dnmt expression in aging bovine fibroblast cultures, which suggest that the Dnmts may play some as yet unidentified role in cellular senescence.

Recently, it has been postulated that the cause of abnormal methylation observed in cloned embryos may be due in part to misexpression of the Dnmt1 $\alpha$  isoform during preimplantation development. Work presented here raises new and significant hypotheses that must be considered both regarding the cadre of DNA methyltransferases that direct epigenetic programming during normal development and regarding the implication of abnormal DNMT expression in cloned embryos.

## **DEDICATION**

To my mother Barbara for her enthusiasm of life and love of academics and to my father David for his unyielding determination and direction. Constants in my life, without which, I would not be who I am today.

## **ACKNOWLEDGEMENTS**

My coming to graduate school would not have been, if it were not for the encouragement of Greg Kelly, Andrew Watson and Gerald Kidder. To this day they encourage and support my academics. I would like to thank Charles Long and Robert Burghardt for their direction, support and close friendship through my Ph.D. In addition to everyone in the Texas A&M Reproductive Sciences Lab, I would especially like to thank Mark Westhusin. His enthusiasm and creativity know no limits. Thanks for the opportunity.

## TABLE OF CONTENTS

	Page
ABSTRACT .....	iii
DEDICATION .....	v
ACKNOWLEDGEMENTS .....	vi
TABLE OF CONTENTS .....	vii
LIST OF FIGURES.....	x
LIST OF TABLES .....	xii
 CHAPTER	
I INTRODUCTION .....	1
Part 1 Transcription.....	4
Regulation of Eukaryotic Gene Expression .....	4
Transcriptional Cycle .....	5
Pre-initiation .....	6
Initiation .....	6
Part 2 Post-transcriptional Gene Regulation .....	12
Splicing.....	12
Alternative Splicing.....	14
Spliceosomal Enhancers.....	16
Spliceosomal Repressors.....	18
Post-transcriptional Gene Silencing.....	20
Part 3 Mammalian Preimplantation Development .....	22
Part 4 Epigenetics and Transcription .....	23
Part 5 Epigenetic Events of Early Development.....	26
Part 6 Epigenetic Events of Early Bovine Development .....	26
Part 7 The DNA (Cytosine 5) Methyltransferases .....	28
Part 8 Structural Domains of the Dnmts .....	31
Catalytic Domain and Mechanism of Action.....	31
Mechanism of Action of the Dnmt Inhibitor	
5-Azacytidine .....	33
Cysteine Rich Zinc Finger.....	34
Domains Unique to Dnmt1 .....	35
Domains Unique to Dnmt3a and Dnmt3b.....	36

CHAPTER	Page
Part 9 Regulation of Genomic Methylation by the Dnmts.....	38
Part 10 Nuclear Transfer .....	45
Part 11 Nuclear Transfer and Epigenetic Reprogramming .....	46
Part 12 Research Project Rationale .....	49
 II ANALYSIS OF DNA (CYTOSINE 5) METHYLTRANSFERASE mRNA SEQUENCE AND EXPRESSION IN BOVINE PREIMPLANTATION EMBRYOS, FETAL AND ADULT TISSUES .....	      53
Analysis of the Bovine DNA Methyltransferase mRNA Sequence .....	 53
Splice Variant Characterization .....	54
Dnmt mRNA Expression in Preimplantation Bovine Embryos .....	 57
Dnmt Real Time Quantification During Bovine Fetal Development .....	 62
Discussion .....	66
 III ANALYSIS OF DNA (CYTOSINE 5) METHYLTRANSFERASE mRNA EXPRESSION IN BOVINE PREIMPLANTATION EMBRYOS PRODUCED FROM <i>IN VITRO</i> FERTILIZATION, SOMATIC CELL NUCLEAR TRANSFER AND PARTHENOGENETIC ACTIVATION.....	      69
Real Time Analysis of Bovine Preimplantation Embryos Produced by <i>in vitro</i> Fertilization, Nuclear Transfer and Parthenogenetic Activation .....	  70
Real Time Analysis of Cultured Donor Fibroblast Cells .....	77
Discussion .....	79
Materials and Methods .....	83
Tissue Collection and RNA Isolation.....	83
Embryo Production and RNA Isolation .....	83
Ribonucleoprotein Fractionation.....	84
Nuclear Transfer.....	85
Reverse Transcriptase Polymerase Chain Reaction .....	86
5' Rapid Amplification of cDNA Ends.....	87
Real Time PCR.....	88

CHAPTER	Page
IV SMALL HAIRPIN RNA MEDIATED KNOCKDOWN OF BOVINE DNA METHYLTRANSFERASE 1: ESTABLISHMENT OF SHRNA MEDIATED FUNCTIONAL GENOMICS IN THE BOVINE MODEL .....	89
RNA Interference in Bovine Cells .....	89
Sequencing of Bovine Dicer Argonaute 2 and Argonaute3 (Ago 2 & 3) .....	89
RNAi Knockdown of GFP in Bovine Embryos .....	91
shRNA Mediated Knockdown of the Prion Protein (PrP) - Methods for RNAi Mediated Gene Silencing in an Agricultural Model .....	93
Large-Scale Design and Synthesis of shRNAs .....	94
<i>In vitro</i> Luciferase Assay - Ligation Independent Cloning of a Reporter Construct .....	96
<i>In vitro</i> Knockdown of the Luciferase-PrP Reporter .....	98
Lenti-viral Mediated Transgenesis .....	101
Discussion .....	102
shRNA Mediated Knockdown of Dnmt1 .....	103
Discussion .....	106
Materials and Methods .....	106
Design and Cloning of Small Hairpin RNAs .....	106
Ligation Independent Cloning of Reporter Constructs ....	108
Cloning of Tagged Proteins .....	109
Cell Culture and Transfection .....	110
Western Blotting .....	110
Luciferase Reporter Assay .....	111
V SUMMARY .....	112
REFERENCES .....	113
APPENDIX A BOVINE DNMT mRNA SEQUENCES AND SEQUENCE ALIGNMENTS .....	126
VITA .....	146

## LIST OF FIGURES

FIGURE	Page
1 Central dogma of biology.....	5
2 Epigenetic events of bovine preimplantation development .....	27
3 Schematic diagram representing the chemical structures of cytosine .....	34
4 Domain structures of the bovine DNA methyltransferase family .....	54
5 Isoform analysis of bovine Dnmt3a .....	55
6 Isoform analysis of bovine Dnmt3b .....	56
7 RT-PCR analysis demonstrating the recruitment of the Dnmt1, Dnmt3a and Dnmt3b mRNAs onto the polysome during the 8-cell to 16-cell transition .....	59
8 RT-PCR analysis demonstrating the recruitment of the Dnmt1, Dnmt2, Dnmt3a and Dnmt3b mRNAs onto the polysome during the 4-cell stage of bovine preimplantation development .....	60
9 RT-PCR analysis of the four known Dnmt gene transcripts during all stages of bovine <i>IVF</i> preimplantation development .....	61
10 Bovine Dnmt2 $\gamma$ .....	63
11 Real time PCR analysis of bovine DNA methyltransferase expression in fetal and adult tissues .....	64
12 RT-PCR analysis of the four known Dnmt gene transcripts during all stages of preimplantation development in embryos reconstructed by somatic cell nuclear transfer .....	71
13 Real time PCR analysis of DNA methyltransferase expression during the preimplantation development of IVF, NT and Parthenote bovine embryos .....	73

FIGURE	Page
14 Comparison of DNA methyltransferase expression over IVF, NT and Parthenote preimplantation development.....	76
15 Real time PCR analysis of bovine DNA methyltransferase expression in the donor cell fibroblasts, cultured for an extended period of time .....	78
16 RT-PCR amplification of the full-length coding sequence for bovine Dicer .....	90
17 RNAi based suppression of GFP expression in bovine embryos .....	92
18 Experimental model .....	97
19 LIC Procedure .....	100
20 Inhibition of Luciferase activity using shRNAs targeting a Luciferase-PrP transcript.....	102
21 Inhibition of Luciferase activity using shRNAs targeting a Luciferase-Dnmt1 target transcript .....	104
22 shRNA mediated knock down of an amino terminal, T7 tagged Dnmt1 protein .....	105

**LIST OF TABLES**

TABLE	Page
1 RT-PCR Primer Sequences .....	87
2 Sequences of Primers Used for Ligation Independent Cloning .....	108

## CHAPTER I

### INTRODUCTION

It is fitting that a work examining the epigenetic events of early bovine development begin with a brief description of the history of this science and its far-reaching impact. Although description of epigenetic phenomena can be identified much earlier, it was the work done on maize by Barbara McClintock that began to crystallize epigenetics into a scientific discipline. Her description of transposable “controlling” elements and their capacity to silence genes based on the proximity of their specific site of integration, suggested that the genome contains transcriptionally permissive and suppressed regions and that their allocation is dynamic.

Some system of gene regulation must be present that is able to order the action of genes in such a manner that these patterns will be produced. Until these problems find some adequate solution, our understanding and our experimental approach to many phenomena will remain obscured. (Cold Spring Harbor Symposium 1951 – Comfort 1999)

This was the first experimental evidence to suggest that genomic loci have the ability to shift between differing transcriptional capacities. Soon after this work, a similar phenomenon was identified in *Drosophila* eye development, which has since been referred to as position effect variegation. Here it was recognized that genes in close proximity to heterochromatic regions of the genome could be silenced, owing not to any

---

This dissertation follows the style of *Genes to Cells*.

gene specific regulatory mechanism but simply due to their proximity to densely packaged DNA. Recognition of this phenomenon and its potential role in regulation of gene expression served as a catalyst for other experiments to demonstrate how widespread and fundamental this epigenetic phenomenon was to the cell's ability to control transcription. However, it was several years before a biochemical mechanism could be envisioned that would allow the cell to dynamically control the establishment and modulation of the epigenetic marks that impart this large-scale control of gene expression. Today, numerous examples of epigenetic gene regulation have been found in virtually all organisms studied and it is likely that we have only just begun to comprehend the breadth and significance of these phenomena.

The two major mechanisms identified to date that impart epigenetic control of mammalian gene expression are DNA methylation and post-translational histone modifications. Both work in concert to provide a general mechanism by which the differing transcriptional states of chromatin are achieved. The modifications to DNA or to the chromatin in which it is packaged serve to compartmentalize the genome into other areas, which are accessible to the transcriptional machinery necessary for gene expression and into other areas, which are not. This compartmentalization is much more efficient and secure than relying on *trans* regulatory factors to control the transcriptional activity of the entire genome. Each mammalian cell has a very specific transcriptional program owing to its function within the specific cell group, tissue and organ system that it is located. Epigenetic marks serve as a fundamental basis for this varying tissue and developmental specific expression pattern. Thus the mechanisms responsible for

generating these patterns have significant consequences on not only the physiology of the cell, but also on the development and metabolism of the entire organism. Moreover, there is evidence to suggest that aberrant epigenetic programming during early mammalian development results in a large number of developmental abnormalities (Young *et al.* 1998; Hill *et al.* 2000a; 2000b; Sinclair *et al.* 2000), in addition to predisposing the organism to early onset of a variety of diseases later in life (Barker, 1990).

This work focuses on the epigenetic events of early bovine development. A large number of studies of bovine embryos produced by somatic cell nuclear transfer have identified abnormalities in the transcriptional control of numerous, seemingly random genes. These large-scale transcriptional disturbances are thought to be the result of abnormal DNA methylation and thus, this aberrant methylation has been hypothesized to be the leading cause of developmental failure of cloned animals. Epigenetic defects observed in these studies are similar on the whole to abnormalities reported in Angelman, Beckwith-Wiedemann, and Prader-Willi syndromes, which have been associated with human assisted reproductive technologies. In order to thoroughly discuss the role of DNA methylation in epigenetic control of gene expression during mammalian embryonic development, a brief review of the enzymatic mechanisms of transcription will be given followed by a detailed discussion of the enzymes that impart genomic methylation, the DNA methyltransferases. This will be followed by a discussion of the function of these enzymes during early mammalian development and their relation to studies of somatic cell nuclear transfer.

## Part 1 Transcription

### Regulation of Eukaryotic Gene Expression

Mammalian biology, be it normal development or disease status, is the sum total of the dynamic regulation of genes encoded by the genome and production of the resultant proteins. Differential patterns of gene expression determine the structural and functional phenotype of the cells, which in turn directly affect the physiology of the organism. Mechanisms that govern which sets of genes are turned on, and those, which are turned off, are fundamental to the processes of development and differentiation. Mechanisms, which modulate gene expression, are diverse and occur at various points in the pathway from gene to protein. This medium can be subdivided into three major areas, transcriptional control of gene expression, post-transcriptional gene regulation and translational control of gene expression. Accordingly, each of the areas listed can further be divided into multiple levels that collectively serve to enhance the ability of the cell to regulate the spectrum of gene expression in a developmentally and tissue specific manner.

The first level of control is transcriptional. It is on this level that the vast majority of control is asserted through the ordered structure of chromatin, DNA modifications, and the assembly and interaction of *trans* components on a multitude of *cis* DNA regulatory elements. Post-transcriptional gene regulation serves to modulate / modify the RNA molecule so as to affect both its physical structure and its capacity to interact with the translational machinery. Translational control of gene expression is the final level of control and centers on the dynamics of the conversion of mRNA to protein.

All of these elements form a functional hierarchy that permits regulation at numerous points thus enabling the cell to quickly respond to the environment or to developmental queues using generic cellular factors modulated in such a way as to give rise to a specific response (Figure 1).

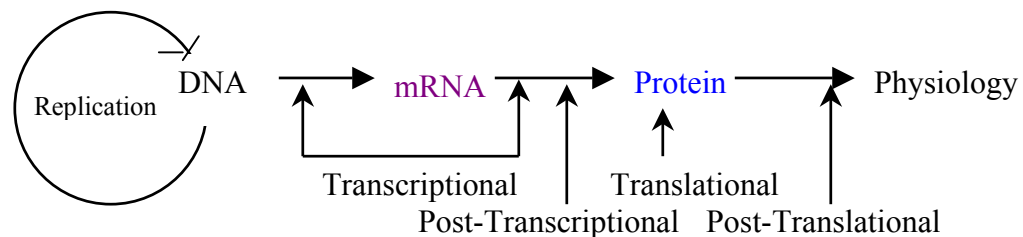


Figure 1 Central dogma of biology. A schematic diagram of the progression of gene to protein response, with the differing control mechanisms applicable to that level listed below.

### Transcriptional Cycle

The majority of our present understanding of how genes are transcriptionally regulated comes from François Jacob and Jacques Monod's work on the bacterial operon. Since these initial studies, it has become evident that the largest degree of control over gene expression occurs at the level of the enzymatic process of transcription itself (Levine and Tjian 2003). The initiation of this process is a key regulatory event and an enormous diversity of proteins and regulatory elements within the DNA sequence cooperate to regulate the dynamics of its initiation and the subsequent production of the nascent RNA transcript. Transcription by the eukaryotic RNA polymerase II is a

multistage process that requires the interaction of hundreds of proteins with varying degrees of post-translational modification. The process is collectively referred to as the transcriptional cycle and occurs sequentially in five different stages: preinitiation, initiation, promoter clearance, elongation and termination. Each stage of the transcription cycle is subject to regulation and thus an enormous diversity of regulatory factors contribute to the production of a single RNA molecule (Ogbourne and Antalis 1998; Dvir *et al.* 2001; Beckett 2001; Shilatifard 1998).

#### Pre-initiation

This stage is defined by the local alteration of DNA via the interaction of regulatory proteins that serve to permit access of the transcriptional machinery to the promoter. This phase is also called the activation phase as regulatory sequences within the promoter are then able to interact with and activate the assembly of their appropriate trans-regulators (Dvir *et al.* 2001; Shilatifard 1998).

#### Initiation

The regulatory factors that function at this stage all serve to position RNA polymerase II (polII) on the promoter and to initiate RNA synthesis. The complete RNA polymerase holoenzyme is a huge complex with a molecular mass in excess of 2500 kilodaltons. RNA pol II is completely dependent on auxiliary factors to initiate transcription. The basal proteins required to initiate transcription are called general transcription factors (GTFs) and these proteins form an ordered complex at the promoter in a regulated and defined order. Assembly of these factors serves as a platform for the recruitment of RNA polymerase and the kinetic events that initiate its function. In most

eukaryotic promoters an element referred to as the TATA box is first recognized by the TATA binding protein. The TATA box is a DNA regulatory element with the consensus sequence 5'-TATAAAA-3' that is located approximately 25 base pairs upstream of the transcriptional start site and is surrounded by GC rich sequences which are subject to regulation through DNA methylation (Ogbourne and Antalis 1998). Activation of a promoter exposes the TATA box and permits binding of the TATA binding protein and thus initiation of the GTF assembly. A group of proteins collectively called TATA-binding associated factors (TAFs) then bind the TATA binding protein and make a complex known as transcription factor IID (TFIID). A central component of TFIID is TAF250. This protein specifically binds acetylated lysines in active euchromatin thus stabilizing the TFIID complex and also seems to have some inherent acetylating activity of its own (Ogbourne and Antalis 1998; Dvir *et al.* 2001; Shilatifard 1998). Thus, exposure of the TATA box and binding of TFIID may be sufficient to induce chromatin priming, at least in the vicinity of the promoter. Binding of TFIID induces a 90° bend in the DNA centered at the TATA box which permits binding of the additional general transcription factors (Beckett 2001).

The assembly of the general transcription machinery continues with recruitment of TFIIA which complexes with TFIID, inducing a conformational change that permits binding of the second GTF. Binding of TFIIB serves as an adaptor molecule that recruits a preassembled TFIIF-RNA polymerase to dock on the assembling protein complex (Ogbourne and Antalis 1998; Dvir *et al.* 2001). The polymerase binds with a large unphosphorylated carboxy-terminal domain. TFIIE serves as an adaptor recruiting

TFIIF and TFIID to the carboxy terminal domain of RNA polymerase. Entry of TFIID into the polymerase complex tightens the DNA around the holoenzyme and induces the unwinding of a short stretch of DNA near the transcriptional start site. TFIIF serves to stabilize the interactions between these large protein subunits and is required for initiation to proceed. At this point, initiation is technically complete, however before transcription can proceed, the carboxy-terminal domain of polymerase must be phosphorylated to permit promoter clearance. It is phosphorylation of this carboxy-terminal domain that imparts the largest degree of regulation over transcriptional initiation. Phosphorylation of the carboxy terminal domain induces a conformational change that promotes the oligomerization of nucleotides within the active site of RNA polymerase. Numerous proteins working in concert catalyze the phosphorylation of this C-terminal domain. TFIID and positive transcription elongation factor (P-TEFb) are two such proteins that display very strong pol II C-terminal domain phosphorylating activity and thus regulate the progression of transcription initiation (Ogbourne and Antalis 1998; Dvir *et al.* 2001; Beckett 2001; Shilatifard 1998).

Some mRNAs have variable 5' untranslated regions indicating that their transcription can begin at multiple sites over a large region; 20 – 200 base pairs in size. These genes often encode proteins involved in intermediary metabolism and are transcribed at relatively low rates (Beckett 2001). The control region for most genes of this type does not contain a classical TATA box initiator site but instead contains a stretch of 20-50 nucleotides composed almost exclusively of cytosine and guanosine. As a dinucleotide, CG is statistically underrepresented in vertebrate genomes and thus the

presence of these long CG repeats just upstream of transcription initiation sites is not a random phenomenon. These CG rich regions are referred to as “CpG islands” given that they occur sporadically in a “sea” of DNA low in this specific repeat. These CpG islands are the binding sites for the SP1 and SP2 transcription factors and thus transcriptional initiation is dependant on the binding of these transcription factors. Sp1 can bind at either an element known as the ETS motif (5'-GGCTTCCTGTCT-3') or another element known as the pyrimidine rich initiator motif (5'-CTCANTCT-3'). Binding of Sp1 to these elements facilitates the stabilization and assembly of the general transcription factors and allows transcription to initiate despite the absence of a defined TATA box (Beckett 2001).

Assembly of the preinitiation complex and its subsequent phosphorylation induced activity are regulated by a host of protein-protein interactions between members of the complex and additional proteins that serve to alter the stability of the complex and its binding to the DNA helix. In fact, regulated assembly is a central component to the control of transcriptional initiation (Ogbourne and Antalis 1998; Dvir *et al.* 2001; Beckett 2001; Shilatifard 1998). Tissue specific and developmental regulation of transcription can thus be achieved via altering the capacity of the GTF to assemble at the promoter. Binding of inhibitors and enhancers as well as the post-translational modification of the GTFs serves to collectively regulate this process. For example, the NtrC transcriptional activator in *E.coli* is a transcription factor that assembles only after phosphorylation. This post-translational modification results in the formation of stable DNA-protein complexes linking two regulatory DNA elements. An additional example

would be the regulation of the SMAD proteins by cellular signaling events.

Phosphorylation of the SMAD proteins results in altered properties of association in that the tendency to form homodimers is shifted to form heterodimers with other members of the SMAD family (Beckett 2001).

Repressors generally alter the binding affinity of the GTFs and thus decrease the rate initiator complex assembly but many have also been identified which can bind to the forming initiation complex and act as a steric block to its formation. An example of this interaction is given by the transcription factor that regulates biotin production in bacteria. This protein is found bound to either a repressor or an enhancer protein.

Repressor binding induces a disorganization of the DNA binding loops as where binding of the enhancer ligand induces a conformational change where the binding loops align to allow DNA-protein complex formation (Shilatifard 1998).

Binding of transcription factors is generally thought to stabilize or enhance the assembly of the initiator complex through protein-protein interactions or delivery of a (or portions of) preassembled initiation complex. Transcription factors may also act by inducing the correct conformation in the DNA helix so as to permit either initiator complex assembly or transcription elongation. Generally speaking though, transcription factors can be thought of as molecules, which induce the appropriate conditions in the DNA, that allow it to serve as a platform for complex assembly (Shilatifard 1998).

A large percentage of vertebrate genes have variable 5' untranslated regions given that their transcription can begin at multiple sites over a large region; 20 – 200 base pairs in size. Often, these genes encode proteins involved in intermediary

metabolism and are transcribed at relatively low rates. The control region for most genes of this type do not contain a classical TATA box initiator site but rather do contain a stretch of 20-50 nucleotides composed almost exclusively of cytosine and guanine. As a dinucleotide, cytosine-guanine (CpG) is statistically underrepresented in vertebrate genomes and thus the presence of long CpG repeats just upstream of transcription initiation sites is not a random phenomenon. These CG rich regions are referred to as “CpG islands” given that they occur sporadically in a “sea” of DNA low in this specific repeat. These CpG islands are the binding sites for the SP1 transcription factor and thus the vast majority of genes following this paradigm are responsive to SP1 initiation. However, CpG islands are also found in the regulatory regions of many if not most other genes and are the sites of a unique regulatory phenomenon, DNA methylation (Bestor 2000).

Methylation of the number five position of cytosine in these CpG islands produces a localized conformational and electrostatic change in the DNA double helix that influences a diverse number of biological processes. These methylated regions have been implicated in transcriptional regulation both in a global and tissue specific manner, X-chromosome inactivation, genomic imprinting, silencing viral retrotransposons, as a mechanism for monitoring cellular ageing and in neoplastic transformation. Disruption or massive alteration of these carefully controlled methylation patterns are incompatible with normal growth and development (Bestor 2000; Li *et al.* 1992).

## Part 2 Post-transcriptional Gene Regulation

While gene silencing mediated by DNA methylation exerts its affect on the level of transcription initiation, it is necessary to review two specific elements of post-transcriptional gene regulation as they pertain to DNA methyltransferase function. These two elements are alternative splicing and post-transcriptional gene silencing or RNA interference (RNAi). Several mechanisms of post-transcriptional gene regulation have been identified in eukaryotic cells and even more have been suggested by recent experimental data. Regulation of RNA stability, regulation of secondary structure, splicing (“normal intronic”, alternative splicing and intergenic splicing), poly(A) tailing, termination, RNA editing, RNA trafficking mRNA localization and post-transcriptional gene silencing represent the major mechanisms identified to date (Akker *et al.* 2001). These methods are diverse in their specific mechanism of action however, they all serve to chemically modify, trim or rearrange the RNA transcript to produce new exon arrangements / translational boundaries that can result in the generation of multiple protein species from a single gene. The end result provides a level of gene regulation beyond the level of transcriptional initiation and thus expands the capacity of the cell to modulate gene expression.

### Splicing

Splicing is the process of removing introns and joining exons to create a coherent coding sequence. Numerous elements common to nearly all eukaryotes work in concert to achieve gene splicing. Splicing occurs via two sequential trans-esterification reactions where one ester bond is exchanged for another. The first reaction forms a

lariat structure where the 5' guanine of the intron is joined in a 2' – 5' phosphodiester bond to an adenine near the 3' end of the intron to produce a “branch point”. Over 100 proteins have been identified in association with the process of splicing and these various proteins are found in differing concentrations depending on developmental stage and cell type thus providing the basis for prospective gene specific regulation of this process. Uricil rich, ribonucleoparticles 1 to 6 (U1,U2,U4,U5 and U6) are some of the major components that participate in the formation of the splicosome and mediate RNA splicing. U1 binds to the universal splice site (5' - GAGGUAAGU - 3') located on the 3' side of the exon-intron boundary through a complimentary sequence in the 5' end of the U1 snRNA. U2 binds the upstream pyrimidine rich site near the 3' end of the intron through the targeted binding action of the U2 snRNP auxiliary factor (U2AF). Docking of the U2 protein with U2AF induces a buldge in the U2 protein allowing the 2' hydroxyl to participate in the first trans-esterification reaction and formation of the lariate structure. U4 and U6 pair up and bind the intron in a sequence independent manor allowing U5 to associate and complete the formation of the splicosome. Spliceosome assembly occurs in a highly ordered and stepwise fashion, upon which U1 and U5 are released via a rearrangement that moves the splice sites into close proximity and allows the second trans-esterification reaction to be completed. Upon completion of the second trans-esterification reaction, the two exons are now joined in frame and the intron is released, still complexed with U2, U4 and U6 (Akker *et al.* 2001; Caceres and Kornblihtt 2002; Lopez 1998; Zhao *et al.* 1999).

The protein machinery necessary to carry out RNA splicing are physically associated with the nuclear matrix, thus splicing occurs in compartmentalized regions and localization of specific factors to these regions contributes to the overall regulation of the process. As well, different splice sites within introns have varying amounts of strength in their capacity to recruit the spliceosomal machinery. Some introns in fact, contain very weak splice sites that require the activity of accessory proteins to activate splicing in these regions. The translational capacity of the transcript can thus be regulated by its propensity to attract these splicing factors, as unspliced RNA molecules are not exported from the nucleus (Lopez 1998; Zhao *et al.* 1999). A wide variety of splicosomal enhancer proteins have been found and shown to mediate these interactions (Caceres and Kornblihtt 2002; Lopez 1998; Zhao *et al.* 1999).

#### Alternative Splicing

With the sequencing of the human genome it has become apparent that the complexity and sheer diversity evident in the proteome cannot be attributed to the limited number of genes identified. This observation highlights the importance of post-transcriptional methods of gene regulation, which are now hypothesized to be crucial to generation of the observed protein diversity. Alternative splicing is a method of generating alternative exon combinations within a single RNA by utilizing alternative 5' splice sites, alternative 3' splice sites, optional exons, mutually exclusive exons, retained introns and alternative poly(A) tail splice sites. All of these serve to change the coding sequence to allow the generation of multiple protein domain combinations from a single primary RNA transcript. A large number of *cis* regulatory elements and *trans*-splicing

factors modulate the alternative selection and omission of exons from specific pre-mRNAs. These methods are diverse in their specific mechanism of action however, they all serve to chemically modify, trim or rearrange the nascent RNA (nRNA) transcript to produce new exon arrangements / translational boundaries that can result in the generation of multiple protein species from a single gene (Akker *et al.* 2001; Caceres and Kornblihtt 2002; Lopez 1998; Zhao *et al.* 1999). The mechanisms by which these heterogeneous RNA transcripts are produced is an expanding area of research and abnormalities identified at many of the steps involved are now being correlated with disease phenotypes.

Since the sequencing of the human genome it has been conservatively estimated that 60% of the genes identified are alternatively spliced (International Human Genome Sequencing Consortium 2001). The methods that mediate alternative exon selection are diverse and many are subject to regulation by extra-cellular signaling pathways. However, a significant portion of alternative splicing can be accounted for by the variations in strengths between 5' and 3' splice sites within a single intron and their relation to the strength of surrounding splice sites. Enhancing or repressing the relative strength of splice sites is the major mechanism by which alternative splicing is asserted. A simple mechanism where by the strong 3' splice site of the downstream exon competes effectively with the weak 3' splice site of the upstream intron can result in exon exclusion. Repetitive di- or tri-nucleotide sequences within or in close to these splicing regulatory sequences strongly influence their strength. Cystic fibrosis, and myotonic dystrophy are all classic examples of how repeats located in intergenic regions

can negatively influence splicing and their resultant protein products (Akker *et al.* 2001; Caceres and Kornblihtt 2002; Lopez 1998; Zhao *et al.* 1999).

Alternative splicing is a highly regulated process where patterns of exon splicing can be modulated in response to developmental and physiological signals, however the majority of splicing events seem to be constitutive, with mRNA variants coexisting at constant ratios cell to cell. This consistency is a reflection of the generic nature of the spliceosome and the involvement of the ubiquitous splicing factors in the majority of alternative splicing reactions. However several trans acting proteins attenuate the recognition of the correct splice sites involved by either acting as enhancers or repressors influencing the coordinated selection of the 5' and 3' splice sites across an exon. Several proteins have been identified in this capacity and more are emerging as our understanding of the molecular nature of splicing increases. Thus, the decisions that govern the alternative splicing of a mRNA can be attributed to competition between and among potential splice sites and as such, any mechanism that alters the relative rate of selection for a splice site serves to regulate the selection or omission of that exon. The proteins that govern the selection or repression of a splice site can basically be separated into two classes, splicing enhancers and splicing repressors (Akker *et al.* 2001; Caceres and Kornblihtt 2002; Lopez 1998; Zhao *et al.* 1999).

#### Spliceosomal Enhancers

The SR family of proteins are a group of eight (or more) proteins characterized by RRM type RNA binding domains in their carboxy-terminal domains and serine-arginine repeats in the amino-terminal domains that mediate protein-protein interactions

with various components of the spliceosomal machinery. Members of this subclass of splicing proteins function via the recognition of exonic splicing enhancer elements and lead to the activation of weak adjacent 3' splice sites. Two SR proteins, termed splicing factor 2 and alternative splicing factor (SF2 and ASF) are two of the main members of this family involved in mammalian alternative splicing. The concentrations of both of these factors and their ratios with respect to each other are important factors in determining the combination of exons during pre-mRNA processing of alternatively spliced genes.

Spliceosomal enhancers function by enhancing U1 and U2AF splice site recognition as well as by stabilizing their binding by forming a protein bridge between the 5' and 3' splice site complexes. The strength of splice site recognition by U1 and U2AF is relative to the similarity of the *cis* splice site to the consensus sequences listed above. Binding of spliceosomal enhancers not only enhances binding of U1 and U2AF, but it also serves to stabilize the interaction between these two components and serves to recruit other spliceosomal components. Several enhancer proteins bind to purine rich elements within the exons of the regulated gene and promote the use of a proximal splice site, whereas others are pyrimidine rich and are located in intergenic regions. The splice sites of RNA molecules subject to alternative splicing often exhibit poor matches to the consensus sequences, or their recognition is hampered by the secondary structure of the proximal region. Additionally, the exon itself may be too large for the protein bridge between U1 and U2AF to form. Spliceosomal enhancer proteins alleviate these impediments by enhancing U1/U2AF binding, removing secondary structure, or by

facilitating the formation of a bridge between U1-U2AF directly. The molar amount and activity of these enhancer proteins is different from cell to cell and changes over development. Thus, spliceosomal enhancer proteins can determine which protein isoforms are present in which cell and at what time (Akker *et al.* 2001; Cáceres and Kornblihtt 2002; López 1998; Zhao *et al.* 1999).

A well-defined example of SR protein mediated alternative splicing is found in studies of *Drosophila* sex determination. There are three main genes involved in sex determination of the fly: sex lethal (*Sxl*) doublesex (*Dsx*) and transformer (*Tra*). Each of these genes produces a pre-mRNA that has two possible splicing patterns, depending upon the sex of the fly. In the male, mRNA production proceeds as normal, which causes the inclusion of two exons in *Sxl* and *tra* that produce mRNAs, which have “premature” stop codons and yield inactive proteins. In females, the *Dsx* protein, a member of the SR family of alternative splicing factors binds the mRNA and induces the alternative splicing of this RNA such that the exons containing the stop codons of *Sxl* and *tra*, are skipped thus producing functional protein products (López 1998; Zhao *et al.* 1999).

### Spliceosomal Repressors

Splice site selection can be blocked by inhibitory *cis* elements that induce the formation of secondary structure, recruiting repressor proteins that sterically block access of U1/U2AF to the splice site or by providing a competitive site for spliceosomal factor binding, thus sequestering the component from the splice site. All of these mechanisms impart a repressive affect upon the selection of the proximal splice site and

thus promote its exclusion from the processed mRNA. In contrast with the SR group of proteins, the hnRNP family of proteins appears to mediate their affect by interfering with U1/U2AF binding to the splice site of the intron and thus shifting the spliceosome to a distal splice site. The hnRP family binds pyrimidine rich regions and are also known as pyrimidine tract binding proteins (PTB). The relative abundance of the SR and hnRNP proteins varies between cell types and developmental stages leading to the production of cell type and developmental specific transcripts. In addition, the activities of these proteins are affected by post-translational phosphorylation, linking alternative splicing to environmental induced or developmental extra cellular signals. The *Drosophila* SXL protein binds pyrimidine rich regions of introns and blocks access of the U2AF protein to the 3' splice site and thus promotes the use of a weaker downstream splice site. SXL is a very strong repressor that physically blocks access of U2AF and can displace molecules already bound to the 3' splice site. Other repressor proteins function either by recruiting U1 to sub-optimal splice sites or by directly binding and blocking the action of spliceosomal components. In summary, spliceosomal repressors function by interfering with spliceosome assembly or by reducing the efficiency of spliceosome assembly such that selection of a downstream splice site is preferred (Lopez 1998; Zhao *et al.* 1999).

The molar range of enhancer and repressor proteins varies over a range of 100 fold between different tissues in the adult rat (Hanamura *et al.* 2002). Variations in the relative concentration or activities of competing and cooperative factors and the strength of their target *cis* elements to recruit them all serve to regulate the spectrum of

alternative splicing for a given cell (Akker *et al.* 2001; Caceres and Kornblihtt 2002; Lopez 1998; Zhao *et al.* 1999).

#### Post-transcriptional Gene Silencing

RNA interference (RNAi) is an evolutionary conserved process, which leads to posttranscriptional suppression of gene expression. Among other functions, it is thought to serve as a natural defense mechanism against a variety of microorganisms, including viruses (Denli and Hannon 2003; Hannon 2002). RNAi was first discovered as a result of experiments on the flat worm *Caenorhabditis elegans*, which demonstrated sequence-specific gene silencing in response to double-stranded RNA (dsRNA). Numerous studies have been performed since, which clearly demonstrate RNAi is operational in many, if not most, eukaryotic organisms (Hannon 2002).

In brief, RNAi seems to involve a two-step process. In the first step, dsRNA is recognized by an RNase III family nuclease termed Dicer. This enzyme cleaves dsRNA into small interfering RNAs (siRNA) containing 21 – 29 nucleotides (Bernstein 2001). The siRNAs are then incorporated into a multicomponent nuclease complex, RNA-induced silencing complex (RISC), which targets specific mRNAs for destruction based on their homology to the siRNA (Denli 2003). The process begins with two anti-parallel dimmers of the Dicer protein utilizing ATP to cleave the long double stranded RNA into oligonucleotides 21-29 base pairs in length. All vertebrates studied today contain a homologue of Dicer and numerous Dicer homologues have been identified in *Arabidopsis* (Bernstein *et al.* 2001). These short interfering RNAs (siRNA) serve as a trigger eliciting a response that culminates in the sequence directed destruction of RNA

molecules containing homologous regions to the siRNA (Hannon 2002). Upon cleavage by Dicer, these RNAs can enter one of three possible pathways. In the first, the siRNA is directed to bind the 3' untranslated region of target mRNAs and inhibit their translation at the level of protein synthesis. The mechanism by which this phenomenon occurs is still unknown as is the reason behind the lack of a requirement for exact complementarity between the siRNA and the target 3' UTR. This pathway has been termed the miRNA pathway as the precursor molecules that initiate it are formed from short 70mers, which are processed by Dicer in the nucleus (Hannon 2002).

The second mechanism by which RNAi directed gene silencing can occur is via the selective destruction of mRNAs complementary to the siRNA which occurs in the cytoplasm. This process is mediated by the RNA induced silencing complex (RISC). This protein complex is formed of various proteins in different combinations depending on species and cell type (Denli and Hannon 2003). The first protein to be identified was Argonaute II, which is a member of highly basic proteins linked to studies of RNAi in plants. Mutations of these proteins result in abnormal shoot apical meristem development that resembles squid like tentacles, *ergo* the French word for squid (Argonaute) was used to describe their appearance. Since the identification of the first Ago proteins, numerous mammalian homologues have been identified including several germ cell specific forms. These proteins can associate with either the siRNA or the miRNA triggers. Upon purification of RISC, several other proteins were identified and many have distinct correlations with human disease. These include the Fragile X Mental Retardation Protein (FMRP), gemin-3 a DEAD Box containing protein and the Tudor

protein, which contains the RNase component of the RISC holoenzyme. The specific interactions that control RISC assembly and function have not been established but it is clear from the wide variety of proteins involved and the numerous isoforms and tissue specific homologues found that this is a highly regulated process. It is likely that similar to transcriptional initiation, the regulated assembly of the RISC complex is subjected to combinatorial control based on the abundance and form of the different RISC components (Bernstein *et al.* 2001; Denli and Hannon 2003; Hannon 2002).

The third mechanism by which RNA mediated post-transcriptional gene silencing mediates its effect is by RNA directed DNA methylation. The first suggestion that this phenomenon existed was the observation that in plants, PTGS was heritable. Later it was found that in yeast, several components of the RNA machinery are required for centromeric silencing (Denli and Hannon 2003). Finally, an experiment where double stranded RNA homologous to the promoter of a gene was able to elicit gene silencing revealed that the mechanism by which this effect is asserted is through DNA methylation. The N-terminal domains of the DNA methyltransferases contain numerous highly conserved protein domains that are separate from their catalytic methyltransferase domain. It is possible that the RNAi directed chromosomal silencing utilizes a specific domain on the Dnmts to carry out this effect, however such a link has not been definitively identified.

### **Part 3 Mammalian Preimplantation Development**

Development of the fertilized zygote through several morphologic changes, ultimately forming a blastocyst occurs via the execution of a preprogrammed

developmental plan that results in the differentiation of two separate cell lineages, one forming the placenta and the other, which will give rise to the embryo proper (Camous *et al.* 1986; Frei *et al.* 1989). In all species, development beyond these early cleavage divisions is completely dependent upon the switch from a reliance on maternal stores of mRNA within the oocyte to transcription of the zygotic genome; a process termed zygotic genome activation (ZGA) (Camous *et al.* 1986; Frei *et al.* 1989; Kopecny *et al.* 1989).

In bovine embryos, zygotic gene activation occurs by the 8-16 cell stage (Camous *et al.* 1986; Frei *et al.* 1989; Kopecny *et al.* 1989) although several studies have revealed transcriptional activity earlier in development (Plante *et al.* 1994; Viuff *et al.* 1996). However, bovine development can proceed only to the 8-cell stage in the presence of a transcriptional inhibitor suggesting that progression beyond this stage is dependent upon gene products derived from the embryo's genome (Liu and Foote, 1997). These findings follow similar observations in other species, including humans (reviewed in Telford *et al.* 1990). During these early stages the mechanisms that govern which sets of genes are turned on and those, which are turned off, are absolutely essential to the processes of development and differentiation. Fundamental to these precise mechanisms of control is the phenomenon of epigenetics.

#### **Part 4 Epigenetics and Transcription**

Mammalian biology, be it normal development or disease status is the sum total of the dynamic regulation of genes encoded by the genome and production of the resultant proteins. Differential patterns of gene expression determine the structural and

functional phenotype of the cell, which in turn directly affect the physiology of the organism. Mechanisms that govern which sets of genes are turned on and those, which are turned off are fundamental to the processes of development and differentiation. The oocyte and early embryo contain a specific preprogrammed developmental plan that once initiated by the process of fertilization, sets in motion the machinery that will build a complete and independent organism from a single cell. Fundamental to this process is the phenomenon of epigenetics.

Epigenetic refers to differential patterns of gene expression based solely on the local physical and biochemical properties of chromatin without a change in DNA sequence. Two major mechanisms appear to be responsible for these specific properties, DNA methylation and post-translational histone modification. (Brown & Strathdee 2002; Bird & Wolffe 1999; Jenuwein & Allis 2001). DNA methylation refers to the addition of a methyl group to cytosine residues at CpG islands (adjacent cytosine and guanine nucleotides) in the double helical structure of DNA. The addition of this side chain results in a local alteration of the DNA double helix reducing the ability of the DNA to be transcribed and thus decreased production of the associated gene product (Hausheer *et al.* 1989). Post-translation covalent histone modifications refer to the addition or removal of phosphate, acetyl and/or methyl groups to the histone proteins in which DNA is packaged. Modification of DNA packaging can act as a local switch resulting in transcription or repression of a specific gene based on whether it lies in an open (acetylated) or closed (phosphorylated, methylated or unmethylated depending on the specific position) conformation (Jenuwein & Allis 2001). The mechanisms listed

above appear to control the vast majority of genes in the genome and, at the very least, influence the rest (Brown & Strathdee 2002; Bird & Wolffe 1999; Jenuwein & Allis 2001). These observations have spawned the creation of the term “epigenome” to describe this intricate control system, which appears to be almost as important as the actual base pair sequence of the genes themselves (Jenuwein & Allis 2001). Cracking this “histone code” is likely to be the next major achievement in molecular biology.

Epigenetics thus represents a heritable mark that can be passed down through progeny but which may also be modified in response to environmental and developmental phenomena thereby modulating the transcriptional program of a cell. Modification of either DNA methylation or histone status that results in differential gene expression is referred to as genetic reprogramming. Patterns of DNA methylation and the resultant histone dynamics have been implicated in transcriptional regulation both in a global and gene specific manner, X-chromosome inactivation, genomic imprinting, silencing viral retrotransposons, as a mechanism for monitoring cellular aging and in neoplastic transformation. A recent wave of research has thus been focused on the identification and characterization of specific factors that both establish these epigenetic marks and control their dynamics. The oocyte and early preimplantation embryo represent perhaps the largest repository of epigenetic factors as it is from here that the epigenetic foundation is laid down, setting the transcriptional basis for the development of mammalian life.

## **Part 5 Epigenetic Events of Early Development**

Mammalian preimplantation development is a critical stage for the generation of the genomic methylation pattern. During the first few cleavage divisions, a genome wide drop in methylation occurs as the nucleus is remodeled to take on the task of supporting embryonic growth and differentiation (Monk *et al.* 1991). This removal of methyl groups is thought to “reset” the genome to a plastic state where it can be reprogrammed to direct embryonic development. A *de novo* wave of methylation then asserts a new pattern during preimplantation or postimplantation development depending on the species. These processes appear to be conserved across mammalian species and are essential for normal development to proceed (Dean *et al.* 2001; Li *et al.* 1992; Okano *et al.* 1999).

## **Part 6 Epigenetic Events of Early Bovine Development**

Relatively little is known about the early epigenetic events of bovine embryonic development. Studies pioneered by Dean *et al.* (2001) demonstrated that during bovine preimplantation development, the genomic methylation pattern is erased during the first few cleavage divisions and then reasserted during the 8cell to 16-cell transition. This developmental pattern is similar to the mouse in that methylation levels begin to drop just prior to syngamy, however *de novo* methylation is seen during the 8cell to 16cell transition, which is in stark contrast to the epigenetic events observed in the mouse (Figure 2). During murine development, methylation levels do not begin to rise until after implantation. In the bovine it would seem that remethylation is initiated during preimplantation development. The biochemical and enzymatic basis for this difference

is presently unknown as are the developmental consequences as they pertain to studies of somatic cell nuclear transfer.

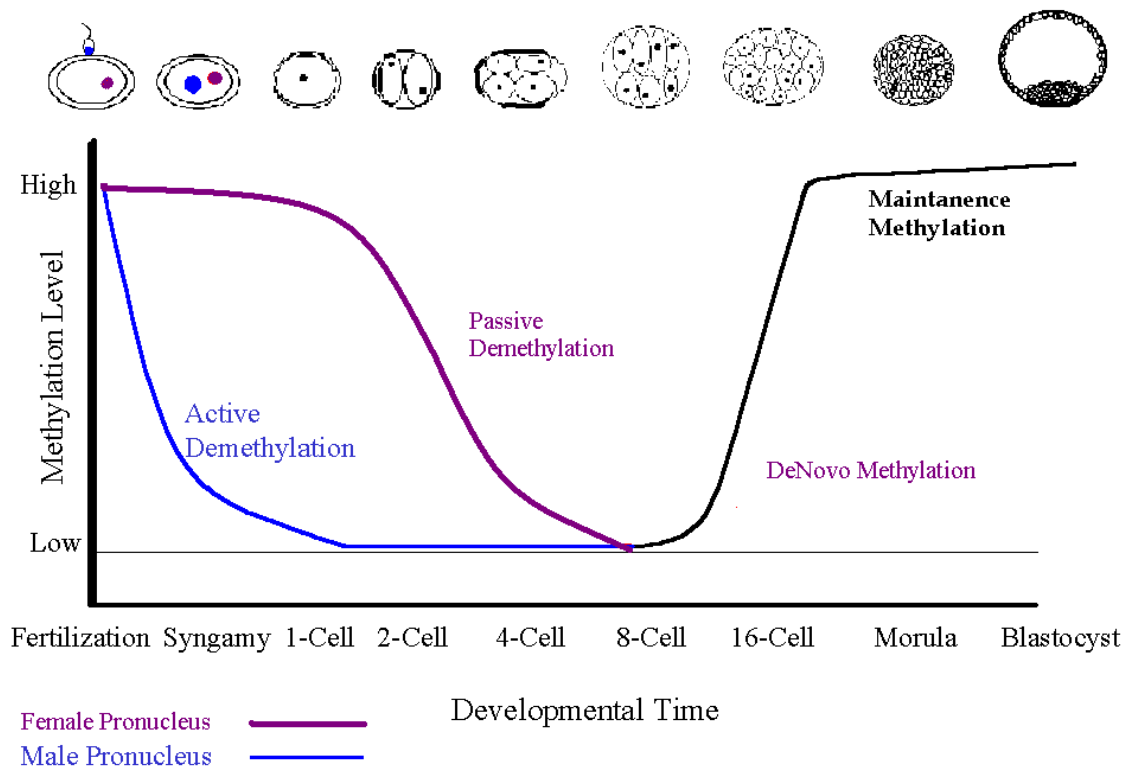


Figure 2 Epigenetic events of bovine preimplantation development. A schematic diagram depicting the timing of the epigenetic alterations to the male and female pronucleus over the course of bovine preimplantation development. The male pronucleus undergoes a rapid, genome wide demethylation immediately following fertilization, whereas the female pronucleus passively demethylates over the course of development to the eight-cell stage. During the eight to sixteen-cell transition, the embryonic genome remethylates. The enzymes responsible for the dynamics depicted in this diagram are currently unknown.

## Part 7 The DNA (Cytosine 5) Methyltransferases

The specific mechanisms by which the dynamics of genomic methylation are controlled during this critical period remain largely unknown, however a family of structurally related proteins termed DNA (cytosine - 5) methyl-transferases (DNMTs) have been identified which catalyze the production and modulate dynamics of the global genomic methylation pattern (Bestor, 2000). DNMT1 is the most abundantly expressed methyltransferase and is responsible for maintaining methylation patterns through DNA replication. This enzyme is constitutively expressed and localizes to the replication foci of actively dividing cells (Leonhardt, 1992). Studies in mice have shown that Dnmt1 employs a stage specific alternatively spliced isoform during preimplantation development to carry out tasks specific to this crucial stage. This isoform has been termed DNMT1o due to its restricted pattern of expression to the oocyte and early preimplantation development (Mertineit *et al.* 1998; Howell *et al.* 2001; Ratnam *et al.* 2002). Murine DNMT1o is the sole isoform of DNMT1 expressed in the preimplantation embryo (Ratnam *et al.* 2002). The Dnmt1o transcript has a unique 5' end that results in the production of a truncated protein. This protein is excluded from the nucleus during the one to eight cell stage, is allowed briefly to enter the nucleus to maintain the maternal imprint patterns, and is then excluded again at the sixteen cell stage until after implantation (Howell *et al.* 2001; Ratnam *et al.* 2002). Knockout studies of DNMT1 in mice are embryonic lethal and fail to progress beyond the 1<sup>st</sup> trimester (Li *et al.* 1992). Whereas, replacing the oocyte specific DNMT1o with the

somatic form of DNMT1 disrupts genomic imprinting and causes developmental failure in the last third of gestation (Howell *et al.* 2001).

DNMT3a and DNMT3b both possess *de novo* methylating ability, transferring methyl groups to previously unmethylated regions (Okano *et al.* 1999; Okano *et al.* 1998a). Inherent in this, these enzymes have the ability to redirect gene expression patterns by altering the topology of DNA to a transcriptionally repressive state. Studies of the epigenetics of preimplantation development to date strongly suggest a model where DNMT3a and 3b cooperate to re-establish the genomic methylation pattern during early development and thus build the epigenetic foundation necessary to direct normal embryonic development. DNMT3a and 3b are abundantly expressed in embryonic tissues and stem cells but appear at low levels in somatic tissue. Both appear to form complexes with histone deacetylase 1 and 2 (HDAC1 and 2) and thus play a key role in recruiting all the known factors necessary to impart transcriptional repression (Fuks *et al.* 2001). Gene knockout studies of both DNMT3a and 3b in mice showed that the embryos arrested shortly after gastrulation and the differentiation of the three embryonic tissue types. Individual knockouts of DNMT3a and 3b were not as profound, however, this is likely due to functional redundancy between these and other members of the methyltransferase family (Okano *et al.* 1999). Recently, several alternatively spliced isoforms have been identified for DNMT3a and 3b, the expression of which is tissue specific and abnormalities therein linked to cancer (Robertson *et al.* 1999; Chen *et al.* 2002; Saito *et al.* 2002). The isoform distribution of DNMT3a and 3b has not been analyzed in the preimplantation embryo and it remains to be seen if these two enzymes

employ stage specific splice variants (isoforms) similar to DNMT1o. DNMT3a and 3b, along with the newly discovered methylation regulatory protein DNMT3L, in which knockout studies involving this gene have also resulted in failed development, are thought to be the major players in the establishment of the epigenome, in particular the unique tissue specific expression patterns and genomic imprints of particular genes (Bourc'his *et al.* 2001a; Hata *et al.* 2002). Further characterization of these *de novo* methyltransferases and their unique splice variants will likely yield great insight into the building of the epigenetic foundation of mammalian life.

DNMT2 is widely expressed in a variety of tissues and although it contains all the conserved methyltransferase motifs, it has only very recently been demonstrated to possess DNA methylating ability (Herman *et al.* 2003, Tang *et al.* 2003 and Kunert *et al.* 2003). In fact, until late this year, DNMT2 had no known function (Okano *et al.* 1998b). No specific study has yet addressed the function of DNMT2 within the oocyte, preimplantation embryo or in the gonads. It is likely that, given that DNMT2 is the most conserved methyltransferase it does play some critical role in the epigenetic control of the genome. Methyl-binding protein 2 is a protein that may or may not have the capacity to strip DNA of methyl groups but does not appear to be involved in preimplantation development specifically. This protein appears to specifically bind methylated DNA and then recruit histone-modifying enzymes to alter the transcriptional activity of the methylated region (Jones *et al.* 1998). Through this mechanism, DNA methylation serves as a mark to signal the post-translational modification of histones and thus modulate gene expression patterns in a tissue specific response to developmental and

environmental cues. Proper function of all the proteins listed above appears to be essential for normal development to proceed.

## **Part 8 Structural Domains of the Dnmts**

### **Catalytic Domain and Mechanism of Action**

Catalytic domains of all the DNA methyltransferases studied to date share ten structural motifs that have been remarkably conserved through evolution. These motifs consist of six highly conserved (motifs I, IV, VI, VIII, IX and X) and four moderately conserved motifs (motifs II, III, V and VII). These form the catalytic domain of the DNA methyltransferase enzymes by folding into two domains; one large domain consisting of motifs I through VIII and most of motif X and a smaller domain consisting largely of motif IX. The DNA double helix fits in the resultant cleft between the large and small domains. The variable region in between motifs VIII and IX confers specificity for the enzyme's binding to the DNA helix (Bestor and Verdine 1994; Kumar *et al.* 1994; Kilmasauskas *et al.* 1994).

The substrate of the DNA methyltransferase is the carbon at the number five position of the base cytosine. In its natural state, the cytosine is deeply buried in the DNA helix and thus not able to allow the reaction to proceed. Binding of DNA methyltransferase to the DNA helix induces a conformational change in motif IV, which is situated on a flexible loop referred to as the "catalytic loop" that results in motif IV coming into contact with the DNA helix. A conserved proline-cysteine dipeptide induces the targeted cytosine residue to release its Watson-Crick base pairing and flip out of the double helical structure into the catalytic pocket of the DNA methyltransferase

enzyme (Kilmasauskas *et al.* 1994). Residues in motif IX form direct hydrogen bonds with the O<sub>2</sub> and N atoms in the targeted cytosine ring giving it the correct orientation (Bestor and Verdine 1994; Kumar *et al.* 1994; Kilmasauskas *et al.* 1994).

During this flip out process, the enzyme itself undergoes a major conformational change where the catalytic loop moves deeper into the cleft and ultimately into the minor groove of the DNA helix. With this conformational change a cysteine in motif IV is now in its proper orientation to serve as the active site for the enzyme. This inducible conformational change brings a nucleophilic cysteine thiol into close proximity to carbon 6 of the cytosine ring. Immediately upon induction of this conformational change a nucleophilic attack of the thiol group on the number six carbon is initiated and results in the formation of a covalent (thioether bond) DNA-protein intermediate. Thus, the binding of the methyltransferase enzyme induces a conformational change on both the protein and the DNA double helix providing a most eloquent example of the induced-fit-mechanism for enzyme-substrate interactions (Bestor and Verdine 1994; Kumar *et al.* 1994; Kilmasauskas *et al.* 1994).

Formation of this stable intermediate permits the methyl-donor, S-adenosyl-L- s-methionine (AdoMet) to be brought into close proximity to both the active site of the enzyme and the targeted carbon five of the cytosine ring. AdoMet binds the large domain of the methyltransferase enzyme through specific interactions with amino acids in motifs I to V. A phenylalanine in motif I specifically interacts with the aromatic ring structure of the adenosyl moiety and serves to hold the cofactor in the correct orientation. The addition of the nucleophile to carbon six in the previous step activates

carbon five and allows the transfer of the methyl group from AdoMet. After the transfer of the methyl group, S-adenosyl-L-homocysteine (adoHyc) is released. The proton now at the 5 position is abstracted by a basic residue on the enzyme which is quickly eliminated via  $\beta$ -elimination (Bestor and Verdine 1994; Kumar *et al.* 1993; Kilmasauskas *et al.* 1994).

The mechanism discussed above suggests a kinetic cycle where the variable domain of the enzyme first binds its target upon which the active site loop clamps down upon the DNA helix, flipping targeted cytosine into the active site via the formation of a C4' exo sugar pucker and thus forming a covalent reaction intermediate. A rapid transfer of a methyl group from AdoMet results in the methylation of carbon 5 of cytosine and the release of AdoHyc. The proton abstraction likely releases the enzyme allowing both the DNA helix to return to the  $\beta$  conformation and the enzyme to proceed to its next substrate.

#### Mechanism of Action of the Dnmt Inhibitor 5-Azacytidine

5-Azacytidine is a nucleoside analogue similar to the natural substrate of the DNA methyltransferases, cytosine with the exception of a nitrogen atom in place of carbon atom number 5 in the ring structure of the base (Figure 3). The initial stages of the methylation cycle are initiated, however the nitrogen atom prevents completion of the reaction and leaves the Dnmt protein covalently bound to the DNA helix via carbon 6 of cytosine, as a reaction intermediate (Santi *et al.* 1984). As a result, Dnmt protein is rapidly depleted from the nucleus and methylation levels drop. However, it is thought

that the removal of the reaction intermediates by the DNA repair machinery results in localized mutations that may be of further harm to the cell.

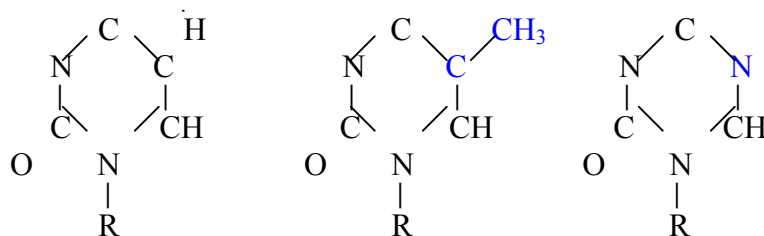


Figure 3 Schematic diagram representing the chemical structures of cytosine. Left to right – cytosine, methyl-cytosine and 5-Azacytidine.

### Cysteine Rich Zinc Finger

The carboxy terminus of all the known DNA methyltransferases contains the catalytic motif discussed above whereas the amino terminal domains of each methyltransferase are unique. With the exception of Dnmt2, which contains no identifiable domain structures besides the catalytic one, all of the remaining Dnmts contain a cysteine rich, zinc-binding region of unknown function. This region is similar in structure to a region in mammalian homologues of the *Drosophila* Trithorax protein, which is involved in maintaining homeotic gene expression. It is hypothesized that the trithorax protein anchors transcriptionally active chromosomal domains to the nuclear matrix, thus playing a role in maintaining their transcriptionally active state. The cysteine rich domain of the Dnmts (with the exception of Dnmt2) physically interacts with DNA and likely other proteins as well which may prevent the enzymes from

methylation transcriptionally active sequences via interacting with other protein factors present at these sites, sequestering the Dnmts from their substrate.

#### Domains Unique to Dnmt1

The amino terminus of Dnmt1 is essential for enzymatic activity and surprisingly, is toxic when expressed independently of the catalytic methyltransferase domain (Tucker *et al.* 1996). Additionally, cleavage of the amino terminus activates *de novo* methylation, but this *de novo* methylation is completely dependant on an interaction with the cleaved amino terminus as alone, the catalytic domain posses no enzymatic activity (Zimmermann *et al.* 1997 and Fatemi *et al.* 2001). Within the amino terminus are four major identifiable domains including the cysteine rich domain common to all other Dnmts (except Dnmt2), a nuclear localization signal, a polybromo domain and a PCNA binding site. The nuclear localization signal is located between amino acids 72-92 and is responsible for shuttling the enzyme to the nuclei in coordination with DNA replication during S-Phase of the cell cycle. A second domain proximal to this localization signal (amino acids 161-174) interacts with the proliferating cell nuclear antigen (PCNA) in a cell cycle coordinated fashion. This region targets Dnmt1 to DNA replication foci via a direct protein-protein interaction with PCNA. The cysteine rich region which is located between amino acids 643-688 has been demonstrated to bind Zinc and is hypothesized to be involved in DNA binding interactions but its exact function has yet to be determined (Bestor 1992 and Chuang *et al.* 1996). The Bromo Adjacent Homology Domain or Polybromo domain is located between amino acids 752-877 and 975-1097 (two adjacent domains) and is involved in

protein-protein interactions with other DNA binding factors, repressors and proteins whose function are currently unknown.

The amino terminus of DNA methyltransferase 1 also interacts with a number of other molecules, however the exact domains and binding sites have yet to be identified. As mentioned above, the N-terminal domain physically interacts with the carboxy-terminal catalytic domain imparting a tertiary structural arrangement that activates the catalytic methyltransferase function. No one sequence motif can be identified that does this solely, rather it is thought multiple sites bind inducing the correct alignment for enzymatic function (Margot *et al.* 2003). The retinoblastoma protein (Rb) binds Dnmt1 somewhere between amino acids 416-913 and an as yet undefined region in the amino terminus (amino acids 1-1,125) interacts with Histone Deacetylase 1 and 2 (HDAC1 & HDAC2) along with the co-repressor DMAP at DNA replication foci. Through these protein-protein interactions Dnmt1 becomes intimately involved with gene specific repression and histone modifying proteins thus tying together DNA methylation and post translational histone modification.

#### Domains Unique to Dnmt3a and Dnmt3b

Domain structures present in the amino termini of DNMT3a and DNMT3b have been less extensively characterized than those of DNMT1, but nonetheless several highly conserved domains have been identified and correlated with diverse functions. The most highly conserved domains present in the Dnmt3 family of methyltransferases are the PWWP and PHD finger domains. The PWWP domain is a 135-residue structure located between amino acids 265 to 321 in DNMT3a and amino acids 231 to 305 in

DNMT3b. This domain is hypothesized to specifically interact with heterochromatin and is probably involved in a physical interaction with the DNA backbone allowing methyl transfer to hemimethylated DNA. This protein domain has been identified in a wide variety of eukaryotic proteins from yeast to mammals, all of which in some way interact with DNA. When expressed alone this domain binds a 12 base pair region of the DNA sugar-phosphate backbone via the positively charged surface of the folded PWWP domain. The PWWP domain shares structural similarity to the SAND domain, which is believed to be involved in chromatin dependent transcriptional regulation. Evidence that both DNMT3a and DNMT3b interact with specific yet distinct pericentromeric heterochromatic loci supports the notion that this domain is involved in protein-chromatin interactions (Bachman *et al.* 2001). Indeed, this hypothesis was further supported by studies by Qiu *et al.* (2002) that demonstrated deletion of the PWWP domain abolished the association of Dnmt3a with heterochromatin. Since DNMT3a and DNMT3b are targeted to differing chromatic regions, there must be some as yet unidentified protein domain(s) responsible for their specific localization. What proteins mediate this specificity and the functional consequences to the alternative splicing of the amino terminal domains of Dnmt3a and 3b remains to be examined.

The Plant Homeodomain finger (PHD) domain is a conserved zinc-binding motif identified in more than 300 eukaryotic proteins. The majority of PHD domain containing proteins localize to the nucleus and are involved in modulating transcription through a variety of biochemical processes. The PHD finger domain of the human Dnmt3 family of proteins has also been called the ATRX like domain as it shares

remarkable identity (97/98 amino acids) with the same domain found in the human ATRX protein. Human ATRX is a chromatin-remodeling complex, mutation of which causes X-linked mental retardation with  $\alpha$ -thalassemia or ATR-X syndrome (Gibbons *et al.* 1995). Examination of PHD domains in numerous proteins using multiple alignments failed to discern any common positioning of this domain with relation to other protein domains or specific termini. Biochemical studies of this domain in other proteins have demonstrated that the PHD finger domain requires  $\text{Zn}^{2+}$  binding for proper folding and that this Zinc binding is mediated by a conserved Cys, Cys, Hys motif common to numerous other proteins across evolutionarily distant species. Recent studies have also revealed that this PDH finger domain is a metal dependent folding motif that mediates numerous protein-protein interactions. It is hypothesized that within this context, the PHD finger domains of the Dnmt3 family of proteins mediate their incorporation into multi-component complexes involved in transcriptional regulation. Recently, biochemical fractionation experiments have demonstrated an association between Dnmt1 and Dnmt3b as well as associations between Dnmt3a and a histone methyltransferase and histone deacetylase 1 (HDAC1) (Datta *et al.* 2003). It is very likely that the PHD finger domain mediates the assembly of the DNMTs into multi-component complexes that are involved in loci specific transcriptional silencing.

### **Part 9 Regulation of Genomic Methylation by the Dnmmts**

Function of the Dnmt family of methyltransferases can be subdivided into three distinct regulatory roles. Their catalytic action during early embryonic development sets up the epigenetic foundation for a given cells transcriptional program. A maintenance

function ensures proper transmission of this cell specific methylation pattern through cell division and finally the Dnmts have the capacity to mediate a certain degree of a cells transcriptional response to the environment, viral infection, neoplastic transformation and aging. None of these characteristics are understood very well but recent work in stem cells and tumors has suggested that each of the Dnmts or, more specifically, each of their individual isoforms plays distinct roles in this capacity. All of the Dnmts identified to date exist as multiple isoforms and each of these contain unique protein domain combinations that likely impart a unique catalytic function. Many of the splice variants identified produce enzymes that possess no methylating ability suggesting that the amino terminus may possess some as yet unidentified catalytic or regulatory role within the cell. However, the isoforms identified that do possess methylating ability all seem to localize to distinct chromosomal regions, have differing preferences for hemi-methylated and non methylated substrates and developmental and tissue specific expression patterns. It is likely that the cell uses alternative splicing to increase the repertoire of DNA methylating enzymes to achieve a more dynamic and precisely controlled system.

During early preimplantation mouse development Dnmt1 $\alpha$  translocates to the nucleus during the S-phase of the 8-cell stage and is essential for maintenance but not establishment of maternal specific imprinting patterns (Howell *et al.* 2001). Further, studies of the human reproductive syndrome that results in a condition known as a hydatidiform mole have discovered that the genomes of these affected individuals suffer from a failure to establish maternal specific imprint patterns and thus the phenotype resembles an androgenome. However, study of the Dnmt1 and Dnmt1 $\alpha$  genes of these

malformed embryos did not reveal any mutations in the Dnmt1/Dnmt1o gene (Hayward *et al.* 2003). Given that Dnmt1o is not required for establishment of either the paternal or maternal gene specific imprint patterns laid down during preimplantation development, and that Dnmt3a and Dnmt3b are not expressed during this developmental stage, there must be some as yet unidentified methyltransferase that mediates the establishment of the gene specific methylation for imprinted loci (Judson *et al.* 2002).

Studies of both human and murine cells suggested that Dnmt1 serves solely as a maintenance methyltransferase and the activity of Dnmt3a and 3b restricted to *de novo* methylation. However, recent studies by Chen *et al.* (2003) revealed that in the absence of the Dnmt3a and Dnmt3b, Dnmt1 alone cannot maintain methylation levels and further that overexpression of Dnmt1 from an artificial plasmid cannot remethylate lost epigenetic marks. Further, these studies revealed that the Dnmt3 family of methyltransferases plays a much more intricate role than first suspected. Examination of Dnmt3a and Dnmt3b activity alone and together in Dnmt3a<sup>-</sup>/Dnmt3b<sup>-</sup> knock out stem cells has revealed specific methylating functions and targets for each of the different isoforms of Dnmt3a and Dnmt3b.

Dnmt3a predominantly localizes to retroviral sequences, major satellite repeats, IAP repeats, non-imprinted genes as well as paternally imprinted genes, and the Xist gene on the X-chromosome. Dnmt3a2 however, appears to specifically methylate paternally imprinted genes whereas Dnmt3a and Dnmt3b1 were not able to remethylate these sequences once the original methyl-mark had been lost. In addition, it appears that over the long term both the Dnmt3a and 3b family of enzymes are required for

maintenance of these paternal imprints and that no amount of overexpression of Dnmt1, Dnmt3a, Dnmt3b or their associated isoforms has been able to restore maternal specific imprint patterns (Chen *et al.* 2003).

Dnmt3b is an enigmatic methyltransferase in that the majority of its isoforms are enzymatically inactive. To date six Dnmt3b isoforms have been identified in the human and eight in the mouse. Of these identified isoforms (dnmt3b1 – Dnmt3b6) only Dnmt3b1 and Dnmt3b2 possess any DNA methylating ability. It has been speculated that alternative splicing of the other isoforms serves in some way to negatively modulate the expression of the enzymatically active ones as over expression of Dnmt3b3 has been correlated with hepatocarcinogenesis (Saito *et al.* 2002). However, there are numerous reports of interactions of domains within the amino-terminus with other methyltransferases, gene specific repressors and histone modifying enzymes. Before any conclusion can be drawn as to the function of these non-methylating isoforms, further characterization of the amino terminal domains and their functional interactions needs to occur.

Human Dnmt1, Dnmt3a and Dnmt3b each have multiple differing types of promoters that regulate their expression, each with differing CpG contents. These multiple promoters are thought to provide a feedback mechanism that serves to regulate the expression of the Dnmts. Human Dnmt1 is regulated by at least four independent promoters; one of which is located in a CpG rich region while the remaining three are CpG poor (Bigey *et al.* 2000). Human Dnmt3a is regulated by two CpG rich promoters and one poor promoter while; Dnmt3b is regulated by one CpG rich and one CpG poor

promoter (Yanagisawa *et al.* 2002). The biological reasoning for control of these genes utilizing differing types of promoters with differing CpG contents is not known for certain however, a negative feedback mechanism that serves to modulate their own expression can be envisioned. Perhaps a threshold level of methylation is obtained during the growth of a cell such that the Dnmts begin to methylate their own promoters and suppress their capacity to initiate transcription. Since these genes are controlled by multiple promoter types, it is likely that methylation of the CpG rich promoters, which are under the control of the SP1 transcription factor and thus constitutively active, shifts transcriptional control to the CpG poor promoters that may be more tightly regulated. Via this hypothesized mechanism the Dnmts could respond to genomic methylation levels by modulating their own expression. Given that Dnmt3a is driven off two CpG rich promoters suggests that it is the lowest abundance methyltransferase in highly methylated adult cells and the most abundantly expressed in embryonic cells or stem cells, which in fact, it is. Conversely, given that Dnmt1 is predominantly regulated by CpG poor promoters would suggest that at the very least this methyltransferase would be the most abundantly expressed methyltransferase in adult cells and intermediary in embryonic, which it appears to be (Chen *et al.* 2003).

Indeed, reported observations of Dnmt expression levels suggest that the postimplantation mouse embryo, developing germ cells and embryonic stem cells all abundantly express Dnmt3a and Dnmt3b but that these genes are only minimally expressed in most somatic tissues. Dnmt1 however is ubiquitously expressed in somatic tissues and equally abundant as the Dnmt3 family during early development. Of further

note is the fact that Dnmt3a and Dnmt3b can be absent in tumor cells but absence of Dnmt1 induces apoptosis (Chen *et al.* 2003). The functional nature of this discrepancy is currently unknown and is the subject of intense investigation.

Taken together, the analysis of Dnmt expression and function to date suggests the following model: During early preimplantation development, Dnmt1 serves to maintain imprint specific methylation patterns established either by some as yet undefined methyltransferase or by the current Dnmts during gametogenesis. How these methylation marks, or the mechanisms that demarcate them pass through the initial wave of embryonic demethylation is presently unknown. Once established however, they are maintained by both Dnmt1 and the Dnmt3a enzymes. During murine postimplantation development, Dnmt3a and Dnmt3b serve to carry out the de novo methylation of the genome establishing heterochromatin, retroviral specific silencing and maintaining the proper expressional paradigm of the given cell. This fine-tuned differential regulation that serves to compartmentalize the genome into heavily methylated areas and hypomethylated areas is likely achieved by the regulation of Dnmt isoform expression through their alternative splicing. Both the Dnmt1 and Dnmt3 families of enzymes are required for establishment and stable maintenance of these specific patterns and further, each cell likely has a specific repertoire of Dnmt splice variants unique to its methylation requirements. These enzymes are all highly transcribed during early development from CpG rich promoters. However, once a cell begins to terminally differentiate, its methylation levels increase to the point where the de novo methylation enzymes are down regulated to basal levels. At this point, Dnmt1 assumes the predominant role as

the maintenance methyltransferase, while Dnmt3a and Dnmt3b serve as proofreaders, filling in the gaps missed by Dnmt1. This minimal expression of Dnmt3a/3b and reliance on Dnmt1 to maintain methylation levels can be thought of as the somatic expression profile. As a cell ages, there is evidence to suggest that methylation levels increase, repressing more and more genes until the cell becomes quiescent and dies. Thus there also appears to be a role in cellular aging.

The Dnmts have also been implicated in mediating a portion of the cells response to the environment. Evidence to support this hypothesis comes from studies examining the transcriptional profiles of cells exposed to environmental insults and viral infection. The nature of this response is as yet unclear and only strong correlative data have been presented. However, given that viral sequences appear to be aggressively methylated and that environmental toxins induce hypermethylation of the genome it is clear that the Dnmts are in some capacity involved in the cell's ability to respond to infection and environmental insult. There is evidence to suggest that Dnmt3a is the lead methyltransferase that carries out methylation specific responses to environmental stimuli as its expression is the most widely varied and has the greatest diversity of target sequences. However, direct experimental evidence to support these hypotheses has yet to be fostered (Suetake *et al.* 2003).

At present, there has been no specific functionality yet ascribed to Dnmt2 within the context discussed. Given that of all the methyltransferases, Dnmt2 is the most highly conserved; it is likely that this enzyme plays some essential function. However, there are currently no experimental reports of this enzyme examining function in

gametogenesis, embryonic development or correlation with any specific gene loci.

Given that we now know that this enzyme has the capacity to methylate DNA (Herman *et al.* 2003; Tang *et al.* 2003 and Kunert *et al.* 2003) it is likely that dissection of the molecular targets and function of DNMT2 will expand our current understanding of the function of this gene family as a whole and of the epigenetics of early mammalian development.

## **Part 10 Nuclear Transfer**

The most popular report of mammalian somatic cell nuclear transfer was by Wilmut *et al.* (1997) which was first to describe the cloning of a mammal from an adult cell. Nuclei of epithelial cells derived from an adult ewe were transferred into enucleated oocytes, activated and transferred to recipients ultimately resulting in the birth of Dolly. Since this milestone was achieved in 1996, cloned mice, goats, cattle, pigs, rats, deer and cats have been reported (Baguisi *et al.* 1999; Cibelli *et al.* 1998; Prather *et al.* 1996; Prather *et al.* 1999; Wilmut *et al.* 1997; Yong & Yuqiang 1998; Shin *et al.* 2002).

Although live animals can be cloned by nuclear transplantation using somatic cells, the efficiency of the technique is very low in comparison to natural mating and current *IVF* technology. In cattle where the vast majority of this work has been done, pre-attachment development appears to proceed normally and does not seem to be the major factor affecting the efficiency. Development to the blastocyst stage *in vitro* is similar to that of embryos produced by *in vitro* fertilization. Maternal recognition and the establishment of pregnancy are also similar between *in vivo* embryos and those

produced by somatic cell cloning. After 35 days of gestation however, the drop in pregnancy rates is dramatic and very few of these fetuses survive to term. Moreover, many of the calves that do make it to term often exhibit developmental abnormalities and die at birth or shortly thereafter (Cibelli *et al.* 1998; Garry *et al.* 1996; Kato *et al.* 1998; Renard *et al.* 1999; Vignon *et al.* 1998; Wilson & Wyatt 1995;). The ones that do survive frequently exhibit cardiovascular abnormalities, immature lung development, a compromised immune system, and diabetes. These diseases are now being correlated with insults during early embryonic development (Hill *et al.* 2000; Pace *et al.* 2002; Barker 1990).

### **Part 11 Nuclear Transfer and Epigenetic Reprogramming**

Some of the most convincing evidence supporting the hypothesis that abnormal epigenetic reprogramming by factors in the oocyte / early embryo results in failed development comes from studies involving somatic cell nuclear transfer. Mammalian development is driven by highly specific temporal and spatial patterns of gene expression. The ability of a transferred nucleus to recapitulate these precise patterns is essential for proper development. However, abnormal patterns of gene expression are the norm for reconstructed embryos and aberrant gene expression can even be seen in adult clones (Wrenzycki *et al.* 2001; Schultz *et al.* 1996; Daniels *et al.* 2001; DeSousa *et al.* 1999; Humpherys *et al.* 2002). Several recent studies in cattle have further demonstrated genomic hypermethylation and aberrant patterns of X-chromosome inactivation in animals produced by nuclear transfer, suggestive of incomplete epigenetic reprogramming (Kang *et al.* 2001, Dean *et al.* 2001, Bourc'his *et al.* 2001b,

Xue *et al.* 2002). This epigenetic instability correlates with the fact that no one particular set or class of genes seems to be specifically and consistently affected by the process of nuclear transfer, rather the transcriptional disruption seems to be random and genome wide. A recent survey of the transcriptional activity of cloned bovine embryos using cDNA microarray analysis failed to identify any consistently misexpressed genes and instead concluded that the abnormalities are unique to each embryo (Donovan *et al.* 2003). Several imprinted genes have been consistently found to be abnormally expressed but again, these differences are tissue specific and highly varied. (Inoue *et al.* 2002; Humpherys *et al.* 2001).

The genetic and epigenetic reprogramming that must occur after nuclear transfer is the critical element in the developmental success of a reconstructed embryo. To date, little research has been directed towards correcting or attributing some causality to the aberrant patterns of gene expression observed in clones. As a cell differentiates from a state of totipotency to a specific functional endpoint, the methylation pattern of the genome changes with it, reflecting a change in the transcriptional program. Different cell types have different transcriptional requirements and thus their methylation patterns are unique to their functional phenotypes. A technique called restriction landmark genome sequencing has recently been used to identify unique patterns of methylation specific to different cell types and developmental stages (Shiota *et al.* 2002 and Rush & Plass 2002). In essence, each cell type has a unique “methylation fingerprint” owing to differences in which specific genes are transcriptionally active as compared to those

which are in a state of repression. This fingerprint can be followed through differentiation and through environmental responses.

The methylation pattern of stem cells is significantly different than that of somatic cells and the pattern of an early preimplantation embryo even more divergent. Each successive stage of development seems to impart a greater level of methylation upon the genome and thus a more restrictive developmental plan. The early preimplantation embryo begins with the erasure of the methylation pattern and then reestablishing gene specific imprint patterns, thus creating embryonic stem cells. These unique cells have been demonstrated capable of generating all the cell types necessary for proper development and as such are the focus of an entire branch of therapeutic medicine. It is hypothesized that as embryonic stem cells develop they are exposed to wave after wave of *de novo* methylation until a “tissue specific stem cell” methylation pattern exists. The final stages of development are achieved when a cell is directed towards a specific functional endpoint by methylation of the genome in such a way as to endorse the transcription of only those genes required of this specific cell type. Via this mechanism the process of differentiation is carried out, placing strict developmental and transcriptional blocks upon a cell.

Owing to the fact that a cloned embryo’s nucleus comes from a somatic cell with its own specific methylation pattern, it is likely that the large portion of the transcriptional and developmental abnormalities observed in clones are attributable to the failure of this methylation pattern to be reset to the demethylated base state. Passage of somatic nuclei through the process of gametogenesis serves to strip the majority of

the methylation pattern and repackage the genome so as to be easily demethylated following fertilization. With the techniques currently employed in somatic cell nuclear transfer, it is extremely unlikely that the nucleus is properly remodeled and the epigenetic state reset. Moreover, it is also likely that the normal methyltransferase activity inherent to early development compounds the issue by the addition of more methyl groups to a genome that should have been stripped but which are in fact already hypermethylated as compared to the natural state. This appears to indeed be the case in cloned embryos as they exhibit a drastic increase in the methylation levels of their genomes as compared *in vivo* controls. However, that several nuclear transfer experiments have been successful strongly suggesting that the capacity exists within the oocyte to redirect this methylation pattern to the point where development can proceed. It is possible that proper function of a few key epigenetic factors can restore developmental potential to a differentiated nucleus and allow proper development.

## **Part 12 Research Project Rationale**

In recent years the study of epigenetic control of gene expression has been recognized for its fundamental role in disease and development and has thus moved to the forefront of developmental biology. This is due simply to the abundance of studies demonstrating the importance of epigenetics as relates not only to normal development but a number of human disease conditions including cancer but also its fundamental role in embryo patterning and development. Aberrant gene expression as a result of improper epigenetic reprogramming by the oocyte following fertilization has become a key target for investigating the causes of failed development. Moreover, it is a primary suspect for

increasing the susceptibility to disease later in life (Brown & Stratthdee 2002). The Barker hypothesis suggests that adverse environmental conditions during fetal development may lead to adult diseases later in life including cardiovascular disease and diabetes (Barker, 1990). Undernutrition in pregnant women leading to low birth weight babies has been associated with an increased incidence of heart attacks, diabetes and high blood pressure as adults (Barker 1990; McCance, 1962; Mott *et al.* 1991; Smart, 1990). It is likely that all these disease phenotypes have strong epigenetic undercurrents.

The key modulators of DNA methylation are the DNMTs. Clearly understanding the factors and mechanisms controlling epigenetic reprogramming during early mammalian development, in particular those involving DNMTs, is important. A recent paper by Ding and Chaillet (2002) demonstrated that overexpression of the embryonic murine DNMT1<sub>o</sub> was tolerated and resulted in the production of live offspring whereas, another paper by Biniszkiewicz *et al.* (2002) demonstrated that overexpression of the somatic form of DNMT1 is embryonic lethal. These observations strongly suggest that the critical element allowing survival of mice overexpressing DNMT1<sub>o</sub> is proper function of the protein during preimplantation development. Somatic tissues are perhaps better able to regulate DNMT1 activity and DNA methylation in general but early embryonic tissues may not be. Indeed, expression of the somatic form of DNMT1, the maintenance methyltransferase, during the preimplantation stages of development would likely impede the genome wide drop in methylation that occurs, resulting in an inability of the nucleus to properly remodel itself to direct embryonic development.

Embryos produced by somatic cell nuclear transfer present a unique opportunity to investigate the intricacies of early epigenetic programming. It is clear that embryos produced by cloning exhibit a state of hypermethylation and this phenomenon has been found in both cattle and mice (Kang *et al.* 2001a, Dean *et al.* 2001, Bourc'his *et al.* 2001b). Owing to the fact that the donor nucleus comes from a somatic cell, it is likely the misexpression due to improper reprogramming or improper regulation of DNA methyltransferase function is responsible for the abnormal patterns of genomic methylation and the inability of cloned embryos to properly recapitulate embryonic transcription. Perhaps the persistence of the somatic form of DNMT1 is responsible for the observed state of hypermethylation as it continues to function, maintaining the patterns of methylation through DNA replication, during a time period when these methyl groups should be diminishing. Improper expression of the *de novo* methyltransferases DNMT3a and 3b may also be to blame for this over methylation, perhaps due to an as yet unexplained inability of a reconstructed embryo to properly regulate these proteins. Whatever the case, studies of nuclear transfer provide an opportunity to study epigenetic function in preimplantation development without the need for a traditional functional mutation.

Somatic cell cloning is not the only reported cause of abnormal methylation as superovulation, alcohol exposure, and *in vitro* culture also result in abnormal methylation patterns in mouse embryos (Shi & Haaf 2002). Environmental conditions involving *in vitro* culture can lead to abnormal gene expression in preimplantation embryos (Wrenzycki *et al.* 2001; Doherty *et al.* 2000; Natale *et al.* 2001) and abnormal

methylation induced by these environmental conditions during early development could well be the cause of this. In addition, stem cells appear to also be epigenetically unstable. Whether this instability is imposed on them by culture techniques or is in fact a property inherent to stem cells in general should be addressed before they are employed in a therapeutic setting.

In order for the full benefits of animal cloning to be realized and made economically viable, the efficiency of the nuclear transfer procedure must be significantly improved upon. Genomic hypermethylation is almost certainly the root of the abnormal gene expression patterns observed in cloned bovine embryos and is thus the underlying cause to the abnormal physiology and developmental failure. Here, we begin to test whether abnormalities in DNMT expression and regulation result in improper epigenetic reprogramming and decreased developmental capacity of NT embryos.

**CHAPTER II**  
**ANALYSIS OF DNA (CYTOSINE 5) METHYLTRANSFERASE mRNA**  
**SEQUENCE AND EXPRESSION IN BOVINE PREIMPLANTATION**  
**EMBRYOS, FETAL AND ADULT TISSUES\***

**Analysis of the Bovine DNA Methyltransferase mRNA Sequence**

Bovine DNA methyltransferase mRNA sequences were cloned by RT-PCR using mRNA obtained from adult testis and the sequences reported to Genbank (Dnmt1 AY244709; Dnmt2 AY244708; Dnmt3a AY271298; Dnmt3b AY244710). Bovine Dnmt cDNA sequences display strong homology to those reported for mouse and human, as would be expected for developmentally essential proteins of this nature (Appendix A). Carboxy-terminal domains share the largest degree of sequence similarity, likely due to the enzymatic activity residing in this region (Bestor 2000). Amino-terminal sequences are less consistent but show conservation of several domain structures including the cysteine rich zinc finger, BAH, PWWP, and ATRX domains, as well as other regions of as yet unknown function (Figure 4).

---

\* Reprinted with permission from “Analysis of DNA (cytosine 5) Methyltransferase mRNA Sequence and Expression in Bovine Preimplantation Embryos, Fetal and Adult Tissues.” by Michael C. Golding and Mark E. Westhusin, 2003. Gene Expression Patterns, 3/5 pages 551-558. Copyright 2003 by Elsevier Science Ireland Ltd.

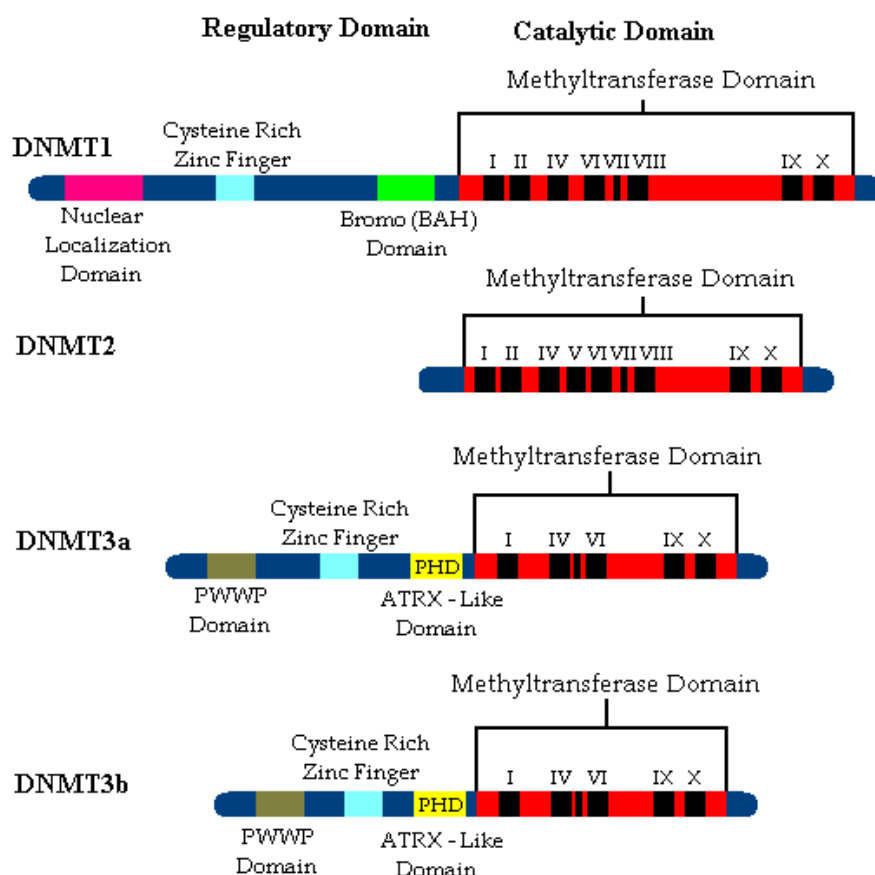
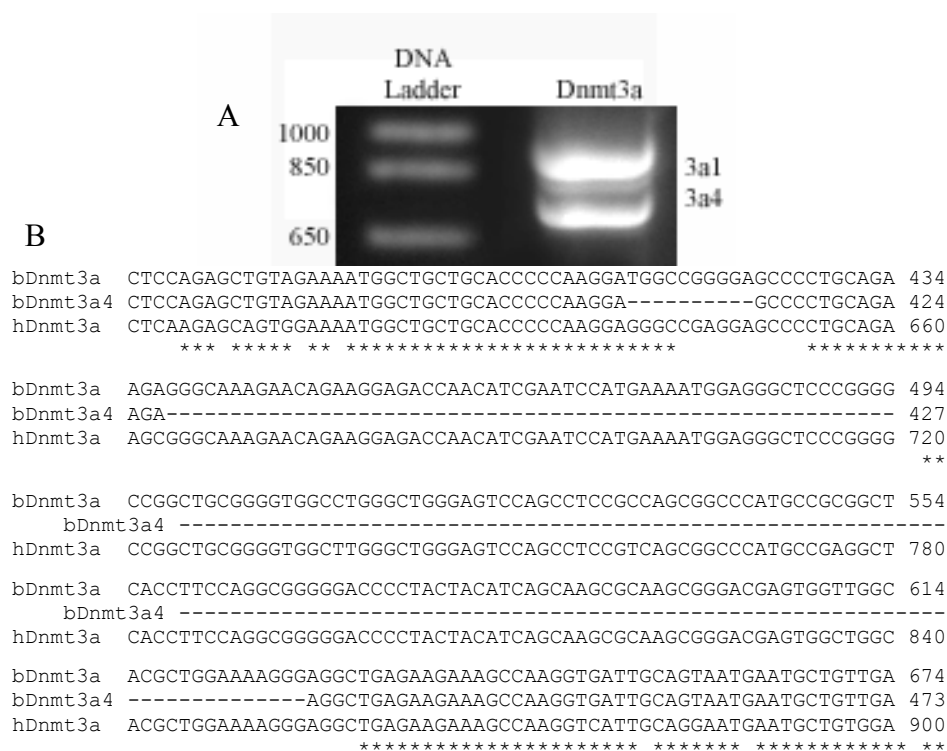


Figure 4 Domain structures of the bovine DNA Methyltransferase family. Sequence analysis finds conserved domain architecture among the bovine Dnmts. All four of the identified methyltransferases possess a C-terminal catalytic methyltransferase domain, which are of similar size and sequence to those of mouse and human. N-terminal domains of Dnmt3a and 3b are distinct from those of Dnmt1 and the proteins share no detectable homology other than in the C-terminal catalytic domain. Dnmt2 appears to lack any N-terminal regulatory domains.

### Splice Variant Characterization

During the course of sequencing the bovine Dnmt family, several alternatively spliced isoforms of Dnmt3a and 3b were identified in adult testis (Figures 5 and Figure 6). The Dnmt3b isoforms (Figure 6) (Dnmt3b1 AY244710; Dnmt3b3 AY244711;



**Figure 5** Isoform analysis of bovine Dnmt3a. (A). A 2% ethidium bromide stained gel showing the migration of both Dnmt3a transcripts, amplified from bovine testis by RT-PCR. Sequence analysis of each of the bands revealed a novel transcript missing 201 base pairs from the central coding region. (B) A sequence alignment of the alternatively spliced regions of bovine Dnmt3a. Exon arrangements for the alternatively spliced regions of Dnmt3a4, as compared to the full-length Dnmt3a sequences of bovine (top) and human (bottom) can be seen here. No other reported splice variants of Dnmt3a exhibit alternative splicing of exons within the central coding region. Matches are designated by a star and missing bases by a dash.

Dnmt3b4 AY244712; Dnmt3b5 AY244713) match perfectly those previously reported in human tissues (Figure 6B), however, the isoform identified for Dnmt3a (Figure 5) is unique. This novel splice variant, which we have termed Dnmt3a4 (AY271299), is missing 201 base pairs (67 amino acids) from the central coding region (Figure 5B). All other reported isoforms for this gene appear to incorporate alternate exons onto the 5' end via the use of alternative promoters (Chen *et al.* 2002; Weisenberger *et al.* 2002). To our knowledge, this is the first reported case of an isoform of Dnmt3a with alternative

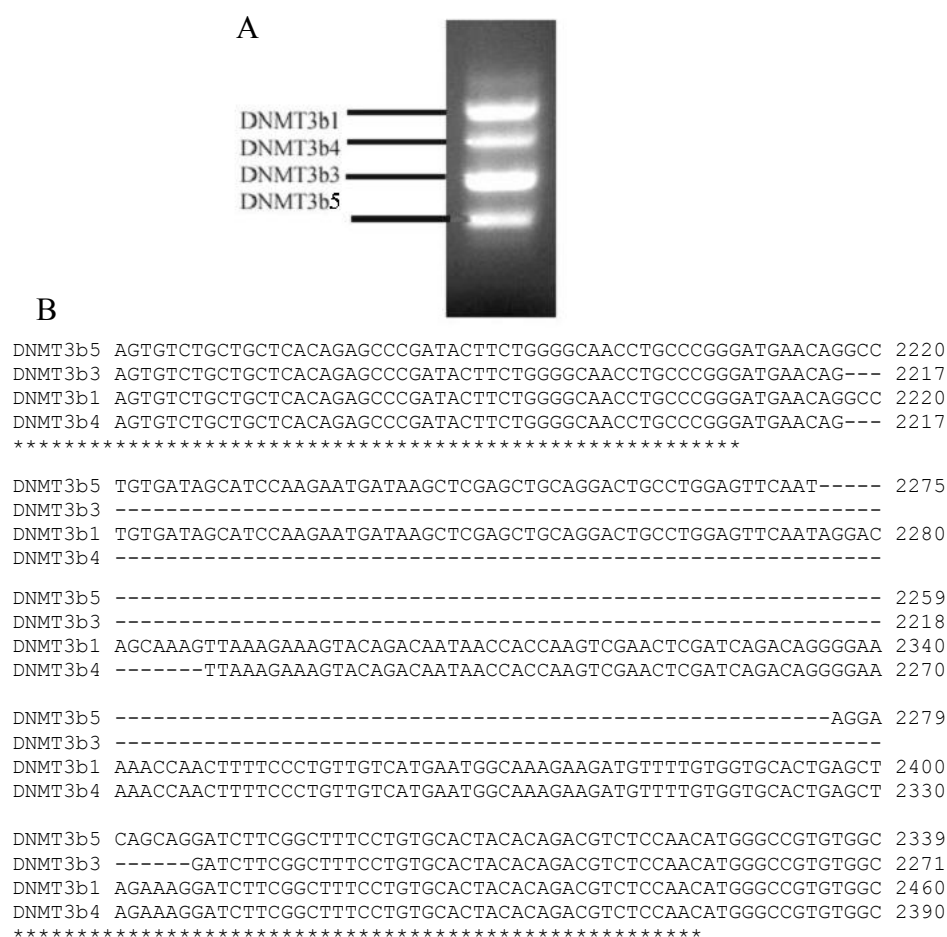


Figure 6 Isoform analysis of bovine Dnmt3b. (A). A 2% ethidium bromide stained gel showing the migration of each of the four Dnmt3b transcripts, amplified from bovine testis by RT-PCR. Sequence analysis of each of the bands demonstrated an exon arrangement remarkably similar to those reported for human Dnmt3b1, 4, 3 and 5 (results not shown). (B) A sequence alignment of the alternatively spliced regions of bovine Dnmt3b. Exon arrangements for each of the alternatively spliced regions of Dnmt3b are compared. Matches are designated by a star and missing bases by a dash.

exon arrangements within the central coding region. BLAST analysis of this

alternatively spliced segment indicates that it shares weak homology with bacterial Rec

G helicase. Further characterization of this splice variant is needed to determine if any

functional differences are imparted by the alternative splicing.

### **Dnmt mRNA Expression in Preimplantation Bovine Embryos**

During bovine preimplantation development, the genomic methylation pattern is erased during the first few cleavage divisions and then reasserted during the 8-cell to 16-cell transition (Dean *et al.* 2001). The specific methyltransferase enzymes responsible for this increase in methylation are currently unknown but work in the mouse would suggest an alternatively spliced isoform of Dnmt1, similar to the murine oocyte/early embryo specific Dnmt1o, might be involved. To identify the alternative transcripts of Dnmt1 present during bovine embryogenesis, rapid amplification of cDNA ends (RACE) was conducted using RNA isolated from bovine oocytes and early *in vitro* produced (IVP) embryos. Early studies of Dnmt1 levels in the mouse suggest that it is 3000 times more abundant in oocytes and embryos than in somatic cells and should thus be readily detectable (Carlson *et al.* 1992). Repeated attempts using a variety of RACE techniques (see methods) readily detected RNA coding for the ubiquitous somatic form (Dnmt1), and yet failed to detect any Dnmt1o transcripts. Sequence analysis of exon four of bovine Dnmt1 mRNA, obtained from somatic tissue clearly identified the reading frame necessary to produce the Dnmt1o protein (Appendix A), however we were unable to identify the 5' oocyte specific exon necessary to move translation to the DNMT1o specific start site.

In order to further investigate the embryonic expression of the somatic isoform of Dnmt1, RNA from *IVF* derived 8-cell to 16-cell embryos was separated into ribosomal and subribosomal subcellular fractions via a ribonucleoprotein fractionation procedure previously described by DeSousa *et al.* (1993). This technique allows the segregation of

RNA that is actively being translated from RNA sequestered in the cytoplasm. Upon separation of the fractions, the RNA was isolated and reverse transcribed into cDNA. Primers designed to amplify regions contained in exons four to five, which are common to all the reported isoforms of Dnmt1 were used to verify its expression. Separate primers that specifically amplify the 5' exon 1 unique to the somatic isoform of Dnmt1, along with exons two through five were used to detect the presence of the Dnmt1 somatic splice variant (Figure 7A). Results from these experiments along with sequencing data of the isolated cDNA amplicons, indicate that unlike the mouse, the cow employs the somatic form of Dnmt1 during the early developmental stages when genomic methylation first begins to rise. Further, Dnmt3a (Figure 7B) and Dnmt3b (Figure 7C) are present on the ribosome at the critical 8 to 16- cell transition.

No specific study has yet addressed the expression of Dnmt2 within the preimplantation mammalian embryo. In order to further investigate the expression of this gene and to determine how early the other Dnmts are expressed during bovine development, RNA from *IVF* derived four-cell stage embryos was separated into ribosomal and subribosomal subcellular fractions via a ribonucleoprotein fractionation procedure described. RNA was again converted into cDNA and used in a PCR reaction using primers specific to the somatic form of Dnmt1, Dnmt2, Dnmt3a and Dnmt3b. Results of these experiments, along with the sequence data of the isolated PCR amplicons demonstrates that expression of all the bovine Dnmts can be detected as early as the four cell stage of IVF development (Figure 8).

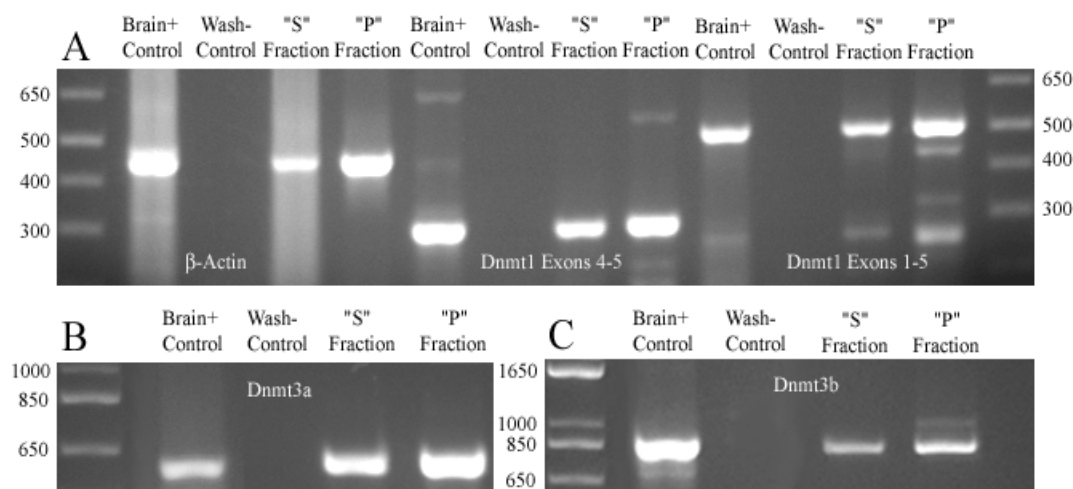


Figure 7. RT-PCR analysis demonstrating the recruitment of the Dnmt1, Dnmt3a and Dnmt3b mRNAs onto the polysome during the 8-cell to 16-cell transition. A 2%, ethidium bromide stained agarose gel showing the migration of the various Dnmt cDNA amplicons. Reactions were conducted on total RNA isolated from 75 8 to 16-cell-stage embryos, fractionated into a subribosomal supernatant ("S") and polyribosomal pellet ("P"). Expression of Dnmt1 (A) was conducted using two primer sets, one to detect exons common to all known splice variants (exons 4-5) and another set to specifically detect the somatic isoform (exons 1-5).  $\beta$ -Actin transcripts were amplified as a positive control. Expression of Dnmt3a (B) and Dnmt3b (C) were also analyzed in both fractions. Bovine brain mRNA was isolated, reverse transcribed into cDNA and then used as a positive control. A small volume of the phosphate buffered saline used in the final embryo wash was passed through the RNA isolation procedure and used as a negative control. This wash sample was reverse transcribed and used as template in a PCR reaction to ensure no contaminating RNA was introduced (standard RT control). Background bands can be seen in some lanes and are the result of excess cDNA template.

Given this apparent divergence from the murine model, a transcriptional profile of the Dnmt family throughout preimplantation development was conducted. RNA was collected from pools of ten embryos for each of the 2-cell, 4-cell, 8-cell, 16-cell, morula and blastocyst stages and reverse transcribed into cDNA. The mRNA for all four

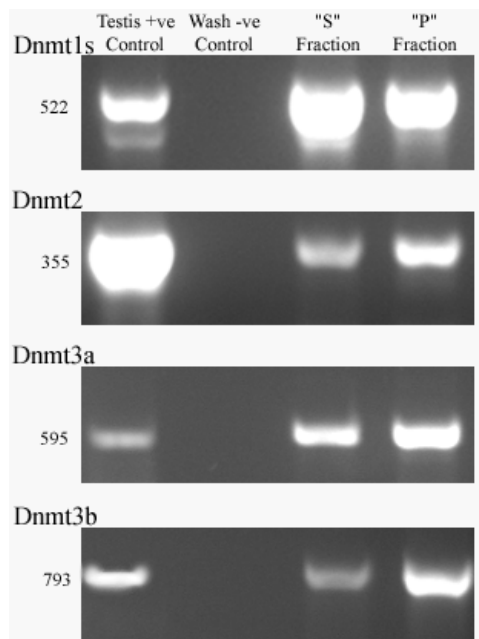


Figure 8 RT-PCR analysis demonstrating the recruitment of the Dnmt1, Dnmt2 Dnmt3a and Dnmt3b mRNAs onto the polysome during the 4-cell stage of bovine preimplantation development. A 2%, ethidium bromide stained agarose gel showing the migration of the various Dnmt cDNA amplicons. Reactions were conducted on total RNA isolated from 4-cell-stage embryos, fractionated into a subribosomal supernatant ("S") and polyribosomal pellet ("P"). Expression of Dnmt1 was conducted using the primer set specifically designed to detect the somatic isoform (exons 1-5), expression of Dnmt2, Dnmt3a and Dnmt3b were also analyzed in both the subribosomal ("S") and Polyribosomal ("P") fractions. Bovine testis mRNA was isolated, reverse transcribed into cDNA and then used as a positive control. A small volume of the phosphate buffered saline used in the final embryo wash was passed through the RNA isolation procedure and used as a negative control. This wash sample was reverse transcribed and used as template in a PCR reaction to ensure no contaminating RNA was introduced (standard RT control).

reported Dnmts is present during each stage of bovine preimplantation development (Figure 9).

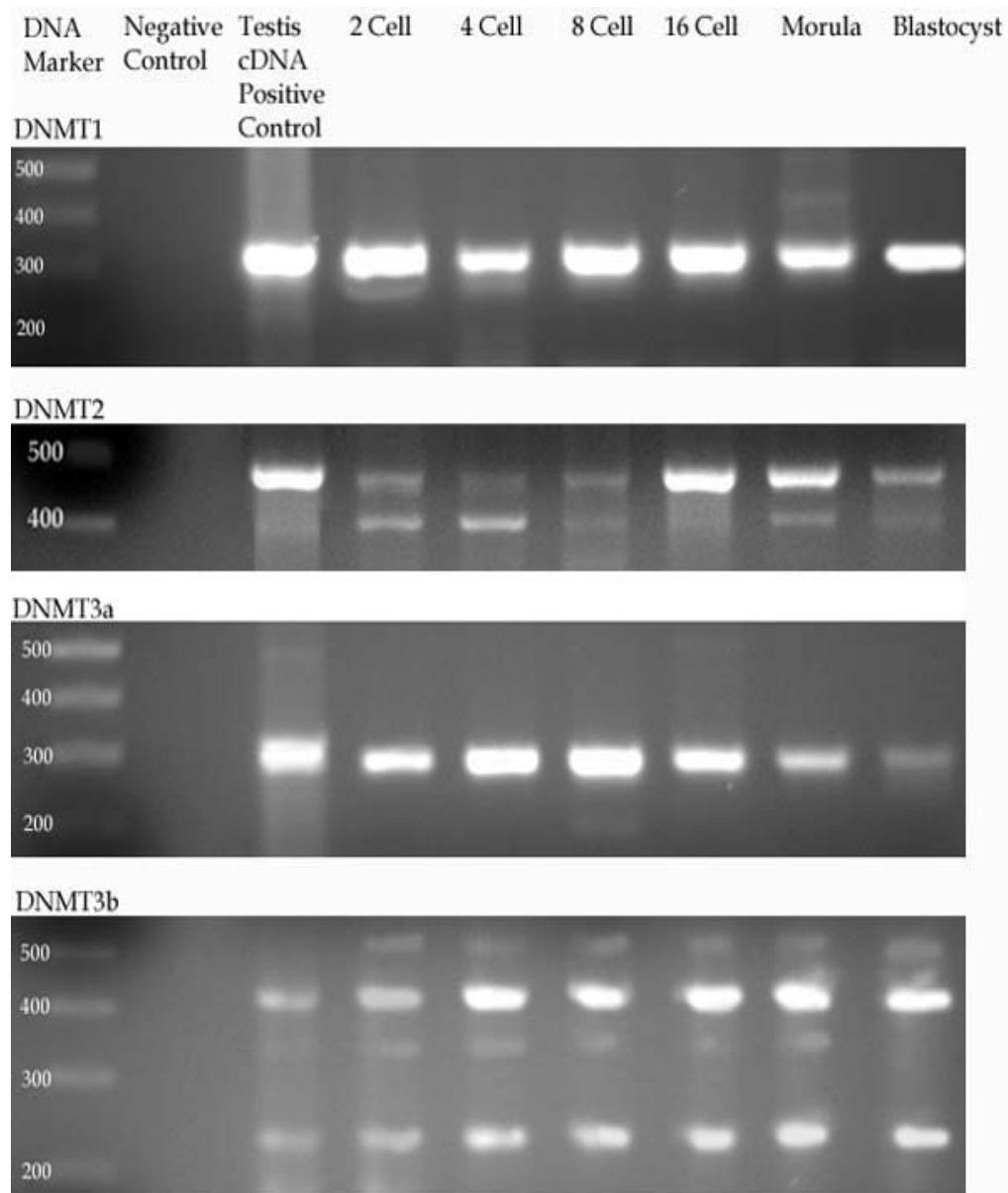


Figure 9. RT-PCR analysis of the four known Dnmt gene transcripts during all stages of bovine *IVF* preimplantation development. Transcripts of Dnmt1, Dnmt2, Dnmt3a and Dnmt3b were amplified by RT-PCR and run on an ethidium bromide stained 2% agarose gel. PCR amplicons of the alternate transcripts for Dnmt2 and Dnmt3b are seen here as multiple bands migrating different distances. Testis cDNA was used as a positive control and again a portion of the last embryo wash served as template for a negative control.

Similar to the testis, the preimplantation embryo contains an abundance of alternatively spliced products, including a novel isoform of Dnmt2 (present as a double band in figure 9), which has been designated Dnmt2-gamma (Dnmt2 $\gamma$ , AY244714). The coding region of Dnmt2 $\gamma$  contains a premature stop codon resulting in the production of a 63 amino acid protein with no identifiable domains and is therefore unlikely to produce a functional protein (Figure 10). Recently, similar non-functional isoforms have been identified for Dnmt3a and 3b and have been implicated in a possible translational regulatory mechanism (Saito *et al.* 2002; Weisenberger *et al.* 2002). Because the amino-termini of each of these proteins contain unique domains that may function independently of the carboxy-terminal catalytic methyltransferase domain, further investigation of the splice variants of each of the Dnmts is necessary to establish this hypothesis.

### **Dnmt Real Time Quantification During Bovine Fetal Development**

No previous study has yet examined tissue specific Dnmt expression profiles in early bovine fetal development. To expand our knowledge base on the developmental expression of this gene family during bovine fetal development, real time quantitation of all four of the known Dnmt transcripts was performed. Primers and probes were designed based on the sequence information reported above. Care was taken to ensure that measurements of a specific Dnmt reflected all the identified splice variants. Total RNA was isolated from fetal (16 week 25-27cm crown rump length - Winters *et al.* 1942) and adult tissues using Trizol reagent.  $\beta$ -actin transcripts were measured as an internal control.

```

hDNMT2      AAGACGATTGAAGGCATTACACTCGAAGAGTTTGACAGATTATCTTTGATATGATTTTA 229
hDNMT2d     AAGACGATTGAAG----- 121
bDNMT2      AAGACAATTGAAGGCATTACACTAGAAGAGTTTGACAGATTATCTTTCAATATGATTTTA 240
bDNMT2g     AAGACAATTGAA----- 192
          *****

hDNMT2      ATGAGCCCTCCCTGCCAGCCATTACAAGGATTGGCCGGCAGGGTGATATGACTGATTCA 289
hDNMT2d     -----ATTGGCCGGCAGGGTGATATGACTGATTCA 151
bDNMT2      ATGAGCCACCTGTCAGCCCTTCACAAGAATTGGCCTGCAAGGTGATGTGACTGATCCA 300
bDNMT2g     -----AATTGGCCTGCAAGGTGATGTGACTGATCCA 223
          ***** **

hDNMT2      AGGACGAATAGCTTCTTACATATTCTAGATATTCTCCCAAGATTACAAAAATTACCAAAG 349
hDNMT2d     AGGACGAATAGCTTCTTACATATTCTAGATATTCTCCCAAGA----- 193
bDNMT2      AGGACAAATAGCTTCTTACATATTCTAGACATTCTCCCAAGATTACAAAAATTACCGAAG 360
bDNMT2g     AGGACAAATAGCTTCTTACATATTCTAGACATTCTCCCAAGATTACAAAAATTACCGAAG 283
          *****

hDNMT2      TATATCTTTTGGAAAATGTTAAAGGTTTGAAGTATCTTCTACAAGAGACCTCTTGATA 409
hDNMT2d     -----GACCTCTTGATA 205
bDNMT2      TATATCTTTTAGAAAACGTTAAAGGTTTGAAGTATCTTCTACAAGAGATCTGTTAATA 420
bDNMT2g     TATATCTTTTAGAAAACGTTAAAGGTTTGAAGTATCTTCTACAAGAGATCTGTTAATA 343
          ** ** **

```

MEPLRALELYSGIGGMHQALRESCIPAQVVAAVDVNTVANEVYKYNFPHT  
QLLAKTIENTWPAR\* **STOP**

Figure 10 Bovine Dnmt2 $\gamma$  A) A sequence alignment of the alternatively spliced regions of bovine Dnmt2 $\gamma$ . Exon arrangements for each of the alternatively spliced regions of Dnmt2 $\gamma$  are compared. Matches are designated by a star and missing bases by a dash. This novel isoform of bovine Dnmt2 is missing an 83 base pair section near the 5' end. B) Translation of Dnmt2 $\gamma$  results in a protein containing a premature stop codon and the production of a 63 AA truncated protein with no identifiable domains.

Significant differences in Dnmt mRNA expression levels were found among different tissue types as well as between fetal and adult stages; (Figure 11) supported by the students T-test. On the whole, fetal expression levels are higher than those observed in adult tissues, specifically fetal brain, heart, rumen (stomach) and lung display the highest expression levels of all tissues examined (Figure 11). During bovine fetal development, the brain and heart form first and continue to steadily grow throughout development where as other tissue types appear to concentrate their growth during specific phases (Salisbury and VanDemark 1961; Winters *et al.* 1942). It is possible that

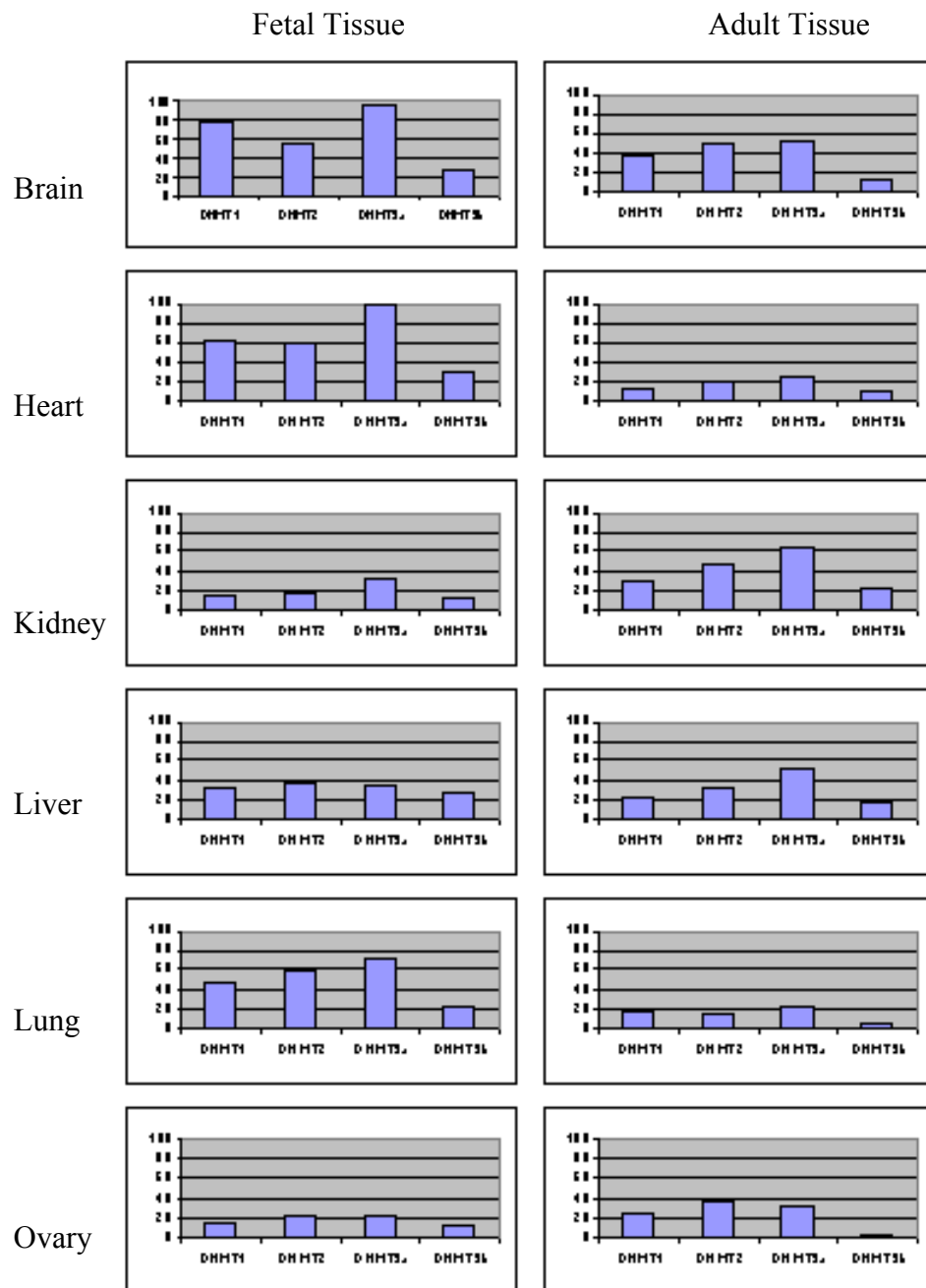


Figure 11 Real time PCR analysis of bovine DNA methyltransferase expression in fetal and adult tissues. Dnmt transcripts were quantified in samples of total RNA isolated from 16 week fetal (left column) and adult tissues (right column). The Dnmt family appears to have a tissue specific and developmentally regulated pattern of expression.

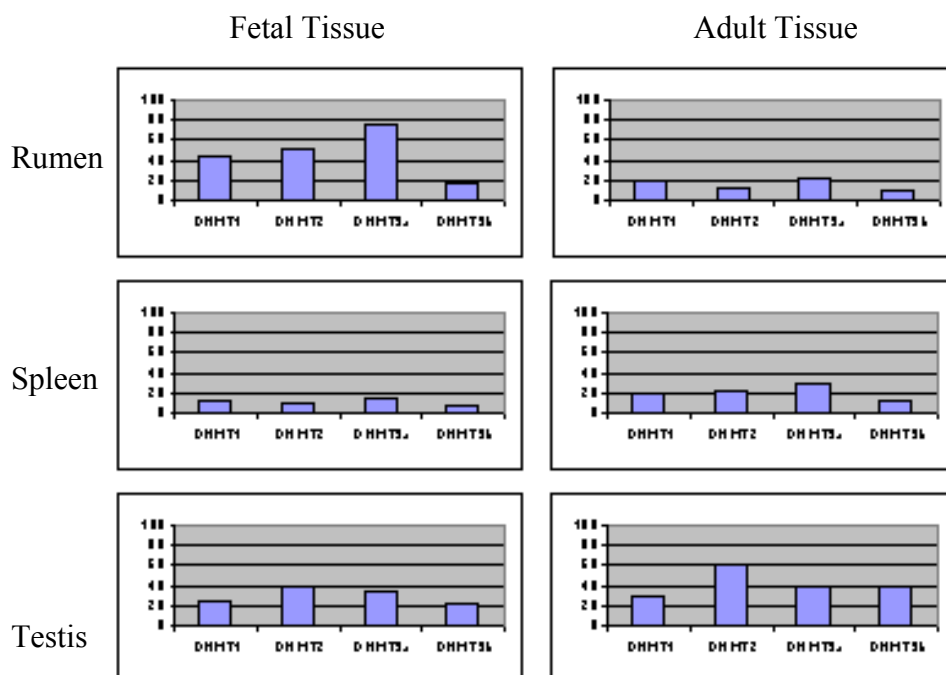


Figure 11 Continued.

during the developmental period examined, rumen and lung tissues are in such a growth phase.

High expression of the Dnmts during fetal development is somewhat expected given the coordinated role of these enzymes in DNA replication and tissue specific gene silencing. However, examination of adult tissues found that the highest expression levels were found in brain, kidney, and testis tissue (Figure 11). Brain and kidney tissues do not have high cellular turnover rates and were expected to have lower expression levels. Additionally stomach tissues that are usually associated with high cellular turnover contained lower levels of Dnmt expression. These observations suggest that

perhaps these enzymes play additional roles in cellular physiology besides mediating DNA methylation.

## **Discussion**

Bovine Dnmt1 mRNA contains the coding sequence necessary to produce the Dnmt1o splice variant but does not appear to specifically utilize this isoform during preimplantation development. It is surprising that something thought to be this fundamental to the early epigenetic events of mammalian development would be so different between the murine and bovine models. Recently, Dnmt1o has been cloned from human oocytes as well as from a South American opossum (Hayward *et al.* 2003 and Ding *et al.* 2003) although involvement of the Dnmt1o protein in the epigenetic events of the early development of these organisms has yet to be demonstrated. Given the high conservation of this gene and the exact spacing of the reading frames encoding the two isoforms, it is likely that bovine Dnmt1o exists but is in fact utilized during a different developmental time than its murine homologue. It is possible that bovine Dnmt1o is transcribed and translated during the earliest stages of oocyte development and utilized during this time to carry out the allele specific imprints observed in the mouse. Whatever role this isoform plays in the embryogenesis of the bovine, it does not appear to be involved in preimplantation development specifically and is thus not likely to be a contributing factor to the epigenetic and developmental abnormalities observed in cloned cattle.

Dnmt2 appears to contain all the catalytic domains necessary to carry out DNA methylation but has only recently been demonstrated to possess any observable

methylation ability (Okano *et al.* 1998a; Herman *et al.* 2003; Tang *et al.* 2003 and Kunert *et al.* 2003). The present study of Dnmt expression between tissues and through development indicate that bovine Dnmt2 mRNA is expressed at significant levels in all tissues and is, in fact, the most abundant methyltransferase found in adult testis and ovary. This in addition to the high level of sequence conservation between species and the multitude of reported splice variants identified for this enzyme would suggest that it might possess some as yet unidentified critical cellular function.

Dnmt3a and Dnmt3b while absent from murine preimplantation development are present during the initial stages of bovine development. This observation correlates with the results presented by Dean *et al.* (2001), which showed that in the bovine, genomic methylation levels begin to rise during the eight to sixteen cell transition. A similar rise in genomic methylation is not seen in mice until after implantation, which also correlates with the expression of murine Dnmt3a and Dnmt3b. Thus it would seem that while the specific enzymes involved in establishing the genomic methylation levels are the same, the timing of the events between the two species is drastically different. The difference in the timing of this increase in genomic methylation and the corresponding expression of the *de novo* methyltransferases may have arisen given the large differences in the timing of implantation. The bovine conceptus does not implant until around day 30, whereas the mouse embryo implants shortly after hatching from the zona pelucida (day 6). Whether the differences in implantation strategies form the basis for the different timing of epigenetic events, or whether it is some as yet undefined embryological process remains to be determined. Additional comparative studies as well as

examination of the Dnmts in embryos produced by nuclear transfer may provide further insight into their function and relevance to disease and development.

**CHAPTER III**

**ANALYSIS OF DNA (CYTOSINE 5) METHYLTRANSFERASE mRNA**

**EXPRESSION IN BOVINE PREIMPLANTATION EMBRYOS PRODUCED**

**FROM *IN VITRO* FERTILIZATION, SOMATIC CELL NUCLEAR TRANSFER**

**AND PARTHENOGENETIC ACTIVATION**

Abnormal patterns of gene expression are the hallmark of cloned embryos and the recent studies in cloned cattle and mice that demonstrate genomic hypermethylation have begun to provide some explanation to these transcriptional abnormalities (Kang *et al.* 2001a, Dean *et al.* 2001, Bourc'his *et al.* 2001b, Xue *et al.* 2002). The key modulators of DNA methylation are the DNMTs and misexpression of these proteins could potentially lead to the observed hypermethylation and the aberrant patterns of X-chromosome inactivation frequently seen in animals produced by nuclear transfer. Examination of the Dnmts and their involvement in controlling epigenetic reprogramming during the early development of reconstructed embryos may provide a clue as to the cause of this hypermethylation phenomenon while at the same time revealing more about the basic biochemical roles of these important enzymes.

Investigation of Dnmt expression during bovine IVF preimplantation development revealed an expression profile divergent from the one previously reported for murine development. Specifically, analysis revealed the presence of mRNA encoding the maintenance methyltransferase Dnmt1, as well as the enigmatic Dnmt2 along with the *de novo* methyltransferases Dnmt3a and Dnmt3b. Given that RNA

encoding all the Dnmts is present during normal IVF development, examination of the entire Dnmt family and expression of their respective isoforms in cloned embryos, was undertaken. RNA was again collected from pools of ten embryos for each of the 2-cell, 4-cell, 8-cell, 16-cell, morula and blastocyst stages and reverse transcribed into cDNA. While transcripts encoding all of the four reported bovine Dnmts are present during each stage of NT preimplantation development, expression of Dnmt1 and Dnmt2 was more difficult to detect than during the initial studies using *in vitro* produced embryos. During analysis of Dnmt1 and 2, as much as three times the amount of cDNA had to be seeded into the PCR amplification before the product could be consistently detected. In addition, while mRNA encoding both Dnmt3a and Dnmt3b could be detected using similar amounts of template, the isoform banding pattern observed for Dnmt3b was much different than the one observed during the analysis of IVF embryos (Figure 12). Whether these abnormalities are an artifact of PCR or are in fact, due to differential expression of the Dnmts during early clone development is unknown.

### **Real Time Analysis of Bovine Preimplantation Embryos Produced by *in vitro* Fertilization, Nuclear Transfer and Parthenogenetic Activation**

In order to more accurately assess Dnmt expression levels during the preimplantation development of cloned embryos, real-time PCR analysis was undertaken. In these experiments, comparisons were made between embryos produced by *IVF*, somatic cell nuclear transfer and by parthenogenetic activation. Embryos activated parthenogenetically contain two female compliments of the genome, can develop as far as the second trimester of pregnancy but lack the ability to produce live

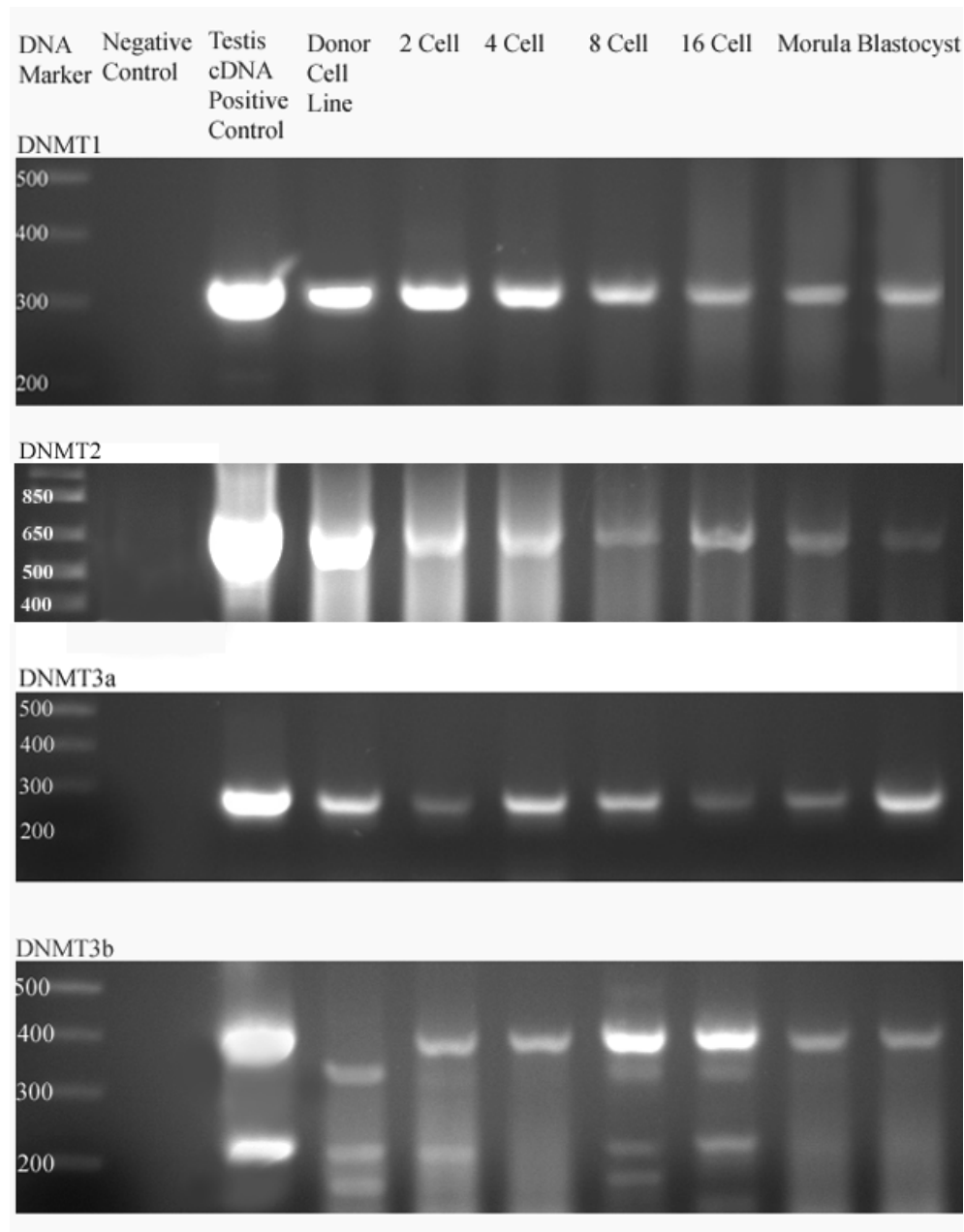


Figure 12 RT-PCR analysis of the four known Dnmt gene transcripts during all stages of preimplantation development in embryos reconstructed by somatic cell nuclear transfer. Transcripts of Dnmt1, Dnmt2, Dnmt3a and Dnmt3b were amplified by RT-PCR and run on an ethidium bromide stained 2% agarose gel. PCR amplicons of the alternate transcripts for Dnmt3b are seen here as multiple bands migrating different distances. Testis cDNA was used as a positive control and again a portion of the last embryo wash served as template for a negative control. RT-PCR amplification of the DnmTs was conducted on mRNA isolated from the fibroblast donor cell line for comparison.

offspring due to abnormal expression of imprinted genes (McGrath and Solter 1984a, McGrath and Solter 1984b). Thus examination of the Dnmt expression profiles in embryos produced using these diverse methods may reveal specific differences imparted either by solely containing a female genome, or by containing a nucleus derived from a somatic cell.

As in the analysis of bovine tissue,  $\beta$ -Actin served as the control. RNA isolated from pools of 10 two-cell stage, 20 eight-cell stage and 10 blastocyst stage embryos, corresponding to approximately 50 ng of total RNA for each pool (Dr. A. Watson personal communication), was seeded into two independent reactions measuring  $\beta$ -Actin expression levels as well as an individual methyltransferase. Examination of smaller pools individually may more accurately reveal minor differences between the experimental groups.

Upon obtaining the  $C_T$  values for each of the Dnmts at each of the examined developmental stages, the numbers were converted to the appropriate scale (a difference of 3.6  $C_T$  units is equal to a 10 fold difference in expression) and the reciprocals applied to an expression scale of 100 (Figure 13).

Overall, results of real time quantitation of the bovine Dnmts revealed significant differences in the expression of Dnmt1 and Dnmt2 between IVF embryos and NT but surprisingly found that transcript levels for Dnmt3a and Dnmt3b were generally the same. Similar differences exist between parthenote embryos and NT embryos but interestingly, differences in the transcript levels of Dnmt3b can be found between parthenote and IVF embryos. During the two-cell stage, no statistical differences in

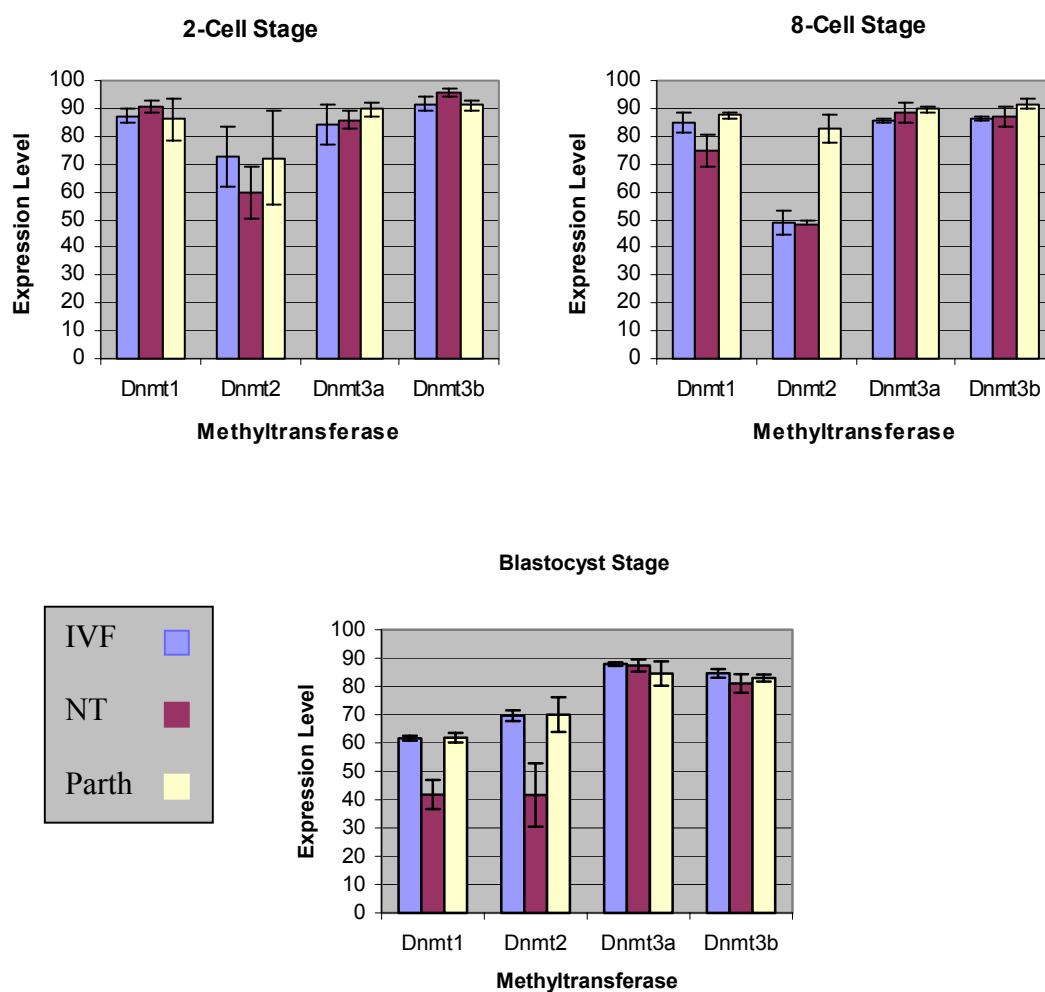


Figure 13 Real time PCR analysis of DNA methyltransferase expression during the preimplantation development of IVF, NT and Parthenote bovine embryos. RNA isolated from pools of 10 two-cell, 25 eight-cell and 10 blastocyst stage IVF, NT and Parthenote embryos was used in a TaqMan one step RT-PCR real time reaction. Measurements for each methyltransferase were normalized to  $\beta$ -Actin expression and compared to each other. A difference of 3.6 CT units represents a ten-fold difference in expression. Measurements were applied to a scale of 100. A minimum of three independent repetitions was carried out for each gene, during each developmental stage. Testis RNA was run as a positive control for each independent experiment (results not shown).

Dnmt expression can be discerned although; this may be due to the rather large standard deviations of the samples. Of note is the difference in the expression of Dnmt3b in NT embryos as compared to IVF, which is close to being statistically different (p-value 0.0867). However, no clear statistical differences exist at this stage.

During the eight-cell stage, a statistically significant difference in the expression of Dnmt1 between NT and IVF embryos was observed (p-value 0.0361). A similar difference in the expression of Dnmt1 between parthenotes and NT embryos may also exist but at the 95% confidence level the values are not significantly different (p-value 0.079). Interestingly, a large difference in the expression of Dnmt2 between parthenotes, as compared to both IVF and NT embryos exists (p-value 0.0116 and 0.0006, respectively). The expression of Dnmt3a is not statistically different during the eight-cell stage while the expression of Dnmt3b is slightly higher in parthenotes than IVF, although not enough to be statistically significant at the 95% confidence interval (p-value 0.0513).

In embryos reconstructed using somatic cell nuclear transfer, both Dnmt1 and Dnmt2 appear to be differentially expressed during the blastocyst stage. Transcript levels of Dnmt1 in NT embryos as compared to both IVF and parthenote development were significantly lower (p-value 0.0231 and 0.028, respectively) and similarly, Dnmt2 transcripts were much lower in NT embryos as compared to measurements of IVF and parthenote embryos (p-values 0.0525 and 0.0324 respectively). Expression of Dnmt3a and Dnmt3b for all experimental groups measured at the blastocyst stage was not significantly different.

To gain a better appreciation for the differences in Dnmt transcript levels,  $C_T$  values were graphed over developmental time for each experimental group (Figure 14). Overall, the expression profiles of the Dnmts look very similar between the different experimental groups. Dnmt2, the least abundantly expressed methyltransferase has the most varied expression profile of all the methyltransferases. The expression of this gene during IVF development is typical for genes during preimplantation development; a decline from fertilization on to activation of the embryonic genome at the eight-cell stage, followed by a resurgence in expression up to the blastocyst stage. Dnmt2 levels in clones appear to decrease steadily over development and fail to rise again while Dnmt2 levels in parthenotes seem to remain high. The expression of this gene during IVF development is typical for genes during preimplantation development; a decline from fertilization on to activation of the embryonic genome at the eight-cell stage, followed by a resurgence in expression up to the blastocyst stage. Dnmt2 levels in clones appear to decrease steadily over development and fail to rise again while Dnmt2 levels in parthenotes seem to remain high.

Of particular interest is the expression profile for Dnmt1. The patterns observed in IVF and parthenote development are remarkably similar, however the values recorded in NT embryos are much different. It appears that Dnmt1 levels of cloned embryos begin at a relatively similar level at the two-cell stage, decrease during the eight-cell stage and then rapidly decline at the blastocyst stage. This pattern is in contrast to the slow decline observed in both IVF and parthenote embryos.

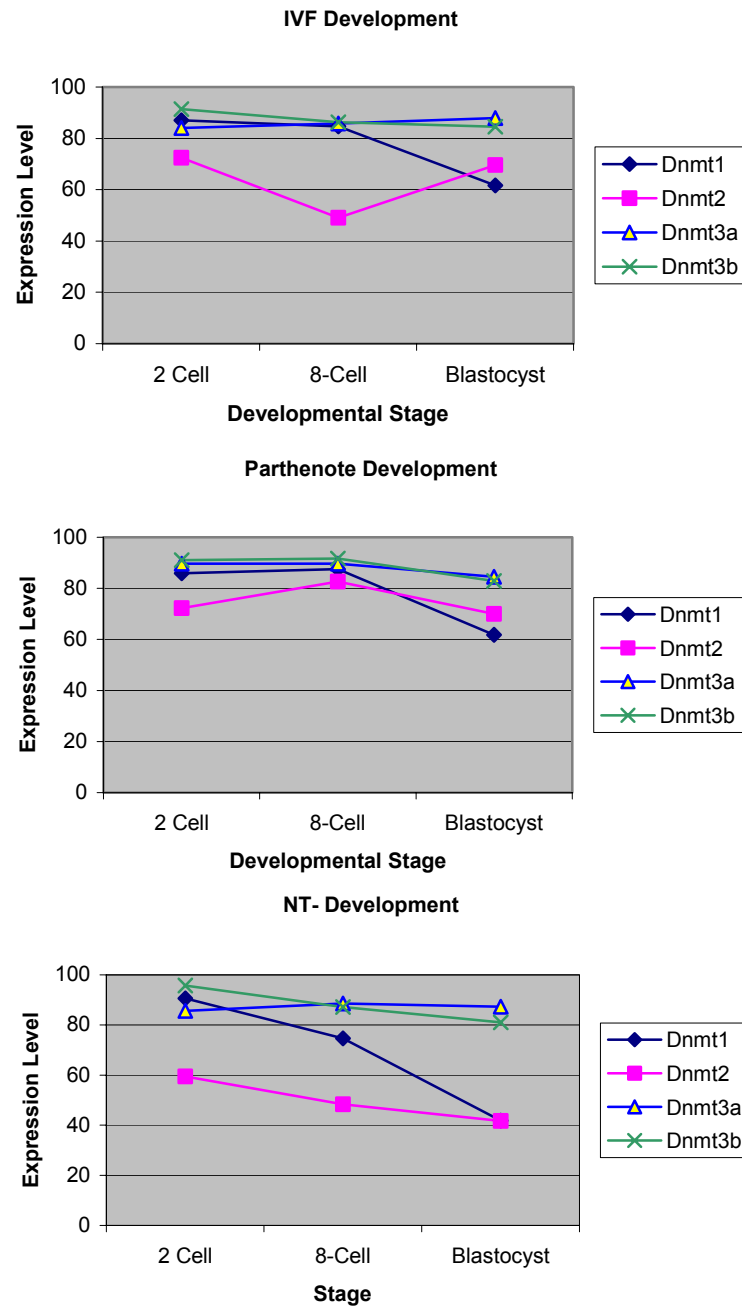


Figure 14 Comparison of DNA methyltransferase expression over IVF, NT and Parthenote preimplantation development. Expression levels of each of Dnmt1, Dnmt2, Dnmt3a and Dnmt3b are tracked over development from the 2-cell stage, through the 8-cell to the blastocyst stage. Differing DNA methyltransferase expression profiles can be noted here.

### **Real Time Analysis of Cultured Donor Fibroblast Cells**

Recently, DNA methylation has been postulated as a mechanism for monitoring cellular ageing (Lopatina *et al.* 2002 and Young and Smith 2001). A correlation between the levels of genomic DNA methylation and cellular age may have a significant impact upon the choice of donor cell line used in studies of nuclear transfer. Further, if this rise in DNA methylation coincides with a concurrent increase in DNA methyltransferase expression, the ability of the oocyte to properly reprogram the reconstructed embryo may be affected by the cellular age of the donor cell line. This differential status of methylation and the potential differences in Dnmt expression may be a significant factor in the observed discrepancies of the ability to clone embryos using fetal and adult cells. In order to further investigate the possibility that Dnmt expression levels are modulated with cellular age, the fibroblast donor cell line used in the experiments herein were cultured as far as the cells would pass before senescence. At several key passages, RNA was isolated from a subset of cells and the remainder stained to examine morphology.

Cells exhibited a gross change in morphology and a decline in gene expression as would be expected for cells entering senescence. As the cells passaged, their morphology shifted from the defined spindle shape typical of fibroblast cells to a more oblong spread out cell body. This change in morphology was accompanied by a steady decline in Dnmt expression (Figure 15). However, between passage 7 and 10, a resurgence in Dnmt expression was observed followed by a rapid decline in methyltransferase expression in all subsequent passages (Figure 15). Of further note is

A

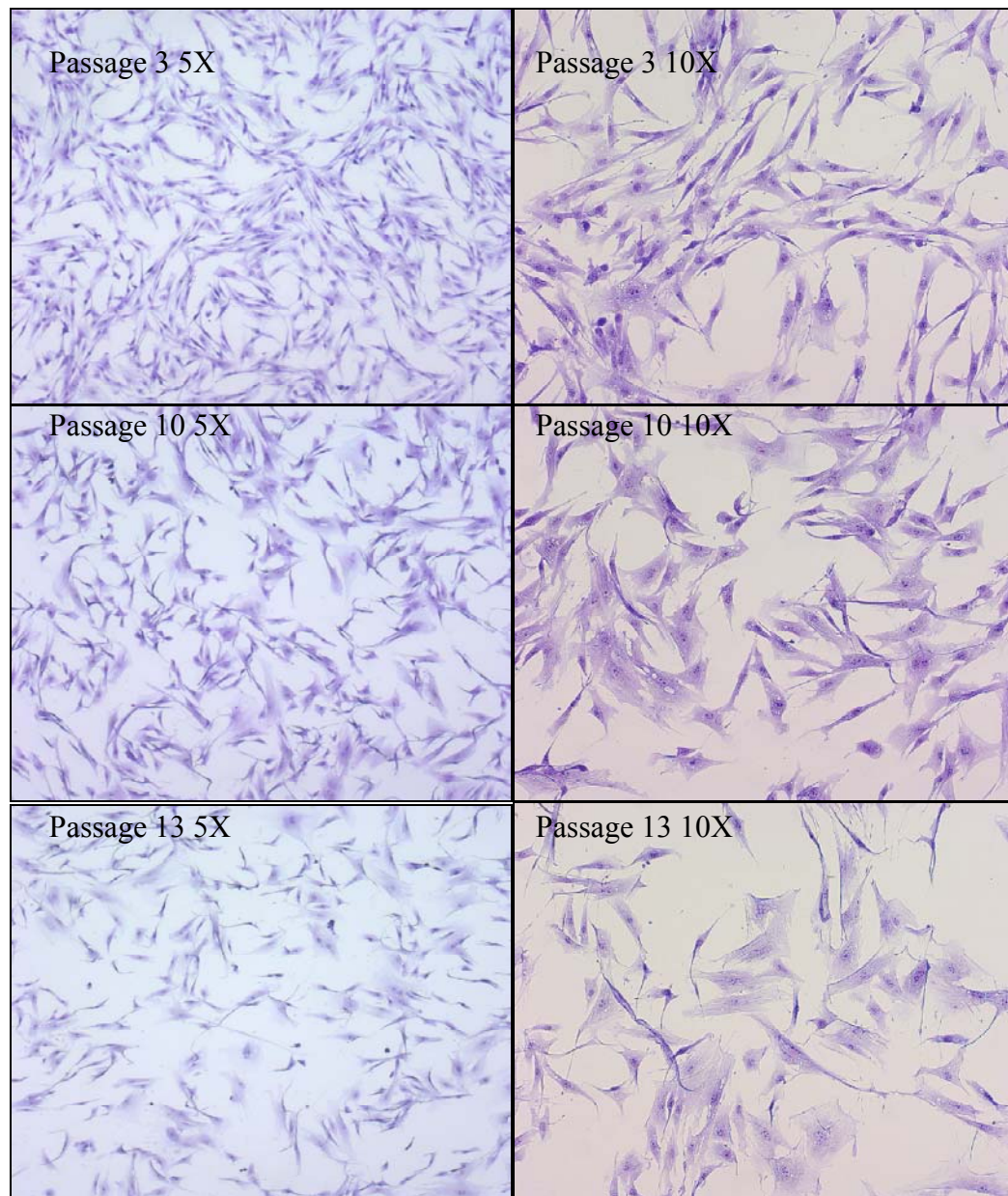


Figure 15 Real time PCR analysis of bovine DNA methyltransferase expression in the donor cell fibroblasts, cultured for an extended period of time. A) Light microscopy of cells cultured for varying time periods. B) (next page) Dnmt transcripts were quantified in samples of total RNA isolated from cultures of the donor fibroblast line taken at various passages seen in the panels of part A. The Dnmt family appears to be briefly upregulated during passages seven and ten, followed by a rapid decline in expression in all subsequent passages.

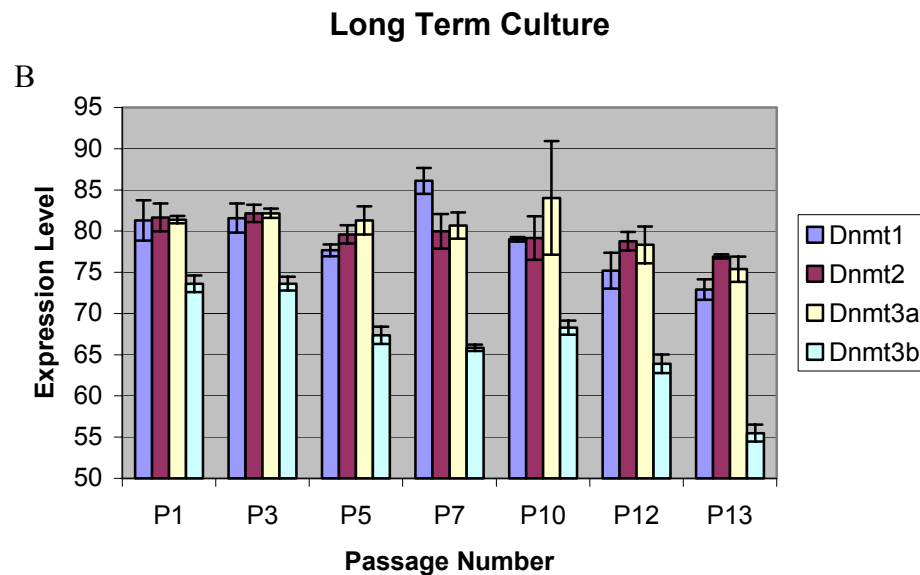


Figure 15 Continued.

the fact that Dnmt2 appears to maintain a steady state of expression while the remaining three methyltransferases show a more pronounced decline. The significance of the spike in Dnmt expression observed during passages 7 and 10 is unknown, but it is conceivable that these enzymes are actively methylating the genome, directing the genome to a transcriptionally silent state, in preparation for cellular senescence. The significance of maintaining Dnmt2 expression levels is unknown.

## Discussion

Real time quantitation of bovine Dnmt transcripts during IVF, NT and parthenote development has revealed Dnmt1 and Dnmt2 as the only consistently abnormally expressed methyltransferases. However, given the lack of *de novo* methylating ability reported for these two enzymes (Bestor 2000 and Tang *et al.* 2003) and the fact that the

expression of these two genes is reduced in NT embryos, it is unlikely that alone, their misexpression results in the observed hypermethylation. It is probable that Dnmt transcription on the whole is properly reprogrammed after transfer of the somatic nucleus into the oocyte but that inappropriate translational and regulatory control of these proteins maintains the genomic methylation levels during a stage when they should be receding. Inappropriate or premature translation of the Dnmts due to the loss or disruption of some as yet unidentified translational control mechanism could indeed account for the proposed abnormalities in DNMT function and the rapid decline in transcript levels for Dnmt1 and Dnmt2. However, why Dnmt3a and Dnmt3b levels are maintained is unknown.

Conversely, it is also conceivable that the drop in transcript levels for these two genes in cloned embryos is the result of their down-regulation due to the embryo repressing their transcription during the eight-cell and blastocyst stages in order to compensate for the state of genomic hypermethylation. Repression of these two genes could occur via the methylation of their promoters in a similar mechanism as outlined in chapter two. Given the hypothesized role of Dnmt3a and Dnmt3b in laying down allele specific imprints and methylating the viral sequences of the genome, expression of these genes may be required at high levels to achieve their developmental function and thus their expression is maintained (Chen *et al.* 2003). Examination of Dnmt expression levels in donor cells during successive increasing passage numbers points to a threshold cellular age, which may be relevant to selection of donor cells in studies of nuclear

transfer. What affect, if any the Dnmt spike observed during passages seven and ten has on an oocytes capacity to reprogram a donor nucleus remains to be examined.

Recent studies of nuclear breakdown and the dynamics of murine Dnmt1o in reconstructed embryos is beginning to allude to inappropriate translation and trafficking of key cellular proteins during the initial cleavage divisions of reconstructed embryos (Gonda *et al.* 2003; and Chung *et al.* 2003). In addition, it has been noted that removal of the female pronucleus from the oocyte precludes any ability of the zygote to demethylate the donor genome as well as removing key modulators of murine Dnmt1o trafficking (Chung *et al.* 2003; Kang *et al.* 2001b and Oswald *et al.* 2000). Taken with the data reported above it is likely that the abnormal methylation seen in cloned embryos is the result of a translational or protein regulatory abnormality rather than a transcriptional one.

To date, the vast majority of anomalies reported in clones have been attributed to the inappropriate transcription of several well defined genes. Given the data presented here in conjunction with the recent evidence of broad scale translational / regulatory abnormalities, it can be hypothesized that the large scale disturbances seen in clones are not due to any single transcriptional phenomenon but rather to a loss of translational and protein regulatory control caused by disruption of the embryonic nuclear architecture.

As biologists, we have a tendency to view the nucleus as a static structure housing the genome when in fact; the nucleus is a dynamic structure intimately connecting transcription and translation. Very little is known of mammalian nuclear architecture for any given cell type let alone the organization and intricacies of the

embryonic nucleus. Studies of the nuclear organization in somatic cells have revealed a close association of nuclear structural proteins with elements of transcriptional, post-transcriptional and translational control (Gonda *et al.* 2003). On a large scale, the nuclear genome can be thought of as a dynamically organized superstructure with elements accessible to exact regulatory factors and other areas that are specifically shielded. The process of nuclear transfer is a violent one and likely causes a massive disruption of the nuclear organization of the zygote, ultimately replacing the intricately organized female pronucleus with a somatic one. Removal of the embryonic nuclear matrix very likely pulls key regulatory factors out with it and conversely, replacement of an embryonic nuclear organization with a somatic one may permit the access of some inappropriate factors while restricting access of other factors needed during this time.

Kang *et al.* (2001a) reported that removal of the female pronucleus drastically reduces the ability of the zygote to demethylate the embryonic genome. It is very likely that structural elements present in the male and female genomes specifically facilitate this process and thus it can further be hypothesized that similar structural elements aid in the regulation of the Dnmts along with several other key regulatory proteins important to the early epigenetic patterning of the mammalian embryo. Thus, the abnormal gene expression thought to be due to the genomic hypermethylation seen in clones is not likely a down stream event of any single given transcriptional disruption *per se* but rather the end result of inappropriate biochemical events due to the loss of basic regulatory elements present in the nuclear matrix.

The above hypothesis would suggest that the structural and biochemical disruptions that cause the epigenetic and transcriptional abnormalities consistent in cloned embryos are imparted during the initial reconstruction procedure and possibly during the first few cleavage divisions. Further investigation of the nuclear structure of the oocyte and its role in directing the enzymatic events of “epigenesis” during early mammalian development will likely reveal a great deal about the patterning of the epigenetic foundation that is so fundamental to directing gene expression during the remainder of the growth and development of the organism.

## **Materials and Methods**

### **Tissue Collection and RNA Isolation**

Fetal and adult tissues were collected from slaughterhouse materials and snap frozen in liquid nitrogen. A Dounce homogenizer (Kontes, Vineland, NJ) was used for tissue disruption, and the Trizol (Gibco Carlsbad, CA) reagent was used to extract total RNA.

### **Embryo Production and RNA Isolation**

Mature bovine ova and preimplantation stage embryos were obtained using standard procedures utilized in our laboratory for *in vitro* embryo production (Winger *et al.* 2000). Briefly, bovine ovaries collected from a local abattoir were aspirated and oocytes placed into maturation medium (composed of TCM 199 supplemented with 10% fetal calf serum (FCS), 5 µg/ml FSH, 5 µg/ml LH, .05 µg/ml EGF and 1% penicillin/streptomycin (100 units of penicillin and 100 µg of streptomycin per ml)) at 39°C in an atmosphere of 5% CO<sub>2</sub> and air for 20-22 hours. Mature oocytes were then placed in

culture wells containing 0.425 ml fertilization medium and semen used to inseminate at a concentration of  $1 \times 10^6$  cells per ml. After 18-20 hours, the presumptive zygotes were removed from fertilization wells, washed in TL Hepes, and placed into 20  $\mu$ l drops of culture medium under oil in a humidified atmosphere of 5% CO<sub>2</sub> / 7% O<sub>2</sub> / 88% N<sub>2</sub>.

Embryos were washed twice through phosphate buffered saline and RNA purified from pools of ten bovine MII oocytes or embryos (2,4,8,16-cell, morula and blastocyst stage) using RNA isolation spin columns commercially available from Qiagen (Valencia, CA), according to the manufacturer's recommendations.

#### Ribonucleoprotein Fractionation

In order to separate RNA into the subribosomal (less than 80S) and polyribosomal (80S and greater) subcellular fractions, a modified protocol based on methods previously reported by De Sousa *et al* (1999) was used. Briefly, embryos were collected and washed once in Ca<sup>+</sup> Mg<sup>++</sup> free PBS and then placed into 300  $\mu$ l of a detergent lysis buffer (1% NP-40, 0.4% sodium deoxycholate, 500 units RNase Out (Gibco Carlsbad, CA), 10 $\mu$ g/ml cycloheximide, and 20 $\mu$ g yeast tRNA (Ambion Austin, TX) in a TSM/EGTA buffer (Kidder & Conlon 1985)). Embryos were homogenized in a mini-Dounce homogenizer on ice and the lysate centrifuged for 3 minutes at 24000g at 4° C. The postmitochondrial supernatant was then layered onto 50 $\mu$ l of 40% sucrose in TSM/EGTA and centrifuged for 40 minutes at 100,000g, at 4° C. The RNA was then isolated from the subribosomal supernatant and polyribosomal pellet using the Qiagen RNA isolation spin columns described above.

## Nuclear Transfer

Methods used to produce NT embryos were taken from Hill et al (2001). Briefly, a cloned fetus was produced by NT using fibroblast cells from an adult bull. The fetus was surgically removed from the uterus and the fetal tissue sliced into 2–5 mm pieces. Tissue pieces were transferred to 25 mm<sup>2</sup> flasks containing Dulbecco's modified Eagles medium (DMEM-F12, Gibco Laboratories Inc., Grand Island, NY)+10% FBS (FBS, Summit, Fort Collins, CO)+1% penicillin/streptomycin (Gibco) and cultured in 5% CO<sub>2</sub> in air.

Recipient oocytes were slaughterhouse derived from predominantly Brahman cross cattle, and matured for 17 h in Medium 199 (Gibco Laboratories Inc., NY) supplemented with 10% (v/v) fetal calf serum (FCS, Gibco) FSH 0.1 units/ml (Sioux Biochem, Sioux City, IA) LH 0.1 units/ml (Sioux Biochem), estradiol 1 µg/ml (Sigma, St Louis, MO), pyruvate 28 µg/ml (Sigma), EGF 0.05 µg/ml (Sigma) and 1% penicillin–streptomycin. The cumulus-oocyte complexes were vortexed 17 h post-maturation for 1–2 min, washed, placed in 0.05% (w/v) pronase E (Sigma) for 3 min, then held in M199+10% FCS.

Oocytes were enucleated at 19 h post-maturation. Prior to enucleation, oocytes were placed for 15 min in Hepes-buffered M199 containing Hanks salts (H-M199, Gibco) with 4 mg/ml fatty acid free BSA (Sigma) that contained 7.5 µg/ml cytochalasin B (Sigma) and 5 µg/ml Hoechst 33342 (Sigma). At this time, oocytes were selected for the presence of a polar body and homogeneous cytoplasm. Suitable oocytes were enucleated in H-M199 with 7.5 µg/ml cytochalasin B using a beveled 25 µm outside

diameter glass pipette. Only oocytes in which the removal of both the polar body and metaphase nucleus was confirmed by observation under UV light were included in the experiment. Oocytes were then randomly allocated to be combined with either early or late passage fetal fibroblasts.

Fibroblasts were combined with enucleated oocytes in 7.5  $\mu$ g/ml cytochalasin B in H-M199 using a 30  $\mu$ m outside diameter glass pipette, then returned to M199+10% FCS. The oocyte-fibroblast couplets were manually aligned and fused in a 3.2 mm fusion chamber that contained Zimmerman cell fusion medium using  $2 \times 20 \mu$ s 1.6 kV/cm dc fusion pulses delivered by a BTX Electrocell manipulator 200 (BTX, San Diego, CA). Oocyte activation was performed 3–5 h after fusion at 27 h post-maturation, by a 4-min incubation in 5  $\mu$ M ionomycin (Calbiochem, San Diego, CA) followed by 4 min in 3% BSA in H-M199.

#### Reverse Transcriptase Polymerase Chain Reaction

Given the highly conserved nature of this gene family, PCR primers were designed against regions of high homology between the mouse and human sequences and then used to amplify the larger portions of the coding regions. The Gibco SuperScript II and Platinum Taq Polymerase system (Carlsbad, CA) was used to amplify the targeted regions. The cDNA amplicons were then isolated and gel purified using a PCR Wizard Prep (Sigma, St Louis, MO) and sequenced using an ABI 3100 Genetic Analyzer (Applied Biosystems Inc, Foster City, CA). Obtained sequences were next used to probe a bovine cDNA library to yield the sequences of the 3' untranslated regions. Sequence analysis was conducted using BLAST software programs.

Expression analysis in preimplantation embryos was conducted using RT-PCR as above and were repeated on a minimum of three independent samples. cDNA amplicons were again isolated and sequenced to verify expression of the gene of interest. Sequences for PCR primers are listed in the table 1. Two independent sets of primers were used to validate the expression of each Dnmt.

Primer Name	Sequence
Dnmt1 Exons 4-5	Fwd - GATGCCTGCCCGAACCG Rev - CCCGTGGGAAATGAGATGTGAT
Dnmt1 Exons 1-5	Fwd - GAGGAGGGCTACCTGGCTAAA Rev - CCCGTGGGAAATGAGATGTGAT
Dnmt2 Set 1	Fwd - CCACCCTGTCAGCCCTTCAC Rev - GGGGATGTTTCAGATTCAGTTTTGG
Dnmt2 Set 2 *	Fwd - GCTCTCAGAGAAAGCTGTAT Rev - GGGGCTTGAAAGGGTAATGG
Dnmt3a Set1	Fwd - GCCCCGAAAGAGCACAACG Rev - GCCCAAGTCCTTCAGCACCAG
Dnmt3a Set2	Fwd - CTGGTGCTGAAGGACTTGGGC Rev - CAGAAGAAGGGGCGGTCATC
Dnmt3b Set1	Fwd - GTGTCCTTCCACCCTCTCTTT Rev - GCTTGTCGCCAACCTTCAT
Dnmt3b Set2	Fwd - GCACGAGGGCAACATCAAA Rev - CTCCAGGACCTTCCCAGCA

Table 1 RT-PCR Primer Sequences

\* Primers used to detect Dnmt2 isoforms based on primers designed by Franchina *et al.* 2001.

### 5' Rapid Amplification of cDNA Ends

Rapid Amplification of Complementary DNA Ends (RACE) techniques were employed to sequence the 5' regions. RNA isolated from adult testis, mature MII oocytes and IVF 8-cell stage embryos was used with the Clontech SMART RACE (Pal

Alto, CA) and Gibco (Carlsbad, CA) RACE kits according to the manufacturer's recommendations. Additionally a modified protocol adapted from Mertineit *et al* (2000) was also employed. Briefly, Superscript 2 reverse transcriptase was used with a gene specific primer to produce cDNA, which was precipitated in ethanol. An anchor primer blocked at the 3' end and phosphorylated at the 5' end (Sigma-Genosys Woodlands, TX) was annealed to the cDNA using T4 RNA Ligase (New England Biolabs Beverly, MA). A primer specific to Dnmt1 exon five was then used with RACE primer 1 in the first PCR reaction using Gibco Taq polymerase. A portion of the first PCR reaction was then seeded into a second PCR reaction with a primer complimentary to Dnmt1 exon 4 and a second RACE primer.

#### Real Time PCR

TaqMan® TAMRA™ Real Time probes (Applied Biosystems Inc, Foster City, CA) were designed based on the bovine nucleotide sequences obtained for the Dnmts. A TaqMan® One-Step RT-PCR kit was used to measure gene expression on a GeneAmp 5700 Sequence Detector (Applied Biosystems).  $\beta$ -actin mRNA expression was used as an internal control and analysis of variance (ANOVA) employed to compare expression levels between tissue types. Experiments involved a minimum of 3 repetitions for each tissue.

**CHAPTER IV**

**SMALL HAIRPIN RNA MEDIATED KNOCKDOWN OF BOVINE DNA  
METHYLTRANSFERASE 1: ESTABLISHMENT OF SHRNA MEDIATED  
FUNCTIONAL GENOMICS IN THE BOVINE MODEL**

In order to further dissect the role bovine Dnmt1 plays in the genomic hypermethylation observed in cloned embryos, a transient disruption of Dnmt1 expression during the initial stages of preimplantation development will be conducted using RNA interference. However, before an RNAi based suppression of gene expression could be attempted it was necessary to determine if an RNAi response could be elicited in bovine cells. Further, reliable methods to design and diagnose the efficacy of individual interfering RNA constructs was needed before attempting a gene knock down in the embryo. To this end, several preliminary experiments and the implementation of protocols to allow easy design and testing of shRNAs were carried out.

**RNA Interference in Bovine Cells**

**Sequencing of Bovine Dicer Argonaute 2 and Argonaute3 (Ago 2 & 3)**

The process of RNA interference (RNAi) begins with recognition of double stranded RNA (dsRNA) by the RNase III family nuclease Dicer. This enzyme cleaves dsRNA into small interfering RNAs (siRNA) 21 – 29 nucleotides in length. The siRNAs are then incorporated into the multicomponent nuclease complex, RNA-induced silencing complex (RISC), which targets specific mRNAs for destruction based on their

homology to the siRNA. Two of the major proteins identified in the RISC complex are Argonaute 2 and Argonaute 3 (Ago 2 and 3) (Carmell 2002).

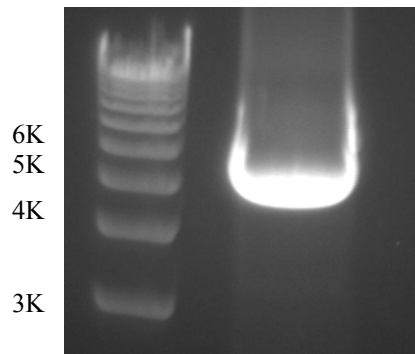


Figure 16 RT-PCR amplification of the full-length coding sequence for bovine Dicer. A 0.5%, ethidium bromide stained agarose gel showing the migration of bovine Dicer amplified from fetal testis cDNA.

To determine if these major components of the RNAi biochemical machinery are present in the bovine model, PCR amplification and sequencing of the bovine Dicer, Ago2 and Ago3 mRNAs was undertaken. Primers designed against homologous regions between the mouse and human cDNA sequences were used in an RT-PCR utilizing bovine testis mRNA (Figure 16). The complete coding sequence for the bovine homologue of Dicer (AY386968) displayed 92% homology on the level of nucleotide sequence and 88% identity on the level of amino acid composition, when compared to the human sequence. Similarly, bovine Ago2 and Ago3 (AY436348) displayed 92% and 94% nucleotide homology and 95% and 99% amino acid similarity, suggesting that these

evolutionarily conserved proteins are present in bovine cells and that the capacity exists to trigger an RNAi based suppression of gene expression.

### **RNAi Knockdown of GFP in Bovine Embryos**

In order to determine if short hairpin mediated RNA interference will effectively suppress gene expression in the bovine model, preliminary trials were conducted using an injection system to deliver a plasmid encoding GFP into 1-cell bovine zygotes in addition to a vector producing an shRNA targeting the GFP mRNA. Both the control and experimental embryos were injected with a plasmid expressing DS Red as a normalization control. In two independent trials, embryos injected with vector carrying GFP alone exhibited the expected green fluorescence, while those injected with vectors carrying both the GFP and shRNAs targeting GFP displayed drastically reduced fluorescence (Figure 17). Results of these experiments strongly suggest that RNAi can be employed in cattle to effectively knock down gene expression using targeted shRNAs.

These experiments also demonstrate that the mouse U6 snRNA promoter can effectively drive the expression of a transgene in bovine cells. Of further interest is the fact that GFP fluorescence was detected as early as the one cell stage. This is significant as transcriptional activity is thought to begin at the eight-cell stage in the bovine (Camous *et al.* 1986; Frei *et al.* 1989; Kopecny *et al.* 1989) although several studies have suggested transcriptional activity as early as the two cell stage (Plante *et al.* 1994; Viuff *et al.* 1996). However, observation of transcription (and translation) of a gene under the control of a snRNA promoter during the one cell stage would suggest that the machinery necessary to achieve transcription is present and active at the one cell stage.

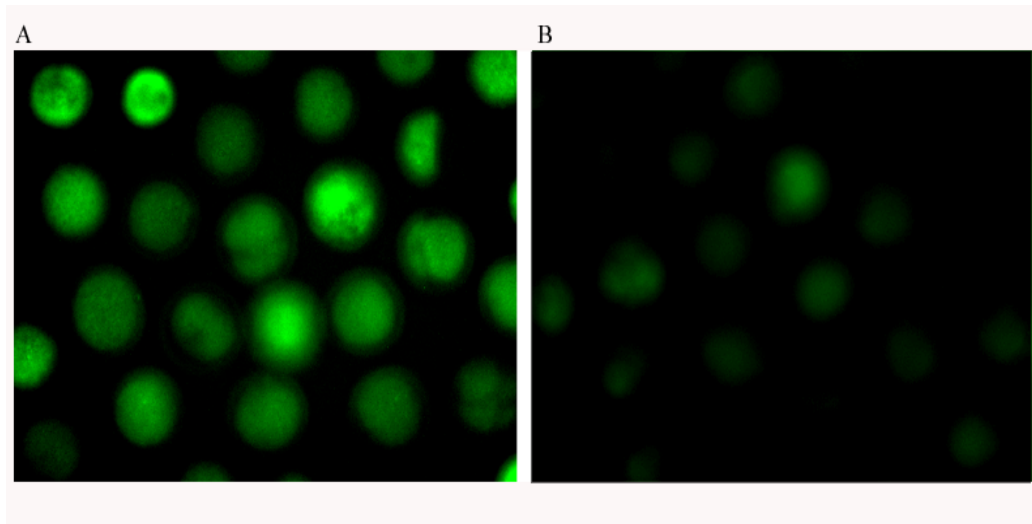


Figure 17 RNAi based suppression of GFP expression in bovine embryos. (A) Bovine 1 and 2-cell stage embryos injected with the Fug-W plasmid expressing GFP under the control of the Mouse U6 promoter. (B) Bovine 1 and 2-cell stage embryos injected with Fug-W and a vector expressing an shRNA homologous GFP under the mouse U6 promoter. A plasmid expressing DS-Red was injected into both experimental groups as a normalization control.

Further, suppression of GFP expression in embryos injected with the shRNA-expressing vector suggests that the mouse H1 promoter is also active during this early stage.

Endogenous H1 RNA forms the RNA component of nuclear RNase P and is transcribed by RNA polymerase III (Pol III) (Myslinski et al 2001). The U6 snRNA gene is also transcribed by Pol III (Domitrovich and Kunkel 2003), which suggests that at least this enzyme is transcriptionally active during this stage. Whether Pol II, the enzyme responsible for the transcription of protein encoding genes is also transcriptionally active will be subject to future investigations.

As development proceeded, GFP fluorescence was eliminated, sometimes as early as the four-cell stage. Whether this is an epigenetic phenomenon or a factor of plasmid degradation remains to be examined.

## **shRNA Mediated Knockdown of the Prion Protein (PrP) – Methods for RNAi Mediated Gene Silencing in an Agricultural Model**

Introduction of shRNAs targeting GFP appeared to effectively suppress the expression of the reporter construct but whether this technique can effectively be used to suppress expression of an endogenous gene remains to be determined. In order to accurately study RNAi based suppression in an agriculturally relevant model, it was desirable to select a target gene that has been well studied, sequenced, is ubiquitously expressed, easily diagnosed on both the level of mRNA and protein and is not essential to cell survival. Based on these criteria, and given the biosecurity issues of working in the bovine, the caprine prion protein was selected due to our experience working with this animal model.

Prion diseases such as mad cow disease, scrapie or Creutzfeldt-Jakob disease are all fatal and are of major agricultural and medical significance. In these diseases the normal cellular prion protein is transformed into an infectious self-propagating disease agent (White *et al.* 2003). The exact nature of this transformation and the natural role of the prion protein have yet to be determined. Recently, studies of prion protein knockout mice revealed that elimination of prion expression conferred disease resistance, and that the animals were phenotypically normal (Sailer *et al.* 1994). Prion proteins (PrP) are expressed in most major cell types and have been extensively studied in numerous model organisms, including the cow and goat. The goat PrP gene has been sequenced and commercial antibodies exist for detection of the protein product. Thus, the PrP gene represents an excellent candidate with which to pioneer studies of RNAi based gene

suppression in an agriculturally relevant model. In addition, creation of a PrP knockdown cell line and animal will likely be of significant research interest.

RNA interference is fast becoming a standard laboratory technique for studies of functional genomics. Using this technique, many genes can be suppressed independently, or in large groups and the biological consequences analyzed. In addition, genes can be suppressed to varying degrees in independent experiments creating epimorphic alleles, which allow simultaneous examination of the biological effects of a partial suppression versus a complete knock down (Hemann *et al.* 2003). Experiments like these may be more informative than simple knockouts based on homologous recombination. Due to the power and versatility of this technique, it is desirable to establish reliable, reproducible methods with which to conduct RNAi based studies of gene expression on multiple genes in a high throughput fashion. To this end, the PrP gene was used to study and establish an RNAi based system in the caprine model that could easily be adapted to study genes of biological interest in other agriculturally relevant species, including cattle. This system allows rapid design and screening of candidate shRNAs, large scale screening of the hairpins utilizing an *in vitro* luciferase reporter assay, *in vitro* suppression of a tagged fusion protein to validate the RNAi molecule and ultimately creation of a transgenic cell line and testing for *in vivo* knock down of gene expression.

### **Large-Scale Design and Synthesis of shRNAs**

Small hairpin RNAs can be designed using a computer program on the Cold Spring Harbor website (<http://www.cshl.edu/public/SCIENCE/hannon.html>) which

searches for sequences that are amenable to forming the hairpin loop structure (Figure 18). The output from this program will contain some base pair changes from the original sequence owing to the fact that it is necessary to change the sequence given that design and propagation of a vector containing an inverted repeat is difficult. Due to the fact that the RNA base uracil will pair with guanine, as stated by the wobble hypothesis, substitutions can be made without altering the functionality of the hairpin. The output from this computer program was taken and modified to include a directional cloning site on the 5' end, PacI restriction sites flanking the entire sequence and a homologous region to the mouse H1 promoter. The modified sequence was then used as a PCR primer to create a short hairpin expressing sequence, under the control of the mouse H1 promoter. Following the initial PCR reaction, sequences were cut from the gel, purified and directionally cloned into the gateway vector pENTR-D. The PacI sites allowed for easy diagnosis of positive clones, which were then sequenced to verify the shRNA and promoter sequence. Positive clones were used in a clonase reaction, transferring the shRNA and promoter sequence to the modified destination vector Fug-W. Fug-W is a lenti-viral vector containing GFP under the U6 promoter and AttI sites allowing recombinatorial insertion of gene sequences via recombination. At the end of this cloning procedure a vector expressing GFP and a shRNA targeted to our gene of interest was produced. All of the above procedures can be done in 96 well format and is thus viable for high throughput studies.

### ***In vitro* Luciferase Assay – Ligation Independent Cloning of a Reporter Construct**

In order to rapidly assess that capacity of a large number of shRNAs to suppress the expression of our genes of interest, a method was needed to create a reporter construct that would easily allow introduction of coding sequences in a high throughput way. To this end, ligation independent cloning was used to insert the coding sequence for the PrP and Dnmt1 gene into a vector expressing firefly luciferase. The coding sequences were inserted into the 3' untranslated region of the luciferase transcript, such that RNAi mediated suppression of our targeted gene sequence would lead to a concurrent suppression of luciferase activity as well. This strategy of utilizing a luciferase fusion transcript to diagnose siRNA activity has been effectively used to identify RNA sequences that elicit maximal response (Yu *et al.* 2003).

Ligation independent cloning utilizes a DNA (or RNA) overhang to join two separate DNA molecules together *via* homologous base pairing between the “sense” and “antisense” overhangs (Coljee *et al.* 2000). Here, DNA overhangs were introduced onto the 3' end of luciferase and the 5' end of PrP, our gene of interest. This was accomplished by the introduction of a 2'-O- methyl RNA base in the primer sequence, proximal to the overhang sequence. This modified RNA base is refractory to RNase H digestion, and causes DNA polymerase to terminate polymerization leaving a single stranded DNA overhang. When incorporated into the final plasmid, the bacterial DNA repair machinery can replace the modified RNA base with a normal DNA base and thus allow normal DNA replication to proceed. Two linker overhang sequences were tested here, one allowing the formation of a luciferase–PrP fusion protein and another that

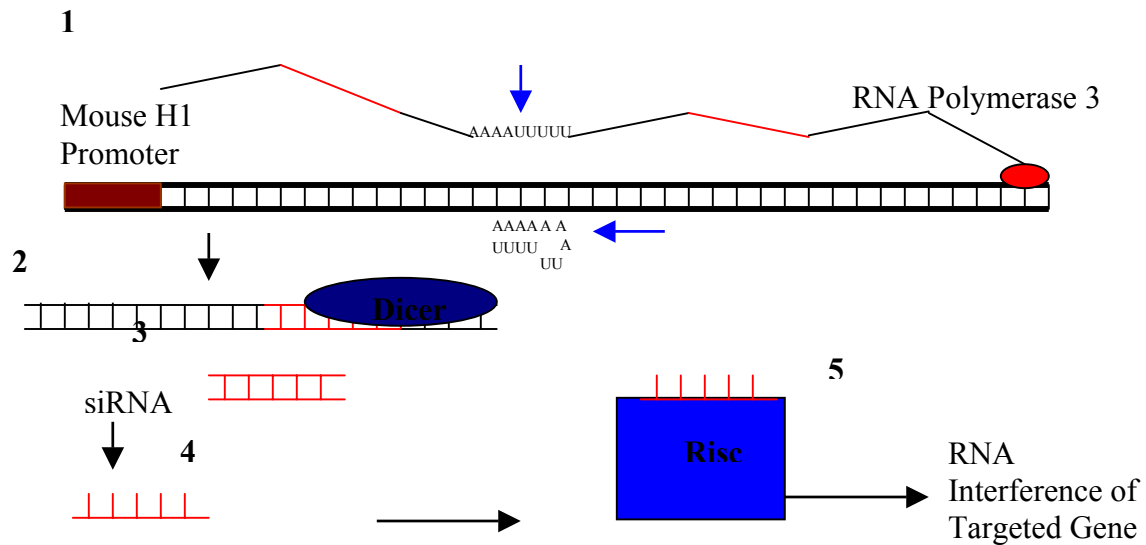


Figure 18 Experimental model. **1)** The murine H1 tRNA promoter will drive the production of an RNA transcript containing a specific sequence, homologous to the targeted gene. This sequence will be arranged in tandem with the “sense” sequence followed by the “antisense” sequence. **2)** In the center of the RNA transcript, between the sense and antisense sequences, there will be a stretch of adenine residues immediately followed by a stretch of uracil residues (indicated by the blue arrow). Base pairing between these residues and the sense and antisense sequences facilitates the formation of a short RNA hairpin loop. **3)** These short hairpins mimic endogenous shRNAs and trigger the RNase III molecule Dicer to attack and cleave the hairpin into 21-29 nucleotide sequences, resulting in the production of small interfering RNAs (siRNAs). **4)** These siRNA molecules will then be taken up by a large protein complex called RNA Induced Silencing Complex (or RISC) that facilitates the selective destruction of the targeted gene. GFP is also expressed from a downstream U6 promoter allowing easy diagnosis of cells expressing the shRNAs.

inserted a stop site between the two gene sequences creating a “fusion transcript” that produced a normal luciferase protein. A directional cloning site was inserted onto the 5’ end of luciferase to allow the plasmid construct to be cloned in the correct orientation. A schematic diagram in figure 19 outlines the procedure discussed.

Using the primers discussed above, PCR reactions were conducted on the pGL3 plasmid (Promega) and goat testis cDNA. The amplified Luciferase and PrP sequences, containing the modified DNA overhangs were gel purified and mixed in equal molar ratios (Figure 19 B and C). DNA was incubated at 94°C for 5 minutes and then allowed to cool to room temperature over the course of two hours. The annealed DNA was then used in a Topo cloning reaction and placed in the gateway vector pENTR-D. Positive clones were selected and sequenced and overall, the cloning reaction had a background of approximately 25% non-positive colonies. From here, the luciferase-PrP fusion was moved to a vector containing the CMV promoter and grown in sufficient quantities for use in the in vitro reporter assay (Figure 19 D).

### ***In Vitro Knockdown of the Luciferase-PrP Reporter***

Using the strategies discussed above, shRNAs targeting the bovine PrP mRNA and a luciferase-PrP reporter construct were created. In order to test whether fusion of the PrP protein sequence to luciferase affected luminescence, two plasmids were constructed, one containing the luciferase-PrP fusion protein coding sequence and another that has been termed a “fusion transcript” containing a stop codon between the two reading frames. Both plasmids were transfected into LynxA cells and luciferase activity read on a luminometer. Renilla luciferase activity was measured and served as a normalization control. Both plasmids displayed luciferase activity however; the luminescence of the fusion protein was greatly reduced when compared to the transcript containing the stop codon. This reduction of luminescence by fusion of the two coding sequences together may affect experimental measurements, thus all further experiments

were conducted using the “fusion transcript” plasmid containing the stop codon between the two coding sequences.

A total of 42 shRNAs were designed against the PrP mRNA and randomly pooled into 6 groups of 7. Co-transfection of the reporter and with a single pool of the shRNA expressing plasmids into Lynx A cells using calcium phosphate, allowed diagnosis of and selection of the pool containing the hairpin eliciting maximal gene suppression. A Small hairpin RNA designed against bovine Dnmt1 served as a control. Luciferase activity as compared to the control was reduced in all pools examined, however, pools 6 and 7 displayed the maximal reduction (Figure 20). Pool 7 exhibited a greater than 5 fold reduction in luciferase activity and was thus selected for expansion into its eight individual plasmids. These individual plasmids were then used in a similar assay to identify the individual shRNA sequence that was most effective in suppressing reporter gene expression.

The shRNA producing maximal reporter gene suppression was then used to knock down a T7 tagged fusion protein. A tagged protein was used to prevent endogenous PrP from interfering with the reporter assay. Co-transfection of the T7 tagged caprine PrP with the vector expressing GFP and the shRNA targeting bovine PrP enabled diagnosis of shRNA effect upon PrP protein translation. Western blot analysis was conducted using a V5 antibody to measure the recombinant PrP protein and an antibody to GFP used to normalize the data allowing for differing transfection efficiencies experiment to experiment.

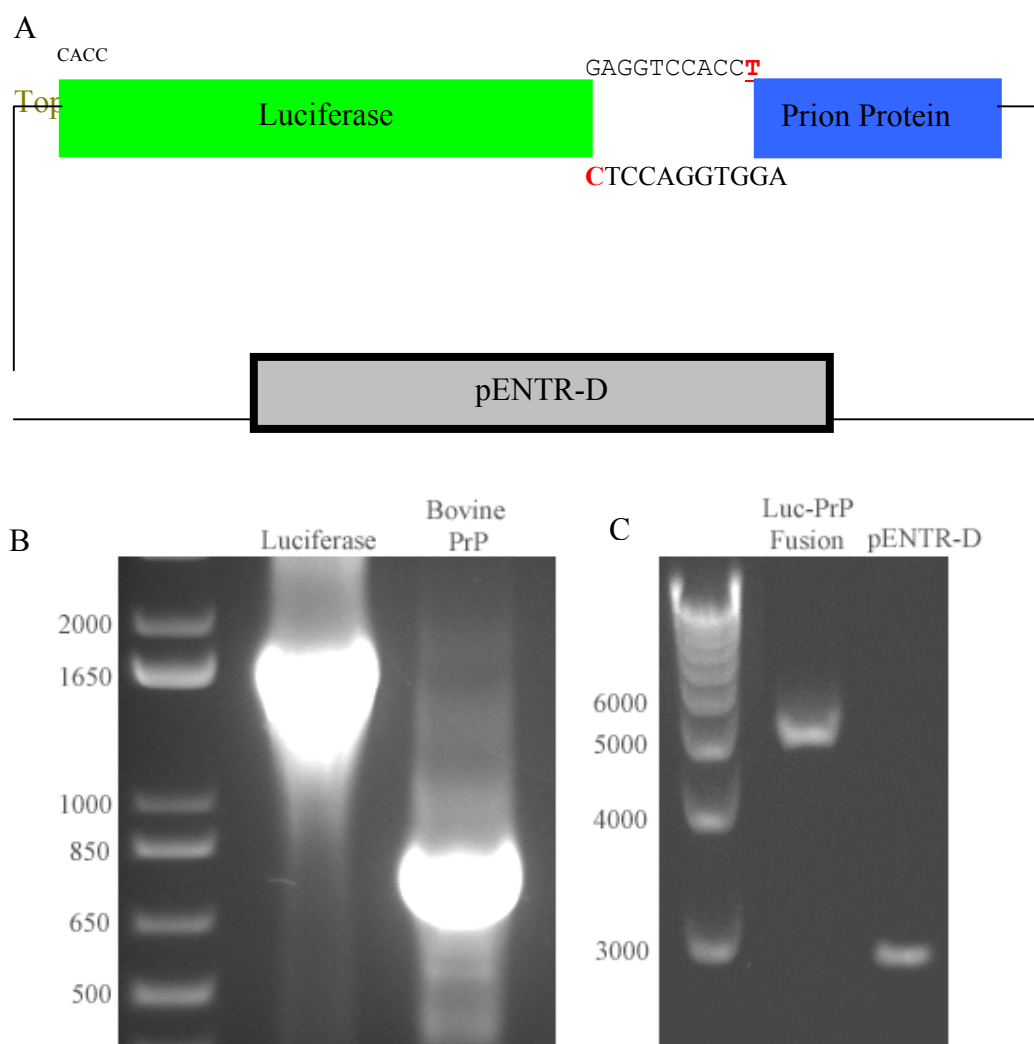


Figure19 LIC Procedure. (A) Schematic diagram of the LIC procedure used to construct the Luciferase-PrP fusion reporter construct. PCR was used to insert a directional cloning site upstream of the firefly luciferase start site and an RNA base (in red) followed by a linker overhang inserted downstream of the luciferase stop site. The complement to the overhang followed by an RNA base (in red) was inserted in front of the PrP coding sequence. (B) PCR amplification of the Luciferase and PrP coding sequences using the LIC primers. (C) LIC cloning of Luciferase-PrP reporter. Lane 1 is the DNA marker, lane 2 the assembled reporter construct and lane 3 the pENTR-D backbone alone. (D) (next page) Plasmid map of the Luciferase-PrP reporter construct.

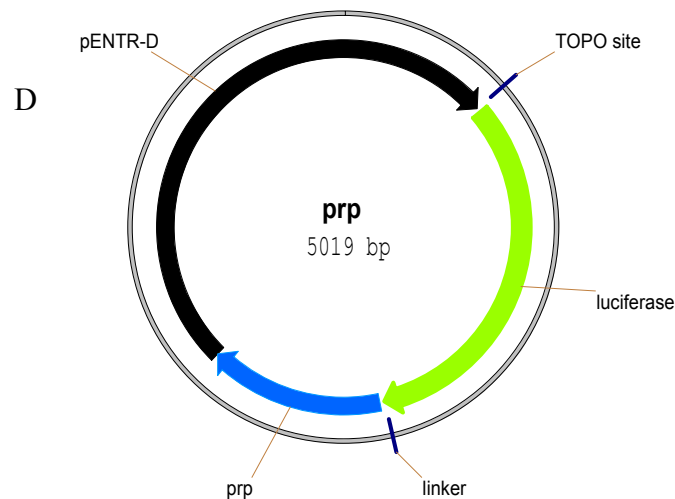


Figure 19 Continued.

### Lenti-viral Mediated Transgenesis

Identification of a PrP suppressing shRNA enables the creation of a transgenic animal that will effectively knock down the PrP protein using an RNAi based approach. Previous work using recombinant lentiviral vectors has demonstrated the efficacy of creating founder animals carrying a transgene in one generation. In these studies, viral vectors carrying GFP were injected into the perivitelline space of 1-cell mouse embryos and then transferred into recipient females. This approach proved to be very efficient with 76% of the resulting pups exhibiting GFP fluorescence (Lois *et al.* 2002). In addition, the transgene was transmitted to progeny of the founder animals. A similar approach has been used to effectively deliver shRNAs into mouse preimplantation embryos that target (knock down) the expression of a specific gene (Rubinson *et al.*

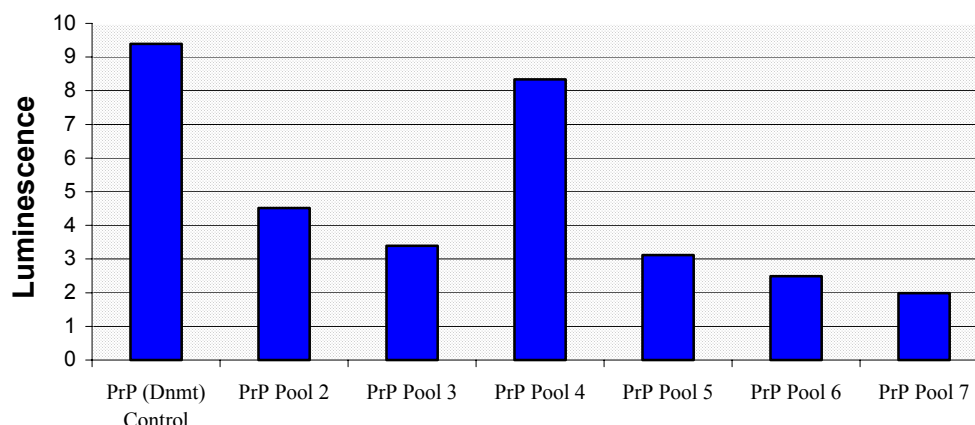


Figure 20 Inhibition of Luciferase activity using shRNAs targeting a Luciferase-PrP transcript. Luciferase activity was assayed 3 days post-transfection using a luminometer. Control cells were transfected with an shRNA molecule targeting bovine Dnmt1. Experimental cells were transfected with pools of shRNAs targeting various regions of the PrP coding sequence. Experiments consisted of three independent trials and six luminescence readings per sample. Cells transfected with Pools 6 and 7 elicited the greatest amount of reporter suppression and were thus selected for further experimentation.

2003). This strategy is currently being employed to create transgenic shRNA expressing animals that suppress some of the genes discussed herein but is not the subject of this dissertation.

## Discussion

The development of reliable methods to accurately diagnose shRNA function is a critical element to establishing this technology for use in functional genomics. Study of shRNA mediated suppression of the goat PrP has revealed that large-scale synthesis and testing of shRNAs targeting a gene of interest is possible utilizing common laboratory

procedures. This greatly extends the scope of molecular studies, which can be utilized in the caprine and other agriculturally relevant species and opens the door for use of RNAi in the creation of disease resistant livestock. Further, it provides a secure platform from which to launch RNAi based studies of bovine Dnmt1.

### **shRNA Mediated Knockdown of Dnmt1**

In order to determine what role, if any the inappropriate expression of Dnmt1 reported above has on the genomic hypermethylation observed in cloned bovine embryos, methods using RNA interference to transiently disrupt Dnmt1 expression were developed. To this end shRNAs homologous to regions of the bovine mRNA sequence were constructed using methods described above. Current literature supports the hypothesis that siRNA sequences designed using sequences near the 5' end of a gene work the best. Four shRNA molecules targeting this region were selected, cloned into Fug-W and tested using the luciferase reporter assay described above. Of these four shRNAs, the first two, targeting bases 5-34 and 623-652 produced a marginal knockdown while shRNA three and four targeting bases 1082-1111 and 1789-1818 produced a much more pronounced effect (data not shown). Subsequently, shRNA3 and shRNA4 were selected for further testing (Figure 21).

In order to determine what affect shRNA3 and shRNA4 have upon translation of the bovine DNMT1 protein, the coding sequence for the first two thirds of the gene was cloned into the pcDNA3 backbone, with a T7 tag appended to the 5' end, just upstream of the natural Dnmt1 start site. Design of this vector specifically containing an incomplete Dnmt1 protein prevents any of the toxic effects associated with

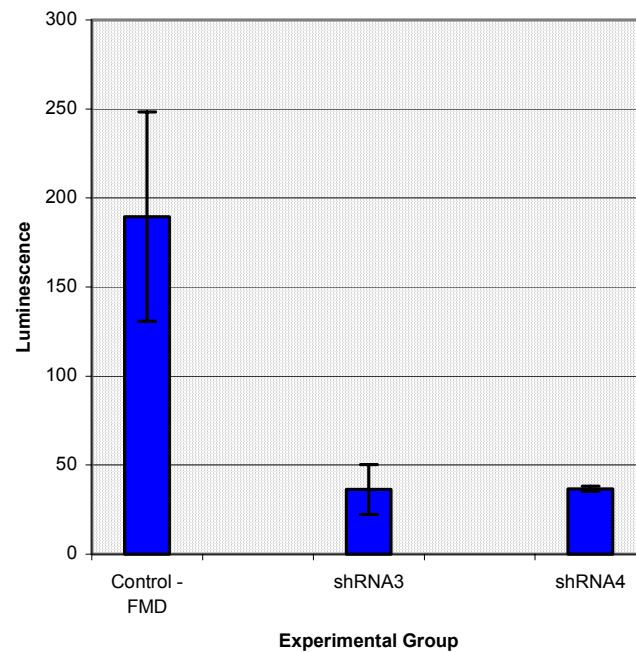


Figure 21 Inhibition of Luciferase activity using shRNAs targeting a Luciferase-Dnmt1 target transcript. Luciferase activity was assayed 3 days post-transfection using a luminometer. Control cells were transfected with an shRNA molecule targeting the 5'UTR of the FMD virus. Experimental cells were transfected with shRNAs targeting various regions of the Dnmt1 coding sequence. Experiments consisted of three independent trials and six luminescence readings per sample. It is likely that the inability of the shRNAs to completely eliminate reporter luminescence is simply due to the difference in the strengths of the CMV and H1 promoters driving reporter and shRNA expression.

overexpression of the full-length protein (Biniszkiewicz *et al.* 2002). This construct was transfected into LynxA cells as above, protein extract collected after 3 days and probed with antibodies recognizing the T7 tag and GFP. Probing the blot with a GFP antibody serves to normalize the data for differing transfection efficiencies between experimental

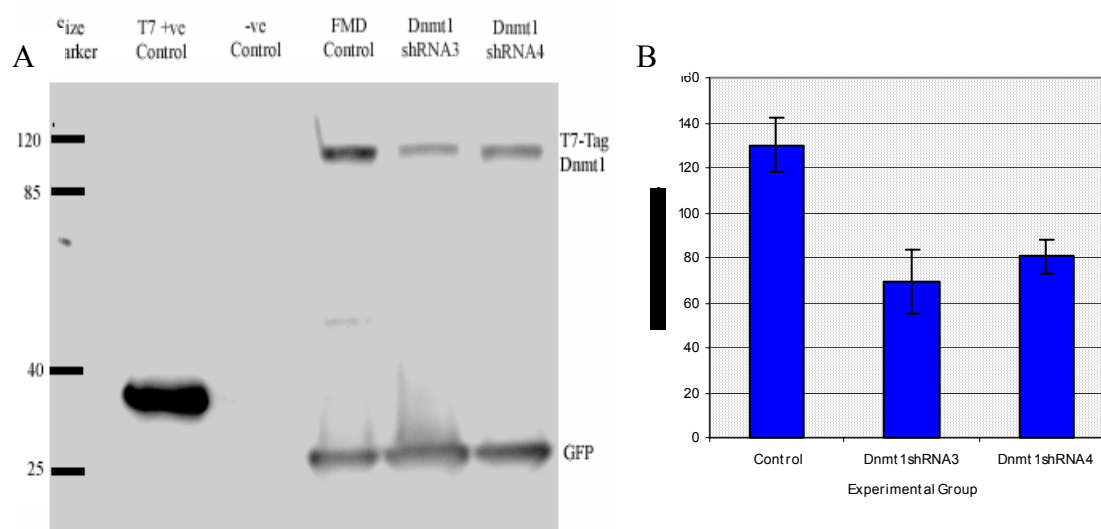


Figure 22 shRNA mediated knock down of an amino terminal, T7 tagged Dnmt1 protein. A) Western blot of protein extracted from Lynx A cells expressing bovine Dnmt1 containing an amino terminal T7 tag, along with the Fug-W plasmid producing shRNAs targeting various regions of the Dnmt1 coding sequence. Blots were probed with an antibody recognizing the T7 antigen and another recognizing GFP (26.8 Kda) to normalize for differing transfection efficiencies. Lane 1 is the molecular marker, lane 2 the 31.1 Kda T7 positive control and lane 3 a negative control using protein extract from untransfected cells. Lane 4 is a control containing protein from cells transfected with T7-tagged Dnmt1 and an unrelated hairpin to the FMD virus. Lanes 5 and 6 contain extract from cells transfected with T7 tagged Dnmt1 as well as plasmids expressing shRNA3 and shRNA4 respectively, both targeting bovine Dnmt1. B) Densitometry analysis of the Western blot showing levels of T7-Dnmt1 expression as a ratio normalized to GFP. Both shRNA3 and shRNA4 produced a reduction in T7-Dnmt1 expression. Four independent trials were conducted and results presented in figure B represent the average densitometry reading. The inability of the shRNAs to completely eliminate the expression of the T7-tagged Dnmt1 is likely a factor of protein stability and again, due to the difference in the strengths of the CMV and H1 promoters driving reporter and shRNA expression.

groups (Figure 22). Results from these experiments clearly show that shRNA3 and shRNA4 are capable of eliciting an RNAi based suppression of Dnmt1.

## Discussion

These experiments show that several of the key biochemical elements necessary to elicit an RNAi based suppression are present in the bovine and caprine models and further that shRNAs can be used to effectively knock down a targeted gene *in vitro*. The next phase of this project will be to test the capacity of these interfering RNAs to modulate the hypermethylation frequently seen in cloned cattle. Given that Dnmt1 is thought to be the predominant methyltransferase and that its inappropriate regulation during the initial stages of clone development could prevent the demethylation necessary to reset the genome, targeting this gene for a transient disruption is a logical choice. Further, given the divergent expression profile for Dnmt expression in the preimplantation bovine embryo, it will be interesting to see if permanent disruption of Dnmt1 elicits the same effect in cattle as it does in the mouse.

## Materials and Methods

### Design and Cloning of Small Hairpin RNAs

Gene sequences were run through a computer algorithm on the world wide web (<http://www.cshl.edu/public/SCIENCE/hannon.html>) and the output modified in the following: the compliment to the mouse H1 promoter was added to the 3' end of the sequence, a directional cloning site and Pac1 restriction site added to the 5' end and the sequence for the U6 promoter removed. These primer sequences were synthesized by Sigma-Genesis (Woodlands TX) and used as the reverse primer in a reaction amplifying the mouse H1 promoter. The sense primer contains a Pac1 restriction site followed by the forward H1 promoter sequence. Reactions were set up using the NEB ThermoPol

reaction buffer (10 mM KCl, 10 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 20 mM Tris-HCl pH 8.8, 2 mM MgSO<sub>4</sub>, 0.1% Triton X-100), 10 mM dNTP, 4% DMSO, 50pmols of each primer and 2 units of *Taq* polymerase. Thirty cycles of 94° C for 30sec, an annealing temperature of 55° C and a 30 second 72° C extension were used to amplify the PCR product. Samples were run on a 2.5% agarose gel, cut and purified using a Quiagen Quiax II gel extraction kit, according to the manufactures recommendations. Purified product (3ul) was then seeded into a Topo cloning reaction and placed into the pENTR-D Gateway vector using the manufactures protocol. Plasmid DNA was used to transform Top10 bacteria, which were grown on kanamycin selective medium. Colonies were selected, grown and plasmid purified using the Quiagen mini-prep system, according to the manufactures recommendations. Samples were digested with the PacI restriction endonuclease and the 180 base pair H1 promoter-shRNA construct separated from the vector backbone on a 2% gel.

Positive clones were sequenced at the Texas A&M Gene Technology Lab on an Applied Biosystems 3100 Gene Analyzer (Applied Biosystems). Sequenced clones were then used in a Clonase reaction (Invitrogen) transferring the H1 Promoter-shRNA construct to a modified Fug-W plasmid (Lois *et al.* 2002) containing the Gateway Recombinase AttI sites, according to the manufactures protocol. Stbl2 bacteria (Invitrogen) were transformed with recombinant Fug-W and grown under ampicillin selection. Colonies were selected, grown and diagnosed as above. Positive clones were pooled or used individually in *in vitro* reporter assays described below. Fug-W expresses the Green florescent protein (GFP) under the control of the U6 promoter and

the shRNA targeting the gene of interest under the H1 promoter. GFP florescence was used to assay transfection efficiency using standard microscopy.

#### Ligation Independent Cloning of Reporter Constructs

Primers homologous to the 3' end of Luciferase and the 5' end of bovine PrP were synthesized with a modified 2'-O-Methyl RNA residue preceding a linker sequence (Sigma-Genesis, Woodlands TX). Sequences given in Table 2. Two linker sequences were tested here; one, which creates a fusion protein between Luciferase and PrP, and another that, contains a stop site and produces the normal luciferase protein. The primers listed below were used to amplify the LIC ready Luciferase and PrP templates using *Taq* polymerase (New England Biosciences). The pGL3 plasmid (Promega) and bovine testis cDNA were used as reaction template for Luciferase and PrP, respectively.

Gene	Primer Sequence
Luciferase 5'	<u>CACCATGGAAGACGCCAAAA</u>
Luciferase 3'	AGGTGGACCTCCCGGAGGACACGGCGATCTTTCCGCCCTT
Fusion	
Luciferase 3' Stop	AGGTGGACCTCCCGGAGGTTACACGGCGATCTTTCCGCC
PrP 5'	CCTCCGGGAGGTCCACCTATGGTGAAAAGCCACATAGG
PrP 3'	ACTATCCTACTATGAGAAAAATGAGG

Table 2 Sequences of Primers Used for Ligation Independent Cloning. Base in Red is the modified 2'-O-Methyl RNA base. Underlined sequence in Luciferase 5' primer is the directional cloning sequence necessary for directional insertion into pENTR-D.

Equimolar amounts of amplified Luciferase and PrP were combined in 50ul of ligation buffer (132 mM Tris-HCl, 20 mM MgCl<sub>2</sub> 2 mM dithiothreitol, 2 mM ATP, 15% Polyethylene glycol pH 7.6) New England BioSciences), heated to 94 degrees for 5

minutes and allowed to cool down to room temperature over the course of two hours. A small aliquot (3-5ul) was used in a Topo-cloning reaction placing the construct into the pENTR-D Gateway vector, per the manufactures recommendations (Invitrogen). Top10 bacteria were transformed, and grown on kanamycin plates. Colonies were selected, grown, and plasmid isolated using the Quiagen mini-prep plasmid isolation system, per the manufactures recommendation. An EcoR1 restriction digest was used to identify positive clones, which were selected and used in a Clonase reaction, moving the reporter gene into the pcDNA 3.1 backbone. Bacteria were grown on ampicillin media and colonies screened as above. A single positive clone was selected for each of the two plasmid designs (fusion protein and fusion transcript) grown and plasmid DNA isolated to a yield of sufficient quantity for *in vitro* testing of the reporters.

#### Cloning of Tagged Proteins

The coding sequence of bovine PrP was amplified using PCR primers containing the directional cloning sequence listed above. PCR product was isolated and used in a directional Topo reaction placing the PrP sequence into the Gateway vector pENTR-D. Positive clones were used in a Clonase reaction transferring the coding sequence to the pCDNA 3.1 – V5 Destination vector. The vector adds an amino terminal V5 fusion tag.

Dnmt1 was amplified in three pieces from an RT-PCR using bovine testis cDNA. An EcoR1 site and T7 Tag were added to the 5' end of the Dnmt1 coding sequence. A Not1 site was added to the 3' terminus and endogenous BamH1 and HindIII restriction sites in the central coding sequence used to assemble the full length Dnmt1 into pcDNA 3.0 (Invitrogen).

## Cell Culture and Transfection

LynX A viral packaging cell lines (Hannon Lab) were grown in DMEM – 5% BSA. Cells were split prior to transfection using a Calcium Phosphate transfection protocol. Briefly, for a 10cm dish, 30ug of total DNA were mixed into 900 ul of H<sub>2</sub>O, 100 ul of 0.25M Calcium Chloride in Hepes and 1 ml of BBS (50 mM BES, pH 6.95, 280 mM NaCl, 1.5 mM Na<sub>2</sub>HPO<sub>4</sub>). Mixed transfection reagent was added drop-wise to the plates and allowed to equilibrate at room temperature for 5 minutes. Cells were grown overnight, then washed with PBS and allowed to grow for two more days before cells were harvested.

For the Luciferase Reporter Assay, 15ug of salmon sperm DNA were transfected with 7.5ug of the shRNA containing Fug-W, 0.75ug Renilla expressing pRL SV40 (Promega) and 6.75ug of the Firefly Luciferase reporter construct (9:1 ratio of Firefly to Renilla). For fusion-protein analysis, 15ug of salmon sperm DNA was transfected with 7.5ug of the shRNA containing Fug-W and 7.5ug of the vector containing the fusion protein.

## Western Blotting

Cells were collected using a standard lysis buffer (50mM Tris pH 7.8, 150 mM NaCl, 1% Nonidet P-40 and 1ml Protease Inhibitor (Calbiochem) per 100 ml of buffer) and protein extracted by centrifugation. Samples were loaded onto a 4-15% polyarylamide gel and separated prior to transfer to PVDF paper (BioRad). Protein blots were blocked in 10% goat serum, 3% milk and washed in TTBS.

T7 tag, (Novagen) V5 tag (Invitrogen) and GFP (AbCam) antibodies were used to probe the blots to examine protein quantities and results analyzed using the SuperSignal Chemiluminescent assay (Pierce).

#### Luciferase Reporter Assay

The Promega Dual-Luciferase Reporter Assay System was used to quantify shRNA mediated reporter gene knock down. Briefly, a passive protein lysis buffer was used to collect protein from the transfected cells. A small aliquot (1-5ul) of the cell lysate was added to 100 ul of Luciferase Assay Reagent II and mixed by pipetting. A reading was then recorded from the luminometer. After the initial reading, 100 ul of Stop and Glo Reagent was added and the sample vortexed. The second reading was then recorded. The Firefly luminescence was then divided by the Renilla luminescence and a ratio used to compare samples. A minimum of 4 independent readings were taken for each sample and students T-test used to affirm significance.

## CHAPTER V

### SUMMARY

The results reported herein demonstrate that the DNA methyltransferase expression profile of the early bovine preimplantation embryo is divergent from the paradigm reported in the mouse. In addition, real time quantitative analysis of the Dnmts in cloned embryos suggests that misexpression of these enzymes is not solely responsible for the hypermethylation consistently seen in cloned embryos.

These experiments demonstrate that several of the key biochemical elements necessary to elicit an RNAi based suppression are present in the bovine model and that shRNAs can be used to effectively knock down a targeted gene *in vitro*. The next phase of this project will be to test the capacity of these interfering RNAs to modulate the hypermethylation frequently seen in cloned cattle. Given that Dnmt1 is thought to be the predominant methyltransferase and that its inappropriate regulation during the initial stages of clone development could prevent the demethylation necessary to reset the genome, targeting this gene for a transient disruption is a logical choice. Further, given the divergent expression profile for Dnmt expression in the preimplantation bovine embryo, it will be interesting to see if permanent disruption of Dnmt1 elicits the same effect in cattle as it does in the mouse.

## REFERENCES

- Akker, S., Smith P., Chew S. (2001) Nuclear post-transcriptional control of gene expression. *J Mol Endocrinol.* **27**:123-31.
- Bachman, K., Rountree, M., Baylin, S. (2001) Dnmt3a and Dnmt3b are transcriptional repressors that exhibit unique localization properties to heterochromatin. *J Biol Chem.* **276**:32282-7.
- Baguisi, A., Behboodi, E., Melican, D., Pollock, J., Destrempe, M., Cammuso, C., Williams, J., Nims, S., Porter, C., Midura, P., Palacios, M., Ayres, S., Denniston, R., Hayes, M., Ziomek, C., Meade, H., Godke, R., Gavin, W., Overstrom, E., Echelard, Y. (1999) Production of goats by somatic cell nuclear transfer. *Nat Biotechnol.* **17**:456-61.
- Barker, D. (1990) The fetal and infant origins of adult disease. *BMJ* **301**:1111.
- Beckett, D. (2001) Regulated assembly of transcription factors and control of transcription initiation. *J Mol Biol.* **314**:335-52.
- Bernstein, E., Denli, A., Hannon, G. (2001) The rest is silence. *RNA* **7**:1509-21.
- Bestor, T. (1992) Activation of mammalian DNA methyltransferase by cleavage of a Zn binding regulatory domain. *EMBO J.* **11**:2611-7.
- Bestor, T. (2000) The DNA methyltransferases of mammals. *Hum Mol Genet* **9**:2395-402.
- Bestor, T. and Verdine, G. (1994) DNA methyltransferases. *Curr Opin Cell Biol.* **6**:380-9.
- Bigey, P., Ramchandani, S., Theberge, J., Araujo, F., Szyf, M. (2000) Transcriptional regulation of the human DNA methyltransferase (dnmt1) gene. *Gene.* **242**:407-18.
- Biniszkiewicz, D., Gribnau, J., Ramsahoye, B., Gaudet, F., Eggan, K., Humpherys, D., Mastrangelo, M., Jun, Z., Walter, J., Jaenisch, R. (2002) Dnmt1 overexpression causes genomic hypermethylation, loss of imprinting, and embryonic lethality. *Mol Cell Biol* **22**:2124-35.
- Bird, A. & Wolffe A. (1999) Methylation-induced repression--belts, braces, and chromatin. *Cell* **99**:451-4.
- Bourc'his, D., Xu, G., Lin, C., Bollman, B., Bestor, T. (2001a) Dnmt3L and the establishment of maternal genomic imprints. *Science* **294**:2536-9.

- Bourc'his, D., Le Bourhis, D., Patin, D., Niveleau, A., Comizzoli, P., Renard, J., Viegas-Pequignot, E. (2001b) Delayed and incomplete reprogramming of chromosome methylation patterns in bovine cloned embryos. *Curr Biol* **11**:1542-6.
- Brown, R. & Strathdee, G. (2002) Epigenomics and epigenetic therapy of cancer. *Trends Mol Med* **8**(4 Suppl):S43-8.
- Caceres, J. & Kornblihtt, A. (2002) Alternative splicing: multiple control mechanisms and involvement in human disease. *Trends Genet.* **18**:186-93.
- Camous, S., Kopecny, V., Flechon, J. (1986) Autoradiographic detection of the earliest stage of [<sup>3</sup>H]-uridine incorporation into the cow embryo. *Biol Cell* **58**:195-200.
- Carlson, L., Page, A., Bestor, T., 1992. Properties and localization of DNA methyltransferases in preimplantation mouse embryos: implications for genomic imprinting. *Genes & Development* **6**:2536-2541.
- Chen, T., Ueda, Y., Xie, S., Li, E. (2002) A novel dnmt3a isoform produced from an alternative promoter localizes to euchromatin and its expression correlates with active de novo methylation. *J Biol Chem* **277**:38746-54.
- Chen, T., Ueda, Y., Dodge, J., Wang, Z., Li, E. (2003) Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. *Mol Cell Biol.* **23**:5594-605.
- Chuang, L., Ng, H., Chia, J., Li, B. (1996) Characterisation of independent DNA and multiple Zn-binding domains at the N terminus of human DNA-(cytosine-5) methyltransferase: modulating the property of a DNA-binding domain by contiguous Zn-binding motifs. *J Mol Biol.* **257**:935-48.
- Chung, Y., Ratnam, S., Chaillet, J., Latham, K. (2003) Abnormal regulation of DNA methyltransferase expression in cloned mouse embryos. *Biol Reprod.* **69**:146-53.
- Cibelli, J., Stice, S., Golueke, P., Kane, J., Jerry, J., Blackwell, C., Ponce de Leon, F., Robl, J. (1998) Cloned transgenic calves produced from nonquiescent fetal fibroblasts. *Science.* **280**:1256-8.
- Coljee, V., Murray, H., Donahue, W., Jarrell, K. (2000) Seamless gene engineering using RNA- and DNA-overhang cloning. *Nat Biotechnol.* **18**:789-91.
- Comfort, N. (1999) "The real point is control": the reception of Barbara McClintock's controlling elements. *J Hist Biol.* **32**:133-62.

- Daniels, R., Hall, V., French, A., Korfiatis, N., Trounson, A. (2001) Comparison of gene transcription in cloned bovine embryos produced by different nuclear transfer techniques. *Mol Reprod Dev* **60**:281-8.
- Datta, J., Ghoshal, K., Sharma, S., Tajima, S., Jacob, S. (2003) Biochemical fractionation reveals association of DNA methyltransferase (Dnmt) 3b with Dnmt1 and that of Dnmt 3a with a histone H3 methyltransferase and Hdac1. *J Cell Biochem.* **88**:855-64.
- 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. *Proc Natl Acad Sci USA* **98**:13734-8.
- Denli, A. & Hannon, G. (2003) RNAi: an ever-growing puzzle. *Trends Biochem Sci.* **28**:196-201.
- De Sousa, P., Valdimarsson, G., Nicholson, B., Kidder, G., 1993. Connexin trafficking and the control of gap junction assembly in mouse preimplantation embryos. *Development* **117**:1355-1367.
- De Sousa, P., Winger, Q., Hill, J., Jones, K., Watson, A., Westhusin, M. (1999) Reprogramming of fibroblast mRNA expression following nuclear transfer in bovine embryos. *Cloning* **1**:63-69.
- Ding, F. & Chaillet, J. (2002) In vivo stabilization of the Dnmt1 (cytosine-5)-methyltransferase protein. *Proc Natl Acad Sci U S A* **99**:14861-6.
- Ding, F., Patel, C., Ratnam, S., McCarrey, J., Chaillet, J. (2003) Conservation of Dnmt1 to cytosine methyltransferase in the marsupial *Monodelphis domestica*. *Genesis*. **36**:209-13.
- Doherty, A., Mann, M., Tremblay, K., Bartolomei, M., Schultz, R. (2000) Differential effects of culture on imprinted H19 expression in the preimplantation mouse embryo. *Biol Reprod* **62**:1526-35.
- Domitrovich, A. & Kunkel, G. (2003) Multiple, dispersed human U6 small nuclear RNA genes with varied transcriptional efficiencies. *Nucleic Acids Res.* **31**:2344-52.
- Donovan, D., Powell, A., Talbot, N., Persel, V., Graninger, P., FosterFrey, J., Becker, K., Dawson, H., Wall, R. (2003) cDNA Microarray Analysis is Not Predictive of Donor Cell Clonability in Bovine Somatic Cell Cloning. *Transgenic Animal Research Conference IV*, University of California Davis, August 10<sup>th</sup> – 14<sup>th</sup>.

- Dvir, A., Conaway, J., Conaway, R. (2001) Mechanism of transcription initiation and promoter escape by RNA polymerase II. *Curr Opin Genet Dev.* **11**:209-14.
- Elbashir S., Harborth J., Lendeckel W., Yalcin A., Weber K., Tuschl T. (2001) Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells, *Nature* **411**: 494-498.
- Elbashir, S., Harborth, J., Weber, K., Tuschl, T. (2002) Analysis of gene function in somatic mammalian cells using small interfering RNAs. *Methods* **26**:199-213.
- Fatemi, M., Hermann, A., Pradhan, S., Jeltsch, A. (2001) The activity of the murine DNA methyltransferase Dnmt1 is controlled by interaction of the catalytic domain with the N-terminal part of the enzyme leading to an allosteric activation of the enzyme after binding to methylated DNA. *J Mol Biol.* **309**:1189-99.
- Fire, A., Xu, S., Montgomery, M., Kostas, S., Driver, S., Mello, C. (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**: 806-811.
- Frei, R., Schultz, G., Church, R. (1989) Qualitative and quantitative changes in protein synthesis occur at the 8-16-cell stage of embryogenesis in the cow. *J Reprod Fertil* **86**:637-41.
- Fuks, F., Burgers, W., Godin, N., Kasai, M., Kouzarides, T. (2001) Dnmt3a binds deacetylases and is recruited by a sequence-specific repressor to silence transcription. *EMBO J* **20**:2536-44.
- Garry, F., Adams, R., Cattell, M., Dinsmore, R. (1996) Comparison of passive immunoglobulin transfer to dairy calves fed colostrum or commercially available colostrum-supplement products. *J Am Vet Med Assoc.* **208**:107-10.
- Gibbons, R., Picketts, D., Villard, L., Higgs, D. (1995) Mutations in a putative global transcriptional regulator cause X-linked mental retardation with alpha-thalassemia (ATR-X syndrome). *Cell* **80**:837-45.
- Gitlin, L., Karelsky, S., Andino, R. (2002) Short interfering RNA confers intracellular antiviral immunity in human cells, *Nature* **418**: 430-434
- Gonda, K., Fowler, J., Katoku-Kikyo, N., Haroldson, J., Wudel, J., Kikyo, N. (2003) Reversible disassembly of somatic nucleoli by the germ cell proteins FRGY2a and FRGY2b. *Nat Cell Biol.* **5**:205-10.

- Hanamura, A., Caceres, J., Mayeda, A., Franza, B., Krainer, A. (2002) Regulated tissue-specific expression of antagonistic pre-mRNA splicing factors. *RNA* **4**:430-44.
- Hannon, G. (2002), RNA Interference, *Nature* **418**:244-251.
- Hata, K., Okano, M., Lei, H., Li, E. (2002) Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. *Development* **129**:1983-93.
- Hausheer, F., Rao, S., Gamcsik, M., Kollman, P., Colvin, O., Saxe, J., Nelkin, B., McLennan, I., Barnett, G., Baylin, S. (1989) Computational analysis of structural and energetic consequences of DNA methylation. *Carcinogenesis* **10**:1131-1137.
- Hayward, B., De Vos, M., Judson, H., Hodge, D., Huntriss, J., Picton, H., Sheridan, E., Bonthron, D. (2003) Lack of involvement of known DNA methyltransferases in familial hydatidiform mole implies the involvement of other factors in establishment of imprinting in the human female germline. *BMC Genet.* **20**:2.
- Hemann, M., Fridman, J., Zilfou, J., Hernando, E., Paddison, P., Cordon-Cardo, C., Hannon, G., Lowe, S. (2003) An epi-allelic series of p53 hypomorphs created by stable RNAi produces distinct tumor phenotypes in vivo. *Nat Genet.* **33**:396-400.
- Hermann, A., Schmitt, S., Jeltsch, A. (2003) The human Dnmt2 has residual DNA-(cytosine-C5) methyltransferase activity. *J Biol Chem.* **278**:31717-21.
- Hill, J., Burghardt, R., Jones, K., Long, C., Looney, C., Shin, T., Spencer, T., Thompson, J., Winger, Q., Westhusin, M. (2000a) Evidence for placental abnormality as the major cause of mortality in first-trimester somatic cell cloned bovine fetuses. *Biol Reprod* **63**:1787-94.
- Hill, J., Winger, Q., Long, C., Looney, C., Thompson, J., Westhusin, M. (2000b) Development rates of male bovine nuclear transfer embryos derived from adult and fetal cells. *Biol Reprod* **62**:1135-40.
- Hill, J., Winger, Q., Burghardt, R., Westhusin, M. (2001) Bovine nuclear transfer embryo development using cells derived from a cloned fetus. *Anim Reprod Sci.* **67**:17-26.
- Howell, C., Bestor, T., Ding, F., Latham, K., Mertineit, C., Trasler, J., Chaillet, J. (2001) Genomic imprinting disrupted by a maternal effect mutation in the Dnmt1 gene. *Cell.* **23**:829-38.

- Humpherys, D., Eggan, K., Akutsu, H., Hochedlinger, K., Rideout, W., Binischkiewicz, D., Yanagimachi, R., Jaenisch, R. (2001) Epigenetic instability in ES cells and cloned mice. *Science*. **293**:95-7.
- Humpherys, D., Eggan, K., Akutsu, H., Friedman, A., Hochedlinger, K., Yanagimachi, R., Lander, E., Golub, T., Jaenisch, R. (2002) Abnormal gene expression in cloned mice derived from embryonic stem cell and cumulus cell nuclei. *Proc Natl Acad Sci U S A* **99**:12889-94.
- Inoue, K., Kohda, T., Lee, J., Ogonuki, N., Mochida, K., Noguchi, Y., Tanemura, K., Kaneko-Ishino, T., Ishino, F., Ogura, A. (2002) Faithful expression of imprinted genes in cloned mice. *Science*. **295**:297.
- International Human Genome Sequencing Consortium (2001) Initial sequencing and analysis of the human genome. *Nature*. **409**:860-921.
- Jenuwein, T. & Allis, C. (2001) Translating the histone code. *Science* **293**:1074-80.
- Jones, P., Veenstra, G., Wade, P., Vermaak, D., Kass, S., Landsberger, N., Strouboulis, J., Wolffe, A. (1998) Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nat Genet* **19**:187-91.
- Judson, H., Hayward, B., Sheridan, E., Bonthron, D. (2002) A global disorder of imprinting in the human female germ line. *Nature* **416**:539-42.
- Kang, Y., Koo, D., Park, J., Choi, Y., Chung, A., Lee, K., Han, Y. (2001a) Aberrant methylation of donor genome in cloned bovine embryos. *Nat Genet* **28**:173-7.
- Kang, Y., Koo, D., Park, J., Choi, Y., Lee, K., Han, Y. (2001b) Influence of oocyte nuclei on demethylation of donor genome in cloned bovine embryos. *FEBS Lett*. **499**:55-8.
- Kato, Y., Tani, T., Sotomaru, Y., Kurokawa, K., Kato, J., Doguchi, H., Yasue, H., Tsunoda, Y. (1998) Eight calves cloned from somatic cells of a single adult. *Science*. **282**:2095-8.
- Kidder, G., & Conlon, R. (1985) Utilization of cytoplasmic poly(A)+RNA for protein synthesis in preimplantation mouse embryos. *J Embryol Exp Morphol*. **89**:223-34.
- Klimasauskas, S., Kumar, S., Roberts, R., Cheng, X. (1994) HhaI methyltransferase flips its target base out of the DNA helix. *Cell* **76**:357-69.

- Kopecny, V., Flechon, J., Camous, S., Fulka, J. (1989) Nucleologenesis and the onset of transcription in the eight-cell bovine embryo: fine-structural autoradiographic study. *Mol Reprod Dev* **1**:79-90.
- Kumar, S., Cheng, X., Klimasauskas, S., Mi, S., Posfai, J., Roberts, R., Wilson, G. (1994) The DNA (cytosine-5) methyltransferases. *Nucleic Acids Res.* **22**:1-10.
- Kunert, N., Marhold, J., Stanke, J., Stach, D., Lyko, F. (2003) A Dnmt2-like protein mediates DNA methylation in *Drosophila*. *Development* **130**:5083-90.
- Lebedeva, L. & Tillib, S. (2003) Trithorax protein, a global factor for maintenance of tissue specific gene activation in *Drosophila melanogaster*, is associated with the nuclear matrix *Genetika* **39**:250-8.
- Leonhardt, H., Page, A., Weier, H., Bestor, T. (1992) A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell* **71**:865-873.
- Levine, M. & Tjian, R. (2003) Transcription regulation and animal diversity. *Nature* **424**:147-51.
- Li, E., Bestor, T., Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69**:915-926.
- Liu, Z. & Foote R. (1997) Effects of amino acids and alpha-amanitin on bovine embryo development in a simple protein-free medium. *Mol Reprod Dev* **46**:278-85.
- Lois, C., Hong, E., Pease, S., Brown, E., Baltimore, D. (2002) Germline transmission and tissue-specific expression of transgenes delivered by lentiviral vectors. *Science*. 295:868-72.
- Lopatina, N., Haskell, J., Andrews, L., Poole, J., Saldanha, S., Tollefsbol, T. (2002) Differential maintenance and de novo methylating activity by three DNA methyltransferases in aging and immortalized fibroblasts. *Journal of Cellular Biochemistry* **84**:324-34.
- Lopez, A. (1998) Alternative splicing of pre-mRNA: developmental consequences and mechanisms of regulation. *Annu Rev Genet.* **32**:279-305.
- Margot, J., Ehrenhofer-Murray, A., Leonhardt, H. (2003) Interactions within the mammalian DNA methyltransferase family. *BMC Mol Biol.* **4**:7.
- McCance, R. (1962) Food, growth and time. *Lancet* **ii**:271-2.

- McGrath, J. & Solter, D. (1984a) Completion of mouse embryogenesis requires both the maternal and paternal genomes. *Cell* **37**:179-83.
- McGrath, J. & Solter, D. (1984b) Maternal Thp lethality in the mouse is a nuclear, not cytoplasmic, defect. *Nature* **308**:550-1.
- Mertineit, C., Yoder, J., Taketo, T., Laird, D., Trasler, J., Bestor, T. (1998). Sex-specific exons control DNA methyltransferase in mammalian germ cells. *Development* **125**:889-897.
- Monk, M., Adams, R., Rinaldi, A. (1991) Decrease in DNA methylase activity during preimplantation development in the mouse. *Development* **112**:189-92.
- Mott, G., Lewis, D., McGill, H. (1991) Programming of cholesterol metabolism by breast or formula feeding. *Ciba Found Symp* **156**:56-66.
- Myslinski, E., Ame, J., Krol, A., Carbon, P. 2001 An unusually compact external promoter for RNA polymerase III transcription of the human H1RNA gene. *Nucleic Acids Res.* **29**:2502-9.
- Natale, D., De Sousa, P., Westhusin, M., Watson, A. (2001) Sensitivity of bovine blastocyst gene expression patterns to culture environments assessed by differential display RT-PCR. *Reproduction* **122**:687-93.
- Ogbourne, S. & Antalis, T. (1998) Transcriptional control and the role of silencers in transcriptional regulation in eukaryotes. *Biochem J.* **331**:1-14.
- Ogonuki, N., Inoue, K., Yamamoto, Y., Noguchi, Y., Tanemura, K., Suzuki, O., Nakayama, H., Doi, K., Ohtomo, Y., Satoh, M., Nishida, A., Ogura, A. (2002) Early death of mice cloned from somatic cells. *Nat Genet* **30**:253-4
- Okano, M., Xie, S., Li, E. (1998a). Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nature Genetics* **19**:219-20.
- Okano, M., Xie, S., Li, E. (1998b). Dnmt2 is not required for de novo and maintenance methylation of viral DNA in embryonic stem cells. *Nucleic Acids Res* **26**:2536-40.
- Okano, M., Bell, D., Haber, D., Li, E. (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**:247-57.
- 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. *Curr Biol.* **10**:475-8.

- Pace, M., Augenstein, M., Betthausen, J., Childs, L., Eilertsen, K., Enos, J., Forsberg, E., Golueke, P., Graber, D., Kemper, J., Koppang, R., Lange, G., Lesmeister, T., Mallon, K., Mell, G., Misica, P., Pfister-Genskow, M., Strelchenko, N., Voelker, G., Watt, S., Bishop, M. (2002) Ontogeny of cloned cattle to lactation. *Biol Reprod* **67**:334-9.
- Pascual, J., Martinez-Yamout, M., Dyson, H., Wright, P. (2000) Structure of the PHD zinc finger from human Williams-Beuren syndrome transcription factor. *J Mol Biol*. **304**:723-9.
- Picketts, D., Tastan, A., Higgs, D., Gibbons, R. (1998) Comparison of the human and murine ATRX gene identifies highly conserved, functionally important domains. *Mamm Genome*. **9**:400-3.
- Plante, L., Plante, C., Shepherd, D., King, W. (1994) Cleavage and 3H-uridine incorporation in bovine embryos of high in vitro developmental potential. *Mol Reprod Dev*. **39**:375-83.
- Prather, R. (1996) Progress in cloning embryos from domesticated livestock. *Proc Soc Exp Biol* **212**:38-43.
- Prather, R., Tao, T., Machaty, Z. (1999) Development of the techniques for nuclear transfer in pigs. *Theriogenology*. **51**:487-98.
- Qiu, C., Sawada, K., Zhang, X., Cheng, X. (2002) The PWWP domain of mammalian DNA methyltransferase Dnmt3b defines a new family of DNA-binding folds. *Nat Struct Biol*. **9**:217-24.
- Ratnam, S., Mertineit, C., Ding, F., Howell, C., Clarke, H., Bestor, T., Chaillet, J., Trasler, J. (2002) Dynamics of Dnmt1 methyltransferase expression and intracellular localization during oogenesis and preimplantation development. *Dev Biol* **245**:304-14.
- Renard, J., Chastant, S., Chesne, P., Richard, C., Marchal, J., Cordonnier, N., Chavatte, P., Vignon, X. (1999) Lymphoid hypoplasia and somatic cloning. *Lancet*. **353**:1489-91.
- Robertson, K., Uzvolgyi, E., Liang, G., Talmadge, C., Sumegi, J., Gonzales, F., Jones, P. (1999). The human DNA methyltransferases (DNMTs) 1, 3a and 3b: coordinate mRNA expression in normal tissues and overexpression in tumors. *Nucleic Acids Research* **27**:2291-8.

- Rubinson, D., Dillon, C., Kwiatkowski, A., Sievers, C., Yang, L., Kopinja, J., Rooney, D., Ihrig, M., McManus, M., Gertler, F., Scott, M., Van Parijs, L. (2003) A lentivirus-based system to functionally silence genes in primary mammalian cells, stem cells and transgenic mice by RNA interference. *Nat Genet.* **33**:401-6.
- Rush, L. & Plass, C. (2002) Restriction landmark genomic scanning for DNA methylation in cancer: past, present, and future applications. *Anal Biochem.* **307**:191-201.
- Sailer, A., Bueller, H., Fischer, M., Aguzzi, A., Weissmann, C. (1994) No propagation of prions in mice devoid of PrP. *Cell* **77**:967-8.
- Saito, Y., Kanai, Y., Sakamoto, M., Saito, H., Ishii, H., Hirohashi, S. (2002) Overexpression of a splice variant of DNA methyltransferase 3b, DNMT3b4, associated with DNA hypomethylation on pericentromeric satellite regions during human hepatocarcinogenesis. *Proc Natl Acad Sci USA* **23**:10060-5.
- Salisbury S. & VanDemark, G. 1961. *Physiology of Reproduction and Artificial Insemination of Cattle*, San Francisco, Freeman & Co.
- Santi, D., Norment, A., Garrett, C. (1984) Covalent bond formation between a DNA-cytosine methyltransferase and DNA containing 5-azacytosine. *Proc Natl Acad Sci U S A.* **81**:6993-7.
- Schultz, G., Harvey, M., Watson, A., Arcellana-Panlilio, M., Jones, K., Westhusin, M. (1996) Regulation of early embryonic development by growth factors: growth factor gene expression in cloned bovine embryos. *J Anim Sci* **74**(Suppl 3):50-57.
- Shi, W. & Haaf, T. (2002) Aberrant methylation patterns at the two-cell stage as an indicator of early developmental failure. *Mol Reprod Dev* **63**:329-34.
- Shin, T., Kraemer, D., Pryor, J., Liu, L., Rugila, J., Howe, L., Buck, S., Murphy, K., Lyons, L., Westhusin, M. (2002) A cat cloned by nuclear transplantation. *Nature.* **415**:859.
- Shilatifard A. (1998) Factors regulating the transcriptional elongation activity of RNA polymerase II. *FASEB J.* **12**:1437-46.
- Shiota, K., Kogo, Y., Ohgane, J., Imamura, T., Urano, A., Nishino, K., Tanaka, S., Hattori, N. (2002) Epigenetic marks by DNA methylation specific to stem, germ and somatic cells in mice. *Genes Cells* **7**:961-9.

- Sinclair, K., Young, L., Wilmut, I., McEvoy, T. (2000) In-utero overgrowth in ruminants following embryo culture: lessons from mice and a warning to men. *Hum Reprod* **15**:68-86.
- Smart, J. (1990) Vulnerability of developing brain to undernutrition. *Ups J Med Sci Suppl* **48**:21-41.
- Staley, J. & Guthrie, C. (1998) Mechanical devices of the spliceosome: motors, clocks, springs, and things. *Cell* **92**:315-26.
- Stec, I., Nagl, S., van Ommen, G., den Dunnen, J. (2000) The PWWP domain: a potential protein-protein interaction domain in nuclear proteins influencing differentiation? *FEBS Lett.* **473**:1-5.
- Suetake, I., Miyazaki, J., Murakami, C., Takeshima, H., Tajima, S. (2003) Distinct enzymatic properties of recombinant mouse DNA methyltransferases Dnmt3a and Dnmt3b. *J Biochem.* **133**:737-44.
- Svoboda, P., Stein, P., Hayashi, H., Schultz, R. (2000) Selective reduction of dormant maternal mRNAs in mouse oocytes by RNA interference. *Development* **127**:4147-56.
- Tamashiro, K., Wakayama, T., Akutsu, H., Yamazaki, Y., Lachey, J., Wortman, M., Seeley, R., D'Alessio, D., Woods, S., Yanagimachi, R., Sakai, R. (2002) Cloned mice have an obese phenotype not transmitted to their offspring. *Nat Med* **8**:262-7.
- Tang, L., Reddy, M., Rasheva, V., Lee, T., Lin, M., Hung, M., Shen, C. (2003) The eukaryotic DNMT2 genes encode a new class of cytosine-5 DNA methyltransferases. *J Biol Chem.* **278**:33613-6.
- Telford, N., Watson, A., Schultz, G. (1990) Transition from maternal to embryonic control in early mammalian development: a comparison of several species. *Mol Reprod Dev.* **26**:90-100.
- Tucker, K., Talbot, D., Lee, M., Leonhardt, H., Jaenisch, R. (1996) Complementation of methylation deficiency in embryonic stem cells by a DNA methyltransferase minigene. *Proc Natl Acad Sci U S A.* **12**:12920-5.
- Vignon, X., Chesne, P., Le Bourhis, D., Flechon, J., Heyman, Y., Renard, J. (1998) Developmental potential of bovine embryos reconstructed from enucleated matured oocytes fused with cultured somatic cells. *C R Acad Sci III.* **321**:735-45.
- Viuff, D., Avery, B., Greve, T., King, W., Hyttel, P. (1996) Transcriptional activity in in vitro produced bovine two- and four-cell embryos. *Mol Reprod Dev.* **43**:171-9.

- Weisenberger, D., Velicescu, M., Preciado-Lopez, M., Gonzales, F., Tsai, Y., Liang, G., Jones, P. (2002) Identification and characterization of alternatively spliced variants of DNA methyltransferase 3a in mammalian cells. *Gene*. **298**:91-9.
- White, A., Enever, P., Tayebi, M., Mushens, R., Linehan, J., Brandner, S., Anstee, D., Collinge, J., Hawke, S. (2003) Monoclonal antibodies inhibit prion replication and delay the development of prion disease. *Nature* **422**:80-3.
- Winger, Q., Hill, J., Watson, A., Westhusin, M. (2000) Characterization of a bovine cDNA encoding citrate synthase, and presence of citrate synthase mRNA during bovine pre-attachment development. *Mol Reprod Dev*. **55**:14-9.
- Wilmut, I., Schnieke, A., McWhir, J., Kind, A., Campbell, K. (1997) Viable offspring derived from fetal and adult mammalian cells. *Nature*. **385**:810-3.
- Wilson, D. & Wyatt, D. (1995) Patterns of lectin binding during mammalian neurogenesis. *J Anat*. **186**:209-16.
- Winters, L., Green, W., Comstock, R., 1942. The prenatal development of the Bovine. Agricultural Experiment Station, University of Minnesota, St. Paul. *Technical Bulletin* **151**.
- Wrenzycki, C., Wells, D., Herrmann, D., Miller, A., Oliver, J., Tervit, R., Niemann, H. (2001) Nuclear transfer protocol affects messenger RNA expression patterns in cloned bovine blastocysts. *Biol Reprod*. **65**:309-17.
- Xue, F., Tian, X., Du, F., Kubota, C., Taneja, M., Dinnyes, A., Dai, Y., Levine, H., Pereira, L., Yang, X. (2002) Aberrant patterns of X chromosome inactivation in bovine clones. *Nat Genet*. **31**:216-20.
- Yanagisawa, Y., Ito, E., Yuasa, Y., Maruyama, K. (2002) The human DNA methyltransferases DNMT3A and DNMT3B have two types of promoters with different CpG contents. *Biochim Biophys Acta*. **1577**:457-65.
- Young, L., Sinclair, K., Wilmut, I. (1998) Large offspring syndrome in cattle and sheep. *Rev Reprod* **3**:155-63.
- Young, J. & Smith, J. (2001). DNA methyltransferase inhibition in normal human fibroblasts induces a p21-dependant cell cycle withdrawal. *J Biol Chem*. **276**:19610-19616.

- Yong, Z. & Yuqiang, L. (1998) Nuclear-cytoplasmic interaction and development of goat embryos reconstructed by nuclear transplantation: production of goats by serially cloning embryos. *Biol Reprod.* **58**:266-9.
- Zhao, J., Hyman, L., Moore, C. (1999) Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiol Mol Biol Rev.* **63**:405-45.
- Zimmermann, C., Guhl, E., Graessmann, A. (1997) Mouse DNA methyltransferase (MTase) deletion mutants that retain the catalytic domain display neither de novo nor maintenance methylation activity in vivo. *Biol Chem.* **378**:393-405.

## APPENDIX A

### BOVINE DNMT mRNA SEQUENCES AND SEQUENCE ALIGNMENTS

#### Bovine Dnmt1 Complete CDS

Green Start Site = Somatic Start

Blue Start Site = Putative Embryonic Start

AAG**ATG**CCTGCCCCGAACCGCCCCGGCGCGGGTGCCTGCGCTGGCCTCCCGG  
 GCCTTCTCACTGCCTGACGATGTCCGCAGGCGGCTCAAAGATTTGGAAAGAG  
 ATAGTTTGACAGAAAAGGAATGTGTGAAGGAGAACTGAATCTCTTGCACG  
 AATTTCTGCGGACAGAAATAAAGAATCAGTTATGTGATTTGGAAACCAAATT  
 GCATAAAGAAGAATTATCTGAGGAGGGCTACCTGGCTAAAGTCAAATCCCTT  
 TTAAATAAAGATTTGTCCTTGGAGAACGGAGCTCATGCTTTCAGTCGGGAAG  
 CGAATGGATGTCTAGAGAACGGGAGCCAGACAAGTGGTGAGGATTGCAGAG  
 TGGTA**ATG**GCAGAGAAAGGCAAGCCCCCAAACCTGTCTCCAGACTTTACA  
 CGCCAGGAGAAGCAAGTCTGATGGAGAAACAAAGTCTGAAGTCTCTTCTA  
 GCCCCAGGATTACAAGGAAGACTACCAGGCAGACCACCATCACATCTCATTT  
 CCCACGGGGCCCTGCCAAACGAAAACCTGAGGAAGAACCTGAAAAAGTGAA  
 GTCAGACGATTCTGTTGATGAAGAAAAAGACCAGGAGGAAAAGAGACGTCG  
 AGTTACATCCAGAGAACGAGTTGCTGGGCTGCTCCCTGCAGAAGAACCAGG  
 AAGAGTAAGACCAGGAACACACATGGAAGAAGAAGGAAGAGATGATAAAG  
 AAGAAAAGAGACTCAGAAGTCAAACCAAAGAACCGACACCTAAACACAAA  
 GCTAAGGAGGAGCCAGACAGAGATGTGAGGCCTGGAGGAGCTCAGGCTGAA  
 ATGAATGAAGGAGAAGACAAAGATGAAAAGAGGCACAGAAGTCAACCCAA  
 AGATCTAGCTAGCAAACGGAGACCAGAAGAAAAAGAACCTGAAAGAGTAA  
 AGCCACAAGTTTCTGATGAGAAAGATGAAGATGAAAAGGAGGAGAAGAGA  
 CGCAGAACTACATACAGAGAACTAACCGAGAAGAAAATGACTCGAACCAAA  
 ATAGCCGTAGTGTCCAAGACCAATCCTCCGAAGTGCACCGAGTGCTTGCAGT  
 ACCTGGACGACCCTGAGCTGAGATACGAGCAGCACCCCCCGATGCGGTGG  
 AAGAGATACAGATACTGACCAACGAGAGGTTGTCCATCTTTGATGCCAACG  
 AATCTGGCTTTGAGAGTTACGAGGATTTGCCTCAGCACAACTAACCTGCTT  
 CAGCGTGTACTGTAAACGCGGTCACCTTTGCCCGATCGACACCGGCCTCATT  
 GAGAAGGATGTCGAGCTCCTCTTTTCTGGTTCAGCAAAGCCGATATATGAGG  
 ATGACCCATCTCCCGAAGGTGGTATTAATGGCAAAAATTTTGGCCCCATAAA  
 CGAATGGTGGATTGCTGGTTTTGATGGAGGTGAAAAGGCTCTTCTTGGCTTT  
 AGCACCTCATTTGCCGAGTATATCTTGATGGATCCCAGCCCAGAGTACGCAC  
 CACTATTCAGCGTGATGCAGGAGAAGATCTATATAAGTAAGATAGTGGTTGA  
 GTTCCTGCAGAGCAACCCTGACTCCACCTACGAAGACCTGATCAATAAGATT  
 GAGACCACCGTTCCCTCCTTGTATGCTCAACTTGAATCGATTCACAGAGGATT  
 CTCTCCTGCGGCATGCCAGTTCGTGGTGGAGCAAGTAGAGAGTTATGATCG  
 GGCTGGGGACAGTGACGAGCAGCCCATCTTCCTGAGCCCCTGCATGAGAGA  
 CCTCATCAAGCTGGCCGGGGTACCCCTGGGAAAAAGGCGAGCCGAGAGGCG  
 GCAGACCATCCGGCAACCCGCCAAAGAGAAGGACAAGGGCCCCACCAAGGC  
 CACCACCACCAAGCTGGTCTACCAGATCTTTGACACTTTCTTTGCGGAGCAA

ATTGAAAAAGATGACAAGGAAGACAAGGAGAATGCCTTCAAGCGCCGGCGC  
TGTGGCGTCTGTGAGATTTGTCAACAGCCCGAGTGTGGAAAGTGTAAGGCCT  
GTAAGGATATGGTTAAATTTGGTGGTAGCGGACGGAGCAAGCAGGCTTGCC  
AAAAGAGGAGGTGTCCCAACATGGCCATGAAGGAGGCAGACGATGACGAG  
GAAGTGGATGACAATATTCCAGAGATGCCATCACCCAAAAAGATGCATCAG  
GGGAAGAAAAAGAAGCAGAATAAGAATCGGATCTCTTGGGTTGGCGATGCC  
GTCAAGACTGACGGGAAGAAGAGTTACTACAAGAAGGTATGCATCGACTCG  
GAAACCCTGGAAGTGGGGGACTGTGTTTCTGTAATTCCAGACGACTCTTCAA  
AACCCTGTATCTAGCAAGGGTCACGGCGCTGTGGGAGGACAGCAGCAATG  
GGCAGATGTTCCATGCCCCACTGGTTCTGTGCTGGGACGGACACGGTCCTCGG  
GGCCACATCGGACCCCCTGGAGCTGTTCTGTTGACGAGTGTGAGGACATG  
CAGCTCTCGTACATCCACAGCAAGGTGCAGGTCATTTATAAGGCGCCCTCAG  
AGAAGTGGGCCATGGAGGGAGGCGTGGACCCCGAGGCCCTGATGTCAGAGG  
ACGACGGGAAGACCTACTTCTACCAGCTGTGGTACGACCAAGACTACGCGA  
GATTTGAGTCCCCTCCGAAAACCTCAGCCGACGGAGGACAACAAGTACAAGT  
TCTGCGCAAGCTGTGCACGTCTGGCCGAAATGAGGCAGAAAGGAAATCCCCA  
GGGTCGTGGAGCAGCTCCAGGACCTGGAAGGCCGCGTCTCTACAGCCTCGC  
CACCAAGAACGGCGTCCAGTACCGGGTGGGCGATGGCGTGTACCTCCCTCCC  
GAGGCCTTCACCTTCAACATCAAGCTGTCCAGTCCTGTGAAACGCCCCCGGA  
AGGAGCCTGTGGACGAAGCTCTGTATCCAGAACACTACCGGAAGTACTCTG  
ACTACATCAAGGGCAGCAACCTGGATGCCCTGAGCCCTACCGTATTGGCCG  
CATAAAGGAGATCTTCTGCAGCAAGAAGAGCAACGGCCGGCCCAATGAGAC  
AGACATCAAGATCAGGGTCAACAAGTTCTACAGGCCGGAGAACACACACAA  
GTCTACCCCAGCCAGTTACCACGCAGACATCAACCTGCTTTACTGGAGCGAT  
GAGGAGGCCGTGGTGGACTTCAAGGCCGTGCAGGGCCGCTGCACCGTGGAG  
TACGGAGAGGACCTGCCTCAGTGCCTCCAGGACTTCTCCGCTGGTGGCCCCG  
ATCGCTTCTATTTTCTCGAGGCCTATAACGCCAAGAGCAAAAGCTTTGAAGA  
TCCTCCGAACCACGCCCGGAGCACCGGAAATAAAGGGAAAGGGAAGGGGA  
AAGGAAAAAACAGGACGAAATCTCAGACGTGTGAGCCGAGTGAAGTGGAG  
ACAGAAATCAAACCTGCCGAAGCTGCGGACCCTGGACGTGTTTTCCGGCTGTG  
GGGGATTGTGGAAGGCTTCCACCAAGCAGGCATCTCGGAAACACTTTGGG  
CCATCGAGATGTGGGACCCTGCGGCCAGGCGTTCCGGTTCAACAACCCTGG  
GTCCACGGTGTTCACAAAGGACTGCAACGTCCTGGTGAAGCTGGTCATGGCC  
GGGGAGGTGACCAACTCCCGCGGCCAGAAGCTGCTTCAAAGGGAGATGTG  
GAGATGTTGTGCGGCGGGCCGCCCTGCCAGGGCTTTAGCGGCATGAACCGCT  
TCAACTCTCGAACCTACTCCAAATTCAAGAAGTCCCTGGTGGTCTCTTTCCTC  
AGCTACTGTGACTACTACCGGCCCCGCTACTTCCTCTTGGAGAACGTTTCGGA  
ACTTCGTCTCCTTCAAGCGCTCCATGGTCCTGAAGCTGACGCTGCGCTGCCTG  
GTCCGCAGGGGGTACCAGTGCACCTTTGGCGTGCTGCAGGCTGGTCAGTACG  
GCGTGGCCAGACTCGGAGGCGAGCCATCATCCTGGCTGCAGCCCCTGGGG  
AGCCACTCCCGCTGTTCCCGGAGCCGTTGCATGTGTTTCGACCCCCGGGCTG  
CCAGCTGAGCGTCGTAGTGGACGACAAGAAGTTTGTGTCAGCAACATCACCAG  
GTTGAGCTCGGGTCCCTTCCGAACCATCACCGTGCGGGACACCATGTCTGAC  
CTCCCTGAGATCCGGAACGGGGCCTCGGCACTGGAGATTTCATACAACCGGG

AGCCCCAGTCCTGGTTCCAGAGGCAGCTCCGGGGCTCGCAGTACCAGCCCAT  
CCTCAGGGATCATATTTGCAAGGACATGAGCGCCTTGGTGGCTGCCCCGATG  
CGGCACATCCCCCTGGCCCCGGGCTCGGACTGGCGTGACCTGCCAACATTG  
AGGTGCGGCTCTCTGACGGCACCTGGCCCCGAAGCTGCGGTACAACCTACCA  
CGACAAGAAGAACGGCTGCAGCAGCAGCGGGCGCCCTCCGTGGGGTCTGCTC  
CTGTGTGGAAGGCAAGCCCTGTGAGCCTGCGGCCCCGACAGTTTAAACACCCTT  
ATCCCCTGGTGCCTGCCCCACACTGGGAACAGGCACAACCACTGGGCGCGGCC  
TCTACGGGCGTCTCGAGTGGGACGGCTTCTTCAGCACAACCTGTCACCAACCC  
CGAGCCCATGGGCAAGCAGGGGCCGCGTGCTCCACCCCGAGCAGCACCGAGT  
GGTGAGCGTCCGGGAGTGCGCCCGCTCCCAGGGCTTCCCCGACACCTATCGG  
CTGTTTCGGCAACATCCTAGACAAGCACCGGCAGGTGGGTAATGCTGTGCCGC  
CGCCACTGGCCAAAGCCATCGGCTTGGAGATCAAGCGCTGCATGTTGGCCAA  
AGCGCGCGAGAGCGCCTCAGCTAAAATCAAGGAGGAGGCTGCCAAGGACTA  
GTTCTCTCCTCCTATCACCCATGTTTCTGCCACCAGAGATCCCCAACGTGCAC  
TGATATTGGTGTATTTTTTCACATGTCAATCAGTCAATTCAGATGTGTCGTATG  
CGGTGTTTGTGGCCTTGGCTGACATGAAACTCTTCAGTGAGATTTGCCTATCG  
GCTAATTTGGACTTANTGATCAAACCTGTGCAGTACTTTGTCCATTCTGGATTT  
TAAAAGTTTTTTTTTACGCATTATATNAAATTTACCACTGTTTGAGTGGNAAT  
TAAGACTTTATGTAGNTTTTATATGTTGNAATATTTCTTCAAAAAATCTCTTC  
TTAAAAACAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

### Bovine DNMT1 Protein Sequence

MPARTAPARVPALASRAFSLPDDVRRRLKDLERDSLTEKECVKEKLNLLHEFLR  
 TEIKNQLCDLETKLHKEELSEEGYLAKVKSLLNKDLSLENGAHAFSREANGCLE  
 NGSQTSGEDCRVVMMAEKKGKPPKPVSRLYTPRRSKSDGETKSEVSSSPRITRKTTR  
 QTTITSHFPRGPAKRKPEEEPEKVKSDDSVDEEKDQEEKRRRVTSRERVAGLLPA  
 EEPGRVRPGTHMEEEGRDDKEEKRLRSQTKEPTPKHKAKEEPDRDVRPGGAQA  
 EMNEGEDKDEKRHRSPKDLASKRRPEEKEPERVKPQVSDEKDEDEKEEKRRR  
 TTYRELTEKKMTRTKIAVVSKTNPCKTECLQYLDDPELRYEQHPPDAVEEIQIL  
 TNERLSIFDANESGFESYEDLPQHKLTCFSVYCKRGHLCPIDTGLIEKDVELLFSG  
 SAKPIYEDDPSPEGGINGKNFGPINWWIAGFDGGEKALLGFSTSFAEYILMDPSP  
 EYAPLFSVMQEKIYISKIVVEFLQSNPDSTYEDLINKIETTVP CMLNLRFTEDSL  
 LRHAQFVVEQVESYDRAGDSDEQIFLSPCMRDLIKLAGVTLGKRRRAERRQTIR  
 QPAKEKDKGPTKATTTKL VYQIFDTFFAEQIEKDDKEDKENAFKRRRCGVCEIC  
 QQPEC GKCKACKDMVKFGGSGRSKQACQKRRCPNMAMKEADDDDEEVDDNIPE  
 MPSPKKMHQGK KKKQKNRISWVGDAVKTDGKKSYYKKVCIDSETLEVGDVCV  
 SVIPDDSSKPLYLARVTALWEDSSNGQMFHAHWFCAGTDTVLGATSDPLELFLV  
 DECEDMQLSYIHSKVQVIYKAPSENWAMEGGVDPEALMSEDDGKTYFYQLWY  
 DQDYARFESPPKTQPTEDNKYKFCASCARLAEMRQKEIPRVVEQLQDLEGRVL  
 YSLATKNGVQYRVGDGVYLPPEAFTFNILSSPVKRPRKEPVDEALYPEHYRKY  
 SDYIKGSNLDAPEPYRIGRIKEIFCSKKSNGRPNETDIKIRVNKFYRPENTHKSTPA  
 SYHADINLLYWSDEEAVVDFKAVQGRCTVEYGEDLPQCLQDFSAGGPDRFYFL  
 EAYNAKSKSFEDPPNHARSTGNKGKGKKGKGNRTKSQTCEPSELETEIKLPKLR  
**TLDVFSGCGGLSEGFHQAGISETLWAIEMWDPAAQAFRFNNPGSTVFTKDC**  
 NVLVKLVMAGEVTNSRGQKLLQKGDVEMLCGGPPCQGFSGMNRFNRSRTYSK  
 FKNSLVVSFLSYCDYYRPRYFLLNVNRFVSFKRSMVLKLT LRCLVRRGYQC  
**TFGV LQAGQY GVAQT RRR AII**LAAAPGEPLPLFPEPLHVFAPRACQLSVVDD  
 KKFVSNITRLSSGPFTITVRDTMSDLPEIRNGASALEISYNREPQSWFQRQLRGS  
 QYQPILRDHICKDMSALVAARMRHIPLAPGSDWRDLPNIEVRLSDGTLARKLRY  
 NYHDKKNGC SSSGALRGVCSCVEGKPCEPAARQFNTLIPWCLPHTGNRHNHWA  
 GLYGRLEWDGFFSTTVTNPEPMGKQGRVLHPEQ**HRVVS VRECARSQGFPD TY**  
**RLFGNILDKHRQVGNAVPPPLAKAIGLEIKRCMLAKARESASAKIKEEAAKD\***  
 Bold sequence = Methyltransferase Domains I, II, IV, VI, VII, VIII, IX, X respectively.

## Sequence Alignment

Bovine	-----AA	2
Human	CCGCGCG--AAAAGCCGGGCGCCTGCGCTGCCGCCGCGCTGCTGAAGCCTCCGA	236
Mouse	CCGCGCGCGCGAAAAAGCCGGGTCTCGTTCCAGAGCTGTTCTGTCGTCTGCAACCTGCAA	191
	*	
Bovine	GATGCCTGCCCCAACC GCCCGCGCGGGTGCCTGCGCTGGCCTCCCGGGCCTTCTCACT	62
Human	GATGCCGCGCGTACCGCCCCAGCCCGGTGCCCACTGGCCGTCCCGGCCATCTCGCT	296
Mouse	GATGCCAGCGCAACAGCTCCAGCCCGAGTGCCTGCGCTTGCTCCCGGCAGGCTCGCT	251
	*****	
Bovine	GCCTGACGATGTCCGCGAGCGGCTCAAAGATTTGGAAAGAGATAGTTTGACAGAAAAGGA	122
Human	GCCCGACGATGTCCGCGAGCGGCTCAAAGATTTGGAAAGAGACAGCTTAACAGAAAAGGA	356
Mouse	CCCGGACCATGTCCGCGAGCGGCTCAAAGACTTGGAAAGAGATGGCTTAACAGAAAAGGA	311
	**	
Bovine	ATGTGTGAAGGAGAACTGAATCTCTGCACGAATTTCTGCGGACAGAAATAAAGAATCA	182
Human	ATGTGTGAAGGAGAAATTGAATCTCTGCACGAATTTCTGCAACAGAAATAAAGAATCA	416
Mouse	ATGTGTGAGGAGAAATTAACTTACTGCATGAATTCCTGCAACAGAAATAAAGGCCA	371
	*****	
Bovine	GTTATGTGATTTGGAAACCAATTGCATAAAGAAGAATTATCTGAGGAGGGCTACCTGGC	242
Human	GTTATGTGACTTGGAAACCAATTACGTAAAGAAGAATTATCCGAGGAGGGCTACCTGGC	476
Mouse	GTTGTGTGACTTGGAAACCAATTACATAAAGAGGAATTATCTGAGGAAGGCTACCTGGC	431
	*****	
Bovine	TAAAGTCAAATCCCCTTTTAAATAAAGATTTGTCTTGGAGAACGGAGCTCATGCTTTCAG	302
Human	TAAAGTCAAATCCCCTTTTAAATAAAGATTTGTCTTGGAGAACGGTGTCTCATGCTTACAA	536
Mouse	TAAAGTCAAGTCCCTCTTAAATAAAGATTTGTCTTGGAGAACGGAAACACACACTCTCAC	491
	*****	
Bovine	TCGGGAAGCGAATGGATGTCTAGAGAACGGGAGCCAGACAAGTGGTGAGGATTGCAGAGT	362
Human	CCGGGAAGTGAATGGACGTCTAGAAAACGGGAACCAAGCAAGTGAAGCCCGTAGAGT	596
Mouse	TCAAAAGCCAACGGTTGTCCCGCAACGGGAGCCGCCAA-----CTGGAGAGC	542
	* **	
Bovine	GGTAATGGCAGAGAAAGGCAAGCCCCCAAACCTGTCTCCAGACTTTACACGCCCAGGAG	422
Human	GGGAATGGCAGATGCCAACAGCCCCCAAACCCCTTTCCAAACCTCGCAGCCCCAGGAG	656
Mouse	AGAAATGGCAGACTCAAATAGATCCCCAAGATCCAGGCCAAGCCTCGGGGACCCAGGAG	602
	*****	
Bovine	AAGCAAGTCTGATGGAGAAACAAAGTC--TGAAGTCTCTTCTAGCCCCAGGATTACAAG	479
Human	GAGCAAGTCCGATGGAGAGGCTAAGCC--TGAACCTTCACCTAGCCCCAGGATTACAAG	713
Mouse	AAGCAAGTCGACAGTGACACCCCTTTCAGTTGAAACTTCACCTAGTTCCGTGGCTACGAG	662
	*****	
Bovine	GAAGACTACCAGGCAGACCACCATCACATCTCATTTCCACGGGGCCCTGCCAAACGAAA	539
Human	GAAAAGCACCAGGCAAAACACCATCACATCTCATTTGCAAAGGGCCCTGCCAAACGGAA	773
Mouse	GAAACCACCAGGCAGACCACCATCACGGCTCACTTACGAAGGGCCCCACTAAACGGAA	722
	**	
Bovine	ACCTGAGGAAGAACCTGAAAAAGTGAAGTCAGACGATTCTGTTGATGAAGAA--AAAGA	596
Human	ACCTCAGGAAGAGTCTGAAAGAGCCAAATCGGATGAGTCCATCAAGGAAGAAGACAAAGA	833
Mouse	ACCCAAGGAAGAGTCGGAAGAGGGAACTCGGCTGAGTCGGCTGCAGAGGAG--AGAGA	779
	*****	
Bovine	CCAGGAGGAAAAGAGACGTGAGTTACATCCAGAGAACGAGTTGCTGGGCTGCTCCCTGC	656
Human	CCAGGATGAGAAGAGACGTAGAGTTACATCCAGAGAACGAGTTGCTAGACCGCTTCTCTGC	893
Mouse	CCAGGAT--AAGAAACGCAGAGTTGTAGACACAGAGAGTGGTGCT--GCAGCTGC-TGT	833
	*****	
Bovine	AGAAGAACCAGGAAGAGTAAGACCAGGAACACACATGGAA--GAAGAAGGAAGAGATGA	713
Human	AGAAGAACCTGAAAGAGCAAAATCAGGAACGCGCACTGAAAAGGAAGAAGAGATGA	953
Mouse	GGAGAACTGGAAGAGGTAACAGCGGAACCCAGCTGGGTCCGGAAGAGCCATGTGAACA	893
	**	
Bovine	TAAAGAAGAAAA--GAGACTCAGAAGTCAAACCAAAGAACCGACACCTAAACACAAAGC	770
Human	AAAAGAAGAAAA--GAGACTCCGAAGTCAAACCAAAGAACCAACCCCAACAGAAACT	1010
Mouse	GGAAGATGACAACAGGAGTCTTCGACGTACACCAGAGAGTATCATTGAGGCGGAAATC	953
	*****	
Bovine	TAAGGAGGAGCCAGACAGAGATGTGAGGCCTGGAGGAGCTCAGGCTGAAATGAATGAAGG	830

Human	GAAGGAGGAGCCGGACAGAGAAGCCAGGGC--AGGCGTGCAGGCTGACGAGGACGAAGA	1067
Mouse	AAAGGAGGATCCAGACAGAGAAGCAAGACC--GGAACCTCACTTGGACGAGGACGAGGA	1010
	***** *	
Bovine	AGAAGAC--AAAGATGAAAAAGAG-----GCACAGAAGTCAACCCAAAGATCTAGCTAG	881
Human	TGGAGACGAGAAAAGATGAGAGAA-----GCACAGAAGTCAACCCAAAGATCTAGCTGC	1121
Mouse	CGAAAA--AAGGATAAAAGAAGTTCAGACCCAGGAGCCAGCCAGAGATCCAGCTGC	1067
	* *	
Bovine	CAAACGGAGACCAGAAGAAAAAGAACCTGAAAGAGTAAAGCCACAAGTTTCTGATGAGAA	941
Human	CAAACGGAGGCCCGAAGAAAAAGAACCTGAAAAAGTAAATCCACAGATTTCTGATGAAAA	1181
Mouse	CAAACGGAGACCCAAGGAAGCAGAGCCAGAGCAGGTAGCTCCAGAGACTCCCGAGGACAG	1127
	***** *	
Bovine	AGATGAAGATGAAAAGGAGGAGAAGAGACGAGAACTACATACAGAGAATAA-----	994
Human	AGACGAGGATGAAAAGGAGGAGAAGAGACGAAAACGACCCCAAGAACCAA-----	1234
Mouse	AGACGAGGATGAGAGGGAGGAGAAGAGACGAAAAACGACACGTAAAAAACTGGAGTCACA	1187
	*** *	
Bovine	-----CCGAGAAGAAAATGACTCGAACCAAAATAGCCGTAGT	1031
Human	-----CGGAGAAAAAATGGCTCGCGCCAAACAGTCATGAA	1271
Mouse	CACCGTTCCCGTTCAGAGCAGATCGAGAGAAAAGCCGCTCAAAGCAAAAG--TGTGAT	1244
	* *	
Bovine	GTCCAAGACCAATCCTCCGAAGTGCACCGAGTGTCTGCAGTACCTGGACGACCCCTGAGCT	1091
Human	CTCCAAGACCCACCCCTCCCAAGTGCATTAGTGCAGGAGTACCTGGACGACCCCTGACCT	1331
Mouse	CCCGAAGATCAACTCACCAGAGTGCAGGAGTGTGGCCAGCACCTAGACGACCCCTAACCT	1304
	* *	
Bovine	GAGATACGAGCAGCACCCCCCGATGCGGTGGAAGAGATACAGATACTGACCAACGAGAG	1151
Human	CAAAATATGGGCAGCACCCACAGACGCGGTGGATGAGCCACAGATGCTGACAAATGAGAA	1391
Mouse	GAAGTACCAGCAGCACCCCTGAGGATGCTGTGGATGAACCCAGATGTTGACCAAGTGAAG	1364
	* *	
Bovine	GTTGTCCATCTTTGATGCCAACGAATCTGGCTTTGAGAGTTACGAGGATTTGCCTCAGCA	1211
Human	GCTGTCCATCTTTGATGCCAACGAGTCTGGCTTTGAGAGTTATGAGGCGCTTCCCAGCA	1451
Mouse	ACTGTCCATCTACGACTCCACCTCGACCTGGTTTGATACTTATGAAGATTCTCCATGCA	1424
	***** *	
Bovine	CAAACCTAACCTGCTTCAGCGTGTACTGTAAACGCGGTACCTTTGCCGATCGACACCGG	1271
Human	CAAACCTGACCTGCTTCAGTGTGTACTGTAAAGCACGGTCACCTGTGTCCATCGACACCGG	1511
Mouse	TAGGTTCACTTCCTTCAGTGTGTACTGCAGTCGCGGGCACCTGTGTCTGTGACACCGG	1484
	* *	
Bovine	CCTCATTGAGAAGGATGTGAGCTCCTCTTTCTGGTTCAGCAAAGCCGATATATGAGGA	1331
Human	CCTCATCGAGAAGAATATCGAACTCTTCTTTCTGGTTCAGCAAAACCAATCTATGATGA	1571
Mouse	TCTCATTGAGAAGAATGTAGAGCTCTACTTTCTGGGTGTGCCAAAGCAATTCATGACGA	1544
	***** *	
Bovine	TGACCCATCTCCCGAAGGTGGTATTAATGGCAAAAATTTTGGCCCCATAAACGAATGGTG	1391
Human	TGACCCGTCTCTTGAGGTGGTGTAAATGGCAAAAATCTTGGCCCCATAAATGAATGGTG	1631
Mouse	GAATCCATCTATGGAAGGTGGTATTAATGGCAAAAACCTCGGGCCAATCAATCAGTGGTG	1604
	* *	
Bovine	GATTGTGGTTTTGTATGGAGGTGAAAAGGCTCTTCTTGGCTTTAGCACCTCATTTGCCGA	1451
Human	GATCACTGGCTTTGTATGGAGGTGAAAAGGCCCTCATCGGCTTCAGCACCTCATTTGCCGA	1691
Mouse	GCTCAGTGGCTTTGTATGGTGGCGAGAAGGTGCTCATTGGCTTCTCCACTGCATTGCTGA	1664
	* *	
Bovine	GTATATCTTGATGGATCCCAGCCCAGAGTACGCACCACTATTCAGCGTGTGACAGGAGAA	1511
Human	ATACATTTCTGATGGATCCCAGTCCCAGATATGCGCCCATATTTGGGCTGATGCAGGAGAA	1751
Mouse	ATACATTTTGTATGGAGCCCAGCAAAGAGTATGAGCCAATATTTGGGCTGTGACAGGAGAA	1724
	* *	
Bovine	GATCTATATAAGTAAGATAGTGGTTGAGTTCCTGCAGAGCAACCCTGACTCCACCTACGA	1571
Human	GATCTACATCAGCAAGATTGTGGTGGAGTTCCTGCAGAGCAATTCGACTCGACCTATGA	1811
Mouse	AATTTACATCAGCAAGATTGTTGTTGAGTTCCTGCAAAACAATCCTGATGCTGTATATGA	1784
	* *	
Bovine	AGACCTGATCAATAAGATTGAGACCACCGTTCCTCCTTGTATGCTCAACTTGAATCGATT	1631
Human	GGACCTGATCAACAAGATCGAGACCACGGTTCCTCCTTCTGGCTCAACTTGAACCGCTT	1871
Mouse	AGACCTGATCAATAAGATTGAGACCACTGTTCCTCCTTCTACCATTAATGTGAACCGGTT	1844
	***** *	
Bovine	CACAGAGGATCTCTCTCGCGCATGCCAGTTCGTGGTGGAGCAAGTAGAGAGTTATGA	1691
Human	CACAGAGGACTCCCTCCTGCGACACGCGCAGTTTGTGGTGGAGCAGGTGGAGAGTTATGA	1931





```

*****
Bovine      TTGGGCCATCGAGATGTGGGACCCCTGCGGCCAGGCGTTCCGGTTCAACAACCCCTGGGTC 3545
Human       GTGGGCCATCGAGATGTGGGACCCCTGCGGCCAGGCGTTCCGGCTGAACAACCCCGGCTC 3788
Mouse       GTGGGCCATCGAGATGTGGGACCCGGCAGCCAGGCATTTCCGGTGAACAACCCCGGCAC 3752
            *****

Bovine      CACGGTGTTCAAAAGGACTGCAACGTCTGGTGAAGCTGGTCATGGCCGGGGAGGTGAC 3605
Human       CACAGTGTTACAGAGGACTGCAACATCCTGCTGAAGCTGGTCATGGCTGGGGAGACCAC 3848
Mouse       CACAGTGTTACAGAGGACTGCAACGTGCTTCTTAAGCTGGTCATGGCTGGGGAGGTGAC 3812
            ***

Bovine      CAACTCCCGCGGCCAGAAGCTGCTTCAAAGGGAGATGTGGAGATGTTGTGCGCGGGGCC 3665
Human       CAACTCCCGCGGCCAGCGGCTGCCCCAGAAGGGAGAGCTGGAGATGCTGTGCGCGGGGCC 3908
Mouse       CAACTCTCTGGGCCAAAGGCTGCCACAGAAGGGCGATGTGGAGATGCTGTGTGGTGGGCC 3872
            *****

Bovine      GCCCTGCCAGGGCTTTAGCGGCATGAACCGCTTCAACTCTCGAACCTACTCCAATTTCAA 3725
Human       GCCCTGCCAGGGCTTCAGCGGCATGAACCGCTTCAATTGCGCACCTACTCCAAGTTCAA 3968
Mouse       ACCCTGCCAGGGCTTCAGTGGCATGAACCGCTTCAACTCCCGCACTTACTCCAAGTTCAA 3932
            *****

Bovine      GAACTCCCTGGTGGTCTCTTCTCCTCAGCTACTGTGACTACTACCGCCCCGCTACTTCCT 3785
Human       AAACCTCTCTGGTGGTTTCCTTCTCCTCAGCTACTGCGACTACTACCGCCCCGGTTCTTCCT 4028
Mouse       AAACCTCCTAGTGGTCTCCTTCTCCTCAGCTACTGTGACTACTACCGCCTCGGTTCTTCCT 3992
            *****

Bovine      CTTGGAGAACGTTTCGGAACCTTCGTCTCCTTCAAGCGCTCCATGGTCTGAAGCTGACGCT 3845
Human       CCTGGAGAATGTCAGGAACCTTGTCTCCTTCAAGCGCTCCATGGTCTGAAGCTCACCT 4088
Mouse       TCTGGAGAACGTCAGGAACCTTCGTGTCTACAGACGCTCCATGGTCTGAAGCTCACACT 4052
            *****

Bovine      GCGCTGCCTGGTCCGCGAGGGGTACCAAGTGACCTTTGGCGTGCTGCAGGCTGGTCAGTA 3905
Human       CCGCTGCCTGGTCCGCATGGGCTATCAGTGACCTTCGGCGTGCTGCAGGCGGGTCAGTA 4148
Mouse       GCGCTGCCTGGTCCGCATGGGCTACCAAGTGACCTTTGGTGTGCTCCAGGCTGGACAGTA 4112
            *****

Bovine      CGGCGTGGCCAGACTCGGAGGCGAGCCATCATCCTGGCTGCAGCCCTGGGGAGCCACT 3965
Human       CGGCGTGGCCAGACTAGGAGGCGGGCCATCATCCTGGCCGCGGCCCTGGAGAGAAGCT 4208
Mouse       TGGCGTGGCCAGACACGAAGGAGGGCCATCATCTTGGCTGCAGCCCGAGGAGAAAAGCT 4172
            *****

Bovine      CCCGCTGTTCCCGGAGCCGTGTCATGTGTTTCGCACCCCGGGCCTGCCAGCTGAGCGTCGT 4025
Human       CCCTCTGTTCCCGGAGCCACTGCACGTGTTTGTCTCCCGGGCCTGCCAGCTGAGCGTGGT 4268
Mouse       GCCTCTGTTCCAGAGCCTCTGCATGTGTTTGCGCCCGCTGCCAGCTGAGCGTTGT 4232
            **

Bovine      AGTGGACGACAAGAAGTTTGTGAGCAACATCACCAGGTTGAGCTCGGGTCCCTTCCGAAC 4085
Human       GGTGGATGACAAGAAGTTTGTGAGCAACATAACCAGGTTGAGCTCGGGTCCCTTCCGGAC 4328
Mouse       GGTGGATGACAAGAAGTTTGTAGCAACATAACGAGGCTGAGCTCGGGGCCCTTCCGAAC 4292
            *****

```

Bovine	CATCACCGTGCGGGACACCATGTCTGACCTCCCTGAGATCCGGAACGGGGCCTCGGCACT	4145
Human	CATCACGGTGCGAGACACGATGTCCGACCTGCCGGAGGTGCGGAATGGAGCCTCGGCACT	4388
Mouse	CATCACCGTGCGAGACACCATGTCTGACCTCCCCGAGATCCAGAATGGAGCCTCGAATTC	4352
	*****	
Bovine	GGAGATTTCATACAACCGGGAGCCCCAGTCTGGTTCCAGAGGCAGCTCCGGGGCTCGCA	4205
Human	GGAGATCTCCTACAACGGGGAGCCTCAGTCCTGGTTCCAGAGGCAGCTCCGGGGCGCACA	4448
Mouse	TGAGATCCCCCTACAATGGAGAGCCACTGTCTGGTTCCAGAGGCAGCTGCGAGGATCACA	4412
	*****	
Bovine	GTACCAGCCCATCCTCAGGGATCATATTTGCAAGGACATGAGCGCCTTGGTGGCTGCCCG	4265
Human	GTACCAGCCCATCCTCAGGGACCACATCTGTAAGGACATGAGTGCAATTGGTGGCTGCCCG	4508
Mouse	CTACCAGCCCATCCTCAGGGACCATACTGCAAGGACATGAGCCCACTGGTGGCTGCCCG	4472
	*****	
Bovine	CATGCGGCACATCCCCCTGGCCCCGGGCTCGGACTGGCGTGACCTGCCCAACATTGAGGT	4325
Human	CATGCGGCACATCCCCCTGGCCCCAGGGTCAGACTGGCGCGATCTGCCCAACATCGAGGT	4568
Mouse	CATGCGGCACATCCCCTGTTCACAGGATCAGATTGGCGTGACCTGCCCAACATACAGGT	4532
	*****	
Bovine	GCGGCTCTCTGACGGCACCCTGGCCCCGAAGCTGCGGTACAACCTACCACGACAAGAAGAA	4385
Human	GCGGCTCTCAGACGGCACCATGGCCAGGAAGCTGCGGTATACCCACCATGACAGGAAGAA	4628
Mouse	GCGGCTGGGAGATGGCGTCATAGCCATAAGCTACAGTACACCTTTCATGATGTAAAAA	4592
	*****	
Bovine	CGGCTGCAGCAGCAGCGGGCCCTCCGTGGGGTCTGCTCCTGTGTGGAAG---GCAAGCC	4442
Human	CGGCCGCAGCAGCTCTGGGGCCCTCCGTGGGGTCTGCTCCTGCGTGGGAAGCCGGCAAGC	4688
Mouse	TGGCTACAGCAGCACCGGTGCCCCGTGGTGAGTCTGTTCCTGTGCAGAAG---GCAAGGC	4649
	***	
Bovine	CTGTGAGCCTGCGGCCGACAGTTTAAACACCCTTATCCCCCTGGTGCCTGCCCCACACTGG	4502
Human	CTGCGACCCCGCAGCCAGGCAGTTCAACACCCTCATCCCCCTGGTGCCTGCCCCACACCGG	4748
Mouse	CTGCGACCCTGAGTCCAGGCAATTACGACCCCTCATCCCCCTGGTGCCTGCCGCACACTGG	4709
	***	
Bovine	GAACAGGCACAACCACTGGGCGGCTCTACGGGCGTCTCGAGTGGGACGGCTTCTTCAG	4562
Human	GAACCGGCACAACCACTGGGCTGGCTCTATGGAAGGCTCGAGTGGGACGGCTTCTTCAG	4808
Mouse	GAACCGGCACAACCACTGGGCTGGCTCTACGGGCGTCTGGAGTGGGATGGCTTCTTCAG	4769
	****	
Bovine	CACAACTGTCAACAACCCCGAGCCCATGGGCAAGCAGGGCCGCGTGTCCACCCCGAGCA	4622
Human	CACAACCGTCAACAACCCCGAGCCCATGGGCAAGCAGGGCCGCGTGTCCACCCAGAGCA	4868
Mouse	CACCACTGTCAACAACCCCTGAGCCCATGGGCAAGCAGGGTGGGTGCTCCACCCGAGCA	4829
	***	
Bovine	GCACCGAGTGGTGAGCGTCCGGGAGTGCGCCGCTCCCAGGGCTTCCCCGACACCTATCG	4682
Human	GCACCGTGTGGTGAGCGTGGGGAGTGTCGCCGCTCCCAGGGCTTCCCTGACACCTACCG	4928
Mouse	GCACCGGTCTGTAGTGTTCGGGAATGTGCCGCTCCCAGGGCTTCCAGATAGCTACCG	4889
	*****	
Bovine	GCTGTTTCGGCAACATCCTAGACAAGCACCGGCAGGTGGGTAATGCTGTGCCCGCGCCACT	4742
Human	GCTCTTCGGCAACATCCTGGACAAGCACCGGCAGGTGGGCAATGCCGTGCCACCGCCCT	4988
Mouse	GTTCTTCGGCAACATCCTGGACAGACACCGGCAGGTGGGTAATGCTGTGCCACCACCCCT	4949
	* *	
Bovine	GGCCAAAGCCATCGGCTTGGAGATCAAGCGTGCATGTTGGCCAAAGCGCGGAGAGCGC	4802
Human	GGCCAAAGCCATTGGCTTGGAGATCAAGCTTGTATGTTGGCCAAAGCCCGAGAGAGTGC	5048
Mouse	GGCCAAAGCCATTGGCTTGGAGATTAAAGCTCTGCCTGCTGTCCAGTGTCTGGGAGAGCGC	5009
	*****	
Bovine	CTCAGCT-----AAAATCAAGGAGGAGG---CTGCCAAGGACTAGTCTCTCCTCCTAT	4853
Human	CTCAGCT-----AAAATAAGGAGGAGGAAGCTGCTAAGGACTAGTCTGCCCTCCCGT	5102
Mouse	ATCAGCTGCAGTTAAGCAAAAGAGGAGGCTGCTACCAAGGACTAG-----	5055
	*****	
Bovine	CACCCATGTTTCTGCCACCAGAGATCCCCAAGCTGCACTGATATTGGTGTATTTTTCACA	4913
Human	CACCCCTGTTTCTGGCACCAGGAATCCCCAAGCTGCACTGATGTTG---TGTTTTTAACA	5159
Mouse	-----	

Human NM\_001379

Mouse NM\_010066

**Bovine DNMT2**

**Bovine Dnmt2 Complete CDS**

Green = Start Site Red = Stop Codon

CCGGGGGGCGGGCGCGGGA**ATG**GAGCCCTTGCGGGGCCCTGGAGCTATACAGC  
GGAATTGGGGGCATGCACCAGGCTCTCAGAGAAAGCTGTATACCTGCACAA  
GTGGTGGCTGCTGTTGATGTAAACACTGTTGCTAATGAAGTATACAAGTATA  
ATTTTCCTCACACACAGTTACTGGCCAAGACAATTGAAGGCATTACACTAGA  
AGAGTTTGACAGATTATCTTTCAATATGATTTTAATGAGCCCACCCTGTCAGC  
CCTTCACAAGAATTGGCCTGCAAGGTGATGTGACTGATCCAAGGACAAATA  
GCTTCTTACATATTCTAGACATTCTCCCAAGATTACAAAAATTACCGAAGTA  
TATTCTTTTAGAAAAACGTTAAAGGTTTTGAAATGTCTTCTACAAGAGATCTGT  
TAATACAAACAATAGAAAATTGTGGTTTTTCAGTATCAAGAATTTCTACTGTC  
TCCAACCTCTCTTGGCATTCCAAATTCAAGATTACGGTACTTCCTTATTGCAA  
AGCTTCAGCCAGAGCCATTCCCTTTTCAGGCCCCCTGGTCAGGTACTGATGGA  
GTTCCCCAAAACCTGAATCTGAACATCCCCCTAAATATGCAATAAATGCAGAA  
AAGAAAACCTGAAGAAAAGAAAACCTGGACCAAAGATTTGCTTTGATAGCAGC  
ACACAGTGTTCTGGAAAAGAGGCCATTCTTTTTTAAGCTTGAAACTGCAGGAG  
AAATTGACAGGAAACATCAACAGGACAGCGATCTCTCTGTGCGAATGCTAA  
AAGATTTTCTTGAAGATGACATTGACAAGCATTCAATCTTTTTTACCACCAA  
GTCATTACTGCGATACGCTCTTTTGTTAGACATTGTTAAACCCACTTCCAGAA  
GATCCATGTGCTTTACAAAAGGTTATGGACGCTACATAGAAGGGACAGGAT  
CTGTGTTACAGACAACAGAGGATGTGCAGATTGAGAATATCTACAAATCCCT  
TACCAGTTTGTCAACAAGAAGAAAAGATAATGAGATTGTCAATGCTTCAACTT  
CGATTTTTTCACTCCTAAAGAAATAGCAAATCTCCTTGGATTTCCCTCCAGAGTT  
TGGATTTCCCTGAGATGACAACCTGTCAAACAGCGTTACCGTCTACTTGGAAT  
AGTCTCAACGTGCATGTTGTAGCTAACTAATCAAAATCCTATGTGAC**TAA**T  
TTTTTAAATAACTCTGAAAGAGGGTCACAGTTTTCTGTGCATATCCATATAGTA  
ACTTTGAAATTCTTTTTTTGAATTAATTTTGACAAAATTTGACTAAATTATTTT  
CTCTTTAATAAGA



Bovine	ACAAACAATAGAAAATGTGGTTTTCTAGTATCAAGAATTTCTACTGTCTCCAACCTCTCT	479
Human	ACAAACAATAGAAAATGTGGCTTTCTAGTACCAAGAATTTCTATTATCTCCAACCTCTCT	468
Mouse	ACAAACAATAGAAAGCCTGTGGCTTTCTAGTATCAAGAGTTTCTATTGTCTCTCTCTCTCT	480
	*****	
Bovine	TGGCATTCCAAATTCAGATTACGGTACTTCTTATTGCAAAGCTTCAGCCAGAGCCATT	539
Human	TGGCATTCCAAATTCAGAGCTACGATATTTCTTATTGCAAAGCTTCAGTCAGAGCCATT	528
Mouse	AGGCATTCCAACTCGAGGCTCCGATATTTCTCATGCAAAGCTTCAGTCAGAGCCATT	540
	*****	
Bovine	CCCTTTTCAGGCCCTGGTCAGGTACTGATGGAGTTCCTTAACTGAATCTGAACATCC	599
Human	ACCCTTTCAAGCCCCTGGTCAGGTACTGATGGAGTTCCTTAACTGAATCTGTACATCC	588
Mouse	CCCCTTCCAGGCCCTGGACAGATACTGATGGAGTTCCTTAACTGTAACTGTTGAGCC	600
	**	
Bovine	CCCTAAATATGCAATAAATGCAGAAAAGAAAAGTGAAGAAAAGAAAAGTGGACCAAAGAT	659
Human	ACAAAAATATGCAATGGATGTAGAAAATAAATTCAGAAAAGAACGTTGAACCAAATAT	648
Mouse	ACAAAAATATGCAGTAGTTGAAGAAAGTACGCCAAGAGTGCAGAGAACTGGACCAAGTAT	660
	*	
Bovine	TTGCTTTGA---TAGCAGCACACAGTGTCTGGAAGAGGCCATTCTTTTAAAGCTTGA	716
Human	TAGCTTTGA---TGGCAGCATACAGTGTCTGGAAGAGATGCCATTCTTTTAAAGCTTGA	705
Mouse	CTGTGCTGAGAGCAGCAGCACACAGAGTCTGGAAGAGATACCATTCTCTTAAAGCTTGA	720
	*	
Bovine	AACTGCAGGAGAAATTCAGAGAAACATCAACAGGACAGCGATCTCTCTGTGCGAATGCT	776
Human	AACTGCAGAGAAATTCACAGGAAATCAACAAGATAGTGATCTCTCTGTGAAATGCT	765
Mouse	GACTGTAGAAGAAAGGACAGGAAACATCAACAAGACAGTGACCTCTCTGTGAGATGCT	780
	****	
Bovine	AAAAGATTTTCTTGAAGATGACATTGACAAGCATTCTTTTACCACCAAAGTCATT	836
Human	AAAAGATTTTCTTGAAGATGACACTGACGTGAACAGTATCTTTTACCACCAAAGTCATT	825
Mouse	GAAAGACTTCTTGAAGATGGC---GACACAGATGAGTACCTTTTACCACCAAAGTATT	837
	*****	
Bovine	ACTGCGATACGCTCTTTTGTAGACATTGTTAAACCCACTTCCAGAAGATCCATGTGCTT	896
Human	GCTGCGATATGCTCTCTGTGTAGACATTGTTACGCCCACTTGTAGAAGGTCCTGTGCTT	885
Mouse	GCTGCGATACGCTCTCTTACTAGATATCGTGAAGCCACGTCAGAGGTCATGTGCTT	897
	*****	
Bovine	TACAAAAGGTTATGGACGCTACATAGAAGGGACAGGATCTGTGTTACAGACAACAGAGGA	956
Human	TACCAAAGGATATGGAAGCTACATAGAAGGGACAGGCTCTGTGTTACAGACTGCAGAGGA	945
Mouse	TACGAAGGGTATGGAGTTACATAGAAGGGACAGGCTCCGTGTTACAGGCTGCAGAGGA	957
	***	
Bovine	TGTGCAGATTGAGAATATCTACAAATCCCTTACCAGTTTGTACACAAGAAGAAAAGATAAT	1016
Human	TGTGCAGGTTGAGAATATCTACAAATCCCTTACCAATTTGTACACAAGAAGAACAGATAAC	1005
Mouse	TGCGCAGATTGAGAATATCTACAAATCTCTTCTGATTGCCACCAGAAGAAAAGATAGC	1017
	**	
Bovine	GAGATTGTCAATGCTTCAACTTCGATTTTCTACTCCTAAAGAAATAGCAAAATCTCCTGG	1076
Human	AAAGCTGTTAATACTTAACTGCGATATTTCTACTCCTAAAGAAATAGCAAAATCTCCTGG	1065
Mouse	TAAATTGTCAATGCTTAACTGCGATATTTACACCCGAAAGAAATGCAAAATCTCCAGGG	1077
	*	
Bovine	ATTTCTCCAGAGTTTGGATTTCCTGAGATGACAACTGTCAAACAGCGTTACCGTCTACT	1136
Human	ATTTCTCCAGAGTTTCGATTTCCTGAGAAGATAACAGTGAAACAGCGTTATCGCGTACT	1125
Mouse	ATTTCTCCAGAAATTGGGTTTCCTGAGAAGACAACAGTGAAACAGCGTTACCGGCTGCT	1137
	*****	
Bovine	TGGAATAGTCTCAACGTGCATGTTGTAGCTAACTAATCAAAATCCTATGTGACTAATT	1196
Human	TGGAATAGTCTCAACGTGCATGTAGTAGCTAACTAATCAAAATCCTATATGAATAATT	1185
Mouse	TGGCAATAGCCTCAACGTGCATGTGGTAGCAAACTCCTACAGTCCTGTGTGAAGGATT	1197
	***	
Bovine	TTTAAATAACTCTGAAAGAGGGTCACA---GTTTCTGTGATATCCATATAGTAACTTTG	1254
Human	TTG-AAATAACTCTGAAAGATGGTCATATGATATTCCTTCAATTTTCAGAGAGTAATCTG	1244
Mouse	TGG-AAATGCCTCTGAGAGCTG-TCACA--AGATGCCGCTAATTCTAGATAGTAATCCA	1253
	*	
Bovine	AAATTCTTTTGAATTAATTTTGACAAAATTTGACTAAATTAATTTT-----CTCTCT	1306
Human	AAATTCTGTTTGAATTAATTTTGAGGAAATTTAACTAAATTAATTTTAAAT---CTGTCC	1300
Mouse	AGATTCTATCTGAATGAATTTATAGAGTTTCAAGTAAATTTTGAATAGCAATTTTC	1313
	*	
Bovine	TTAATAAGA-----	1315

Human	TTATTAAGAAATTTGGATTTTATTAATAAAAAATCCATGTGTTCATCAAATTTATATTACT	1360
Mouse	TGGGTAATTCAGCAGGAACTTAAATGTGTACATCCAATTGTCCTCACATTATGTCACT	1373
	*       ***	

Human AF012128  
Mouse AF012129

## Bovine DNMT3b Protein Sequence

MKGVDSLINEDKHANRREDSVITDGAVIAQCCDSKQSPSPRILQSISTLE  
 IIGARGVRGRSSSRLSKREVSSLLSYTQDLTGDGDGEDGDGSDTPVM  
 PKLFRETRRSESPAVRTRNNSSTSTRERHRPSLRSTQGRQARNHVDESP  
 VAFSTTRSLRRRTGSSAGTPWSPASPYLTIDLDEDVVPQSSSTPYARL  
 GQDSQQESMESSQLDADGRDADSTEYQDGKEFGIGDLVSCGGKIKGFSWW  
 PAMVVSWKATSKRQAMSGMRWVQWFGDGKFSEIPADKLVALGLFSQHFNL  
 ATFNKLVSYRKAMYHALEKARIRAGKMFSSPGDSLEDQLKPMLEWAHGG  
 FKPTGVEGLKPNKQPENKTRRRTADDSATSDYCPPPKRLKTNCYNNGKD  
 RGEEDQSREQMASDVASNKGNLEDSCLSCGRKNPVSFHPLFEGGLCQTCR  
 DRFLELFYMYDDDGYQSYCTVCCEGRELLCSNTSCRCFCVECLEVLVG  
 AGTAAEAKLQEPWSCYMCLPQRCHGILRRRKDWSVRLQAFFTSDPGLEYE  
 APKLYPAIPANRRRPIRVLSLFDGIATGYLVLKELGIKVEKYVASEVCEE  
 SIAVGTVKHEGNIKYVNDVRNITKKNIEEWGPFDLVIGGSPCNDLSNVP  
 ARKGLYEGTGRLFFEFYHLLNYTRPKEGEDRPFFWMFENVVAMKVGDKRD  
 ISRFLCNPVMIDAIKVSAAHRARYFWGNLPGMNRPVIAKNDKLELQDC  
 LEFNRTAKLKKVQTITTKSNSIRQGKNQLFPVVMNGKEDVLWCTELERIF  
 GFPVHYTDVSNMGRVARQKLLGRSWR

## Bovine Dnmt3b Isoform Sequence Allignment

DNMT3b5	GNNTGAAGTAAGCATGAAGGGAGTCGACAGCCTAATCAATGAAGACAAGCAGGCCAACAG	60
DNMT3b3	GNNTGAAGTAAGCATGAAGGGAGTCGACAGCCTAATCAATGAAGACAAGCAGGCCAACAG	60
DNMT3b1	GNNTGAAGTAAGCATGAAGGGAGTCGACAGCCTAATCAATGAAGACAAGCAGGCCAACAG	60
DNMT3b4	GNNTGAAGTAAGCATGAAGGGAGTCGACAGCCTAATCAATGAAGACAAGCAGGCCAACAG	60
*****		
DNMT3b5	AAGGGAAGACTCCGTCATCACCGACGGGGCCGTCATCGCCAGTGTGTGACTCCAAGCA	120
DNMT3b3	AAGGGAAGACTCCGTCATCACCGACGGGGCCGTCATCGCCAGTGTGTGACTCCAAGCA	120
DNMT3b1	AAGGGAAGACTCCGTCATCACCGACGGGGCCGTCATCGCCAGTGTGTGACTCCAAGCA	120
DNMT3b4	AAGGGAAGACTCCGTCATCACCGACGGGGCCGTCATCGCCAGTGTGTGACTCCAAGCA	120
*****		
DNMT3b5	GTCTCCTTCACCCCGGATCCTGCAGTCTATCAGCACCTGGAGATCATAGGTGCCCGAGG	180
DNMT3b3	GTCTCCTTCACCCCGGATCCTGCAGTCTATCAGCACCTGGAGATCATAGGTGCCCGAGG	180
DNMT3b1	GTCTCCTTCACCCCGGATCCTGCAGTCTATCAGCACCTGGAGATCATAGGTGCCCGAGG	180
DNMT3b4	GTCTCCTTCACCCCGGATCCTGCAGTCTATCAGCACCTGGAGATCATAGGTGCCCGAGG	180
*****		
DNMT3b5	TGTCAGAGGCCGAGATCAAGCTCAGACTGTCCAAGAGGGAGGTCTCCAGCCTGCTAAG	240
DNMT3b3	TGTCAGAGGCCGAGATCAAGCTCAGACTGTCCAAGAGGGAGGTCTCCAGCCTGCTAAG	240
DNMT3b1	TGTCAGAGGCCGAGATCAAGCTCAGACTGTCCAAGAGGGAGGTCTCCAGCCTGCTAAG	240
DNMT3b4	TGTCAGAGGCCGAGATCAAGCTCAGACTGTCCAAGAGGGAGGTCTCCAGCCTGCTAAG	240
*****		
DNMT3b5	TTATACTCAGGACCTGACGGGTGATGGAGATGGCGAGGGAGAAGACGGGGATGGCTCCGA	300
DNMT3b3	TTATACTCAGGACCTGACGGGTGATGGAGATGGCGAGGGAGAAGACGGGGATGGCTCCGA	300
DNMT3b1	TTATACTCAGGACCTGACGGGTGATGGAGATGGCGAGGGAGAAGACGGGGATGGCTCCGA	300
DNMT3b4	TTATACTCAGGACCTGACGGGTGATGGAGATGGCGAGGGAGAAGACGGGGATGGCTCCGA	300
*****		
DNMT3b5	CACTCCAGTGATGCCAAAGCTCTTCCGAGAAACCAGGACTCGGTCTGAAAGCCCAGCTGT	360
DNMT3b3	CACTCCAGTGATGCCAAAGCTCTTCCGAGAAACCAGGACTCGGTCTGAAAGCCCAGCTGT	360
DNMT3b1	CACTCCAGTGATGCCAAAGCTCTTCCGAGAAACCAGGACTCGGTCTGAAAGCCCAGCTGT	360
DNMT3b4	CACTCCAGTGATGCCAAAGCTCTTCCGAGAAACCAGGACTCGGTCTGAAAGCCCAGCTGT	360
*****		

```

DNMT3b5      CCGAACCCGAAATAACAGCAGTACCTCCACCCGGGAGAGGCACAGGCCCTCCCTACGTTT 420
DNMT3b3      CCGAACCCGAAATAACAGCAGTACCTCCACCCGGGAGAGGCACAGGCCCTCCCTACGTTT 420
DNMT3b1      CCGAACCCGAAATAACAGCAGTACCTCCACCCGGGAGAGGCACAGGCCCTCCCTACGTTT 420
DNMT3b4      CCGAACCCGAAATAACAGCAGTACCTCCACCCGGGAGAGGCACAGGCCCTCCCTACGTTT 420
              *****

DNMT3b5      CACCCAAGGCCGGCAGGCCCGCAACCACGTGGATGAATCCCCTGTGGCGTTCTCAACTAC 480
DNMT3b3      CACCCAAGGCCGGCAGGCCCGCAACCACGTGGATGAATCCCCTGTGGCGTTCTCAACTAC 480
DNMT3b1      CACCCAAGGCCGGCAGGCCCGCAACCACGTGGATGAATCCCCTGTGGCGTTCTCAACTAC 480
DNMT3b4      CACCCAAGGCCGGCAGGCCCGCAACCACGTGGATGAATCCCCTGTGGCGTTCTCAACTAC 480
              *****

DNMT3b5      CAGGTCCCTGAGGCGAAGGACGGGATCCTCTGCAGGCACGCCATGGCCGTCCTCCGCCAG 540
DNMT3b3      CAGGTCCCTGAGGCGAAGGACGGGATCCTCTGCAGGCACGCCATGGCCGTCCTCCGCCAG 540
DNMT3b1      CAGGTCCCTGAGGCGAAGGACGGGATCCTCTGCAGGCACGCCATGGCCGTCCTCCGCCAG 540
DNMT3b4      CAGGTCCCTGAGGCGAAGGACGGGATCCTCTGCAGGCACGCCATGGCCGTCCTCCGCCAG 540
              *****

DNMT3b5      CCCCTACCTCACCATCGACCTCACAGACGAGGATGTGGTGCCGAGAGCAGCAGCAGGCC 600
DNMT3b3      CCCCTACCTCACCATCGACCTCACAGACGAGGATGTGGTGCCGAGAGCAGCAGCAGGCC 600
DNMT3b1      CCCCTACCTCACCATCGACCTCACAGACGAGGATGTGGTGCCGAGAGCAGCAGCAGGCC 600
DNMT3b4      CCCCTACCTCACCATCGACCTCACAGACGAGGATGTGGTGCCGAGAGCAGCAGCAGGCC 600
              *****

DNMT3b5      CTACGCCCCGCTGGGCCAGGACAGCCAGCAGGAGAGCATGGAGTCTCGCAGCTGGACGC 660
DNMT3b3      CTACGCCCCGCTGGGCCAGGACAGCCAGCAGGAGAGCATGGAGTCTCGCAGCTGGACGC 660
DNMT3b1      CTACGCCCCGCTGGGCCAGGACAGCCAGCAGGAGAGCATGGAGTCTCGCAGCTGGACGC 660
DNMT3b4      CTACGCCCCGCTGGGCCAGGACAGCCAGCAGGAGAGCATGGAGTCTCGCAGCTGGACGC 660
              *****

DNMT3b5      AGACGGCAGAGATGCAGACAGCACTGAGTATCAGGATGGGAAGGAGTTTGAATAGGAGA 720
DNMT3b3      AGACGGCAGAGATGCAGACAGCACTGAGTATCAGGATGGGAAGGAGTTTGAATAGGAGA 720
DNMT3b1      AGACGGCAGAGATGCAGACAGCACTGAGTATCAGGATGGGAAGGAGTTTGAATAGGAGA 720
DNMT3b4      AGACGGCAGAGATGCAGACAGCACTGAGTATCAGGATGGGAAGGAGTTTGAATAGGAGA 720
              *****

DNMT3b5      TCTTGTGTCGTGTGGGGGAAAGATCAAGGGTTTCTCCTGGTGGCCTGCCATGGTGGTGTC 780
DNMT3b3      TCTTGTGTCGTGTGGGGGAAAGATCAAGGGTTTCTCCTGGTGGCCTGCCATGGTGGTGTC 780
DNMT3b1      TCTTGTGTCGTGTGGGGGAAAGATCAAGGGTTTCTCCTGGTGGCCTGCCATGGTGGTGTC 780
DNMT3b4      TCTTGTGTCGTGTGGGGGAAAGATCAAGGGTTTCTCCTGGTGGCCTGCCATGGTGGTGTC 780
              *****

DNMT3b5      CTGGAAGGCCACCTCCAAGCGGCAGGCAATGTCTGGCATGCGGTGGGTCCAGTGGTTTGG 840
DNMT3b3      CTGGAAGGCCACCTCCAAGCGGCAGGCAATGTCTGGCATGCGGTGGGTCCAGTGGTTTGG 840
DNMT3b1      CTGGAAGGCCACCTCCAAGCGGCAGGCAATGTCTGGCATGCGGTGGGTCCAGTGGTTTGG 840
DNMT3b4      CTGGAAGGCCACCTCCAAGCGGCAGGCAATGTCTGGCATGCGGTGGGTCCAGTGGTTTGG 840
              *****

DNMT3b5      TGATGGCAAGTTCTCCGAGATTCAGCAGATAAGTTGGTGGCATTGGGATTGTTTCAAGCA 900
DNMT3b3      TGATGGCAAGTTCTCCGAGATTCAGCAGATAAGTTGGTGGCATTGGGATTGTTTCAAGCA 900
DNMT3b1      TGATGGCAAGTTCTCCGAGATTCAGCAGATAAGTTGGTGGCATTGGGATTGTTTCAAGCA 900
DNMT3b4      TGATGGCAAGTTCTCCGAGATTCAGCAGATAAGTTGGTGGCATTGGGATTGTTTCAAGCA 900
              *****

```

DNMT3b5	GCAC TTAACTTGGCGACCTTCAATAAGCTGGTCTCTTACAGGAAGGCCATGTACCATGC	960
DNMT3b3	GCAC TTAACTTGGCGACCTTCAATAAGCTGGTCTCTTACAGGAAGGCCATGTACCATGC	960
DNMT3b1	GCAC TTAACTTGGCGACCTTCAATAAGCTGGTCTCTTACAGGAAGGCCATGTACCATGC	960
DNMT3b4	GCAC TTAACTTGGCGACCTTCAATAAGCTGGTCTCTTACAGGAAGGCCATGTACCATGC	960
	*****	
DNMT3b5	TCTGGAGAAAGCCAGGATCCGGGCTGGCAAGATGTTCCCCAGCAGCCCTGGAGACTCACT	1020
DNMT3b3	TCTGGAGAAAGCCAGGATCCGGGCTGGCAAGATGTTCCCCAGCAGCCCTGGAGACTCACT	1020
DNMT3b1	TCTGGAGAAAGCCAGGATCCGGGCTGGCAAGATGTTCCCCAGCAGCCCTGGAGACTCACT	1020
DNMT3b4	TCTGGAGAAAGCCAGGATCCGGGCTGGCAAGATGTTCCCCAGCAGCCCTGGAGACTCACT	1020
	*****	
DNMT3b5	GGAGGATCAGCTGAAGCCCATGTTGGAGTGGGCCCATGGAGGCTTTAAGCCCACTGGGGT	1080
DNMT3b3	GGAGGATCAGCTGAAGCCCATGTTGGAGTGGGCCCATGGAGGCTTTAAGCCCACTGGGGT	1080
DNMT3b1	GGAGGATCAGCTGAAGCCCATGTTGGAGTGGGCCCATGGAGGCTTTAAGCCCACTGGGGT	1080
DNMT3b4	GGAGGATCAGCTGAAGCCCATGTTGGAGTGGGCCCATGGAGGCTTTAAGCCCACTGGGGT	1080
	*****	
DNMT3b5	CGAGGGTCTCAAACCTAACACAAGCAACCAGAGAATAAGACGCGGAGACGCACAGCTGA	1140
DNMT3b3	CGAGGGTCTCAAACCTAACACAAGCAACCAGAGAATAAGACGCGGAGACGCACAGCTGA	1140
DNMT3b1	CGAGGGTCTCAAACCTAACACAAGCAACCAGAGAATAAGACGCGGAGACGCACAGCTGA	1140
DNMT3b4	CGAGGGTCTCAAACCTAACACAAGCAACCAGAGAATAAGACGCGGAGACGCACAGCTGA	1140
	*****	
DNMT3b5	CGACTCAGCCACCTCTGACTACTGCCCCCACCAGCGCCTCAAGACAAATTGTTACAA	1200
DNMT3b3	CGACTCAGCCACCTCTGACTACTGCCCCCACCAGCGCCTCAAGACAAATTGTTACAA	1200
DNMT3b1	CGACTCAGCCACCTCTGACTACTGCCCCCACCAGCGCCTCAAGACAAATTGTTACAA	1200
DNMT3b4	CGACTCAGCCACCTCTGACTACTGCCCCCACCAGCGCCTCAAGACAAATTGTTACAA	1200
	*****	
DNMT3b5	CAACGGCAAGGACCGAGGAGAGGAGGACCAGAGTCGAGAACAAATGGCTTCGGATGTTGC	1260
DNMT3b3	CAACGGCAAGGACCGAGGAGAGGAGGACCAGAGTCGAGAACAAATGGCTTCGGATGTTGC	1260
DNMT3b1	CAACGGCAAGGACCGAGGAGAGGAGGACCAGAGTCGAGAACAAATGGCTTCGGATGTTGC	1260
DNMT3b4	CAACGGCAAGGACCGAGGAGAGGAGGACCAGAGTCGAGAACAAATGGCTTCGGATGTTGC	1260
	*****	
DNMT3b5	CAGCAACAAAGGCAATCTGGAAGATAGCTGTTTGTCTGTGGTAGGAAAAACCCCGTGTC	1320
DNMT3b3	CAGCAACAAAGGCAATCTGGAAGATAGCTGTTTGTCTGTGGTAGGAAAAACCCCGTGTC	1320
DNMT3b1	CAGCAACAAAGGCAATCTGGAAGATAGCTGTTTGTCTGTGGTAGGAAAAACCCCGTGTC	1320
DNMT3b4	CAGCAACAAAGGCAATCTGGAAGATAGCTGTTTGTCTGTGGTAGGAAAAACCCCGTGTC	1320
	*****	
DNMT3b5	CTTCCACCCCTCTCTTTGAGGGTGGGCTCTGCCAGACATGCCGGGACCGCTTCCTCGAGCT	1380
DNMT3b3	CTTCCACCCCTCTCTTTGAGGGTGGGCTCTGCCAGACATGCCGGGACCGCTTCCTCGAGCT	1380
DNMT3b1	CTTCCACCCCTCTCTTTGAGGGTGGGCTCTGCCAGACATGCCGGGACCGCTTCCTCGAGCT	1380
DNMT3b4	CTTCCACCCCTCTCTTTGAGGGTGGGCTCTGCCAGACATGCCGGGACCGCTTCCTCGAGCT	1380
	*****	
DNMT3b5	CTTCTACATGTACGACGACGACGGCTACCAGTCGTACTGCACCGTGTGCTGCGAGGGCCG	1440
DNMT3b3	CTTCTACATGTACGACGACGACGGCTACCAGTCGTACTGCACCGTGTGCTGCGAGGGCCG	1440
DNMT3b1	CTTCTACATGTACGACGACGACGGCTACCAGTCGTACTGCACCGTGTGCTGCGAGGGCCG	1440
DNMT3b4	CTTCTACATGTACGACGACGACGGCTACCAGTCGTACTGCACCGTGTGCTGCGAGGGCCG	1440
	*****	
DNMT3b5	CGAGCTGCTCCTGTGCAGCAACACGAGCTGCTGCCGGTGCTTCTGCGTGGAGTGTCTGGA	1500
DNMT3b3	CGAGCTGCTCCTGTGCAGCAACACGAGCTGCTGCCGGTGCTTCTGCGTGGAGTGTCTGGA	1500
DNMT3b1	CGAGCTGCTCCTGTGCAGCAACACGAGCTGCTGCCGGTGCTTCTGCGTGGAGTGTCTGGA	1500
DNMT3b4	CGAGCTGCTCCTGTGCAGCAACACGAGCTGCTGCCGGTGCTTCTGCGTGGAGTGTCTGGA	1500
	*****	

DNMT3b5	GGTGTCTGGTGGGCGCGGGCACGGCGGCAGAGGCCAAGCTGCAGGAGCCCTGGAGTTGCTA	1560
DNMT3b3	GGTGTCTGGTGGGCGCGGGCACGGCGGCAGAGGCCAAGCTGCAGGAGCCCTGGAGTTGCTA	1560
DNMT3b1	GGTGTCTGGTGGGCGCGGGCACGGCGGCAGAGGCCAAGCTGCAGGAGCCCTGGAGTTGCTA	1560
DNMT3b4	GGTGTCTGGTGGGCGCGGGCACGGCGGCAGAGGCCAAGCTGCAGGAGCCCTGGAGTTGCTA	1560
	*****	
DNMT3b5	CATGTGTCTCCCGCAGCGTTGCCACGGCATCCTGCGGCGCCGCAAGGACTGGAGTGTGCG	1620
DNMT3b3	CATGTGTCTCCCGCAGCGTTGCCACGGCATCCTGCGGCGCCGCAAGGACTGGAGTGTGCG	1620
DNMT3b1	CATGTGTCTCCCGCAGCGTTGCCACGGCATCCTGCGGCGCCGCAAGGACTGGAGTGTGCG	1620
DNMT3b4	CATGTGTCTCCCGCAGCGTTGCCACGGCATCCTGCGGCGCCGCAAGGACTGGAGTGTGCG	1620
	*****	
DNMT3b5	TCTGCAGGCCCTTCTTACCAGCGACCCGGGCTCGAATATGAAGCCCCAAGTTATACCC	1680
DNMT3b3	TCTGCAGGCCCTTCTTACCAGCGACCCGGGCTCGAATATGAAGCCCCAAGTTATACCC	1680
DNMT3b1	TCTGCAGGCCCTTCTTACCAGCGACCCGGGCTCGAATATGAAGCCCCAAGTTATACCC	1680
DNMT3b4	TCTGCAGGCCCTTCTTACCAGCGACCCGGGCTCGAATATGAAGCCCCAAGTTATACCC	1680
	*****	
DNMT3b5	TGCGATTCTTGCAAACCGAAGGCGGCTATTCGAGTCTTGTCACTGTTTGATGGAATTGC	1740
DNMT3b3	TGCGATTCTTGCAAACCGAAGGCGGCTATTCGAGTCTTGTCACTGTTTGATGGAATTGC	1740
DNMT3b1	TGCGATTCTTGCAAACCGAAGGCGGCTATTCGAGTCTTGTCACTGTTTGATGGAATTGC	1740
DNMT3b4	TGCGATTCTTGCAAACCGAAGGCGGCTATTCGAGTCTTGTCACTGTTTGATGGAATTGC	1740
	*****	
DNMT3b5	AACAGGGTACTTGGTCTCAAAGAACTGGGCATCAAAGTGGAGAAATACGTGGCCTCCGA	1800
DNMT3b3	AACAGGGTACTTGGTCTCAAAGAACTGGGCATCAAAGTGGAGAAATACGTGGCCTCCGA	1800
DNMT3b1	AACAGGGTACTTGGTCTCAAAGAACTGGGCATCAAAGTGGAGAAATACGTGGCCTCCGA	1800
DNMT3b4	AACAGGGTACTTGGTCTCAAAGAACTGGGCATCAAAGTGGAGAAATACGTGGCCTCCGA	1800
	*****	
DNMT3b5	AGTGTGTGAAGAGTCCATTGCCGTTGGCACCGTTAAGCACGAGGGCAACATCAAATACGT	1860
DNMT3b3	AGTGTGTGAAGAGTCCATTGCCGTTGGCACCGTTAAGCACGAGGGCAACATCAAATACGT	1860
DNMT3b1	AGTGTGTGAAGAGTCCATTGCCGTTGGCACCGTTAAGCACGAGGGCAACATCAAATACGT	1860
DNMT3b4	AGTGTGTGAAGAGTCCATTGCCGTTGGCACCGTTAAGCACGAGGGCAACATCAAATACGT	1860
	*****	
DNMT3b5	GAATGACGTCAGGAATATCAGAAAGAAAAACATTGAAGAATGGGGCCCATTTGACTTGGT	1920
DNMT3b3	GAATGACGTCAGGAATATCAGAAAGAAAAACATTGAAGAATGGGGCCCATTTGACTTGGT	1920
DNMT3b1	GAATGACGTCAGGAATATCAGAAAGAAAAACATTGAAGAATGGGGCCCATTTGACTTGGT	1920
DNMT3b4	GAATGACGTCAGGAATATCAGAAAGAAAAACATTGAAGAATGGGGCCCATTTGACTTGGT	1920
	*****	
DNMT3b5	GATTGGTGGAAGCCCATGCAATGATCTCTCCAATGTGAACCCTGCCAGAAAAGGCCTGTA	1980
DNMT3b3	GATTGGTGGAAGCCCATGCAATGATCTCTCCAATGTGAACCCTGCCAGAAAAGGCCTGTA	1980
DNMT3b1	GATTGGTGGAAGCCCATGCAATGATCTCTCCAATGTGAACCCTGCCAGAAAAGGCCTGTA	1980
DNMT3b4	GATTGGTGGAAGCCCATGCAATGATCTCTCCAATGTGAACCCTGCCAGAAAAGGCCTGTA	1980
	*****	
DNMT3b5	TGAGGGCACAGGCCGGCTCTTCTTTGAGTTCTACCACCTGCTGAATTACACTCGCCCCAA	2040
DNMT3b3	TGAGGGCACAGGCCGGCTCTTCTTTGAGTTCTACCACCTGCTGAATTACACTCGCCCCAA	2040
DNMT3b1	TGAGGGCACAGGCCGGCTCTTCTTTGAGTTCTACCACCTGCTGAATTACACTCGCCCCAA	2040
DNMT3b4	TGAGGGCACAGGCCGGCTCTTCTTTGAGTTCTACCACCTGCTGAATTACACTCGCCCCAA	2040
	*****	
DNMT3b5	GGAGGGTGAAGACCGGCTTTCTTCTGGATGTTTGAGAATGTGGTGGCCATGAAGGTTGG	2100
DNMT3b3	GGAGGGTGAAGACCGGCTTTCTTCTGGATGTTTGAGAATGTGGTGGCCATGAAGGTTGG	2100
DNMT3b1	GGAGGGTGAAGACCGGCTTTCTTCTGGATGTTTGAGAATGTGGTGGCCATGAAGGTTGG	2100
DNMT3b4	GGAGGGTGAAGACCGGCTTTCTTCTGGATGTTTGAGAATGTGGTGGCCATGAAGGTTGG	2100
	*****	

```

DNMT3b5      CGACAAGCGGGACATCTCTCGGTTTTTGGAGTGTAACCCAGTGATGATTGATGCCATCAA 2160
DNMT3b3      CGACAAGCGGGACATCTCTCGGTTTTTGGAGTGTAACCCAGTGATGATTGATGCCATCAA 2160
DNMT3b1      CGACAAGCGGGACATCTCTCGGTTTTTGGAGTGTAACCCAGTGATGATTGATGCCATCAA 2160
DNMT3b4      CGACAAGCGGGACATCTCTCGGTTTTTGGAGTGTAACCCAGTGATGATTGATGCCATCAA 2160
*****

DNMT3b5      AGTGTCTGCTGCTCACAGAGCCCGATACTTCTGGGGCAACCTGCCCGGGATGAACAGGCC 2220
DNMT3b3      AGTGTCTGCTGCTCACAGAGCCCGATACTTCTGGGGCAACCTGCCCGGGATGAACAG--- 2217
DNMT3b1      AGTGTCTGCTGCTCACAGAGCCCGATACTTCTGGGGCAACCTGCCCGGGATGAACAGGCC 2220
DNMT3b4      AGTGTCTGCTGCTCACAGAGCCCGATACTTCTGGGGCAACCTGCCCGGGATGAACAG--- 2217
*****

DNMT3b5      TGTGATAGCATCCAAGAATGATAAGCTCGAGCTGCAGGACTGCCTGGAGTTCAAT----- 2275
DNMT3b3      -----
DNMT3b1      TGTGATAGCATCCAAGAATGATAAGCTCGAGCTGCAGGACTGCCTGGAGTTCAATAGGAC 2280
DNMT3b4      -----

DNMT3b5      ----- 2259
DNMT3b3      ----- 2218
DNMT3b1      AGCAAAGTTAAAGAAAGTACAGACAATAACCACCAAGTCGAACTCGATCAGACAGGGGAA 2340
DNMT3b4      -----TTAAAGAAAGTACAGACAATAACCACCAAGTCGAACTCGATCAGACAGGGGAA 2270

DNMT3b5      -----AGGA 2279
DNMT3b3      -----
DNMT3b1      AAACCAACTTTTCCCTGTTGTCATGAATGGCAAAGAAGATGTTTTGTGGTGCACTGAGCT 2400
DNMT3b4      AAACCAACTTTTCCCTGTTGTCATGAATGGCAAAGAAGATGTTTTGTGGTGCACTGAGCT 2330

DNMT3b5      CAGCAGGATCTTCGGCTTTCCCTGTGCACTACACAGACGTCTCCAACATGGGCCGTGTGGC 2339
DNMT3b3      -----GATCTTCGGCTTTCCCTGTGCACTACACAGACGTCTCCAACATGGGCCGTGTGGC 2271
DNMT3b1      AGAAAGGATCTTCGGCTTTCCCTGTGCACTACACAGACGTCTCCAACATGGGCCGTGTGGC 2460
DNMT3b4      AGAAAGGATCTTCGGCTTTCCCTGTGCACTACACAGACGTCTCCAACATGGGCCGTGTGGC 2390
*****

DNMT3b5      CCGTCAGAAGCTGCTGGGAAGGTCCTGGAGA 2370
DNMT3b3      CCGTCAGAAGCTGCTGGGAAGGTCCTGGAGA 2302
DNMT3b1      CCGTCAGAAGCTGCTGGGAAGGTCCTGGAGA 2491
DNMT3b4      CCGTCAGAAGCTGCTGGGAAGGTCCTGGAGA 2421
*****

```

## VITA

Michael Cameron Golding  
11181 Charlton Dr.  
Denfield Ontario, Canada  
N0M 1P0  
Born: 10 March 1977

### Education

- 2003 Ph.D. Veterinary Physiology  
Texas A&M University, College Station TX USA
- 2000 H.B.Sc. Genetics  
University Of Western Ontario, London, ON Canada

### Selected Publications

1. D. Holstead Jones, Michael C. Golding, Kevin J. Barr, Guo-Hua Fong and Gerald M. Kidder. (2001) The mouse Na<sup>+</sup> -K<sup>+</sup> -ATPase  $\gamma$ -subunit gene (Fxyd2) encodes three developmentally regulated transcripts. *Physiological Genomics* **6**: 129-135.
2. Michael C. Golding and Mark E. Westhusin. (2003) Analysis of DNA (cytosine 5) Methyltransferase mRNA Sequence and Expression in Bovine Preimplantation Embryos, Fetal and Adult Tissues. *Gene Expression Patterns* **5**:551-558