2009

DNA methylation analysis and identification of a novel antisense transcript in the Peg3 imprinted domain

Jennifer Marian Huang

Louisiana State University and Agricultural and Mechanical College

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_dissertations

Recommended Citation

https://digitalcommons.lsu.edu/gradschool_dissertations/1070
DNA METHYLATION ANALYSIS AND IDENTIFICATION OF A NOVEL ANTISENSE TRANSCRIPT IN THE PEG3 IMPRINTED DOMAIN

A Dissertation

Submitted to the Graduate Faculty of the Louisiana State University and Agricultural and Mechanical College in partial fulfillment of the requirements for the degree of Doctor of Philosophy

in

The Department of Biological Sciences

By
Jennifer Marian Huang
B.S., Louisiana State University, 2004
December 2009
ACKNOWLEDGEMENTS

I am pleased to express my gratitude to my advisor, Dr. Joomyeong Kim, for his constant support and guidance during the past five years. This has been an incredible journey, and I cannot imagine having a better mentor to guide me through it. I would also like to thank my committee members: Dr. Mark Batzer, Dr. Stephania Cormier, Dr. David Donze, and Dr. Arthur Penn. My graduate career has been greatly improved by their thoughtful suggestions.

I am thankful to my colleagues in the Kim lab for making this such a great place to pursue a Ph.D. I would like to thank Dr. Jeong Do Kim and Dr. Jungha Choo for helping me learn not only how to do science but how to think about science. I would also like to thank Christopher Faulk, Hana Kim, Keunsoo Kang, and Muhammad Ekram for their contributions to my research and critiques of all of my manuscripts and talks.

The research contained within this dissertation would not have been possible without the contributions of many people. For chapter two, I would like to thank Michelle Aucoin, Uduak Udoh, Letitia Lacour, Dr. Jeong Do Kim, and Anne Bergmann for their technical assistance. In chapter three, Dr. Mark Batzer and Jerilyn Walker provided the chimpanzee and rhesus macaque DNA samples. I would also like to thank Priscilla Milligan for answering my many questions about the Ph.D process and making sure my paperwork was always in order.

I must thank my friends and family for their encouragement and for helping me find perspective on my work. I appreciate the efforts of my siblings Amy Thompson, Michael Thompson, Joseph Thompson, and Robert Thompson to understand what I spent all this time doing. I also thank Natalie Vance, Sarahfaye Mahon, Eve McCulloch, and Deepa Srikanta for all of their support. Last but not least, I would like to thank my husband Charles Huang for his constant love, support, and supply of chocolate throughout my graduate studies.
TABLE OF CONTENTS

ACKNOWLEDGEMENTS .......................................................................................................................... ii
LIST OF TABLES ........................................................................................................................................ iv
LIST OF FIGURES ....................................................................................................................................... v
ABSTRACT .................................................................................................................................................. vi
CHAPTER ONE: BACKGROUND .............................................................................................................. 1
CHAPTER TWO: DNA METHYLATION ANALYSIS OF THE MAMMALIAN PEG3 IMPRINTED DOMAIN .............................................................................................................. 11
CHAPTER THREE: IDENTIFICATION OF AN ANTISENSE TRANSCRIPT TO ZIM2 IN THE PRIMATE LINEAGE ........................................................................................................ 34
CHAPTER FOUR: CONCLUSIONS .......................................................................................................... 50
APPENDIX A: SUPPLEMENTAL DATA ................................................................................................. 60
APPENDIX B: LETTERS OF PERMISSION ............................................................................................ 67
VITA ............................................................................................................................................................ 70
LIST OF TABLES

2.1 Analysis of CpG islands and repeat content in the PEG3 imprinted region ........................................ 15
LIST OF FIGURES

1.1 The DNA methylation reaction ........................................................................................................................................... 4

1.2 The bisulfite conversion reaction ........................................................................................................................................ 6

2.1 Organization of the PEG3 genomic region ....................................................................................................................... 14

2.2A Bisulfite sequencing and COBRA analyses of selected CpG islands in the mouse Peg3 domain ................................................................................................................................................ 17

2.2B Summary of methylation at different developmental time points in each of the four loci tested. ............................................................................................................................................... 19

2.3 Bisulfite sequencing and COBRA analyses of the cow Peg3 domain ................................................................................................................................................ 20

2.4 COBRA analysis of the CpG islands located in the human PEG3 imprinted region ................................................................................................................................................ 22

2.5 Bisulfite sequencing of the CpG islands located in the human USP29 and PEG3 promoter regions ................................................................................................................................................ 24

2.6 Summary of the methylation status of CpG islands in the PEG3 region ................................................................................................................................................ 26

3.1 Organization of the region between ZFP28 and PEG3 in the mouse, rhesus macaque, chimpanzee, and human genomes ................................................................................................................................................ 37

3.2 Structure of the ZIM2as transcript and the surrounding sequence ................................................................................................................................................ 38

3.3 Expression analysis of genes in the vicinity of ZIM2as ................................................................................................................................................ 40

3.4 DNA methylation analysis of ZIM2as ................................................................................................................................................ 41

3.5 Primate-specific changes in the PEG3 region ................................................................................................................................................ 44

4.1 Expression patterns of Peg3, Usp29, and Zim3 ................................................................................................................................................ 54

4.2 Expression patterns of ZNF835, ZIM2as, and PEG3/ZIM2 ................................................................................................................................................ 57
ABSTRACT

Genomic imprinting is a process that leads to the silencing of one allele of a gene in a parent of origin specific manner. Genes that are involved in this process are often regulated in clusters, one of which is the Peg3 (Paternally expressed gene 3) imprinted domain. We investigated this region for both CpG islands and long antisense transcripts, two common features of imprinted gene clusters.

First, we performed a systematic survey of DNA methylation status of the CpG islands in this region of the mouse, cow, and human genomes. We identified two previously unreported differentially methylated regions (DMR): one in the promoter region of mouse Zim3 and another in the promoter region of human USP29. The PEG3-CpG island is the only DMR that is conserved among these three species. PEG3 has been implicated in several types of cancer, so we examined the methylation status of several CpG islands in this region using human tumor derived DNA. The CpG islands near PEG3 and USP29 both showed hypermethylation in DNA derived from breast and ovarian tumors. Second, we identified an antisense transcript to ZIM2 (zinc finger imprinted gene 2) called ZIM2as in the human, chimpanzee, and orangutan. In non-primate mammals, the 5’ side of the Peg3 imprinted domain is bounded by a cluster of olfactory receptor (OLFR) genes which may curtail the spread of imprinting. We report the presence of two previously unreported DMRs near the ZIM2as promoter region. The CpG island distal to ZIM2as was methylated allele-specifically in the human testis, while the CpG island proximal to the ZIM2as promoter showed a mosaic methylation pattern in the chimpanzee. Two CpG islands near the promoter region of ZIM2as showed different methylation patterns in these three species.

Overall, this work provides a firm foundation for future studies of the Peg3 imprinted domain. It represents the first systematic study of DNA methylation in the Peg3 imprinted region. It also describes an antisense transcript that has formed in the great ape PEG3 imprinted domain which may control the extension of this imprinted domain.
CHAPTER ONE: BACKGROUND
Genomic Imprinting

Most autosomal genes are expressed equally from two parental alleles, but up to 200 mammalian genes are only transcribed from one allele based on the parent of origin due to the genomic imprinting process (Ideraabdullah et al., 2008). Genomic imprinting is found only in marsupials and placental mammals, organisms that utilize a unique reproductive strategy in which young offspring develop inside females’ wombs. The phenomenon of genomic imprinting was discovered in 1984 after the failure of nuclear transfer experiments to produce either androgenetic or parthenogenetically derived mice (Surani et al., 1984). The most prevalent explanation for the evolution of genomic imprinting is the conflict theory (Moore and Reik, 1996). To maximize fitness, it is in the male’s best interest to ensure that his offspring receive the most nutrients, but the female’s interest lies in maximizing the number of offspring (spreading the resources among her children). Consistent with this, many imprinted genes are involved in controlling fetal growth rates and nurturing behaviors (Tilghman, 1999). Maternally expressed genes are often growth suppressors, and paternally expressed genes are often growth inducers. The proper dosage of imprinted genes is critical for the survival of mammals, and abnormalities in the dosages quite often manifest as genetic diseases in humans. Imprinting-related diseases include Beckwith-Wiedemann, Prader/Willi, Angelman, and Silver-Russell syndromes as well as autistic spectrum disorders (Ferguson-Smith et al., 2004; Ideraabdullah et al., 2008).

Imprinted genes often exist in clusters that are regulated by one or a few imprinting control regions (ICRs) (Edwards and Ferguson-Smith, 2007), which control the monoallelic expression of multiple genes within the cluster. One common feature of imprinted domains is the expression of antisense transcripts that regulate imprinting (Pauler et al., 2007). Antisense transcripts are a frequent part of the regulatory machinery of imprinted domains. For example, Air, an antisense transcript to Igf2r, is required to maintain monoallelic expression of Igf2r. Disruption of Air causes Igf2r to lose imprinting and become expressed from both alleles (Sleutels, et al., 2002). The mechanisms by which antisense
transcripts silence genes are mostly unknown, but *Air* has been shown to interact with G9a, a histone 3 lysine 9 methyltransferase (Nagano et al., 2008). G9a has been shown to interact with DNA methyltransferases (Esteve et al., 2006). Thus, antisense transcripts in imprinted regions may target repressive modifiers to silence transcription from one allele.

Another common feature of imprinted domains is the presence of CpG islands with allele-specific DNA methylation patterns (also called differentially methylated regions: DMR) (Ferguson-Smith et al., 2004). Some of these DMRs inherit their methylation as a gametic signal from the previous generations, and these DMRs play critical roles for maintaining the imprinting and transcription of a given domain (Edwards and Ferguson-Smith, 2007). Abnormal methylation levels of these DMRs, either hypermethylation or hypomethylation, are also often associated with many types of human diseases as ‘epimutations’ (Hatchwell and Greally, 2007). There are two periods of DNA demethylation and remethylation that occur during mammalian development. The first occurs around the time of implantation, and the second occurs in the primordial germ cells that will eventually produce gametes (Lees-Murdoch and Walsh, 2008). This second round of DNA methylation setting ensures that the future gametes will have the correct regulation of imprinted genes for their gender.

**DNA Methylation**

During the early studies of DNA, several modified bases were found to be present in low concentrations in addition to the four that make up the majority. In mammals, the most prevalent minor base is 5-methyl cytosine, which only occurs within the context of a CpG (cytosine-phosphodiester bond-guanine) dinucleotide. DNA methylation of a gene’s promoter generally leads to transcriptional inactivation. The formation of 5-methyl cytosine is achieved by the transfer of a methyl group from S-adenosyl methionine to a cytosine by a DNA methyltransferase enzyme (Bird, 2002). (Fig 1.1) There are two related *de novo* methyltransferases: Dnmt3a and Dnmt3b, and a related non-catalytic protein, Dnmt3l that is highly expressed in the male germline. Dnmt3l forms a complex with Dnmt3a and/or Dnmt3b and
appears to act as an adapter to target DNA methylation to specific regions. Disruption of Dnmt3l causes hypomethylation of transposable elements and prevents spermatocytes from completing development (Bourc’his and Bestor, 2004). Knockout of either Dnmt3a or Dnmt3b is lethal, but knockout of Dnmt3b specifically in the germline shows no obvious phenotype; mice with a germline knockout of Dnmt3a are healthy but cannot produce live offspring (Kaneda et al., 2004).

![Fig 1.1. The DNA methylation reaction.](image)

Dnmt1, the DNA maintenance methyltransferase, has a higher affinity for DNA that is methylated on one strand only, as a product of the semiconservative replication of DNA (Bird, 2002). In this case, the parent strand received methylation prior to the initiation of replication (either from a de novo methyltransferase or from a maintenance methyltransferase in a previous round of replication), and the daughter strand is methylation-free. Dnmt1 is also necessary for maintaining the correct methylation pattern in imprinted regions (Hirasawa et al., 2008).

Proper DNA methylation is essential for survival in mammals and the majority of DNA in the mammalian genome is methylated. The major exception to this rule is the CpG islands, which feature a higher than expected prevalence of the CpG dinucleotide (cytosine followed immediately by guanine). Since methylated cytosine has a high rate of mutation to thymine, CpG sites that are constitutively
methylated will decay to TpG or CpA, reducing the prevalence of CpG dinucleotides (Bird, 1996). Loss of DNA methylation leads to global derepression of transposable elements and chromosomal rearrangements. This effect suggests that DNA methylation may have evolved as a defense against transposable elements (Yoder et al., 1997). In addition to the repression of transposable elements, DNA methylation is also important in the monoallelic expression characteristic of the genomic imprinting process. Without expression of DNA methyltransferases, imprinted regions revert to biallelic expression (Kaneda et al., 2004; Hirasawa et al., 2008).

**Techniques for DNA Methylation Analysis**

There are several different methods used to study the location and extent of DNA methylation. One of the most sensitive methods uses a Southern blot in which a DNA sample is split and each fraction is digested with one of two restriction enzymes (Southern, 2006). One of these enzymes is inhibited by the presence of CpG methylation and the other is not inhibited by this modification. The blot is then probed with DNA complementary to the region of interest and the size of the band reveals the methylation status. If the fragment is the same size in both fractions, the DNA is unmethylated, but if the fragment is larger in the methylation-sensitive digest, the DNA is methylated. This method is very sensitive, but has the limitation of detecting methylation only at the specific site recognized by the restriction enzyme.

Several other techniques are based on the bisulfite conversion method. This chemical conversion method changes any unmethylated cytosines to uracil, while leaving methylated cytosines intact (Clark et al., 1994). After bisulfite conversion, the DNA is amplified by PCR and the uracil is replaced with thymine (Fig. 1.2). The resulting bisulfite PCR product can be analyzed by using a restriction enzyme that will differentiate between the methylated and unmethylated DNA based on the sequence differences that exist after bisulfite conversion (Xiong and Laird, 1997). An enzyme that has a CpG within its recognition site is generally used. The recognition site will be preserved only in methylated DNA, so only methylated DNA will be digested. This method, the combined bisulfite restriction analysis
(COBRA), is powerful and quick, but only allows analysis at sites recognized by restriction enzymes and can be confounded by polymorphisms in the DNA sequence. In contrast, bisulfite sequencing allows identification of the methylation status of each CpG site in the bisulfite-modified PCR product and also allows detection of most polymorphisms, although cytosine to thymine transitions in the genomic sequence will be undetected (Grunau et al., 2001).

Figure 1.2. The bisulfite conversion reaction. A) Bisulfite treatment of an unmethylated cytosine converts it to uracil, which is replaced by thymine during PCR. B) The bisulfite reaction proceeds much more slowly on methylated cytosines.

The preceding DNA methylation study methods are relatively labor-intensive and not amenable to high-throughput efforts, but recently several new methods of analyzing DNA methylation have been developed. These include immunoprecipitation of methylated DNA using antibodies against 5-methylcytosine (ME-DIP) and reduced representation bisulfite sequencing (RRBS), both of which are useful for studying DNA methylation on a genomic scale (Suzuki and Bird, 2008; Meissner et al., 2005).
Peg3 Imprinted Domain

The PEG3 protein contains a SCAN box followed by twelve zinc fingers, suggesting that it may function as a transcription factor (Li et al., 1999). Peg3 knockout mice show a phenotype with several effects including alterations in maternal behavior, smaller pup size, and higher body fat percentage. Peg3 has also been linked to the p53 and apoptosis pathways since it promotes Bax translocation to the cytosol (Deng and Wu, 2000). Recent studies have found that human PEG3 expression is often absent in several types of cancers, and DNA hypermethylation on the promoter region appears to be a prime cause for this loss of PEG3 expression (Maegawa et al., 2001).

Peg3 is located within a cluster of imprinted genes that is well conserved evolutionarily—it is found in the mouse (proximal chr 7), cow (chr 18), and human (chr 19), as well as in all other mammals. In the mouse, this 500-kb genomic region contains 6 additional imprinted genes, including Usp29/Mim1 (Mer-repeat containing imprinted transcript 1), Zim1 (Zinc finger gene imprinted 1), Zim2, Zim3, Zfp264 (Zinc finger protein gene 264), and APeg3 (Antisense to Paternally expressed gene 3) (Kim and Stubbs, 2005). Several members of this domain are also imprinted in human and cow (Kim et al., 2007). Other studies suggest that the epigenetic abnormalities of this domain may be associated with other human diseases (Van den Veyver et al., 2001). Despite this close linkage to human diseases, this domain has not been systematically studied so far in humans and other mammals.

In chapter two, we describe a survey of the DNA methylation status of the Peg3 imprinted domain. This region has been studied in the adult mouse, and selected genes have been studied in the human, but we have performed the most comprehensive analysis of DNA methylation in this domain. We analyzed 750 kb of sequence surrounding Peg3 in the mouse, cow, and human genomes for the possible presence of CpG islands as well as other characteristics. To provide a reference for this analysis, we also analyzed a 2 Mb region of sequence containing a cluster of non-imprinted genes for the same characteristics as above. Imprinted domains commonly show different patterns of sequence structure than
nonimprinted regions of the genome, e.g. SINEs (short interspersed nuclear elements) are depleted and LINEs (long interspersed nuclear elements) are enriched in these regions. We hypothesize that the Peg3 domain will show some unique sequence characteristics, especially in the distribution of repetitive sequence and CpG islands. Finally, we analyzed the methylation status of DNA obtained from the mouse (germline, early development, placenta, and adult tissues), adult cow, and human (both normal and tumor tissues). We expect that the methylation status of the Peg3-CpG island will be well conserved among all of the species studied, but that the surrounding CpG islands may show some differences.

In chapter three, we describe a transcript antisense to ZIM2 (called ZIM2as) in the great apes. This transcript spans the region from the first intron of ZNF835 to the fourth from last intron of ZIM2 and is expressed in human brain and testis. This region is the site of a cluster of olfactory receptor genes in all the non-primate mammals for which we could obtain genome sequences, including the mouse, cow, and dog. However, the primates have lost this cluster of genes and ZIM2as is present in three great ape genomes. We analyzed expression patterns of ZIM2as and surrounding genes in the human. We also tested the DNA methylation status of several CpG islands surrounding ZIM2as in three different primates: rhesus macaque, chimpanzee, and human. Since olfactory gene clusters elsewhere in the genome act as boundaries between differently regulated domains, the cluster present at the 5’ side of the Peg3 imprinted domain may function to limit the spread of imprinting. Long noncoding RNAs in imprinted domains often function to maintain imprinting of the surrounding genes and disruption of these transcripts causes biallelic expression of the surrounding genes. Thus, the genomes with (human and chimpanzee) and without (rhesus) ZIM2as may differ in terms of the DNA methylation status of the CpG islands near this transcript, which could affect expression of the associated genes.

References

Bourc'his D., Bestor T.H., 2004. Meiotic catastrophe and retrotransposon reactivation in male


9
Oshimura M. Epigenetic silencing of PEG3 gene expression in human glioma cell lines. Mol Carcinog. 31, 1-9.


CHAPTER TWO: DNA METHYLATION ANALYSIS OF THE MAMMALIAN
PEG3 IMPRINTED DOMAIN*

*Reprinted with permission of *Gene
Introduction

The majority of autosomal genes are expressed equally from two parental alleles. However, in fewer than 200 mammalian genes, transcription from one allele is epigenetically repressed based on its parent of origin due to a process called genomic imprinting (Wagschal and Feil, 2006). Genomic imprinting is found only in marsupials and placental mammals that utilize a unique reproductive strategy in which young offspring develop inside females’ wombs. Most imprinted genes are also involved in controlling fetal growth rates and nurturing behaviors (Tilghman, 1999). The proper dosage of imprinted genes is very critical for the survival of these mammals, and abnormal dosages quite often manifest as genetic diseases in humans. The imprinting-related diseases include Beckwith-Wiedemann, Prader/Willi, Angelman, and Silver-Russell syndromes as well as autistic spectrum disorders (Ferguson-Smith et al., 2004; Ideraabdullah et al., 2008). In addition, several imprinted genes are involved in many types of cancers (Jirtle and Skinner, 2007). Thus, genomic imprinting is regarded as a gene dosage control mechanism for this subset of genes critical for mammalian-specific reproductive scheme (John and Surani, 2000).

The promoter regions of imprinted genes are usually associated with CpG-rich regions, termed CpG islands. Some of these CpG islands are differentially methylated in an allele-specific manner, and thus called Differentially Methylated Regions (DMRs). Besides monoallelic expression, the allele-specific methylation pattern of these CpG islands is another molecular signature that is used to define imprinted genes (Ferguson-Smith et al., 2004). Some of these DMRs also inherit their methylation as a gametic signal from the previous generations, and these DMRs play critical roles for maintaining the imprinting and transcription of a given domain (Edwards and Ferguson-Smith, 2007). Abnormal methylation levels of these DMRs, either hyper or hypomethylation, are also often associated with many types of human diseases as ‘epimutations’ (Hatchwell and Greally, 2007).
An evolutionarily conserved imprinted region, the PEG3 domain, is found on human chromosome 19q13.4/proximal mouse chromosome 7/cow chromosome 18. In the mouse, this 500-kb genomic region contains 6 additional imprinted genes, including Usp29/Mim1 (Mer-repeat containing imprinted transcript 1), Zim1 (Zinc finger gene imprinted 1), Zim2, Zim3, Zfp264 (Zinc finger protein gene 264), and APEG3 (Antisense to Paternally expressed gene 3) (Kim and Stubbs, 2005). Several members of this domain are also imprinted in human and cow (Kim et al., 2007). According to recent studies, human PEG3 expression is often missing in several types of cancers and DNA hypermethylation on the promoter region appears to be a prime cause for this loss of PEG3 expression (Maegawa et al., 2001). Other studies suggest that the epigenetic abnormalities of this domain may be associated with other human diseases (Van den Veyver et al., 2001). Despite this close linkage to human diseases, this domain has not been systematically studied so far in humans and other mammals. Thus, the current study sought to analyze the genome sequence and DNA methylation status of this evolutionarily conserved imprinted domain. This study revealed that the PEG3 domain of humans and mice contains at least two DMRs, but contains only one DMR in cows. In addition, the methylation status of the two human DMRs is often affected in the ovary and breast tumor DNA.

Results

Characterization of the CpG Islands Located in the PEG3 Domain

The overall organization of the 500-kb genomic region of the PEG3 domain is well conserved among mammals, but the neighboring regions show large-scale genomic changes, such as the loss of an Olfactory Receptor (OLFR) cluster as well as a Vomeronasal Organ Receptor (VNO) cluster in the human genome (Fig. 2.1). Since the exact boundaries of this imprinted domain are unknown, we analyzed the 750 kb region surrounding PEG3 from mouse chromosome 7, cow chromosome 18, and human chromosome 19. We identified 50, 34, and 19 CpG islands in the mouse, cow, and human PEG3 regions, respectively (Table 2.1). A 2 Mb nonimprinted region that is syntenic between human, mouse, and cow
Figure 2.1. Organization of the PEG3 genomic region.

Outline of the 750 kb genomic region surrounding PEG3 in the mouse, cow, and human. Directions of arrows indicate the direction of transcription. Maternally expressed genes are indicated by bold red text; paternally expressed genes are indicated by underlined blue text. The positions of the olfactory receptor gene cluster (OLFR) and the vomeronasal gene cluster (VNO) are indicated by boxes. Dotted yellow lines indicate the approximate regions in the mouse and cow genomes that contain the OLFR and VNO clusters as well as the approximate region from which these clusters were lost from the human genome.
Table 2.1. Analysis of CpG islands and repeat content in the PEG3 imprinted region.

<table>
<thead>
<tr>
<th></th>
<th>Mouse</th>
<th></th>
<th>Cow</th>
<th></th>
<th>Human</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Peg3 Region</td>
<td>Peg3</td>
<td>Control</td>
<td>Peg3</td>
<td>Control</td>
<td>Peg3</td>
<td>Control</td>
</tr>
<tr>
<td>Sequence size</td>
<td>750 kb</td>
<td>2 Mb</td>
<td>750 kb</td>
<td>2 Mb</td>
<td>750 kb</td>
<td>2 Mb</td>
</tr>
<tr>
<td>GC content</td>
<td>43.66%</td>
<td>47.31%</td>
<td>47.85%</td>
<td>39.86%</td>
<td>44.36%</td>
<td>40.86%</td>
</tr>
<tr>
<td>CpG islands</td>
<td>36 (15)*</td>
<td>41</td>
<td>34 (13)</td>
<td>16</td>
<td>19 (9)</td>
<td>26</td>
</tr>
<tr>
<td>Tandem Repeats in</td>
<td>9 (3.75)</td>
<td>10</td>
<td>1 (0.75)</td>
<td>2</td>
<td>10 (0)</td>
<td>0</td>
</tr>
<tr>
<td>entire region</td>
<td>173 (104)</td>
<td>279</td>
<td>223 (129)</td>
<td>343</td>
<td>49 (31.75)</td>
<td>85</td>
</tr>
<tr>
<td>Total masked bases</td>
<td>19.79%</td>
<td>31.46%</td>
<td>35.49%</td>
<td>49.42%</td>
<td>58.05%</td>
<td>55.78%</td>
</tr>
<tr>
<td>Total interspersed repeats</td>
<td>14.13%</td>
<td>28.89%</td>
<td>34.36%</td>
<td>48.39%</td>
<td>55.68%</td>
<td>54.53%</td>
</tr>
<tr>
<td>SINEs</td>
<td>3.04%</td>
<td>10.93%</td>
<td>15.34%</td>
<td>15.17%</td>
<td>21.57%</td>
<td>14.45</td>
</tr>
<tr>
<td>LINEs</td>
<td>6.75%</td>
<td>7.89%</td>
<td>14.60%</td>
<td>28.32%</td>
<td>18.87%</td>
<td>25.50%</td>
</tr>
<tr>
<td>LTR</td>
<td>3.70%</td>
<td>8.99%</td>
<td>2.65%</td>
<td>3.78%</td>
<td>10.67%</td>
<td>11.71%</td>
</tr>
</tbody>
</table>

*number in parentheses indicates expected number, based on number in control region

and contains a cluster of zinc finger genes was used as a basis for analysis of the sequence characteristics of this region. The PEG3 region and its associated CpG islands show several interesting differences when compared to the reference region. First, although the overall GC content of the PEG3 region and the reference region are similar, the PEG3 region contains more CpG islands than expected. Second, both the PEG3 region and the CpG islands contain more tandem repeats than expected. These tandem repeats
were detected using the Tandem Repeat Finder program, and consist of simple sequence repeats ranging in size from approximately 15 bp to several hundred bp. Third, the repeat content of the PEG3 domain deviates greatly from the reference region. Lower levels of SINEs, LINEs, and LTRs were found in the mouse, and reduced levels of LINEs and LTRs were found in both the human and cow sequences. Fourth, the CpG islands in this region are mainly found in regions near genes: 50% (mouse), 58% (cow), and 84% (human) of the CpG islands are within 10 kb of an annotated gene. In addition, 60% (mouse), 45% (cow), and 63% (human) of the above CpG islands overlap at least one exon. Although most of the CpG islands are not conserved among the three species in terms of sequence and position, one CpG island surrounding the first exon of PEG3 appears to be well conserved.

**DNA Methylation Analysis of the CpG Islands from the Mouse Peg3 Domain**

We analyzed the methylation status of the computationally predicted CpG islands along with several other regions associated with genes in the Peg3 imprinted domain using genomic DNA from the mouse, cow, and human. Briefly, this genomic DNA was treated via bisulfite conversion, and then the bisulfite-converted DNA was amplified using PCR. The resulting bisulfite PCR products were analyzed by COBRA (COmbined Bisulfite Restriction Analysis) (Xiong and Laird, 1997) and/or subcloning and sequencing (Fig. 2.2A).

Within the mouse Peg3 domain, we targeted four CpG islands (Zim1, Zim2, Zim3, and Peg3) using genomic DNA isolated from various Mus musculus samples as well as from the livers of F1 and F2 offspring of interspecific crosses between M. musculus and M. spretus. The use of DNA from both F1 and F2 animals allowed us to exclude the possibility that the genetic background of the mice affected our methylation analysis. In addition, since the F1 and F2 samples are from mice of different ages (three months and two weeks old, respectively), we are able to determine if age has an effect on the methylation status of the analyzed regions. Analysis of DNA from placenta, germ cells, blastocyst stage embryos, ES cells, and adult mice gave us a developmental stage-specific view of the methylation status of this domain. Sequence polymorphisms between the two parental species were first identified within the
Mouse genomic DNA was obtained from Mus musculus sperm, oocytes, blastocysts, embryonic stem cells, placenta, and from the liver tissues of the F1 (3 months old) and F2 (2 weeks old) offspring of interspecific crossing of M. musculus and M. spretus.

A) COBRA and bisulfite sequencing analysis. The gene associated with each PCR product is shown to the left of the figure. Each DNA sample was treated with sodium bisulfite and used in the PCR reaction. Next, each PCR product was cloned and sequenced and/or analyzed by the COBRA analysis in which each PCR product was digested with the enzyme shown to the right of each picture to assess the methylation level of the region (*: the Peg3-CpG island was digested with BstUI in the sperm DNA sample). Total digestion by this enzyme indicates methylation of the region and a lack of digestion indicates the absence of methylation in the region. For each COBRA panel, column U contains DNA that was not exposed to the selected restriction enzyme, and column C contains DNA that was exposed to the restriction enzyme. The arrow labelled U indicates the position of undigested, unmethylated DNA, and the arrow labelled M indicates the position of digested, methylated DNA. Repeated analyses (at least three times) yielded similar results, so a representative picture is shown for each COBRA. For the bisulfite sequencing results, each row represents a different clone, and each column represents a different CpG site. Filled circles indicate methylated cytosines and open circles indicate unmethylated cytosines.

studied CpG regions, and later used to differentiate the parental alleles in the F1 and F2 samples. The Zim1-CpG island was not digested by the enzyme HhaI in any of the tested samples, showing a lack of methylation in the region at all of these stages, although we were unsuccessful in obtaining data on this
region in the blastocyst stage embryo. Bisulfite sequence analysis of this region confirmed this result. As demonstrated in earlier studies (Li et al., 2000), the CpG island near the promoter of Peg3 displayed a maternal allele-specific methylation pattern based on results from both COBRA and bisulfite sequencing. The Peg3-CpG island was almost completely methylated in oocyte DNA and lacked methylation in the sperm DNA, as expected for a maternally methylated DMR, while ES cell, placenta, and both F1 and F2 adult tissues derived DNA showed the expected DMR pattern. At the Zim2-CpG island, a small amount of PCR product was digested after incubation with the enzyme TaqI in the sperm-derived DNA. This pattern could either be a result of somatic contamination of the sperm-derived DNA or could indicate that there are low levels of methylation in sperm DNA. In contrast, the placenta, ES cell, F1, and F2 samples all show digestion of approximately 50% of the PCR products. This pattern indicates about half of these PCR products are methylated. Bisulfite sequencing of the PCR products confirmed this result, but the methylation was not allele-specific. The first four CpG sites were all methylated in most clones, while the next four CpG sites were unmethylated in most clones, and the ninth CpG site (which is part of the TaqI enzyme site used for COBRA) was methylated in approximately half of the clones. The PCR product from the Zim3-CpG island showed approximately 50% digestion with the enzyme BstUI in DNA derived from blastocyst, placenta, F1, and F2 mice, indicating that half of the PCR products were methylated (Fig. 2.2A). Sequencing revealed that the unmethylated DNA was mostly derived from the maternal allele (Fig. 2.2A). This is consistent with the maternal allele-specific expression of Zim3 (Kim et al., 2001). However, the Zim3-CpG island is unmethylated in sperm DNA and ES cell DNA, suggesting that the methylation on this CpG island is not inherited as a gametic signal, but that allele-specific methylation at this CpG island may be a somatic signal established during early development of the mouse. Overall, the Zim3-CpG island appears to be the second DMR (Differentially Methylated Region) discovered in the mouse Peg3 domain. In addition, we show that the CpG island near Zim1 is unmethylated, based on the lack of digestion of bisulfite PCR products by HhaI and analysis of the
sequence of individual \textit{Zim1} PCR products. The Zim2-CpG island showed a pattern of DNA methylation that is not allele-specific, but rather stage- or tissue-specific (Fig. 2.2B).

**DNA Methylation Analysis of the CpG Islands from the Cow Peg3 Domain**

We performed a similar series of analyses using DNA prepared from the liver of F1 hybrid offspring of \textit{Bos taurus} and \textit{B. indicus}. We analyzed CpG islands associated with \textit{Peg3}, \textit{Usp29}, \textit{Ast1}, \textit{Zim2}, and \textit{Zim3}. The COBRA results from the CpG islands near \textit{Zim2} and \textit{Zim3} showed digestion of nearly all of the PCR products, and bisulfite sequencing confirmed that most of the CpG sites are methylated. The CpG island associated with \textit{Zim2} showed a similar pattern of total digestion indicating complete methylation. We also analyzed the CpG island located close to the promoter of cow \textit{Usp29} since this has been shown to be an imprinted gene with paternal allele specific expression (Kim et al., 2007).

**Figure 2.2B.** Summary of methylation at different developmental time points in each of the four loci tested. Open circles indicate unmethylated regions, completely filled black circles indicate methylated regions, half-filled black regions indicate probable DMRs, and gray circles indicate time points for which there is no data available.
However, the results from the DNA methylation analyses revealed that the CpG island near this gene is almost completely methylated in the cow (Fig. 2.3). It is important to note that although cow Usp29 is imprinted, this gene has lost its ORF (Open Reading Frame) capability in recent evolutionary time, suggesting some changes in the bovine lineage (Kim et al., 2007). Thus, the hypermethylation of the Usp29-CpG island of cow might be related to the loss of its ORF during evolution. In contrast, lack of digestion by HhaI shows that the CpG island near Ast1 is not methylated at all. As predicted, the CpG island close to the Peg3 promoter is methylated in an allele-specific manner (Fig. 2.3).

Figure 2.3. Bisulfite sequencing and COBRA analyses of the cow Peg3 domain
Cow DNA was obtained from the liver of offspring of interspecific crossing of Bos taurus and Bos indicus. The gene associated with each PCR product is shown to the left of the figure. Each DNA sample was treated with sodium bisulfite, used in the PCR reaction, and then cloned and sequenced and/or analyzed by the COBRA analysis. For each COBRA panel, column U contains DNA that was not exposed to the selected restriction enzyme, and column C contains DNA that was exposed to the restriction enzyme listed on the right side. On the right side of the COBRA panel, the arrow labelled U indicates the position of undigested, unmethylated DNA, and the arrow labelled M indicates the position of digested, methylated DNA. Repeated analyses (at least three times) yielded similar results, and a representative picture is shown for each COBRA. For the bisulfite sequencing results, each row represents a different clone, and each column shows a different CpG site. Filled circles indicate methylated cytosines and open circles indicate unmethylated cytosines.

Bisulfite sequencing results further confirmed maternal allele-specific methylation at this CpG island (Fig. 2.3), consistent with the paternal allele-specific expression of cow Peg3. In summary, this survey indicated that the Peg3-CpG island is the only DMR known at this time to exist in the cow Peg3 region.

DNA Methylation Analysis of the CpG Islands from the Human PEG3 Domain
We also performed a series of DNA methylation analyses of the CpG islands identified from the
human PEG3 domain (Fig. 2.4). These include the CpG islands located close to ZNF71 (Zinc finger protein 71), ZNF835 (Zinc finger protein 835), PEG3, USP29, DUXA (double homeobox A), and ZNF264 (Zinc finger protein 264). In addition, we analyzed the promoter region of ZIM3 to allow for comparison of the methylation status of this region between the cow, mouse and human, and also a CpG island that was predicted by MethPrimer to exist in the first intron of DUXA (DUXA-5/6). We used human genomic DNA isolated from the brain, testis, lung and liver of 4 normal individuals. Representative results from the adult brain and testis of different individuals are shown in Fig. 2.4. The methylation status of the CpG islands in the human PEG3 domain can be summarized as follows. The PCR products derived from the CpG islands of ZNF71, ZIM3, ZNF264, and DUXA (DUXA-3/4) were not digested by their respective restriction enzymes, indicating that these CpG islands were unmethylated in the tested DNA from adult brain, testis, liver and lung.

In contrast, the PCR products from the CpG island in the first intron of DUXA (DUXA-5/6) showed the opposite pattern in the COBRA analysis: the majority of this region was digested, indicating methylation. Digestion of the ZNF835 PCR product by HpyCH4IV revealed a tissue-specific methylation pattern: less DNA was digested in the sample derived from testis, indicating hypomethylation compared to the brain sample.

Finally, the CpG islands of PEG3 and USP29 showed a digestion pattern characteristic of a differentially methylated region: only half of the PCR products were digested while the other half remained undigested (Fig. 2.4). We could not determine allele-specific methylation in either locus since the tested regions of both PEG3 and USP29-CpG islands lack sequence polymorphisms. However, the results from bisulfite sequencing of the PEG3-CpG island clearly indicated that half of the clones were methylated while the other half were unmethylated, the typical pattern of a DMR (Fig. 2.5). In contrast, the bisulfite sequencing results of the USP29-CpG island showed a less clear pattern than the PEG3-CpG island. About half of the clones from testis DNA (4 out of 10 clones, 26.3% of the CpGs) showed methylation whereas a much greater number of the clones from brain DNA showed methylation but with high levels
of mosaicism (54.5% of the CpGs were methylated). It is important to note that human USP29 is expressed only in testis (Kim et al., 2000). Thus, it is possible that the observed higher levels of methylation at brain might reflect the transcriptional activity of human USP29, but this remains to be tested in the near future.

Figure 2.4. COBRA analysis of the CpG islands located in the human PEG3 imprinted region. Human DNA was derived from normal brain, normal testis, breast tumor, and ovary tumor. Each DNA was converted using sodium bisulfite and used in the PCR product. Each PCR product was incubated with the enzyme indicated to the right of the figure. Repeated analyses (at least three times) yielded similar results, so a representative picture is shown for each COBRA. The gene associated with each PCR product is shown to the left of the figure. Each lane contains DNA that was incubated with the appropriate restriction enzyme. The arrow labelled U indicates the position of undigested, unmethylated DNA, and the arrow labelled M indicates the position of digested, methylated DNA.

Since several independent reports previously indicated the hypermethylation of human PEG3-DMR in cancers (Maegawa et al., 2001), we also performed a similar series of DNA methylation analyses using 4 representative cancer DNAs of different tissue origin, including ovary, breast, lung and liver. According
to the results from this survey (Fig. 2.4), the two DMRs of the human PEG3 domain were affected in two cancer types, ovary and breast, but not in lung and liver (Appendix A).

In the case of the PEG3-DMR, half of the PCR products from breast cancer DNA were digested by \textit{HpyCH4IV}, which is similar to the pattern observed from normal brain and testis DNA. However, the same analysis revealed that more than half of the PCR products from ovarian cancer DNA were digested by \textit{HpyCH4IV}, suggesting hypermethylation in this cancer DNA (Fig. 2.4, lanes marked by $\blacklozenge$). About 80% of the ovarian cancer DNA showed methylation at the PEG3-DMR although three other samples show much lower levels of methylation. This result was further confirmed through sequencing the PCR products as shown in Fig. 2.5. In addition, the previously undiscovered USP29-DMR showed a similar pattern-the majority of the PCR products from both breast and ovarian cancer DNA were digested by \textit{TaqI}, suggesting that this CpG island is also hypermethylated in these two types of cancer DNA. Our sequencing analyses indeed confirmed the hypermethylation of USP29-DMR in breast and ovarian cancer DNAs (Fig. 2.5). The hypermethylation levels in the ovarian cancer DNA appears to be much greater (81.8\% of CpGs are methylated) than those in the breast cancer DNA (65.5\% of CpGs are methylated). Also, it is interesting to note that the methylation pattern at the breast cancer DNA is somewhat mosaic, which is similar to the pattern seen in normal brain DNA but with much more methylation. For this series of analyses, we have also included the human H19-DMR as a control. As shown in Fig. 2.4, this region showed a DMR pattern in the normal DNA as well as two types of cancer DNA. This also suggests that the observed DNA hypermethylation may be specific to the two DMRs of the human PEG3 domain. In sum, the data presented above is consistent with the previous observations revealing the hypermethylation of human PEG3 in cancer (Maegawa et al., 2001), and further indicates that this hypermethylation probably occurs on the DMRs of both PEG3 and USP29 in cancer DNA.
Figure 2.5. Bisulfite sequencing of the CpG islands located in the human *USP29* and *PEG3* promoter regions.

Results obtained from analysis of normal brain and testis derived DNA are shown along with breast and ovary tumor-derived DNA. Each row indicates a different clone, and each column indicates a different CpG site. Filled circles indicate methylated cytosine and open circles indicate unmethylated cytosine. The analyzed CpG islands are indicated with the name of the gene with which they are associated.

**Discussion**

In this study, we surveyed methylation status of the CpG islands of the *PEG3* imprinted domain in the mouse, cow, and human genomes. This survey led to the discovery of two previously unreported differentially methylated regions: mouse *Zim3* and human *USP29*. In addition, we examined the methylation status of the CpG islands in this region using human tumor derived DNA. The CpG islands near *PEG3* and *USP29* both showed hypermethylation in DNA derived from breast and ovarian tumors.

We performed a comprehensive analysis of the sequence in the 750 kb region containing the *PEG3* imprinted domain. This analysis showed that the sequence structure in this region was different
from a non-imprinted region containing similar types of genes (Table 2.1). Both tandem repeats and CpG islands were over-represented in both the entire sequence and the CpG island sequence of the PEG3 region. Tandem repeats have been reported to be associated with imprinted genes, and may play a role in setting up DNA methylation for the CpG islands of imprinted genes during gametogenesis (Hutter et al., 2006). It is well known that tandem repeat sequences can attract DNA methylation although it is still unclear how only one allele of tandem repeats become methylated in the case of imprinted genes, such as the DMRs of imprinted genes. At the same time, the tandem repeats may also play a role in the genesis and maintenance of CpG islands during evolution. Since methylated cytosines are prone to mutation, CpG islands (which contain high numbers of cytosines) must be protected from this mutation by some mechanism. Tandem duplication of CpG islands, which would increase the overall size of the island, could be one way to prevent attrition of cytosines. These two conflicting needs might have contributed to increasing the number of tandem repeats in the CpG islands of mammalian imprinted genes.

In the mouse, we analyzed four CpG islands associated with Zim1, Peg3, Zim3, and Zim2, respectively. We assessed methylation status of each CpG island in DNA obtained from sperm, blastocyst, embryonic stem cell (ES cell), placenta, and in somatic tissue from mice of two different ages (two weeks and three months) (Fig. 2.2). Methylation status of other loci is known to show differences at different stages of development, and also with age, but a broad survey such as this has not been previously performed in the PEG3 imprinted region. The Zim2- and Zim3-CpG islands showed different patterns between blastocysts and ES cells, which are at a similar stage of development: Zim2 was unmethylated in the blastocyst and showed a DMR-type digestion pattern in the COBRA results from the ES cell, and Zim3 had a DMR pattern in the blastocyst and was unmethylated in the ES cell. While the Peg3-DMR has been shown to be stable in ES cells, other regions have not, so this change in methylation at these regions can probably be attributed to the effects of ES cell culture (Chang et al., 2009). Peg3, Zim2, and Zim3 all show approximately 50% digestion of DNA derived from the placenta. This may indicate that the methylation pattern of these regions in the placenta is similar to that found in the adult
somatic tissue, which could indicate that these genes are regulated similarly in the placenta and somatic tissues, although it is also possible that this pattern is due to the fact that the placenta is derived from both maternal and fetal cells. The current study identified two additional DMRs, Zim3-DMR in the mouse and USP29-DMR in the human (Fig. 2.6).

**Figure 2.6. Summary of the methylation status of CpG islands in the PEG3 region**
Outline of the 750 kb genomic region surrounding PEG3 in the mouse, cow, and human. Directions of arrows indicate the direction of transcription. Maternally expressed genes are indicated by bold red text; paternally expressed genes are indicated by underlined blue text. The boxes show the approximate position of the olfactory (OLFR) and vomeronasal (VNO) gene clusters. Dotted yellow lines indicate the approximate regions in the mouse and cow genomes that contain the OLFR and VNO clusters as well as the approximate region from which these clusters were lost from the human genome. Below each chromosome, black boxes indicate methylated regions, empty boxes indicate unmethylated regions, half-filled boxes indicate differentially methylated regions, and grey boxes containing the letter T indicate potential tissue-specific methylation patterns.

However, the DMR status of these CpG islands appears to be lineage-specific. First, although the Zim3-CpG island is methylated in an allele-specific manner in the mouse, both alleles of the human ZIM3-CpG island are unmethylated, and both alleles of the cow Zim3-CpG island are methylated (Fig. 2.6). Second, the human USP29-CpG island appears to be a DMR in a tissue-specific manner in the
testis, but the homologous region in cow showed hypermethylation, indicating that both alleles may be methylated. These differences might be an indication of the presence of some species-specific changes in the imprinting status of the surrounding genes. In this regard, it is interesting to point out the presence of a genomic rearrangement in the rodent lineage: the mouse genome does not contain the *Duxa* gene between *Zim3* and *Zfp264* although two other lineages, human and cow, have this gene, suggesting that a lineage-specific deletion event occurred during rodent evolution. Also, the mouse genome contains a very long (over 300 kb) *Usp29* transcript while this long transcript has been truncated in the human and cow lineages (Fig. 2.6). This transcript could be one of the mechanisms that maintains the imprinted status of the surrounding genes, similar to the functions of *Air* and *Kcnq1ot1* in their respective imprinted domains (Ideraabdullah et al., 2008). Once we obtain the imprinting status of human and cow *ZIM3*, it is possible that the relationship between this genomic deletion and the imprinting status of the surrounding genes will be clarified.

*PEG3* expression is silenced in various tumor types, including gliomas, choriocarcinomas, and ovarian tumors (Maegawa et al., 2001; Van den Veyver et al., 2001). This silencing of human *PEG3* was found to be a result of DNA methylation (Murphy et al., 2001). Our results also confirm the hypermethylation of cytosines at the PEG3-DMR in ovarian tumor-derived DNA (Fig. 2.4 & 2.5). Although some regions show tissue-specific methylation patterns, the PEG3-DMR is methylated on the maternal allele only over a range of normal tissues. In each normal tissue, we expect to see a pattern in which approximately half of the bisulfite clones are unmethylated and half are methylated. Since this CpG pattern has such a consistent methylation pattern, any deviation from this in abnormal tissues is probably meaningful. This study only analyzed a limited number of samples, but other studies of human cancers have shown hypermethylation of this region and reduced *PEG3* expression (Maegawa et al., 2001; Feng et al., 2008). Our study showing hypermethylation in the PEG3-CpG island in ovarian cancer DNA adds to the evidence supporting the hypothesis that *PEG3* functions as a tumor suppressor. We also found an increase in the DNA methylation level at the USP29-DMR in breast and ovarian tumor DNA.
Since this CpG island tends to show mosaicism in its DNA methylation pattern, it is unclear whether the observed increase in the DNA methylation levels in the two types of cancer DNA truly represents abnormal DNA methylation or is simply a tissue-specific methylation pattern. However, it is relevant to note that human USP29 was previously discovered through a screen of a tumor-derived expression library (Tureci et al., 2002). Although human USP29 is known to be expressed only in the testis, the expression of this gene has been seen in several types of cancers. This further suggests a potential role of human USP29 in human cancer. Also, it is well known that several genes in a given imprinted domain tend to be co-regulated in terms of its expression and epigenetic modifications. Thus, it is likely that the observed DNA methylation changes in the USP29-DMR along with PEG3-DMR might reflect together the abnormal status of DNA methylation in the two tissues tested in this study. This could suggest a link between abnormal expression of USP29 and PEG3 in human cancers.

**Methods**

**CpG Island Prediction and Sequence Analysis**

A Perl script was used to analyze the genomic sequences surrounding the PEG3 imprinted domain (Chr 19: 61750000-62500000, 750 kb for human; Chr. 7: 6293901-7043900, 750 kb for mouse; Chr. 18: 6398699-6473700, 750 kb for cow) and a nonimprinted region containing similar types of genes that was used to provide a basis for comparison for sequence analysis (Chr 1: 244543476-246543476 for human; Chr. 11: 58303940-60303939 for mouse; Chr. 7: 38923539-40923539 for cow). For the cow and human sequences, this Perl script was set to recognize a sequence as a CpG island only if three conditions were met: length greater than 500 bp, C+G content greater than 55%, and observed/expected CpG ratio at least 0.65 (Takai and Jones, 2002). An initial CpG island prediction using these criteria resulted in very few predicted islands in mice, so the minimum length parameter was reduced to 200 bp for this species only. To test evolutionary conservation, the sequence of each CpG island was analyzed using BLAST (Altschul et al., 1990) and the ECR browser of the dcode website (http://www.dcode.org/) (Ovcharenko et al.,
The CpG islands predicted by this program were also analyzed for the presence of repetitive elements using RepeatMasker and Tandem Repeat Finder (Smit 1996-2004, Benson 1999). The default parameters and appropriate species were used for RepeatMasker, and the parameters for Tandem Repeat Finder were adjusted as follows (Hutter et al., 2006): match score 2, mismatch score 5, indel score 7, match probability 80, indel probability 10, minscore to report 100, maxperiod 2000. The sequence of each CpG island and related information regarding repeat contents and evolutionary conservation are available upon request.

**COBRA (COmbined Bisulfite Restriction Analysis) and Bisulfite Sequencing**

Mouse genomic DNA was isolated from the liver tissues of the F1 (3 months old) and F2 (2 weeks old) offspring of interspecific crossing of *Mus musculus* and *M. spretus* (Kim et al., 2001). Mouse placentas were isolated from 17 day embryos. Mouse sperm DNA was isolated from the epididymis of 3 month old male mice according to a previously established protocol (Bunch and Saling, 1991). Briefly, the epididymides were incubated in sperm elution buffer (130 mM NaCl, 20 mM Tris, 2 mM EDTA pH 7.4) for ten minutes at 37°C. The epididymides were then removed and the solution was centrifuged for 30 seconds at 800 rpm. Then, the sperm were washed twice more with the sperm elution buffer. The isolated sperm were examined under a microscope, and only samples that did not display somatic cell contamination were used for the methylation analyses. The sperm from a single mouse (~10^5-10^6 sperm) was pooled and subjected to bisulfite conversion. In preparation for isolation of blastocyst-stage embryos, female mice were superovulated (Eppig and Telfer, 1993; Horgan et al., 1994). First, 5 IU of Pregnant Mare Serum (PMS) (Cat. G4877, Sigma) was injected subcutaneously. Then, the same mice were injected with 5 IU of human Chorionic Gonadotropin (hCG) hormone (Cat. C1063, Sigma) 48 hours after the PMS injection. The treated mice were mated with male littermates, and sacrificed 3.5 days later. The embryos were flushed from each uterus, and the isolated blastocysts were examined under the microscope to assess their developmental stages and purity. Only embryos at the blastocyst stage were used for methylation analysis. Seven blastocysts were pooled for bisulfite treatment of the DNA. To
isolate oocytes, female mice were superovulated as described above, but the mice were sacrificed 12
hours after the second injection. Mature eggs were isolated from the swollen ampulla of the oviducts,
incubated in hyaluronidase solution (Cat. H3506, Sigma) for several minutes to separate the cumulus
cells, and subsequently washed three additional times to remove potential somatic tissue contamination.
DNA isolated from approximately 400 oocytes obtained from ten females (8 weeks old) was pooled and
used for bisulfite conversion.

Cow genomic DNA was also isolated from the liver of the hybrid offspring of interspecific
crossing of *Bos taurus* and *B. indicus*. These hybrid animals have been previously used to test imprinting
of several genes in the *PEG3* domain (Kim et al., 2001). Human genomic DNAs derived from normal
and tumor tissues were obtained from a commercial firm (Biochain).

Each DNA (2 μg) was modified with the bisulfite conversion reaction according to the
manufacturer protocol (EZ DNA methylation kit, Zymo Research). The converted DNA was eluted with
15 μl of TE. Each converted DNA (1 μl) was used as a template for PCR with primers designed using the
MethPrimer program (Li and Dahiya, 2002). The PCR amplification was performed using the Maxime
PCR premix kit (Intron Biotech, South Korea). Information regarding the primer sequences and detailed
PCR conditions for each tested region are available in Appendix A.

The amplified PCR products were analyzed using restriction enzyme digestion (COBRA) (Xiong
and Laird, 1997). Each PCR product was analyzed with two sets of restriction enzymes. First, the
efficiency of the bisulfite conversion reaction was monitored with a set of enzymes that contain non-CpG
cytosines in their recognition sites (Ddel, HpaII; data not shown). Any digestion by these enzymes
indicates that the conversion reaction was incomplete. A second set of enzymes that distinguish between
unmethylated and methylated DNA were chosen to analyze the degree of methylation in each tested
region. The recognition site of each of these enzymes contains a CpG site (TaqI, BstUI, HhaI and
*HpyCH4IV*) or aTpG site (*HphI*). Since methylation inhibits the conversion of cytosines into thymidines
During the bisulfite conversion, bisulfite-treated DNA will be digested by the first group of enzymes if the DNA is methylated. The recognition site for HphI is GGTGA(N)₈, so it will only digest DNA in which the CpG site is unmethylated in vivo. Each of these restriction digestion reactions was repeated at least three times.

Some of the PCR products amplified from the bisulfite-treated DNA were further analyzed through cloning and sequencing. Each of the selected PCR products was purified using the MEGA-Spin agarose gel purification kit (Intron), and then individually cloned into the pGEM-tEasy vector (Promega). At least 10 different clones were randomly selected for DNA sequencing for each PCR product. Due to a cloning bias toward methylated fragments in the mouse Zim2 and Zim3 regions, methylated and unmethylated fragments were separated using COBRA, gel purified, and the unmethylated fragment was ligated into the pGEM-tEasy vector. The purified plasmid DNA was sequenced using BigDye v3.1 (Applied Biosystems). Unincorporated primers and dye terminators were removed via ethanol precipitation, and an ABI 3130 XL was used to analyze the results. To determine methylation status, the resulting electropherograms were visually inspected in BioEdit (Hall, 1999).

References


31


CHAPTER THREE: IDENTIFICATION OF AN ANTISENSE TRANSCRIPT TO ZIM2 IN THE PRIMATE LINEAGE*

*reprinted by permission of *Gene*
**Introduction**

Most autosomal genes are expressed equally from two parental alleles, but up to 200 mammalian genes are only transcribed from one allele based on the parent of origin due to a process called genomic imprinting (Ideraabdullah et al., 2008). Genomic imprinting is found only in marsupials and placental mammals, organisms that utilize a unique reproductive strategy in which young offspring develop inside females’ wombs. Genomic imprinting is a critical gene dosage control mechanism for a subset of genes involved in this strategy (John and Surani, 2000). Most imprinted genes are involved in controlling fetal growth rates and nurturing behaviors (Tilghman, 1999). Proper dosage of imprinted genes is critical for the survival of mammals, and abnormalities in the dosages quite often manifest as genetic diseases in humans. Imprinting-related diseases include Beckwith-Wiedemann, Prader/Willi, Angelman, and Silver-Russell syndromes as well as autistic spectrum disorders (Ferguson-Smith et al., 2004; Ideraabdullah et al., 2008).

Although imprinted genes are found only in mammals, their imprinting status of these genes is not always conserved among all mammals. For example, *Igf2r* (Insulin-like growth factor type-2 receptor) is imprinted in the mouse, but is biallelically expressed in the human (Kalscheur et al., 1993). This domain is an example of one common feature of imprinted domains: the expression of antisense transcripts that regulate imprinting (Pauler et al., 2007). Another common feature of imprinted domains is the presence of CpG islands with allele-specific methylation patterns (differentially methylated regions: DMR) (Ferguson-Smith et al., 2004). Some of these DMRs inherit their methylation as a gametic signal from the previous generations, and these DMRs play critical roles for maintaining the imprinting and transcription of a given domain (Edwards and Ferguson-Smith, 2007). Abnormal methylation levels of these DMRs, either hyper or hypomethylation, are also often associated with many types of human diseases as ‘epimutations’ (Hatchwell and Greally, 2007).
An evolutionarily conserved imprinted region, the PEG3 (Paternally expressed gene 3) domain, is found on human chromosome 19q13.4/chimpanzee chromosome 19/rhesus macaque chromosome 19/proximal mouse chromosome 7. In the mouse, this region contains several imprinted genes including Peg3 and Zim2 (Zinc finger imprinted gene 2). The overall structure of the PEG3 imprinted domain is generally well conserved among mammals, but there have been several lineage specific changes in this region. For example, a cluster of olfactory receptor genes that is present in most mammals has been lost from the primate lineage (Huang and Kim, 2009). Since this deletion of olfactory receptor genes is adjacent to the PEG3 imprinted domain, we analyzed the effects of this change in terms of sequence structure of the region and methylation and expression levels of nearby genes. This study revealed the presence of an antisense transcript to ZIM2, called ZIM2as, in this region of the human, orangutan and chimpanzee genomes.

Results

Olfactory Receptor Deletion in Primates

The region between Zfp28 (Zinc finger protein 28) and Zim2 contains a cluster of olfactory receptor genes in the cow, dog, mouse, and several other mammals (Fig. 3.1). However, this location is devoid of olfactory receptors in the rhesus macaque, orangutan, chimpanzee, and human. This deletion was accompanied by a loss of approximately 90 kb between the mouse and rhesus macaque regions. Comparison of the rhesus, orangutan, chimpanzee, and human sequences revealed the presence of a cluster of tandem repeats in this region in the human, orangutan, and chimpanzee genomes only, which are responsible for an approximately 70 kb size difference between this region in the rhesus macaque and the same region in humans, orangutans and chimpanzees (Fig. 3.2). Analysis of ESTs suggested the possibility of a transcript spanning this region (AI829612.1, DA176232.1, BG707577, AL554662, BX365418.2, BM710959, AW850989), and cDNA cloning and sequencing confirmed the identity of the
Figure 3.1. Organization of the region between ZFP28 and PEG3 in the mouse, rhesus macaque, chimpanzee, and human genomes.

Since the size of this region is different in each organism, this figure is drawn to the scale indicated in the upper right corner. The location for the sequences shown in this figure are human (hg18)-chr19:61,742,129-62,050,982; chimpanzee (panTro2)-chr19:62,386,300-62,704,467; and rhesus (rheMac2)-chr19:62,502,756-62,749,741. Arrows show the direction of transcription. The mouse region is shown as a representative of the non-primate mammals that possess a cluster of olfactory receptor genes in this region. The tree is only representative of the relative relationship between these four animals, and branch lengths are not to scale.
Figure 3.2. Structure of the ZIM2as transcript and the surrounding sequence.

Dot plot comparison of the sequence structure of the region between ZNF71 and PEG3 in the rhesus macaque and chimpanzee against human. The locations of the sequences used for this analysis are human (hg18)-chr19:61,797,100-62,047,400; chimpanzee (panTro2)-chr19:62,451,000-62,701,000; orangutan (ponAbe2)-chr19:58,643,223-58,893,596; and rhesus macaque (rheMac2)-chr19:62,498,300-62,748,400. Each identity between the two sequences is plotted as a dot, so the diagonal line seen in each graph shows a region of sequence similarity. The human sequence is on the x axis of each graph. In the diagrams on the right, two CpG islands in the promoter region of ZIM2as are indicated by asterisks (*) and labeled with the names of the primers used to analyze methylation in this region.

A) Rhesus vs. human. The corresponding rhesus macaque sequence is on the y axis of the dot plot, and the diagram on the right shows the structure of the genes in the indicated region of the rhesus genome. Exons are indicated by boxes, introns by lines, and arrows indicate the direction of transcription.

B) Orangutan/chimpanzee vs. human. Similar results were seen using chimpanzee and orangutan sequence, and the chimpanzee sequence is shown on the y axis of the dot plot. The diagram on the right shows the structure of the genes in the studied region in the human, orangutan, and chimpanzee genomes, including ZIM2as. The ZIM2as transcript is 568 bp long, has five exons, and spans over 118 kb of genomic region. The first exon is found within the first intron of ZNF835 and the final exon is found within an intron of ZIM2. Exons are indicated by boxes, introns by lines, and arrows indicate the direction of transcription.
**ZIM2as Transcript Formation**

The ZIM2as transcript is composed of five exons and is 568 bp long, but lacks an ORF (Appendix A). It spans over 100 kb between ZNF835 (Zinc finger protein 835) and ZIM2, beginning within the first intron of ZNF835 and ending in the fourth from last intron of ZIM2 (Fig. 3.2). The 5’-end of ZIM2as is located in a nuclease accessible site (Boyle et al., 2008) that shows enrichment in histone 3 lysine 4 dimethylation, a mark associated with transcriptional start sites (Mikkelsen et al., 2007), suggesting that the transcription of ZIM2as likely starts at this site. We performed RT-PCR-based expression analyses using primers specific to the ZNF835, ZIM2as, ZIM2, and PEG3 transcripts (Fig. 3.3). Expression of each of these transcripts was assayed using a commercially available panel containing normalized cDNA from human brain, heart, kidney, liver, and placenta as well as human testis. Out of all the tissues studied, ZIM2as expression was detected in the brain and testis (Fig. 3.3). Expression of the ZNF835 transcript was detected in all tested tissues, most strongly in the heart. Both the ZIM2 and PEG3 transcripts were expressed in all tissues tested with relatively high levels in brain and placenta.

**DNA Methylation Analysis**

Since long antisense transcripts are often associated with the maintenance of imprinted domains, we tested the methylation status of several genes near ZIM2as. Briefly, genomic DNA was converted using the bisulfite method, and analyzed via restriction enzyme digestion (COBRA, Xiong and Laird, 1997) and/or cloning and sequencing. We analyzed the methylation levels of the DNA that were derived from
Figure 3.3. Expression analysis of genes in the vicinity of ZIM2as.

Expression patterns of ZIM2as and several nearby transcripts, ZNF835, ZIM2, and PEG3 were analyzed by RT-PCR. GAPDH was used as an internal control. The tissue origin of each cDNA is indicated on the top, and the transcript analyzed is indicated on the left. The two additional lanes shown for the ZIM2as transcript represent results from the cDNA used to clone and sequence this transcript. These two cDNA samples are not normalized with respect to the lanes containing results from the commercial cDNA.

human brain and testis, rhesus and chimpanzee fibroblast cell lines. We first analyzed the two CpG islands near the promoter region of ZIM2as. The one farther from ZIM2as (named ZIM2as-a) was totally digested by HpyCH4 IV in chimpanzee, rhesus macaque, and human brain DNA, indicating complete methylation (Fig. 3.4A).

However, only 50% of the PCR product was digested in the human testis. Bisulfite sequencing and SNP analysis of this region showed allele-specific methylation in the testis (Fig. 3.4B, T vs. G). However, we are unable to determine the parent-of-origin of the methylated allele. The CpG island closer to ZIM2as (named ZIM2as-b) showed species-specific methylation patterns (Fig. 3.4A): the human samples showed no digestion (indicating a lack of methylation), the rhesus macaque sample showed complete digestion, but the chimpanzee sample showed approximately 50% digestion after incubation with HpyCH4 IV. Subsequent bisulfite sequencing of the same chimpanzee sample revealed a mosaic
Figure 3.4. DNA methylation analysis of ZIM2as
The methylation status of CpG islands associated with several genes upstream of the ZIM2as transcript was assayed by bisulfite conversion of each DNA sample followed by COBRA and/or cloning and sequencing. Each analyzed CpG island is indicated by the name of the associated gene.
A) COBRA. The type of DNA sample is indicated above each column, along with (-) for DNA that was incubated without enzyme, and (+) for DNA that was incubated with the appropriate enzyme. The enzyme used for each COBRA is indicated to the right of each gel image. Arrows labeled U and M indicate the position of the unmethylated and methylated DNA bands, respectively. Each reaction was performed at least three times, and a representative image is shown. The bottom portion of the figure contains a summary of the COBRA results in which open boxes indicate a lack of methylation, filled boxes indicate complete methylation, and half-filled boxed indicate a DMR region. The regions that show different methylation patterns among the human, chimpanzee, and rhesus macaque are indicated by diamonds above each lane.
B) Bisulfite sequencing. The associated gene is shown above and the type of DNA sample is shown on the left. The T and G on the right of the human testis bisulfite sequencing result indicate the SNP used to separate the two alleles. Each row of the bisulfite sequencing results indicates a different clone, and each column is an individual cytosine within a CpG dinucleotide. Filled circles indicate methylation, and empty circles denote a lack of methylation.
A) Human brain and testis compared to Chimpanzee fibroblast and Rhesus macaque fibroblast for ZNF71, ZNF470, ZIM2as-a, ZIM2as-b, PEG3, and ZIM3. Each band is labeled with restriction enzymes TaqI, HpyCH4 IV, and BstUI.

B) Genomic blotting images of hZIM2as-a and ptZim2as-b for brain and testis.
pattern of methylation, with no clone being entirely methylated (Fig. 3.4B). In sum, the two CpG islands of ZIM2as displayed species-specific variations among primates, but one of these CpG islands (ZIM2as-a) showed an allele-specific methylation pattern in human testis DNA.

To investigate the possibility that the ZIM2as transcript formation coincided with the expansion of this imprinted region, we also tested methylation status at several upstream genes, including ZNF71 (zinc finger gene 71) and ZNF470 (zinc finger gene 470). The CpG islands associated with each of these genes were unmethylated in the human brain and testis as well as in the chimpanzee and rhesus macaque fibroblast, based on lack of digestion when incubated with the appropriate restriction enzyme (Fig. 3.4A). This indicates an overall similar pattern between different primates. We also analyzed two CpG islands on the downstream side of ZIM2as. The PEG3-CpG island showed a typical DMR pattern in all three species. The ZIM3 (Zinc finger gene imprinted 3)-CpG island showed a similar pattern of non-digestion (unmethylation) between human and chimpanzee. However, some fraction of the ZIM3-CpG island from rhesus was digested, indicating some levels of DNA methylation, which warrants further investigation in the near future.

In summary, the above data show that an antisense transcript has formed in the region upstream of the PEG3 imprinted domain in the human and chimpanzee. The two CpG islands close to the promoter of this primate-specific transcript gene showed species-specific variations in terms of their DNA methylation status. This is in stark contrast to the stable and conserved DNA methylation pattern observed from the nearby zinc finger genes. One of these CpG islands, Zim2as-a, showed allele-specific methylation in the human testis, suggesting that ZIM2as is probably imprinted in the human (Fig. 3.4B).

Discussion

In this study, we report the formation of a new primate-specific antisense transcript, ZIM2as, in a region containing a cluster of olfactory receptor genes in non-primate mammals (Fig. 3.2). In the human, this transcript is expressed in the brain and testis (Fig. 3.3), and a CpG island within its promoter region
shows allele-specific methylation in the testis, suggesting that this transcript may be imprinted (Fig. 3.4). The presence and possible imprinting of this antisense transcript gene suggest that the PEG3 imprinted domain may have expanded in the great apes.

The overall structure of the PEG3 imprinted domain is generally well conserved among mammals, but there have been several lineage specific changes in this region. Figure 3.5 shows these changes in the structure of the area covered by *ZIM2as* and highlights the probable timing of two major events. First, a cluster of olfactory receptor genes that is present in most mammals has been lost from the primate lineage (Fig. 3.1). Second, an antisense transcript to *ZIM2*, called *ZIM2as*, is found in the homologous regions of the human, chimpanzee and orangutan genomes. These two changes might be related to each other.

Clusters of olfactory genes have been shown to be involved as insulators between differently regulated chromatin domains (Valenzuela and

**Figure 3.5. Primate-specific changes in the PEG3 region**
The tree is only representative of the relative relationship between these four animals, and branch lengths are not to scale. For simplicity, only selected genes are shown in this figure. The directional arrows show the direction of transcription of selected genes in this region, and the boxes show the positions of the olfactory receptor (OLFR) and vomeronasal organ receptor (VNO) gene clusters that flank the Peg3 imprinted domain in the mouse. The approximate timings of the olfactory receptor cluster deletion and *Zim2as* transcript formation are shown on the tree.
Kamakaka, 2006). This cluster may have originally acted as a boundary to prevent spreading of imprinting outside of the PEG3 domain, and its loss might have triggered the formation of an antisense transcript gene, \textit{ZIM2as}, in the primates. So far, all the primate genomes analyzed lack this olfactory cluster. On the other hand, sequence analysis and EST searches confirmed the presence of \textit{ZIM2as} only in the great apes, but not in the rhesus genome. This suggests that the loss of the OLFR cluster may predate the formation of \textit{ZIM2as}, and further predicts that the formation of this antisense gene was around 15 million years ago. This antisense transcript gene is most likely imprinted in the humans based on the allele-specific DNA methylation observed from the human testis sample (Fig. 3.4B). We also predict that this antisense transcript gene is imprinted in other primates based on the patterns observed from similar antisense transcripts in other imprinted domains (Ideraabdullah et al., 2008). In particular, many antisense transcripts associated with imprinted domains also begin in intronic regions, e.g \textit{Kcnq1ot1} and \textit{91H} (Ideraabdullah et al., 2008; Bertaux et al., 2008). Following this pattern, the beginning of \textit{ZIM2as} is located within the first intron of \textit{ZNF835}. If it is the case that \textit{ZIM2as} is imprinted, it will be interesting to test if the PEG3 imprinted domain has indeed expanded in the primate lineage. In particular, it would be of great interest to investigate the potential imprinting of several zinc finger genes located in the region immediately adjacent to \textit{ZIM2as}, in particular \textit{ZNF835}.

It is interesting to note that DNA methylation patterns are not conserved between great apes. The methylation status of two CpG islands adjacent to this end of the transcript differs between the chimpanzee and human. The chimpanzee lacks methylation at the first CpG island (Zim2as-a) and has a mosaic methylation pattern at the second (Zim2as-b), while in the human ZIM2as-a is a tissue specific DMR and ZIM2as-b is completely methylated (Fig. 3.3B). This result is consistent with previous reports that human and chimpanzee methylation often differs (Enard et al., 2004). Several possibilities could account for this result. First, the chimpanzee DNA was derived from a cell line, and cell lines have been reported to show altered methylation patterns (Chang et al., 2009). However, four other CpG islands (ZNF470, ZNF471, PEG3, and ZIM3) showed the same methylation pattern between chimpanzee and
human. Second, Zim2as-a could represent a testis-specific DMR. Without access to chimpanzee testis tissue, we are unable to determine if this is the case. Finally, the lineage-specific differences in methylation could be due to the recent appearance of this transcript. There has probably not been enough evolutionary time for its regulation to undergo selection pressure, so the imprinting status of this transcript might not have been fixed due to the very young age of this gene.

In conclusion, both the human and chimpanzee genomes have lost a cluster of olfactory receptors from the same region in which the ZIM2as transcript has formed, and the ZIM2as promoter region shows allele-specific methylation in the human testis. Although the loss of the olfactory cluster and the formation of ZIM2as were not simultaneous, the methylation data and the fact that this transcript is antisense to an imprinted gene suggest that ZIM2as may be imprinted. Since OLFR clusters often function as boundary elements, the loss of this OLFR cluster may have allowed the expansion of the PEG3 imprinted domain in the primates.

Methods

Sequence Analysis

The sequences covering the 250 kb upstream of PEG3 were obtained from the UCSC genome browser for four species: human (hg18-chr19: 61,742,129-62,050,982), chimpanzee (panTro2-chr19: 62,386,300-62,704,467), orangutan (ponAbe2-chr19:58,643,223-58,893,596), and rhesus macaque (rheMac2-chr19: 62,502,756-62,749,741). Pairwise comparisons between human and chimpanzee, human and orangutan, and human and macaque were made using PipMaker (Schwartz et al., 2000). The genomic DNA was analyzed with a Perl script to predict the location of CpG islands. CpG islands were defined as a region of sequence at least 500 bp long with greater than 55% C+G content and an observed/expected CpG dinucleotide ratio of at least 0.65 (Takai and Jones, 2002).
COBRA (COmbined Bisulfite Restriction Analysis) and Bisulfite Sequencing

Human genomic DNA derived from normal tissues were obtained from a commercial firm (Biochain). Chimpanzee (*Pan troglodytes*) and rhesus macaque (*Macaca mulatta*) fibroblast genomic DNA were the generous gift of Dr. Mark Batzer. Each DNA (2 μg) was modified with the bisulfite conversion reaction according to the manufacturer protocol (EZ DNA methylation kit, Zymo Research). The converted DNA was eluted with 15 μl of TE. Each converted DNA (1 μl) was used as a template for PCR with primers designed using the MethPrimer. The PCR amplification was performed using the Maxime PCR premix kit (Intron Biotech). Information regarding the primer sequences and detailed PCR conditions for each tested region are provided as Appendix A.

The amplified PCR products were analyzed using restriction enzyme digestion (COBRA) (Xiong and Laird, 1997). Each PCR product was analyzed with two sets of restriction enzymes. First, the efficiency of the bisulfite conversion reaction was monitored with a set of enzymes that contain non-CpG cytosines in their recognition sites (*DdeI*, *HpaII*). Any digestion by these enzymes indicates that the conversion reaction was incomplete. A second set of enzymes that distinguish between unmethylated and methylated DNA were chosen to analyze the degree of methylation in each tested region. The recognition site of each of these enzymes contains a CpG site (*TaqI*, *BstUI*, and *HpyCH4IV*). Since methylation inhibits the conversion of cytosines into thymidines during the bisulfite conversion, digestion by these enzymes indicates methylation on a given CpG site *in vivo*. Each of these restriction digestion reactions was repeated at least three times.

Selected PCR products amplified from the bisulfite-treated DNA were further analyzed through cloning and sequencing. Each of the selected PCR products was purified using the MEGA-Spin agarose gel purification kit (Intron), and then individually cloned into the pGEM-tEasy vector (Promega). At least 10 different clones were randomly selected for DNA sequencing for each PCR product. Unincorporated primers and dye terminators were removed via ethanol precipitation, and an ABI 3130 XL was used to
analyze the results. To determine methylation status, the resulting electropherograms were visually inspected in BioEdit (Hall, 1999).

**RT PCR**

Normalized cDNA from the human brain, heart, kidney, liver, and placenta was obtained from Biochain. RT-PCR was performed using the Maxime PCR premix kit (Intron Biotech) using an annealing temperature of 56 °C and 38 cycles.

**cDNA Cloning**

Human brain and testis RNA were obtained from a commercial firm (Biochain). Brain and testis RNA (5 µg) were converted to cDNA using the Superscript III First-Strand Synthesis System (Invitrogen). PCR products amplified using the Zim2as-a and Zim2as-b2 primers were cloned into the pGEM-T Easy vector (Promega). Seven colonies were picked, plasmid DNA was isolated using the DNA-spin kit (Intron Biotech), and sequenced using BigDye v3.1 (Applied Biosystems). Unincorporated primers and dye terminators were removed via ethanol precipitation, and the ABI 3130 XL was used to analyze the results.

**References**


CHAPTER FOUR: CONCLUSIONS
Most genes are expressed from two parental alleles, but a few are expressed only from one allele due to the genomic imprinting process, which selectively silences one allele based on the parent of origin (Ideraabdullah et al., 2008). Imprinted genes are mainly found in clusters, where the imprinting of several genes is coordinately controlled by an imprinting control region (ICR) (Edwards and Ferguson-Smith, 2007). Imprinted clusters have a few features in common: differentially methylated regions (DMRs), regions of sequence methylated only on one parental allele, and antisense transcripts (Ferguson-Smith et al., 2004; Pauler et al., 2007). Both of these features play a role in regulating expression of the genes in the imprinted cluster. Imprinting can have an immediate effect on health because the genes are effectively hemizygous, so mutations in the expressed allele have the same effect as a dominant mutation. In addition, many imprinted genes have been implicated in cancer formation, consistent with the fact that many of these genes are involved in growth (Tilghman, 1999). The Peg3 imprinted domain is a region specific to placental mammals, and its overall organization is conserved among all mammals (Kim and Stubbs, 2005). Several genes in this region have been shown to be imprinted in the mouse and cow (Kim et al., 2007). The imprinting and expression pattern of the Peg3 gene itself has been studied in detail in several species including human. Here we present the first systematic analysis of DNA methylation over this entire region. This analysis of the “normal” methylation patterns in this domain provides a foundation for future studies of possible changes in the epigenetic status of this region in disease or after exposure to various environmental factors.

In chapter two, we describe a survey of the DNA methylation status of the CpG islands in the Peg3 imprinted domain in human normal and tumor samples, several developmental stages of the mouse, and the adult cow. Our study revealed several interesting features of this domain. In terms of sequence characteristics, we found that this domain contains more CpG islands and tandem repeats (both within CpG islands and in the greater region) than a nonimprinted region containing similar types of genes. The prevalence of these two features may be linked. CpG islands have a high mutation rate due to spontaneous deamination of methylated cytosines to produce thymine. Tandem repeats are often present
within CpG islands and are especially enriched in imprinted regions (Hutter et al., 2006). Creating tandemly repeated arrays of a CpG island would preserve its functionality if there is some threshold size required for a CpG island to be relevant to gene expression. Also, we found fewer LINEs in this region than the reference region. This is the opposite result from that found by Allen et al., which found enrichment of LINEs in imprinted regions. This result cannot be explained simply by base composition, since the Peg3 imprinted region and the reference region are similar in this respect. Thus, this may be a specific feature of the Peg3 imprinted domain. On the other hand, this could be an effect of the relatively small (compared to the entire genome) region of sequence studied.

After analyzing the sequence characteristics of this region, we analyzed the DNA methylation status of the predicted CpG islands that were associated with genes since these were most likely to be relevant to their expression. We found that the Peg3-CpG island was a DMR in adult tissues of all three species studied, and it was hypermethylated in human ovarian tumor DNA. We also identified two previously unreported DMRs in the Peg3 imprinted domain. The Zim3-CpG island is a DMR in the mouse, but this methylation status is not shared with human (unmethylated) or cow (completely methylated). In addition, we discovered a new possible tissue-specific DMR at the human USP29-CpG island, which showed a mosaic pattern of methylation in the human brain and was comparatively hypomethylated in normal testis DNA, but was hypermethylated in breast and ovarian tumor DNA. The Zim1-CpG island lacked methylation in all of the samples studied (sperm, ES cell, placenta, and adult liver).

This was an unexpected result, since Zim1 is a maternally expressed gene (Kim et al., 1999). Since Zim1 is most highly expressed around embryonic day 14, differential methylation might be present at this stage. The lack of methylation in the adult could be explained by the lack of maintenance of the mark past this stage, effectively passively demethylating this region. However, the imprint could also be regulated through other epigenetic mechanisms, such as histone modifications. To discover what
mechanism controls the imprinting of this gene, samples from this developmental stage should be analyzed for both DNA methylation levels and chromatin modifications.

The Peg3-CpG island was a DMR in adult tissues of all three species studied. As expected based on the results of previous studies (Li et al., 2000), the mouse germline samples showed that methylation was present only on the maternal allele at this stage. This pattern continued throughout the blastocyst, ES cell, placenta, and adult samples tested. In the human, we saw the same pattern in DNA derived from normal brain and testis tissues as well as in breast tumor DNA. However, the PEG3-CpG island was hypermethylated (80% methylated) in human ovarian tumor DNA. PEG3 expression is reduced by the hypermethylation of its promoter CpG island in gliomas and ovarian tumors (Maegawa et al., 2001; Feng et al., 2008). Since Peg3 knockout mice have not been reported to show more tumor development than wild type mice, Peg3 is probably not involved in the initiation of cancer (Li et al., 1999). It is probably either methylated as an indirect effect of an oncogene, or is shut off in the later phases of the establishment of a tumor. Inactivation of Peg3 could potentially give cancer a growth advantage since this protein has been shown to promote apoptosis subsequent to p53 activation (Deng and Wu, 2000).

We also identified two previously unreported DMRs in the Peg3 imprinted domain. The Zim3-CpG island is also a DMR in the mouse, but this methylation status is not shared with human (unmethylated) or cow (completely methylated). This species-specific DNA methylation pattern indicates that the Zim3-CpG island is possibly less constrained, or that Zim3 plays a different role in each animal. Changes in DNA methylation status could possibly happen faster than sequence divergence, so these species-specific changes could play a role in the evolution of the species. Compared to the Peg3-CpG island, the Zim3-CpG island showed an interesting pattern of methylation acquisition. This CpG island is an interesting example of a region with parental allele-specific methylation that is not set in the germline. The methylation on the Peg3-CpG island was established on the maternal allele and remained there throughout development (germline DMR). However, the Zim3-CpG island was unmethylated in the sperm sample, but methylation was found specifically on the paternal allele in the adult. This swap
implies that the DNA methylation was reset after fertilization and then directed to the paternal allele (somatic DMR). This is consistent with the fact that Zim3 is maternally expressed (Kim et al., 2001). The methylation pattern on the Zim3-DMR is probably established directly by the ICR of the Peg3 imprinted domain, but may be an indirect effect of the methylation pattern of Usp29. In other words, the Peg3-ICR may be responsible for setting the methylation status of Usp29, which in turn sets the pattern of Zim3. In mice the Usp29 and Zim3 transcripts overlap and are antisense to each other (Fig. 4.1). Thus, the presence of transcription machinery on one of these genes would physically block transcription of the other. Also, the antisense transcript might direct repressive modifications to the sense promoter.

![Diagram of mouse chromosome 7 between Peg3 and Zim3](Image)

**Figure 4.1 Expression patterns of Peg3, Usp29, and Zim3.** A schematic of mouse chromosome 7 between Peg3 and Zim3. The maternal and paternal copies are displayed separately and indicated by their respective symbols. The direction of each arrow indicates the direction of transcription, and an x over an arrow indicates that that transcript is not expressed.

In addition, we discovered a new possible tissue-specific DMR at the human USP29-CpG island, which showed a mosaic pattern of methylation (54.5% methylated) in the human brain and was hypomethylated (26.3% methylated) in normal testis DNA. This CpG island showed hypermethylation in both breast (81.8% methylated) and ovarian (65.5% methylated) tumor DNA. Overall, two CpG islands in the Peg3 imprinted domain were found to be hypermethylated in two types of human cancer genomic DNA. This would most likely result in repression of the associated genes. Future studies of this region should survey a greater range of cancer types to see if the hypermethylation and/or repression of PEG3
and USP29 only occurs in certain types of cancers or if it is a general pattern. In addition, it would be interesting to determine at what point of cancer progression PEG3 and USP29 become hypermethylated. If this can be determined, hypermethylation of these genes could be used as a marker for the progression of a cancer.

In chapter three, we describe a transcript antisense to ZIM2 in the great apes. This transcript spans the region from the first intron of ZNF835 to the fourth from last intron of ZIM2 and is expressed in human brain and testis. We tested the DNA methylation status of several CpG islands surrounding ZIM2as, including two CpG islands in its promoter region. This investigation revealed a species-specific methylation pattern. The CpG island farthest from the start site of ZIM2as is a human testis-specific DMR, but is methylated in the chimpanzee. The CpG island proximal to ZIM2as shows a mosaic pattern of methylation in the chimpanzee, but it was unmethylated in the human. The differences between DNA methylation in the human and chimpanzee represent the ability of DNA methylation to provide a source of variation beyond nucleotide sequence, since there is 99% sequence identity between these two species. Both CpG islands associated with ZIM2as were methylated in the rhesus macaque. The complete methylation of both of these CpG islands in the rhesus macaque is most likely because this region is simply an intron sequence in this species. Several of the CpG islands on the 5’ side of ZIM2as (those associated with ZNF 470, ZNF71, ZNF471, ZNF28) were unmethylated in all three species. We tested two CpG islands on the 3’ side of ZIM2as and found that the PEG3-CpG island was a DMR in all three species. However, the ZIM3-CpG island showed a DMR-like pattern in the rhesus, but was unmethylated in both human and chimpanzee. This DMR pattern is similar to the pattern found in the mouse, which may indicate that the presence of ZIM2as in the great apes can affect some of the nearby genes in the imprinted domain. An alternative explanation for the methylation changes seen here is the fact that both the rhesus and chimpanzee DNA samples are from cultured cell lines. Prolonged growth in cell culture has been shown to alter methylation patterns (Chang et al., 2009). However, it is rather unlikely that this
is the source of these differences since the methylation patterns of DNA from two different human tissues and both of the primate cell lines are identical in four surrounding regions.

In most mammals, the Peg3 imprinted domain is flanked by a cluster of olfactory receptor genes on one end and a cluster of vomeronasal organ receptor genes on the other. This cluster of olfactory receptor genes is present in all the non-primate mammals for which we could obtain genome sequences, including the mouse, cow, and dog. Since these types of gene clusters often have insulator or barrier functions (Valenzuela and Kamakaka, 2006), the clusters flanking the Peg3 region probably delineate the boundaries of this imprinted domain in most mammals. However, these gene clusters are not found in the rhesus macaque, orangutan, chimpanzee, or human (the only primates for which there are genome sequences of sufficient coverage available). It is possible that the loss of this particular cluster on the 5’ side of the primate PEG3 imprinted domain may result in spreading of imprinting to more distant regions. Another interesting difference in the primate Peg3 region is the presence of a transcript oriented antisense to \textit{ZIM2} (\textit{ZIM2as}) in the human, chimpanzee, and orangutan genomes. The majority of its coding sequence is absent from the rhesus macaque genome. The fact that all primates analyzed lack the olfactory receptor cluster while only humans, chimpanzees, and orangutans have the antisense transcript suggests that \textit{ZIM2as} formed after the deletion. Alternatively, the olfactory cluster deletion and \textit{ZIM2as} formation might have been simultaneous, and the rhesus may have undergone a lineage-specific deletion of \textit{ZIM2as}. Completion and assembly of the marmoset genome sequence should provide some insight into the true sequence of the genomic rearrangements in the primate PEG3 imprinted domain.

Overall, it seems that the formation of \textit{ZIM2as} in the great apes is associated with several changes in DNA methylation, including allele-specific methylation in the human testis. This antisense transcript gene is most likely imprinted in the humans based on the allele-specific DNA methylation observed from the human testis sample. We also predict that this antisense transcript gene is imprinted in other primates based on the patterns observed from similar antisense transcripts in other imprinted domains (Ideraabdullah et al., 2008). If it is the case that \textit{ZIM2as} is imprinted, it will be interesting to test if the
PEG3 imprinted domain has indeed expanded in the primate lineage. In particular, it would be of great interest to investigate the potential imprinting of several zinc finger genes located in the region immediately adjacent to ZIM2as, in particular ZNF835. Since ZIM2as spans both this gene and ZIM2, it may regulate the expression of either or both of these transcripts (Fig 4.2).

Figure 4.2. Expression patterns of ZNF835, ZIM2as, and PEG3/ZIM2.
A schematic of allele-specific expression on human chromosome 19 between ZNF835 and PEG3. The maternal and paternal copies are displayed separately and indicated by their respective symbols. The direction of each arrow indicates the direction of transcription, and an x over an arrow indicates that that transcript is not expressed.

If ZIM2as does function similarly to antisense transcripts found in other imprinted domains, one of its roles would be the establishment and maintenance of imprinting of this region (Ideraabdullah et al., 2008; Bertaux et al., 2008). It might be also involved in maintaining the imprinting of the expanded Peg3 domain. In the future it will be interesting to analyze changes in chromatin modifications and expression status of the genes surrounding ZIM2as.

This dissertation represents the first systematic study of DNA methylation in the Peg3 imprinted region in the adult human as well as in several developmental stages in the mouse (sperm, oocyte, blastocyst, ES cell, placenta, and adult liver). It also describes an antisense transcript that has formed in the great ape PEG3 imprinted domain which may control the extension of this imprinted domain. The results of these studies provide a firm foundation upon which future studies of this imprinted domain can rest. Although it was only possible to study DNA methylation in relatively short stretches of sequence in
these studies, advances in technology should soon make the study of DNA methylation across this domain feasible and aid in the determination of the exact extent of the imprinted domain.

References


APPENDIX A: SUPPLEMENTAL DATA
Bisulfite primer information for analysis of the PEG3 imprinted region in the human, mouse, and cow.

<table>
<thead>
<tr>
<th>Organism</th>
<th>Primer</th>
<th>Sequence</th>
<th>Annealing temp (°C)</th>
<th>Restriction enzyme to analyze methylation</th>
<th>Restriction enzyme to monitor bisulfite conversion</th>
</tr>
</thead>
<tbody>
<tr>
<td>human</td>
<td>hZNF264-bis-1</td>
<td>gttttgttgggttttgatttggaaggaaag</td>
<td>60</td>
<td>HpyCH4IV</td>
<td>HpaII</td>
</tr>
<tr>
<td></td>
<td>hZNF264-bis-2.1</td>
<td>aaccctcaataaactacaacctctcaaac</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>hUSP29-bis-5</td>
<td>gtttaggaaggtagatttagtagggtaggtgtgt</td>
<td>60</td>
<td>TaqI</td>
<td>HpaII</td>
</tr>
<tr>
<td></td>
<td>hUSP29-bis-6.1</td>
<td>ctaaaccctcctaactatatcatctcaaac</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>hZIM3-bis-1</td>
<td>gtttaggaaggtagatttagtagggtaggtgtgt</td>
<td>60</td>
<td>BstUI</td>
<td>HpaII</td>
</tr>
<tr>
<td></td>
<td>hZIM3-bis-2.1</td>
<td>ctaaaccctcctaactatatcatctcaaac</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>hDUXA-bis-5</td>
<td>gtttaggaaggtagatttagtagggtaggtgtgt</td>
<td>60</td>
<td>HpyCH4IV</td>
<td>DdeI</td>
</tr>
<tr>
<td></td>
<td>hDUXA-bis-6</td>
<td>aattcaatacatcatctcctcaacctcc</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>hDUXA-bis-3</td>
<td>atgttagaaaggtagatttagtagggtaggtgtgt</td>
<td>60</td>
<td>TaqI</td>
<td>HpaII</td>
</tr>
<tr>
<td></td>
<td>hDUXA-bis-4</td>
<td>ccatcaattaactaaaaacttaaatttta</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>hPEG3-bis-3</td>
<td>gtttttttgtgtaataaagttttaggtgtgt</td>
<td>58</td>
<td>HpyCH4IV</td>
<td>HpaII</td>
</tr>
<tr>
<td></td>
<td>hPEG3-bis-4</td>
<td>ccatcaataactcacttactctatctccta</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>hZNF835-3</td>
<td>tggtaggtttaggttacgtttggtgtga</td>
<td>60</td>
<td>HpyCH4IV</td>
<td>HpaII</td>
</tr>
<tr>
<td></td>
<td>hZNF835-4</td>
<td>ccatcaataactcactctatctatctccta</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>hH19-bis-3</td>
<td>tggtaggtttaggttacgtttggtgtga</td>
<td>60</td>
<td>TaqI</td>
<td>HpaII</td>
</tr>
<tr>
<td></td>
<td>hH19-bis-4</td>
<td>actctataaatctctattctcctaataac</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>hZNF71-bis-5</td>
<td>agttttttgtgagatggaggtttagttg</td>
<td>62</td>
<td>TaqI</td>
<td>HpaII</td>
</tr>
<tr>
<td>Species</td>
<td>Name</td>
<td>Sequence</td>
<td>Size</td>
<td>Enzymes</td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>---------</td>
<td>-------------------</td>
<td>------</td>
<td>------------------</td>
<td></td>
</tr>
<tr>
<td>cow</td>
<td>cPeg3-bis-5</td>
<td>tattttgttttataaggagttgtttttt</td>
<td>60</td>
<td>BstUI, Ddel</td>
<td></td>
</tr>
<tr>
<td></td>
<td>cPeg3-bis-6</td>
<td>ctcatactataataaactacataaa</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>cZim3-bis-3</td>
<td>gtaagagttgaggtgtgtataggtagg</td>
<td>60</td>
<td>ClaI, HpaII</td>
<td></td>
</tr>
<tr>
<td></td>
<td>cZim3-bis-4</td>
<td>ccaatataacccctataataatcatac</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>cZim2-bis-5</td>
<td>gagaatttttgagtttttttagttagtaaga</td>
<td>58</td>
<td>BstUI, HpaII</td>
<td></td>
</tr>
<tr>
<td></td>
<td>cZim2-bis-6</td>
<td>ttttccttaaaccacaactatattaact</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>cAst1-bis-5</td>
<td>agggatttaagatataatatatttagata</td>
<td>58</td>
<td>HhaI, HpaII</td>
<td></td>
</tr>
<tr>
<td></td>
<td>cAst1-bis-6</td>
<td>atacraaaattacttaaaaaataaaaaac</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>cUsp29-bis-5</td>
<td>gtgtataaaaaagagatataatttttagttagtaaga</td>
<td>56</td>
<td>TaqI, HpaII</td>
<td></td>
</tr>
<tr>
<td></td>
<td>cUsp29-bis-6</td>
<td>aacctcataaaaaataatcctaaaaatct</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mouse</td>
<td>mPeg3-met-15</td>
<td>aggaagagttagaggagttatatttataga</td>
<td>60</td>
<td>HphI, BstUI, Ddel</td>
<td></td>
</tr>
<tr>
<td></td>
<td>mPeg3-met-16.1</td>
<td>cctattacaacacacacaataaaatca</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>mPeg3-met-15.2</td>
<td>tagtgtttattttattttttagttgtag</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>mZim1-met-3</td>
<td>gttagtagaagttagttttaagtagttgtaggttta</td>
<td>58</td>
<td>HhaI, HpaII</td>
<td></td>
</tr>
<tr>
<td></td>
<td>mZim1-met-4</td>
<td>cttactaaacactttccccacaaatcc</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>mZim2-met-3</td>
<td>gttataatatatttagttatattttatttttt</td>
<td>55</td>
<td>TaqI, Ddel</td>
<td></td>
</tr>
<tr>
<td></td>
<td>mZim2-met-4</td>
<td>atataacaactttcttaataataaaacct</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>mZim2-met-5</td>
<td>gatatatttattatttttttttgggtaaa</td>
<td>55</td>
<td>TaqI, Ddel</td>
<td></td>
</tr>
<tr>
<td></td>
<td>mZim2-met-6</td>
<td>aattactttcatcaatctactctttaaaac</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>mZim3-met-3</td>
<td>tatattttgaagaaattttttttttttaa</td>
<td>58</td>
<td>BstUI, Ddel</td>
<td></td>
</tr>
<tr>
<td></td>
<td>mZim3-met-4</td>
<td>accactaacttccttaactactacaaact</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
DNA Methylation analysis of Peg3 and Usp29 in human matched-pair lung and liver samples

COBRA data from human matched pair DNA. Lanes labeled U indicate undigested DNA, and lanes labeled C indicate digested DNA. The enzyme with which each sample was incubated is indicated to the right of the figure. The arrows labeled U and M indicate the position of unmethylated and methylated DNA, respectively. There is no difference in the methylation levels of the PEG3-CpG island and the USP29-CpG island between the different tissues or between normal and tumor tissues.
### Zim2as Exon Structure

<table>
<thead>
<tr>
<th>Feature</th>
<th>Size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exon 1</td>
<td>50</td>
</tr>
<tr>
<td>Intron 1</td>
<td>11870</td>
</tr>
<tr>
<td>Exon 2</td>
<td>74</td>
</tr>
<tr>
<td>Intron 2</td>
<td>25430</td>
</tr>
<tr>
<td>Exon 3</td>
<td>53</td>
</tr>
<tr>
<td>Intron 3</td>
<td>80072</td>
</tr>
<tr>
<td>Exon 4</td>
<td>151</td>
</tr>
<tr>
<td>Intron 4</td>
<td>977</td>
</tr>
<tr>
<td>Exon 5</td>
<td>240</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Transcript Length</th>
<th>Genomic DNA Span</th>
</tr>
</thead>
<tbody>
<tr>
<td>568</td>
<td>118917</td>
</tr>
</tbody>
</table>

### RT-PCR primer information for analysis of transcripts surrounding ZIM2as in the human.

<table>
<thead>
<tr>
<th>Locus</th>
<th>Primer Name</th>
<th>Primer Sequence</th>
<th>Annealing Temp</th>
</tr>
</thead>
<tbody>
<tr>
<td>PEG3</td>
<td>hPEG3rt-1</td>
<td>aagagaagtgcctacccaaage</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>hPEG3rt-2</td>
<td>aacetaagctccctaatg</td>
<td></td>
</tr>
<tr>
<td>ZNF835</td>
<td>hZNF835rt-3</td>
<td>gttggaagaaactgaaacac</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>hZNF835rt-4</td>
<td>tgggttaagatgacgctgaac</td>
<td></td>
</tr>
<tr>
<td>GAPDH</td>
<td>hGAPDHrt-1</td>
<td>gcaggggggagccaaaggg</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>hGAPDHrt-2</td>
<td>tcacgtgccccacggctaaagag</td>
<td></td>
</tr>
<tr>
<td>ZIM2</td>
<td>ZIM2rt-1</td>
<td>catttagacattttgggtgtag</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>ZIM2rt-2</td>
<td>gtttggggtcattatgtaag</td>
<td></td>
</tr>
<tr>
<td>ZIM2as</td>
<td>ZIM2as-a</td>
<td>gagttcagaggggagccttctg</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>ZIM2as-b2</td>
<td>ccaggagacagggcataaagaggt</td>
<td></td>
</tr>
</tbody>
</table>
Bisulfite primer information for analysis of the *ZIM2as* region in the human, chimpanzee, and rhesus macaque.

<table>
<thead>
<tr>
<th>Organism</th>
<th>Primer Name</th>
<th>Primer Sequence</th>
<th>Annealing Temp</th>
<th>Restriction enzyme to monitor bisulfite conversion</th>
<th>Restriction enzyme to analyze methylation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human</td>
<td>hPEG3-bis-3</td>
<td>gtttttctgggtgaataaagtttttggttag</td>
<td>58</td>
<td>HpaII</td>
<td>HpyCH4 IV</td>
</tr>
<tr>
<td></td>
<td>hPEG3-bis-4</td>
<td>cccaccttaaatcaccatatcttaacraaac</td>
<td>58</td>
<td>HpaII</td>
<td>BstUI</td>
</tr>
<tr>
<td></td>
<td>ptZim3-bis-1</td>
<td>ttttagattttggaggtaaggttagg</td>
<td>58</td>
<td>HpaII</td>
<td>TaqI</td>
</tr>
<tr>
<td></td>
<td>ptZim3-bis-2</td>
<td>caactacttcaaatctataattctacttactc</td>
<td>58</td>
<td>HpaII</td>
<td>TaqI</td>
</tr>
<tr>
<td></td>
<td>hZNF71-bis-5</td>
<td>agtttttttttttttttaatgagttagttagtg</td>
<td>60</td>
<td>TaqI</td>
<td>TaqI</td>
</tr>
<tr>
<td></td>
<td>hZNF71-bis-6</td>
<td>cctctactcctaatctataactaactctacta</td>
<td>60</td>
<td>TaqI</td>
<td>TaqI</td>
</tr>
<tr>
<td></td>
<td>hZNF470-bis-1</td>
<td>gggataatttttaaggggaattttagtagt</td>
<td>60</td>
<td>TaqI</td>
<td>TaqI</td>
</tr>
<tr>
<td></td>
<td>hZNF470-bis-2</td>
<td>aaccacttaaatctacttactctacttactcc</td>
<td>60</td>
<td>TaqI</td>
<td>TaqI</td>
</tr>
<tr>
<td>Chimpanzee</td>
<td>ptZnf71-bis-1</td>
<td>gttttttttattttttttttttttttttttttttt</td>
<td>60</td>
<td>TaqI</td>
<td>TaqI</td>
</tr>
<tr>
<td></td>
<td>ptZnf71-bis-2</td>
<td>ccccaaatattactataactaactactaactaca</td>
<td>60</td>
<td>TaqI</td>
<td>TaqI</td>
</tr>
<tr>
<td></td>
<td>ptZnf470-bis-1</td>
<td>gattataatttttaataggttgttttagtg</td>
<td>60</td>
<td>TaqI</td>
<td>TaqI</td>
</tr>
<tr>
<td></td>
<td>hZNF470-bis-2</td>
<td>aaccacttaaatatcaactaactactacaactaca</td>
<td>60</td>
<td>TaqI</td>
<td>TaqI</td>
</tr>
<tr>
<td></td>
<td>hZNF470-bis-3</td>
<td>gttttttttattttttttttttttttttttttttt</td>
<td>60</td>
<td>TaqI</td>
<td>TaqI</td>
</tr>
<tr>
<td></td>
<td>hZNF470-bis-4</td>
<td>ccccaaatattactataactaactactaactaca</td>
<td>60</td>
<td>TaqI</td>
<td>TaqI</td>
</tr>
<tr>
<td></td>
<td>ptZim2as-bis-1</td>
<td>gtattttttttttttttttttttttttttttttttt</td>
<td>60</td>
<td>TaqI</td>
<td>TaqI</td>
</tr>
<tr>
<td></td>
<td>ptZim2as-bis-2</td>
<td>acaccttaaatatcaactaactaactacta</td>
<td>60</td>
<td>TaqI</td>
<td>TaqI</td>
</tr>
<tr>
<td>Rhesus Macaque</td>
<td>Primer</td>
<td>Sequence</td>
<td>T (bp)</td>
<td>Enzyme</td>
<td>Restriction Enzyme</td>
</tr>
<tr>
<td>---------------</td>
<td>--------</td>
<td>----------</td>
<td>--------</td>
<td>--------</td>
<td>-------------------</td>
</tr>
<tr>
<td>rhPeg3-bis-5</td>
<td>ttagataataataggggaaggggtgag</td>
<td>58</td>
<td>HpyCH4 IV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rhPeg3-bis-6</td>
<td>ccrctaaatatatatattccacaaaaa</td>
<td>58</td>
<td>HpaI</td>
<td>BstUI</td>
<td></td>
</tr>
<tr>
<td>rhZim3-bis-5</td>
<td>aatgaaaggtatgtgtagtaggattagt</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rhZim3-bis-6</td>
<td>aaccaattaatitaacctctctcaac</td>
<td>58</td>
<td>HpaI</td>
<td>BstUI</td>
<td></td>
</tr>
<tr>
<td>rhZnF71-bis-3</td>
<td>tttgggagtagaggagtttagttggtttttg</td>
<td>62</td>
<td>TaqI</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rhZnF71-bis-4</td>
<td>aaccaattaaaatttttaaccatctctcaac</td>
<td>62</td>
<td>TaqI</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rhZnF470-bis-1</td>
<td>ttttagttttgtttttgttttatatttt</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rhZnF470-bis-2</td>
<td>tactacctatctcttttaacactc</td>
<td>60</td>
<td>HpyCH4 IV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rhZnF835-bis-1</td>
<td>gagttttttggttaaggtttttatatt</td>
<td>60</td>
<td>HpyCH4 IV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rhZnF835-bis-2</td>
<td>cccctactactacaaccaatac</td>
<td>60</td>
<td>HpyCH4 IV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rhZm2as-bis-1</td>
<td>ttaggagaatgttagttgttgtaagtt</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rhZm2as-bis-2</td>
<td>ataccaaccacactatcatctacaca</td>
<td>60</td>
<td>HpyCH4 IV</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX B:

LETTERS OF PERMISSION
This is a License Agreement between Jennifer M Huang ("You") and Elsevier ("Elsevier") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

Supplier Elsevier Limited
The Boulevard, Langford Lane
Kidlington, Oxford, OX5 1GB, UK

Registered Company Number 1982084
Customer name Jennifer M Huang
Customer address 107 Life Sciences Building
Baton Rouge, LA 70803

License Number 2218850314179
License date Jun 30, 2009
Licensed content publisher Elsevier
Licensed content publication Gene
Licensed content title DNA methylation analysis of the mammalian PEG3 imprinted domain
Licensed content author Jennifer M. Huang and Joomyeong Kim
Licensed content date 1 August 2009
Volume number 442
Issue number 1-2
Pages 8
Type of Use Thesis / Dissertation
Portion Full article
Format Both print and electronic
You are an author of the Elsevier article Yes
Are you translating? No
Order Reference Number
Expected publication date Aug 2009
Elsevier VAT number GB 494 6272 12
Permissions price 0.00 USD
Value added tax 0.0%

Total 0.00 USD
This is a License Agreement between Jennifer M Huang ("You") and Elsevier ("Elsevier") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

| Supplier | Elsevier Limited  
The Boulevard, Langford Lane  
Kidlington, Oxford, OX5 1GB, UK |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Registered Company Number</td>
<td>1982084</td>
</tr>
<tr>
<td>Customer name</td>
<td>Jennifer M Huang</td>
</tr>
</tbody>
</table>
| Customer address | 107 Life Sciences Building  
Baton Rouge, LA 70803 |
| License Number | 2235391462698 |
| License date | Jul 24, 2009 |
| Licensed content publisher | Elsevier |
| Licensed content publication | Gene |
| Licensed content title | Identification of an antisense transcript to ZIM2 in the primate lineage |
| Licensed content author | Jennifer M. Huang, Sungryul Yu and Joomyeong Kim |
| Licensed content date | 15 September 2009 |
| Volume number |  |
| Issue number |  |
| Pages |  |
| Type of Use | Thesis / Dissertation |
| Portion | Full article |
| Format | Both print and electronic |
| You are an author of the Elsevier article | Yes |
| Are you translating? | No |
| Order Reference Number |  |
| Expected publication date | Aug 2009 |
| Elsevier VAT number | GB 494 6272 12 |
| Permissions price | 0.00 USD |
| Value added tax 0.0% | 0.00 USD |
| Total | 0.00 USD |
VITA

Jennifer Marian Huang was born in New Orleans, Louisiana, in June 1981. She graduated from Baton Rouge High School in 1999. Jennifer graduated from Louisiana State University with a Bachelor of Science degree in biological sciences in 2004. After a year of nonmatriculating study, she began her doctoral research in 2005 under the guidance of Dr. Joomyeong Kim. Mrs. Huang will graduate with the degree of Doctor of Philosophy in December 2009.