2016

PEG3 Functions as a Repressor for its Downstream Genes

An Ye
Louisiana State University and Agricultural and Mechanical College, yeanww1987@gmail.com

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

Recommended Citation
Ye, An, "PEG3 Functions as a Repressor for its Downstream Genes" (2016). LSU Doctoral Dissertations. 2261.
https://digitalcommons.lsu.edu/gradschool_dissertations/2261

This Dissertation is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Doctoral Dissertations by an authorized graduate school editor of LSU Digital Commons. For more information, please contact gradetd@lsu.edu.
PEG3 PROTEIN FUNCTIONS AS A REPRESSOR FOR ITS DOWNSTREAM GENES

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

An Ye

M.D., Hebei Medical University, 2010
August 2016
ACKNOWLEDGEMENTS

Graduation means another start of a new journey. However, it would not even be possible to reach the end of my graduate career without help and support from many people. I would like to take this chance to sincerely thank them for all help they have given me over my four-year graduate study. First and foremost, I would like to express my sincere appreciations and many thanks to my major professor, Dr. Joomyeong Kim, who has made enormous contributions on the path of my Ph.D study. He is an outstanding scientist with great enthusiasms, immerse knowledge and tremendous motivations. Meanwhile, I would like to thank him a lot for being patients on my first paper writing and my dissertation. His guidance and help on my research have been invaluable throughout my graduate studies.

Besides my major advisor, I would like to thank my committee members Dr. Anne Grove, Dr. David Donze and my dean’s representative Dr. Gary Byerly for their time, insightful comments on my work and recommendations. My special thanks to Dr. David Donze for his permission to use the sonicator in his lab, and for welcoming me at his lab parties. Also, I would like to thank my first research mentor, Dr. Lanjuan Zhao, who invited me and helped me settle down in the United State.

I also want to thank my lab members, Dr. Hongzhi He, Dr. Hana Kim, Dr. Wesley Frey, Bambarendage Pinithi Perera, Cory Bretz, and Arundhati Bakshi for their helps in my experiment design and insightful suggestions on my research. I especially thank Dr. Hongzhi He and Dr. Hana Kim who have constantly guided and supported me on my Ph.D study. I also want to take this chance to thank Bambarendage Pinithi Perera for her kind suggestions and help on my research and personal life.
Last but not least, I would like to thank my family for their endless support. I thank my parents who always encourage me and support me to pursue greater things. I thank my dear wife, Dr. Qing Wang, for being part of my life. I thank her for her patience, support, consideration, and for providing unconditional love to my whole family. Also, I am thankful that I have an adorable daughter, Anna Ye, who brings a lot of joy and happiness to my family.
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS ........................................................................................................... ii  

LIST OF TABLES .................................................................................................................. v  

LIST OF FIGURES .................................................................................................................. vi  

ABSTRACT .............................................................................................................................. viii  

CHAPTER 1 INTRODUCTION ............................................................................................ 1  

CHAPTER 2 PATERNALLY EXPRESSED PEG3 CONTROLS MATERNALLY EXPRESSED ZIM1 AS A TRANS FACT ................................................................. 26  

CHAPTER 3 PEG3 BINDS TO H19-ICR AS A TRANSCRIPTIONAL REPRESSOR ............................................................................................................. 47  

CHAPTER 4 CONCLUSIONS AND DISCUSSIONS ........................................................... 84  

APPENDIX: AUTHORIZATION FOR USE OF PUBLISHED MATERIAL ....................... 90  

VITA ...................................................................................................................................... 91
LIST OF TABLES

Table 2.1 Primer sequences used in this study .................................................................30
Table 3.1 Final output of MACS2 describing ChIP-seq peaks derived from WT MEF cells........51
Table 3.2 Final output of MACS2 describing ChIP-seq peaks derived from KO MEF cells........53
Table 3.3 Primer sequences used in this study .................................................................56
Table 3.4 Downstream genes of PEG3 in MEF cells ........................................................61
LIST OF FIGURES

Figure 1.1  Roles of H19, Zac1, and Peg3 in the imprinted gene network..........................8

Figure 1.2  The overview of the Peg3 domain.................................................................11

Figure 1.3  KRAB-containing ZFPs repress gene expression through interacting with SETDB1 and HP1.................................................................15

Figure 1.4  The amino acid sequence of KRAB domain in PEG3 protein......................16

Figure 2.1  Removal of the PEG3 protein results in the up-regulation of Zim1.............33

Figure 2.2  PEG3 binds to the zinc finger exon of Zim1..............................................36

Figure 2.3  PEG3 binds to the zinc finger exon of Zim1 using custom-made anti-PEG3 antibody...37

Figure 2.4  PEG3 binds to both alleles of Zim1..........................................................38

Figure 2.5  Reduced levels of H3K9me3 in the mutant cells lacking PEG3..................39

Figure 2.6  Restoring the protein levels of PEG3 down-regulates Zim1.........................41

Figure 2.7  Overexpression of the protein levels of PEG3 down-regulates Zim1.............42

Figure 2.8  Paternally expressed Peg3 controls maternally expressed Zim1 as a trans factor involving H3K9me3.................................................................44

Figure 3.1  PEG3 ChIP-seq using WT and KO MEF cells..............................................59

Figure 3.2  Expression level changes of the identified downstream genes of Peg3.........62

Figure 3.3  Expression level changes using 13-dpc embryos........................................63

Figure 3.4  Expression level changes of H19-related genes.........................................64

Figure 3.5  PEG3 binding to the H19-ICR.................................................................66
Figure 3.6  Quantitative measurement of the immunoprecipitated DNA by polyclonal antibodies against PEG3

Figure 3.7  Individual ChIP experiment confirming the binding of PEG3 to the identified targets by ChIP-Seq.

Figure 3.8  Imprinting test of Igf2 expression and other imprinted genes’ expression level changes after Peg3 knockout.

Figure 3.9  Electromobility shift assay for the DNA-binding sites of PEG3 within the H19-ICR.

Figure 3.10  Mutational analyses of the PEG3 binding site of the H19-ICR.

Figure 3.11  Expression level analyses of H19 using the MEF cells with the restored expression of Peg3.

Figure 3.12  Testing the repressor role of PEG3 through reporter assays.
ABSTRACT

Paternally expressed gene 3 (Peg3) is an imprinted gene that encodes a protein with twelve C2H2 zinc finger domains. Thus, PEG3 protein is predicted to serve as a transcription factor that may regulate other gene expression. Chromatin immunoprecipitation (ChIP) experiment revealed a list of genes that are bound by PEG3, supporting PEG3 is a DNA-binding protein. Genome wide expression analysis using wild type and Peg3 knockout mouse models further suggested that a large number of genes were up-regulated in Peg3 knockout model, including imprinted genes and some tissue-specific genes, suggesting that PEG3 may mainly function as a repressor for its downstream genes.

It is well known that most imprinted genes are organized into clusters (imprinted domains) to share cis-acting elements which can regulate imprinted gene expression. However, in Peg3 imprinted domain, one of my studies suggested that two oppositely imprinted genes in Peg3 imprinted domain could interact through their gene products rather than shared cis regulatory elements. According to the results, paternally expressed Peg3 controls maternally expressed Zim1 as a trans factor. Removing PEG3 resulted in elevated expression of Zim1, thus PEG3 should serve as a repressor for Zim1. ChIP experiment further suggested that PEG3 might repress Zim1 expression through SETDB1/KAP1-driven H3K9me3 mechanism.

Subsequent ChIP-seq analysis using mouse embryonic fibroblast (MEF) cells further identified 16 target genes as PEG3 downstream genes. Interestingly, most of these genes showed high level of expression in oocyte, in which Peg3 is not expressed. Furthermore, qRT-PCR analysis confirmed that PEG3 mainly functions as a repressor for these downstream genes. In
addition, electromobility shift assay (EMSA) and the luciferase reporter assay further demonstrated that PEG3 can directly bind to H19-ICR to repress \( H19 \) expression.

Overall, the research presented in this dissertation advances our understanding of the repression function of PEG3 protein and its potential tumor suppressor function linked to its downstream genes.
1.1 Overview of genomic imprinting

Genomic imprinting is an epigenetic process that results in a small subset of genes showing parental-specific expression patterns (Barlow, 2011). So far, forms of genomic imprinting have been described in fungi, insects, plants and animals (Martienssen and Colot, 2001). In mammals, genomic imprinting is a rare phenomenon, and only a small number of genes are identified as imprinted genes (Barlow and Bartolomei, 2014). Based on our current knowledge, most known imprinted genes are organized into clusters to share a common cis-acting element named imprinting control region (ICR) (Barlow and Bartolomei, 2014; Kim et al., 2012). ICR is a key element that controls imprinted gene expression in a cluster of imprinted genes. One main feature of ICR is that all known ICRs are germline DMRs (differentially methylated regions), which show DNA methylation on one allele but not on the other (Kim et al., 2012). A series of studies also suggested that some ICRs tended to show tandem repeat sequence structure, which contains DNA-binding sites for specific transcriptional factors, such as CTCF and YY1 (Kim, 2008). Subsequent research on ICR deletion further demonstrated that ICR played critical roles to maintain the imprinted gene expression in their respective domains (Bartolomei, 2009; Kim, 2008).

At present, there are less than 200 mouse genes and a small number of human genes shown imprinted gene expression (Barlow, 2011; Weksberg, 2010). Compared to the mammalian gene catalogue, imprinted genes only make up less than 1% of this catalogue. However, imprinted genes are indispensable genes for embryonic growth and development (Ivanova and Kelsey, 2011; Miyoshi et al., 2006). One major supporting research was performed by Solter and
Surani laboratories in the 1980s, where they reconstructed mouse embryos with two maternal (gynogenotes) or two paternal (androgenotes) copies using elegant pronuclear transplantation experiments (McGrath and Solter, 1983; Surani and Barton, 1983; Surani et al., 1984). According to the results, both maternal uniparental embryos (known as gynogenetic embryos) and paternal uniparental embryos (known as androgenetic embryos) failed to survive during embryogenesis, whereas only embryos reconstructed with one paternal and one maternal pronuclei were viable and led to fertile offspring. These outcomes led to a conclusion that mammalian embryonic development required both maternal and paternal contribution to proceed normally. Also, this independent study was the first to suggest that mammalian genome possessed imprinted genes that played critical roles for embryonic development. Human diseases related to genomic imprinting were also reported decades ago, including Prader-Willi syndrome (partial paternal chromosome deletion), Angelman syndrome (partial maternal chromosome deletion) (Butler and Palmer, 1983; Nicholls et al., 1989). In addition to these syndromes, imprinted genes are also linked to certain cancers. For instance, $H19$ and $Igf2$ are two imprinted genes, where only one maternal copy of $H19$ and one paternal copy of $Igf2$ are normally expressed (Steenman et al., 1994). In Wilms’s tumor cells, however, both $H19$ and $Igf2$ typically lose their imprinting, indicating the involvement of imprinted genes in tumorigenesis (Moulton et al., 1994; Steenman et al., 1994). Cell lines derived from Human patients with ovarian and breast cancers also showed notable down-regulation of paternally expressed imprinted gene $Peg3$ (Chen et al., 2011; Dowdy et al., 2005; Feng et al., 2008).

To sum up, even though only a few of genes have been identified as imprinted genes, genomic imprinting plays a critical role in embryonic development and growth. In some cases, loss of imprinted gene expression results in imprinting disease or certain type of cancer.
However, research on genomic imprinting remains relatively small, and more studies are required to explore the molecular mechanism of imprinting gene network and epigenetic codes that are essential for imprinted gene regulation and establishment.

1.2 Epigenetic modifications

Genomic imprinting is an epigenetic phenomenon resulting in a mono-allelic expression pattern. Epigenetics is referred to as the heritable changes in gene activity and expression that occur without alteration in DNA sequence (Goldberg et al., 2007). According to our current knowledge, epigenetics is considered as a bridge between genotype and phenotype (Bernstein et al., 2007; Goldberg et al., 2007). Furthermore, epigenetics research can also explain how identical cells differentiate into different cell types and maintain differentiated cellular states (Jaenisch and Young, 2008). At present, it is known that these non-genetic alternations are tightly regulated by two major epigenetic modifications: DNA methylation and histone modification (Bernstein et al., 2007; Goldberg et al., 2007). Moreover, epigenetic modifications are crucial for DNA accessibility and chromatin structure, thereby regulating gene expression patterns (Jaenisch and Bird, 2003; Kouzarides, 2007; Ozanne and Constancia, 2007).

1.2.1 DNA methylation

One well-characterized epigenetic modification is DNA methylation, which is usually associated with transcriptional silencing and gene regulation (Esteller, 2008; Feinberg and Tycko, 2004). In mammals, DNA methylation occurs at the 5’ position of the cytosine ring within a CpG dinucleotide to create a 5-methylcytocine. According to current results, this modification is mainly catalyzed by three DNA methyltransferases (DNMTs), including DNMT1, DNMT3a, and DNMT3b (Bestor, 2000; Chen and Li, 2004). Subsequent knockout targeting DNMTs in
mouse model resulted in embryonic lethal, indicating the importance of DNA methylation for mammals (Li et al., 1992; Okano et al., 1999). Normally, methylated CpGs were primarily found in heterochromatic region, whereas euchromatic CpGs remained un-methylated. According to our current knowledge, DNA methylation is critical for the control of gene silencing and chromosomal stability (Esteller, 2008; Feinberg and Tycko, 2004; Jones and Baylin, 2007). However, this conclusion has been challenged by a recent study, which illustrated that methylated FoxA2 CpG islands displayed high levels of methylation in expressing tissues but not in non-expression tissue (Bahar Halpern et al., 2014). According to the results, this research group illustrated a novel but unexpected role for DNA methylation in activation of FoxA2 gene expression.

1.2.2 Histone modification

Another well-characterized epigenetic modification is histone modifications, which are recognized as a post-translational modification (PTM). Nowadays, histone modifications are well known to regulate chromatin structure and function, which subsequently affects many DNA-related processes, such as transcription, DNA repair and replication (Consortium et al., 2007; Groth et al., 2007; Koch et al., 2007). It is also known that histone can be both a positive and negative regulator of gene expression based on the histone tail modification differences. So far, histone modifications include methylation, acetylation, phosphorylation, and ubiquitination (Kouzarides, 2007; Ruthenburg et al., 2007). Among these, acetylation and methylation of lysine at histone tails are the two most common PTMs that closely related to my current study.
1.2.2.1 Histone lysine Methylation

Notably, the two most studied methylations of lysine at histone are located at histone H3 lysine 4 (H3K4) and lysine 9 residue (H3K9). Both H3K4 and H3K9 can be mono, di, or tri methylated, and each has slightly different distributions. According to the involved methyl group difference, methylation on H3 lysine 4 or lysine 9 can modulate chromatin structure and regulate gene expression with positive or negative effect.

H3K9me1 is enriched mainly at the transcriptional start site to activate gene expression, whereas H3K9me2 and H3K9me3 are often involved with gene silencing (Barski et al., 2007). For instance, H3K9me2 is the marker of the inactivated X chromosome and it mediates X-inactivation by recruitment of protein CDYL (Escamilla-Del-Arenal et al., 2013). So far, the most studied methylation on H3K9 is H3K9me3. This histone modification marks for heterochromatin, which is the condensed, transcriptionally inactive state of chromatin. It mediates heterochromatin structure mainly through recruitment of heterochromatin protein 1 (HP1), which is responsible for transcriptional repression, and formation and maintenance of heterochromatin. Furthermore, according to these results, HP1 also recruits DNMT3b (DNA methyltransferase), which provides one good examples of the interplay between histone methylation and DNA methylation (Lehnertz et al., 2003).

Current research on methylation of H3K4 residue supports that this modification is closely associated with transcriptional start sites of actively transcribed genes to serve as transcriptional activator (Barski et al., 2007). Nowadays, the most commonly studies methylation on H3K4 is H3K4me3, which can activate transcription by promoting the binding of positive transcription factors and blocking the negative ones. For instance, H3K4me3 can recruit the
chromatin remodeling factors CHD1 and BPTF, which will subsequently open the chromatin structure (Flanagan et al., 2005; Li et al., 2006). Meanwhile, H3K4me3 can prevent the binding of the repressive NuRD and INHAT complexes, which are involved in closing the chromatin states that lead to silencing of gene (Nishioka et al., 2002; Schneider et al., 2004).

1.2.2.2 Histone lysine acetylation

Histone acetylation was first reported by Allfrey and it is one of the most extensively studied histone modifications (Allfrey et al., 1964; Johnson and Turner, 1999). Two families of enzymes, histone acetyltransferases (HATs) and histone deacetylases (HDACs), mediate writing and erasing acetyl groups from histones lysine residues, respectively. HATs can use acetyl CoA as cofactor and catalyze the transfer of an acetyl group to the ε-amino group of lysine side chains (Hodawadekar and Marmorstein, 2007). Attachment of acetyl group results in neutralization of the positive charge on lysine residues, which can further disrupt higher order chromatin folding and facilitate transcription (Garcia-Ramirez et al., 1995; Nightingale et al., 1998; Tse et al., 1998). Thus, acetylation on histone lysine residues is associated with activating gene expression. In contrast to HATs, recruitment of HDACs can erase these acetyl groups and lead to histone deacetylation, which further restores higher order of chromatin structure to achieve gene silencing (Davie and Chadee, 1998).

1.3 Trans-regulatory effects of imprinted genes and imprinted gene network

Since imprinted genes play critical roles in controlling embryonic growth and development, a more systematic search for imprinted genes were made to gain insights into co-regulation of imprinted genes (Gabory et al., 2009; Varrault et al., 2006). Imprinted genes
encode either proteins or long non-coding RNA (lncRNA) that can act in *trans* to regulate many loci across multiple chromosomes, including at imprinted gene loci (Patten et al., 2016).

One example of how imprinted lncRNA affects other imprinted genes is the gene *IPW*, which regulates imprinted gene loci in *trans* through interacting with G9A (histone methyltransferase). *IPW*, a paternally expressed imprinted gene in the PWS imprinted domain of human chromosome 15, which encodes for a lncRNA that has inhibitory effects on the expression of several maternally expressed genes in the imprinted DLK1-KIO3 domain on chromosome 14 (Stelzer et al., 2014). Pluripotent stem cells deficient of *IPW* exhibited elevated expression of the maternally expressed non-coding RNA genes at the DLK1-DIO3 locus on chromosome 14 (Stelzer et al., 2014). In order to explain how the paternally expressed lncRNA *IPW* affected expression of all maternally expressed ncRNAs of this domain, subsequent research studies demonstrated that *IPW* appeared to modulate histone H3 lysine-9 methylation at the ICR of the DLK1-DIO3 domain through recruitment of G9A (histone methyltransferase) (Patten et al., 2016; Stelzer et al., 2014).

Another well-characterized imprinted lncRNA is encoded by the gene *H19*. This gene is expressed only from the maternal allele in mouse and human genome, and was among the first imprinted genes identified more than 20 years ago (Bartolomei et al., 1991; Zhang and Tycko, 1992). Results derived from a *H19* deletion mouse model demonstrated that loss of the imprinted *H19* lncRNA results in an upregulation of a network which has at least six imprinted genes, including paternally expressed genes and maternally expressed genes (Figure 1.1B) (Gabory et al., 2009). Restoration of *H19* expression in this *H19* deletion mouse model resulted in wild-type expression levels of this network, indicating that the lncRNA itself is important to regulate these imprinted genes, rather than its transcription (Gabory et al., 2009; Monnier et al., 2013). Similar
to the regulatory mechanism of IPW, H19 lncRNA achieves its repressive effects through interacting with methyl-CpG-binding domain protein MBD1 to form H19 lncRNA-MBD1 complex, which can further recruit histone H3K9 methyltransferase SETDB1 (Monnier et al., 2013). Taken together, imprinted lncRNAs can function in trans to regulate many other imprinted genes.

Figure 1.1: Roles of H19, Zac1, and Peg3 in the imprinted gene network. (A) The wild type expression level of the imprinted gene network. Maternally expressed genes are shown in red and paternally expressed genes are shown in blue. (B) Deletion of the maternally expressed gene H19 in fetal mouse muscle. Compared with the wild type expression level, enlarged circles show elevated expression and shrunk circles show reduced expression after H19 deletion. (C) Deletion of the paternally expressed gene Zac1 in fetal mouse liver. (D) Deletion of the paternally expressed Peg3 in mouse neonatal brain.
Imprinted genes also use their protein products to achieve trans-regulatory effects. Murine \textit{Zac1} (also called Plagl1), a paternally expressed imprinted gene, encodes a protein that binds to DNA through its zinc-finger motif (Varrault et al., 2006). Studies using \textit{Zac1}-deficient mice demonstrated that ZAC1 protein can promote expression of an imprinted network, including \textit{Igf2}, \textit{H19}, \textit{Cdkn1c} and \textit{Dlk1} (Figure 1.1C). Chromatin immunoprecipitation (ChIP) experiments further suggest that ZAC1 binds to \textit{H19/Igf2} shared enhancers, which results in transactivation of \textit{Igf2} and \textit{H19} promoters. Another imprinted gene using its protein products to achieve trans-regulatory effects is \textit{Peg3}. Murine \textit{Peg3} encodes a protein that binds to DNA through its zinc-finger motifs (Iuchi, 2001; Kuroiwa et al., 1996; Relaix et al., 1996). And numerous PEG3 binding sites in the mouse genome, including at imprinted genes, have been reported by our lab in the past ten years (Thiaville et al., 2013a; Ye et al., 2014). One interesting discovery is that PEG3 binds to its adjacent maternally expressed \textit{Zim1} gene, and \textit{Peg3} knockout mouse models exhibited elevated expression of \textit{Zim1}, indicating that PEG3 protein may function as a repressor to repress \textit{Zim1} expression (Ye et al., 2014). Furthermore, ChIP-seq analysis and genome wide expression analysis (DNA-microarray) suggested that PEG3 protein can directly bind to promoter or zinc-finger regions of imprinted genes to repress these imprinted genes expression (Figure 1.1D). Although the precise repressive mechanism of PEG3 remains unclear, a series of current results suggest that PEG3 protein, as a KRAB-domain-containing zinc-finger protein, seems to achieve its repression function through interacting with another co-factor named KAP1. This co-factor will further recruit histone methyltransferase or histone deacetylase (HDAC) to modulate chromatin structure. Taken together, these results suggest that some imprinted genes using their protein products to regulate other imprinted genes in trans.
In summary, recent studies in mammals reveal that imprinted genes are not solo-existing genes, and they achieve interplay mainly through their trans-acting elements.

1.4 Peg3 imprinted domain

*Peg3* imprinted domain is a 500 kb region located on the human chromosome 19q13.4/proximal and mouse chromosome 7 (Kim et al., 2012). Since most imprinted genes are organized into clusters, as predicted, current knowledge of *Peg3* imprinting domain demonstrates that seven imprinted gene are clustered in this domain. This includes four paternally expressed genes (*Peg3, Usp29, A Peg3, Zfp264*) and three maternally expressed genes (*Zim1, Zim2, Zim3*) (Figure 1.2) (He and Kim, 2014). Based on the presence of zinc finger motifs within their open reading frames, most of these imprinted genes are predicted to be DNA-binding transcription factors (He and Kim, 2014). Interestingly, among mammalian species, all of these imprinted genes except *Zim1* are well conserved (Kim et al., 2007). Within the 500-kb region of the *Peg3* domain, these seven imprinted genes are unevenly distributed as shown in figure 1.2. The middle 250-kb region lacks any obvious open reading frame (ORF), however, about 20 small genomic fragments, distributed in this 250-kb region, show high levels of sequence conservation among multiple placental mammals. These small conserved regions, termed evolutionarily conserved regions (ECRs), are known to be involved in controlling the transcription and imprinting of the *Peg3* domain in cis (Shen et al., 2012; Thiaville et al., 2013b)

DNA methylation analyses further demonstrated that three DMRs were identified in this domain, including DMRs of *Peg3, Zim2* and *Zim3*. However, only Peg3-DMR was a germline DMR, which indicated that *Peg3*-DMR should be an ICR for this imprinted domain (Huang and Kim, 2009; Lucifero et al., 2002). Subsequent mouse KnockOut (KO) model targeting Peg3-
DMR further illustrated that partial deletion of this region resulted in a global impact on the imprinting and transcription of the Peg3 domain, including changes in the imprinting status of Zim2 and Zfp264 and also transcriptional level of the Peg3 domain (Kim et al., 2012). These outcomes further confirm that Peg3-DMR is the ICR of this imprinted domain.

Figure 1.2: The overview of the Peg3 domain. Each imprinted gene in this domain is indicated with an arrow. The genes with red are paternally expressed genes and the genes with blue are maternally expressed genes. Three DMRs are indicated with gray boxes under the genes.

Imprinted genes located within the Peg3 domain also control fetal growth rate and nurturing behavior, which is similar to that of other mammalian imprinted genes (Champagne et al., 2009; Curley et al., 2004). Current studies on Peg3 domain demonstrate that most genes in this domain are found to express in early embryogenesis (Peg3, Zim1) and brain (Peg3, Zim1, Zim2, APEG3), indicating that these genes may play important roles in early embryogenesis and brain development (Kim et al., 1999; Kuroiwa et al., 1996; Li et al., 1999). Our lab’s Peg3 mutation mouse model tend to have two consistent phenotypic effects, milk-suckling defects in pups and various reproductive and behavioral problems in adult mice (Kim et al., 2013). Moreover, Usp29 mutation in bovine caused stillbirth of calves (Flisikowski et al., 2010). Taken together, all of our current results supported that Peg3 domain should be involved in controlling the growth and development of the fetus during the gestational period.

1.5 Overview of zinc finger protein
Zinc finger proteins (ZFPs) are proteins which contain zinc finger domains that play an essential role in the recognition and binding of DNA, protein and RNA (Laity et al., 2001). Many ZFPs contain DNA-binding domains that are included in their zinc finger domains, and also effective domains such as KRAB, SCAN, BTB/POZ, PLAG domains, which are linked to the zinc finger domains.

ZFPs were first discovered in a study of transcription factor TFIIIA in the African clawed frog, *Xenopus laevis* (Miller et al., 1985). Since then, more proteins containing zinc finger motifs have been identified as ZFPs. Based on the number and order of the highly conserved Cys and His residues, which coordinate a zinc ion, generally, zinc fingers are classified into different types—Cys2His2 (C2H2), Cys4(C4), Cys6(C6), Cys2Cys2(C2C2), Cys4His3 (C4H3), Cys3HisCys4 (C3HC4) and so on. Among these, the C2H2 group is by far the best-characterized class of ZFPs as 2% of human genome encodes them (Lander et al., 2001). Typically, a C2H2 motif is comprised of 28-30 amino acid residues, including two conserved cysteine and two conserved histidine residues holding a zinc ion (Krishna et al., 2003). The four conserved residues hold one of the antiparallel β-sheet to one end of the α-helix. The α-helix is known as the recognition helix as it mediates DNA binding (Lee et al., 1989).

In addition to the zinc finger motif, the zinc finger gene family encodes conserved regulatory domains (KRAB domain, SCAN domain, and BTB/POZ domain) for subcellular localization (Sander et al., 2003). To date, we are aware that C2H2 ZFPs with KRAB domain function as transcriptional repressors through recruiting a co-repressor--KRAB-associated Protein-1 (KAP-1) (Peng et al., 2000). C2H2 ZFPs with BTB/POZ domains can recruit another protein to form a complex, which can also repress gene expression (Williams et al., 1995). In contrast to the BTB/POZ and KRAB domains, the isolated SCAN domain is not associated with
either transcriptional activation or repression, but mainly functions as a protein interaction domain (Williams et al., 1999).

Overall, ZFPs are the most abundant proteins in eukaryote. Because of the conserved regulatory domain, such as KRAB domain, SCAN domain and BTB/POZ domain, their functions are extraordinarily diverse, including DNA recognition, transcription regulation and protein/protein interactions.

1.6 KRAB domain containing zinc finger proteins

Since PEG3 protein is predicted as a KRAB domain-containing ZFP, I will mainly describe the function and structure of KRAB-containing ZFPs based on our current knowledge. KRAB domain was first discovered in 1991 when one research group made up approximately one third of the 799 different zinc-finger proteins present in the human genome (Bellefroid et al., 1991). According to their results, they demonstrated that this group of proteins is present in almost one third of the analyzed ZFPs and is also the largest family of transcriptional regulator in mammals. Biochemical analysis of KRAB domain demonstrated that KRAB domain is a 75 amino acid transcriptional repressor module commonly found in eukaryotic ZFPs and is predicted to fold into two amphipathic helices that are involved in protein/protein interaction (Peng et al., 2000). The minimal repression module for this domain is approximately 45 amino acid residues, and substitutions for conserved residues abolish repression (Peng et al., 2000). The KRAB domain was originally identified as a conserved motif at NH2 terminus of ZFPs (Bellefroid et al., 1991). However, subsequent studies suggest that it is a potent, DNA binding-dependent transcriptional repression module (Margolin et al., 1994; Peng et al., 2000). It consists of two sub-regions, A and B, encoded by two separate exons respectively. Based on the structure
of the KRAB domain, KRAB-containing ZFPs can be classified into three subfamilies: the KRAB A subfamily consists of those containing only the classical A box, the KRAB A+B subfamily is comprised of those with both the A and B boxes, and the KRAB A+b subfamily being the class of C2H2 ZFPs having the classical KRAB A box combined with a divergent B box, named “b” (Urrutia, 2003).

Recent studies on protein function suggest that KRAB-containing ZFPs bind to genes in a DNA sequence-specific manner and mainly serve as repressors for the bound gene. To gain insights into KRAB-ZFP-mediated repression mechanism, current research studies demonstrate that KRAB-containing ZFPs repress transcription of the bound gene by interacting with KAP1 which subsequently recruits histone deacetylases (HDACs), SETDB1 (histone H3 lysine 9 methyltransferase) to achieve gene silencing (Urrutia, 2003).

However, repression through histone modification is just part of this repression mechanism, as the KAP1 co-repressors can also directly interact with heterochromatin protein (HP1) to form a quaternary complex among DNA, KRAB-ZFP, KAP1, and HP1, which result in the formation of a heterochromatin-like complex on the target gene, ultimately leading to gene silencing (FIGURE 1.3) (Urrutia, 2003). Besides gene silencing function, KRAB/KAP1 pathway also plays an essential role in DNA methylation events taking placing during the early embryonic period. Research studies on mouse knockout model targeting Zfp57, which encodes a KRAB-containing ZFP, exhibited a failure to establish maternal methylation imprints at specific loci, indicating the involvement of ZFP57 protein in DNA methylation establishment (Li et al., 2008; Li and Leder, 2007).
Taken together, ZFPs containing KRAB domain are the largest family of transcriptional regulator in mammal and they mainly serve as transcriptional repressor to silence gene expression.

![Diagram of KRAB-containing ZFPs](image)

**Figure 1.3:** KRAB-containing ZFPs repress gene expression through interacting with SETDB1 and HP1. The schematic model illustrated that KRAB-containing ZFPs bind to their corresponding DNA sequence through their Zinc finger domain; subsequently, KAP1 forms a scaffold recruiting HP1, SETDB1, and HDACs, silences gene expression by forming a facultative heterochromatin environment on a target promoter. This heterochromatin complex can compact DNA and decrease gene expression.

### 1.7 PEG3 protein function

The paternally expressed gene *Peg3* is the main protein-coding gene located in the middle region of *Peg3* domain. It is predicted to encode a C2H2 Kruppel-type zinc finger protein with the capability of binding to DNA (Iuchi, 2001; Kuroiwa et al., 1996; Relaix et al., 1996). Comparison of amino-acid sequence of KRAB domain in different mammal species, PEG3
protein includes all of the critical amino acid for KRAB domain, which indicates that PEG3 protein is a KRAB-containing zinc finger protein (Figure 1.4).

In murine models, expression of Peg3 is mainly found in the ovary, placenta and hypothalamus (Curley et al., 2004; Kim et al., 1997). Mouse knockout model targeting Peg3 demonstrated that PEG3 protein is responsible for maternal caring behavior, olfactory alteration, lower birth weight, higher percentage of body fat, problems regulating core body temperature, and lower metabolic activity. Furthermore, Peg3 is also known as a potential tumor suppressor gene because the expression level of Peg3 is notably reduced in human patients with ovarian and breast cancer (Chen et al., 2011; Dowdy et al., 2005; Feng et al., 2008).

Figure 1.4: The amino acid sequence of KRAB domain in PEG3 protein. Exon 7 of Peg3 is predicted to express KRAB domain in mouse, human, bovine and horse. Comparison of amino-acide sequence of KRAB domain, the key and highly conserved positions are designated in “*” at the bottom of the figure. And mouse PEG3 homology structure of KRAB domain includes all of the critical points for KRAB domain.

At present, PEG3 is a protein known to be present in both nucleus and cytoplasm (Deng and Wu, 2000). Many studies have explored its cytoplasmic interactions and functions. For instance, PEG3 can interact with SIAH1A to induce P53-mediated cell apoptosis (Relaix et al., 2000; Yamaguchi et al., 2002); PEG3 can also inhibit Wnt signaling by promoting β-catenin
degradation (Jiang et al., 2010); furthermore, PEG3 regulates TNF responsive pathway by activating NFκβ through interacting with TRAF2 (Relaix et al., 1998). Nowadays, our group has put lots of efforts to identify PEG3’s downstream genes, potential binding motif of PEG3, and its potential regulatory mechanism. Based on gel shift assay results and MEME motif analysis, we report the DNA-binding motif of PEG3, 5'-GTGGCAGT-3' (Lee et al., 2015). PEG3 ChIP-seq analysis identified 16 target genes as downstream genes of PEG3, which show oocyte high expression pattern. ChIP-PCR and qRT-PCR analysis demonstrated that PEG3 mainly served as a repressor for these downstream genes (Thiaville et al., 2013a; Ye et al., 2014). Since PEG3 is a KRAB-containing ZFP, we predict that it may repress downstream genes expression through interacting with KAP1 and SETDB1 as we described above. Subsequent H3K9me3 ChIP-PCR results supported that losing PEG3 protein resulted in reduced amount of H3K9me3 repressor marker at PEG3 binding site, indicating its repression function may be mediated by H3K9me3 (Ye et al., 2014).

Taken together, current results on PEG3 protein suggest that PEG3 protein is a DNA-binding ZFP containing KRAB domain. Its protein function is extraordinarily diverse, including gene silencing through H3K9me3 mediated mechanism, tumor suppressor gene in ovary and breast cancer and regulate biological pathway or signaling through protein/protein interaction.

1.8 References


CHAPTER 2
PATERNALLY EXPRESSED PEG3 CONTROLS MATERNALLY EXPRESSED ZIM1
AS A TRANS FACTOR

2.1 Introduction

In mammalian imprinted domains, two genes with opposite imprinting are quite often localized right next to each other, and such examples include maternally and paternally expressed H19/Igf2 and Gtl2/Dlk1. This genomic layout is related to the fact that two genes are usually co-regulated through shared cis elements, such as Imprinting Control Regions (ICRs) (Barlow and Bartolomei, 2014). As such, one gene is very closely associated with the other gene in terms of their transcription levels and allele-specific expression patterns. This has been well demonstrated through a series of mouse mutagenesis experiments. For instance, mutating an endoderm-specific enhancer located in the 3’-side of H19 caused down-regulation for both H19 and Igf2 (Leighton et al., 1995), yet repositioning this enhancer from the downstream region of H19 to an intergenic region between the two genes resulted in the reactivation of the silenced maternal allele of Igf2, causing biallelic expression of Igf2 (Webber et al., 1998). This has been a basis for identifying an enhancer-blocking function of the ICR that is located immediate upstream of H19 (Hark et al., 2000). It is most likely that similar regulatory mechanisms operate in the other imprinted domains with this genomic layout, the juxtaposition of two adjacent genes with opposite imprinting.

A similar genomic layout is also found in the Peg3 domain, which contains 7 imprinted genes: paternally expressed Peg3, Usp29, APeg3, Zfp264 and maternally expressed Zim1, Zim2, Zim3. Among these genes, paternally expressed Peg3 and maternally expressed Zim1 are localized right next to each other, suggesting potential co-regulation of these two genes through shared cis elements. As expected, this domain is regulated through an ICR, termed the Peg3-DMR (Differentially Methylated Region), the 4-kb genomic region surrounding the 1st exons of Peg3 and Usp29 (Kim et al., 2003). Deleting part of this ICR, the 2.5-kb genomic region harboring multiple YY1 binding sites, caused global changes in the expression levels and imprinting status of several genes within this domain (Kim et al., 2012). In particular, the expression levels of Peg3 and Zim1 were affected in a dosage-dependent manner: 4-fold down-regulation of Peg3 coinciding with 4-fold up-regulation of Zim1. Interestingly, the observed up-regulation of Zim1 was still derived from the maternal allele although the mutation causing down-regulation of Peg3 was on the paternal allele (Kim et al., 2012). This trans-allelic outcome by a mutation has not been observed before, and thus suggests the presence of different regulatory mechanisms involving possible trans factors rather than the known mechanisms involving shared cis-regulatory elements as seen from the H19/1gf2 pair.

According to recent studies, Peg3 encodes a DNA-binding protein with transcriptional repression function (Thiaville et al., 2013). Given the observed tight correlation between Peg3 and Zim1, it is possible that Peg3 may control directly the transcription of Zim1 as a trans factor. In this case, the absence or reduced protein levels of PEG3 might be responsible for the observed up-regulation of Zim1. To further test this possibility, the Peg3/Zim1 pair was analyzed using a new mutant model targeting Peg3. In this new model, the mutation truncates the transcription of
Peg3, thus removing the PEG3 protein (Kim et al., 2013). The results revealed that the removal of PEG3 through the mutation on the paternal allele caused up-regulation of Zim1 without disrupting its maternal-specific expression. PEG3 also binds to the Zim1 locus as a trans factor, yet this binding by PEG3 is closely associated with the histone modification mark H3K9me3, suggesting a potential repression mechanism for PEG3. More detailed results have been described in the following sections.

2.2 Materials and methods

Ethics Statement

All the experiments related to mice were performed in accordance with National Institutes of Health guidelines for care and use of animals, and also approved by the Louisiana State University Institutional Animal Care and Use Committee (IACUC), protocol #13-061.

Derivation of MEF (Mouse Embryonic Fibroblast) cells

Two litters of 14.5-dpc embryos were harvested through timed mating of the male and female mutant mice heterozygous for the KO allele with the female and male wild-type littermates, respectively. The mutant allele of Peg3 used for the current study has been previously reported and characterized in detail (Kim et al., 2013). The head portion and the red tissues were removed from the embryos, and the remaining portions were minced with razor blades. These minced tissues were transferred to a 15 mL conical tubes containing 1 mL trypsin (Invitrogen, Cat. No. 25300062). After 5 min incubation at 37°C, the cells were harvested with centrifugation, and later resuspended in 15 mL media (Life technologies, Cat. No.10566024). Finally, the resuspended cells were plated onto a T-75 flask. MEF from each embryo was first genotyped using the following primer set: Peg3-for (5’-ATGAGTCTCGATCCCAGGTATGCC-
3’) and LoxR (5’-TGAACTGATGGCGAGCTCAGACC-3’). The gender of each MEF was also determined using the following primer set: mSry-F (5’-GTCCCGTGAGTAGGACACAAG-3’) and mSry-R (5’-GCAGCTCTACTCCAGTCTTGCC-3’).

**Chromatin ImmunoPrecipitation (ChIP) analyses**

Chromatins were prepared from two different types of samples, MEF and neonatal brains, according to the method previously described (Thiaville et al., 2013). In brief, the homogenized samples were first cross-linked with 1% formaldehyde for 20 mins, and then lysed with the buffer containing protease inhibitor cocktail (Millipore, Cat. No. 539131). The released nuclei were fractionated with sonication to a pool of DNA fragments size-ranging from 300 to 1,000 bp in length. The prepared chromatin was immunoprecipitated with the following two antibodies: PEG3 (Abcam, Cat. No. ab99252) and H3K9me3 (Abcam, Cat. No. ab8898). Each immunoprecipitated DNA was dissolved in 100 μl of TE for either PCR or qPCR analyses.

**Transfection experiments**

MEF cells were transfected with the following two constructs, GFP (pIRES-puro-GFP) and FLP (pIRES-puro-FLP), using the GenJect transfection reagent (Cat. No. SL100489-MEF) according to the manufacturer’s protocol. Transfection efficiency was monitored through GFP expression under a fluorescence microscope after 24 hours. The transfected cells were harvested at 72-hour post transfection for RNA and DNA isolation. The reverted allele of Peg3 by FLP was detected through PCR with the following primer set: Flpko-F (5’-CCCTCAGCAGAGCTTCTGACC-3’) and Flpko-R (5’-AAGCTACCTGGGAAATGAGTGTGG-3’).

**RNA isolation and qRT-PCR analyses**
Total RNA was isolated from either MEF or the brains of one-day-old neonates using a commercial kit (Trizol, Invitrogen) according to the manufacturer’s protocol. The total RNA was then reverse-transcribed using the M-MLV kit (Invitrogen), and the subsequent cDNA was used as a template for quantitative PCR. The qRT-PCR analysis was performed with SYBR Select Master Mix (Applied Biosystems, Life Technologies) using the iCycler iQTM multicolor real-time detection system (Bio-Rad). All qRT-PCR reactions were carried out for 40 cycles under standard PCR conditions with internal controls (28S and β-actin). The results derived from qRT-PCR were further analyzed using the threshold (Ct) value. The ΔCt value was initially calculated by subtracting Ct value of a testing replicate of a given gene from the average Ct value of the internal control (28S and β-actin). The fold difference for each replicate was then calculated by raising the ΔΔCt value as a power of 2 (Winer et al., 1999). The average and standard deviation for each sample were then calculated by compiling the normalized values. The information regarding individual primer sequences is also available (Table 2.1).

Table 2.1 Primer sequences used in this study

<table>
<thead>
<tr>
<th>ChIP Primers;</th>
<th>Region 1</th>
<th>Region 2</th>
<th>Region 3</th>
<th>Region 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fp------------</td>
<td>CTCGTCCTCGCGCTTGGTTAT</td>
<td>TCCCCGAGACACATTCCAGCTTCT</td>
<td>CATCACAGGCTCCAGCACATTCA</td>
<td>GGGTGTTGGTGGTACATGCTTTA</td>
</tr>
<tr>
<td>RP------------</td>
<td>CCTAGAGCTGGGGACTTGTTT</td>
<td>GACGCTGTTGACGGCATCCTAA</td>
<td>CCCTTCATGGGGTGACTCCAAG</td>
<td>GACCAGGCTGTCTTTGAACTCTCA</td>
</tr>
</tbody>
</table>
(Table 2.1 continued)

<table>
<thead>
<tr>
<th>Genomic primers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zim1 F1/F2</td>
</tr>
<tr>
<td>FP----GGATTCACACGGGAGTGAGA</td>
</tr>
<tr>
<td>RP----CTTGCAACCGGTACCTGGAGT</td>
</tr>
<tr>
<td>Zim1 a/b</td>
</tr>
<tr>
<td>FP----CCCAGACCACTGTCTCTCTACAGAG</td>
</tr>
<tr>
<td>RP----CCTCAGCGTGTACATACATAATC</td>
</tr>
<tr>
<td>PGM2l1</td>
</tr>
<tr>
<td>FP----ACCAAGCCCTGAGGCGAAGA</td>
</tr>
<tr>
<td>RP----TGCCGTTTCCTCAGCTGTCTTTTC</td>
</tr>
<tr>
<td>H19</td>
</tr>
<tr>
<td>FP----GTGTGGCTCACTATAGGAAGGCA</td>
</tr>
<tr>
<td>RP----CCTATGCGCGTCTGCGAGCAAA</td>
</tr>
<tr>
<td>FlpKO</td>
</tr>
<tr>
<td>FP----CCCTCAGCAGCGCTGTTCTCTGCC</td>
</tr>
<tr>
<td>RP----AAGCTACCTGGAAATGAGTGTGG</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RT-Primers;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zim1</td>
</tr>
<tr>
<td>FP----GATCACCAGGTTGGAGCAAGGAGT</td>
</tr>
<tr>
<td>RP----AGCGCTCTGTGTTGTGTAGTTG</td>
</tr>
<tr>
<td>28s</td>
</tr>
<tr>
<td>FP----TTGAAAAATCCCGGGGGAGAG</td>
</tr>
<tr>
<td>RP----ACATTGTTCACACATGCCAG</td>
</tr>
<tr>
<td>Beat-actin</td>
</tr>
<tr>
<td>FP----GAGCACCCTGTGTGCCACAGGA</td>
</tr>
<tr>
<td>RP----CTCTTTGTGTACCGACGATTTTC</td>
</tr>
<tr>
<td>Peg3 3-6</td>
</tr>
<tr>
<td>FP----ATCCCTGAAACGCTCAAGCCCT</td>
</tr>
<tr>
<td>RP----AAGATCCCGTTGAGGCAGCC</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
2.3 Results

2.3.1 Removal of the PEG3 protein results in the up-regulation of Zim1

According to the results from the previous study, deletion of part of the Peg3-DMR derived a concurrent 4-fold down and up-regulation of Peg3 and Zim1, respectively (Kim et al., 2012). However, it is currently unknown the causal relationship between the observed down and up-regulation of the two genes since the mutation also caused other changes within the Peg3 domain.

To further clarify the observed effects on the Peg3/Zim1 pair, a new mutant model targeting Peg3 was used for the current study (Figure 2.1 A). This model was originally constructed with a combinatory scheme of knock-in/knock-out, thus will be referred to as a KO model hereafter for the simplicity.

In this model, the mutant allele carries an expression cassette containing two ORFs (Open Reading Frames) within its 5th intron of Peg3: the promoterless β-galactosidase gene and the neomycin resistance gene driven by the human β-actin promoter (Kim et al., 2013). Because of the two poly(A) signals within the cassette, the mutant allele truncates the transcription of Peg3, thus removing the PEG3 protein (Kim et al., 2013).

Subsequent global gene expression analyses revealed that a large number of genes were affected by the removal of PEG3 protein in both the embryos and placentas of 14.5 d.p.c. (days post coitum) (Kim et al., 2012).
As expected, *Zim1* was also found to be up-regulated in this survey: 2-fold up-regulation in both tissues, which is consistent with the results from the mutant allele deleting part of the Peg3-DMR (Kim et al., 2012).

Figure 2.1 Removal of the PEG3 protein results in the up-regulation of *Zim1*. (A) Schematic representation of the genomic structure of paternally expressed *Peg3* and maternally expressed *Zim1*. In the mutant allele, a 7.1-kb cassette containing a promoterless β-galactosidase (β-Gal) and human β-actin promoter-driven neomycin resistant gene (NeoR) has been inserted between exon 5 and 6 of *Peg3*. The inserted cassette is flanked by two FRT sites (open ovals), thus can be removed through FLP-mediated recombination. (B) Expression analyses of *Zim1* using a set of female mouse embryonic fibroblast (MEF) cells that had been prepared through breeding female and male heterozygotes with their wild-type littermates. The MEF cells with the wild-type and the heterozygote with the paternal transmission of the mutant allele were used for qRT-PCR. (C) Imprinting tests of *Zim1* using the neonatal brains of the F1 hybrid derived from the crossing of a male heterozygote C57BL/6J (B6) and a female PWD/PhJ (PWD) breeder. The RT-PCR products from *Zim1* were digested with *Dra*I, showing two parental patterns (lane 1 and 2) as well as the maternal-specific expression pattern from the two neonates with WT and KO (lane 3 and 4), respectively. Schematic representation of this imprinting test was shown above the gel picture.

To further follow-up this initial result, we have derived a set of mouse embryonic fibroblast (MEF) cells from the mutant animals. Two litters of 14.5-dpc embryos were prepared
through timed mating of the female and male heterozygotes for the mutant allele with male and female wild-type littersmates, respectively. The first litter inheriting the mutant allele maternally does not have any mutational effects on Peg3 since Peg3 is expressed mainly from the paternal allele. In contrast, the second litter inheriting the mutant allele paternally has an effect on Peg3, removing the PEG3 protein completely (Kim et al., 2013). After genotyping and gender determination, each MEF line from a given embryo was individually cultured, and subsequently used for isolating total RNA for RT-PCR analyses. According to the results from the female set (Figure 2.1 B), the expression levels of Zim1 were 2.5-fold greater in the paternal heterozygote cell (+/-p) than the wild type cell (+/+). Since this mutant allele disrupts the transcription and translation of Peg3 only, the observed up-regulation of Zim1 is most likely an outcome of the removal of PEG3. We repeated this series of expression analyses using another set of MEFs, which also showed a consistent up-regulation of Zim1. The allele-specific expression of the observed up-regulation of Zim1 was also tested using the F1 hybrid animals obtained from the inter-specific crossing of C57BL/6J and PWD/PhJ (Figure 2.1 C). The results from the total RNA of the neonatal brain indicated that the up-regulated expression of Zim1 was still derived from the maternal allele. Overall, this series of expression analyses derived a consistent outcome as seen from the other mutant allele, the up-regulation of Zim1 on the maternal allele coinciding with the mutation on the paternal allele of Peg3. Since both mutant alleles share one common feature, the reduced protein levels of PEG3, the observed up-regulation of Zim1 is most likely caused by the changed protein levels of PEG3 in both mutant models.

2.3.2 PEG3 binds to the zinc finger exon of Zim1

Since Peg3 encodes a DNA-binding protein with repression activity, we tested the possibility that the protein PEG3 may control directly the transcription of Zim1 as a trans factor.
As an initial step, the binding of PEG3 to the Zim1 locus was investigated using Chromatin ImmunoPrecipitation (ChIP) technique (Figure 2.2). For this series of ChIP experiments, 4 genomic regions were selected to scan the entire locus of mouse Zim1: Region 1 (promoter), 2 (intron), 3 (zinc finger exon) and 4 (3’-UTR) (Figure 2.2 A).

The primer set amplifying the promoter of Pgm2l1 (phosphoglucomutase 2-like 1) was also included as a positive control since this region has been shown to be a target locus of the PEG3 protein (Thiaville et al., 2013). Three different sets of cross-linked chromatins were immunoprecipitated with anti-PEG3 polyclonal antibodies. First, the immunoprecipitated DNA from the two MEF cells representing wild-type (WT) and mutant (KO) cells were individually analyzed with PCR (Figure 2.2 B).

As expected, the enrichment at the Pgm2l1 locus was detected only in the WT cells but not in the KO cells lacking PEG3, confirming the binding of PEG3 to this locus and also the specificity of our ChIP experimental system. The same set of immunoprecipitated DNA was further tested using the 4 primer sets of Zim1.

Although the two sets, Region 2 and 4, detected some levels of the enrichment, the detections were not specific to the WT cells, indicating non-specific binding of the anti-PEG3 antibody to other unknown proteins. On the other hand, the promoter region of Zim1 (Region 1) did not show any level of the enrichment, indicating no binding of PEG3 to this region.

In contrast, the zinc finger exon (Region 3) showed much higher levels of the enrichment in the WT cells than in the KO cells, indicating that the observed enrichment likely represents the genuine binding of PEG3 to this region.
This has been further confirmed through a set of independent ChIP experiments with a custom-made antibody against PEG3 (Figure 2.3).

Figure 2.2 PEG3 binds to the zinc finger exon of Zim1. (A) Genomic structure of Zim1 and the relative positions of the primer sets used for ChIP experiments. The 8 exons of Zim1 are indicated with vertical lines and boxes. (B) PEG3-ChIP experiments using the two sets of chromatin prepared from WT and KO (+/-p) MEF cells. The DNA from Inputs, Negative controls (Neg), and Immunoprecipitates with anti-PEG3 antibody (PEG3 IP) was used for PCR amplification. This series of ChIP experiments also included another locus, Pgm2l1, as a positive control since this locus is known to be bound by PEG3.

Another set of ChIP experiments was also performed using the second set of chromatin prepared from the brains of the WT and KO (+/-p) neonates (Figure 2.4 B). According to the results, the enrichment at the promoter of Pgm2l1 was also detected higher levels in the WT than in the KO sample, confirming again the specificity of the anti-PEG3 antibody and the ChIP system. A similar pattern of WT-specific enrichment was observed in both Region 1 and 3,
indicating the potential binding of PEG3 to these regions. However, the enrichment at Region 3 is much more obvious than Region 1. Also, the observed enrichment at Region 3 is consistent with the results from the MEF cells, thus confirming the genuine binding of PEG3 to this region in the neonatal brain. Finally, the third set of ChIP experiments were performed using the chromatin prepared from the F1 hybrid between C57BL/6J and PWD/PhJ (Figure 2.4 C). Restriction enzyme digestion scheme was employed to differentiate the two alleles of the immunoprecipitated DNA.

Figure 2.3 PEG3 binds to the zinc finger exon of Zim1 using custom-made anti-PEG3 antibody. (A) PEG3-ChIP experiments using the two sets of chromatins prepared from WT and KO (+/-p) MEF cells. The DNA from Inputs, Negative controls (Neg), and Immunoprecipitates with anti-PEG3 antibody (PEG3 IP) was used for PCR amplification. This series of ChIP experiments also included another locus, Pgm2l1, as a positive control since this locus is known to be bound by PEG3. (B) The relative enrichment of the ChIP samples in different regions were also quantified by qPCR. (C) The quality of the custom-made anti-PEG3 antibody was tested through western blot.

According to the results, the immunoprecipitated DNA at Region 1 and 3 both were derived equally from the two parental alleles, indicating that PEG3 likely binds to both alleles of these two regions. Taken together, this series of ChIP experiments concluded that PEG3 binds to
the zinc finger exon of Zim1 on both alleles, further supporting the possibility that PEG3 may control directly the transcription of Zim1 as a trans factor.

2.3.3 Reduced levels of H3K9me3 in the mutant cells lacking PEG3

The mouse locus of Zim1 was carefully examined using the genome browser (genome.ucsc.edu) to obtain hints regarding the potential functional outcomes of the observed
PEG3 binding. Histone modification levels of H3K4me3 and H3K9me3 derived from ES cells and whole brain tissues are presented as Figure 2.5 A. This examination revealed that the zinc finger exon of Zim1 is marked with the histone modification H3K9me3 (Figure 2.5 A). Although this histone mark is relatively rare in gene-rich euchromatic regions, it is well known that this modification is quite often associated with the genomic regions encoding the Kruppel-type zinc finger gene family (Frietze et al., 2010).

Figure 2.5 Reduced levels of H3K9me3 in the mutant cells lacking PEG3. (A) Histone modification profiles on the Zim1 locus. The histone modification profiles of H3K4me3 and H3K9me3 derived from ES (upper) and whole brains (lower) are presented along with the exon structure of Zim1. (B) H3K9me3-ChIP using the two sets of chromatins prepared from WT and KO (+/-p) MEF cells. The DNA from Inputs, Negative controls (Neg), and Immunoprecipitates with anti-H3K9me3 antibody (H3K9me3 IP) was used for PCR amplification. This series of ChIP experiments included another locus, the imprinting control region of H19, as a positive control since the paternal allele of this ICR is known to be modified with H3K9me3. (C) The reduced levels of H3K9me3 on Zim1 were further analyzed with qPCR. Shown are the relative values of the immunoprecipitated DNA to the negative controls derived from MEF cells. Region 1 does not show any difference whereas Region 3 shows reduced levels of H3K9me3 in KO compared to those from WT.
Zim1 is a member of this gene family (Kim et al., 1999), thus the detection of this histone mark at the Zim1 locus is consistent with the pattern observed from previous studies. This modification at Zim1 is most obvious in ES cells but some levels of this modification are also detected in other cell types, including adult whole brain.

To further test a potential connection between PEG3 binding and H3K9me3, the modification levels of this repression signal were compared between the WT and KO cells. According to the results (Figure 2.5 B), the modification levels of H3K9me3 at Region 3 were 2 fold lower in the KO cells than those in the WT cells although the modification levels at other genomic regions, such as the ICR of H19/Igf2, were similar between the two MEF cells (Figure 2.5 B&C). This indicates that the observed reduction of H3K9me3 in the KO cells is specific to the zinc finger exon of Zim1. Since H3K9me3 is known to be a repression mark for transcription, the reduced levels of H3K9me3 are also consistent with the observed up-regulation of Zim1 in the KO cells lacking PEG3. Overall, the genomic region of Zim1 bound by PEG3 is closely associated with H3K9me3, further suggesting that PEG3 might repress Zim1 through H3K9me3.

2.3.4 Restoring the protein levels of PEG3 down-regulates Zim1

To further test PEG3’s repressor role in the transcription of Zim1, we performed the following set of in vitro transfection experiments (Figure 2.6). In the MEF KO (+/-p) cells, the transcription of Peg3 is disrupted by the inserted cassette, which is flanked by two FRT (Flippase Recombination Target) sites. The vector construct expressing Flippase (FLP) was transiently transfected into the KO cells to remove the cassette, restoring the transcription and translation of Peg3. This pool of cells was used for measuring the expression levels of Zim1 along with a set of
control cells: the cells transfected with no vector (Mock) and a Green Florescent Protein vector (GFP) (Figure 2.6). As shown in figure 2.6 B, the transient expression of FLP indeed removed the inserted cassette based on the detection of a genomic fragment without the inserted cassette. This removal of the cassette subsequently restored the expression of Peg3 based on RT-PCR. In these cells with the restored PEG3, the transcriptional levels of Zim1 was 2.5 and 1.5-fold reduced as compared to those observed from the two control cells (Figure 2.6 C). It is also prudent to mention that the transcriptional levels of Zim1 were further reduced in the set transfected with FLP than in the set with GFP.

Figure 2.6 Restoring the protein levels of PEG3 down-regulates Zim1. (A) Genomic structure of the mutant allele of Peg3 and FLP-mediated recombination scheme to restore the expression of Peg3. The inserted cassette is flanked by two FRT sites, thus can be removed by Flippase (FLP). (B) Three pools of KO MEF cells were transfected with the following constructs: no vector as a mock control (lane 1), a Green Fluorescent Protein (GFP) expression vector as a negative control (lane 2), and a FLP expression vector (lane 3). The total RNA isolated from these cells were analyzed with RT-PCR to measure the expression levels of β-actin, Zim1, and Peg3. The bottom panel shows genotyping results confirming FLP-mediated recombination of the mutant allele (Rev KO) and endogenous allele (WT) of Peg3. (C) The observed down-regulation of Zim1 was further analyzed using qRT-PCR.
These results demonstrated that the restored expression of \textit{Peg3} is responsible for the down-regulation of \textit{Zim1}, confirming again the inverse correlation between \textit{Peg3} and \textit{Zim1}. Down-regulation of \textit{Zim1} was further tested with the over-expression of \textit{Peg3} (Figure 2.7 A). The results confirmed the down-regulation of \textit{Zim1} (Figure 2.7 B). Overall, this series of transfection experiments confirmed that PEG3 functions as a repressor for the transcription of \textit{Zim1}.

Figure 2.7 Overexpression of the protein levels of PEG3 down-regulates \textit{Zim1} (A) Three pools of KO MEF cells were transfected with the following constructs: no vector as a mock control (lane 1), a Green Fluorescent Protein (GFP) expression vector as a negative control (lane 3), and a \textit{Peg3} expression vector (lane 2). The total RNA isolated from these cells were analyzed with RT-PCR to measure the expression levels of \textit{β-actin}, \textit{Zim1}, and \textit{Peg3}. (B) The observed down-regulation of \textit{Zim1} was further analyzed using qRT-PCR.
2.4 Discussion

In the current study, we tested the possibility that paternally expressed Peg3 may control the transcription of maternally expressed Zim1 as a trans factor. According to the results, the reduced protein levels of PEG3 is indeed responsible for the up-regulation of Zim1. The PEG3 protein also binds to the zinc finger exon of Zim1 that is marked with the repression mark H3K9me3. Furthermore, the KO cells lacking PEG3 have the reduced levels of H3K9me3 at the zinc finger exon of Zim1, suggesting that PEG3 might control Zim1 through H3K9me3. In vitro transfection experiments further demonstrated that reintroducing the PEG3 protein into the KO cells restores the down-regulation of Zim1. Overall, these results confirm that paternally expressed Peg3 controls maternally expressed Zim1 as a trans factor.

The imprinted gene pair of Peg3/Zim1 is unique based on the following reasons. As seen in the other pairs of oppositely imprinted genes, the transcriptional level of Peg3 is also inversely correlated with those of Zim1. However, this inverse correlation does not involve any change in their allele-specific expression pattern, which is quite different from the other pairs of oppositely imprinted genes (Barlow and Bartolomei, 2014). This suggests the involvement of some unknown trans factors in the inverse correlation between Peg3 and Zim1. The results form the current study further indicated that the protein encoded by Peg3 itself is likely this unknown trans factor. According to previous studies, Peg3 encodes a DNA-binding protein with repression activity (Thiaville et al., 2013). Also, there is a very tight inverse correlation between Peg3 and Zim1: 4-fold down regulation of Peg3 coinciding with 4-fold up-regulation of Zim1 (Kim et al., 2012). This line of evidence supports the idea that Peg3 controls the transcription of Zim1 as a trans factor. Nevertheless, the results from the current study also confirm another important aspect of Zim1 that PEG3 is unlikely involved in regulating the allele-specific expression of
since the removal of PEG3 does not have any effect on the maternal expression of Zim1. The observed PEG3-mediated regulation of the transcriptional levels of Zim1 appears to be separate from some unknown mechanisms controlling the maternal-specific expression of Zim1. In summary, the transcription of Zim1 is regulated through two separate mechanisms: one controlling the allele-specific expression and the other controlling the expression level through PEG3 (Figure 2.8).

Figure 2.8 Paternally expressed Peg3 controls maternally expressed Zim1 as a trans factor involving H3K9me3. Schematic representation for Peg3’s functional role in transcriptional control of Zim1. The gene product of paternally expressed Peg3 binds to the zinc finger exon of maternally expressed Zim1 on both alleles, resulting in transcriptional repression through H3K9me3. The protein PEG3 may interact with some unknown proteins, such as KAP-1, to recruit SETDB1 for the H3K9me3 modification on the Zim1 locus. This role of Peg3 is independent of the maternal-specific expression of Zim1, thus the observed up-regulation of Zim1 is still derived from the maternal allele of the Peg3 mutant animals.

The binding of PEG3 to the zinc finger exon of Zim1 is consistent with several known facts about the evolutionary origin of Peg3 as well as the repression mark H3K9me3. First, Peg3 is localized in the middle of a Cys2His2-type zinc finger gene cluster (Kim et al., 1997), yet Peg3 itself encodes a DNA-binding protein with C2H2 zinc finger motifs (Thiaville et al., 2013).
This suggests that Peg3 may have originated from this type of zinc finger genes (ZNFs). The C2H2-type ZNFs are known to interact the H3K9 methylase SETDB1 via the co-repressor protein KAP-1 (Schultz et al., 2002). Thus, the binding of PEG3 to the genomic region with the H3K9me3 modification makes sense given the evolutionary origin of Peg3, and further implies that PEG3 might still recruit SETDB1, possibly through the interaction with KAP-1 (Figure 2.8). Second, it is well known that ZNFs, such as Zim1, are usually regulated through H3K9me3 (Frietze et al., 2010). This is further supported by the fact that the zinc finger-coding region of Zim1 is indeed modified by this repression mark in ES and other somatic cells. Yet, the KO cells lacking PEG3 have the reduced levels of H3K9me3. This further suggests that PEG3 likely controls the transcription of Zim1 through H3K9me3. According to the results from the mutant mouse model targeting Peg3, many placenta-specific gene families are also up-regulated in the brains of KO mice, yet all of these gene families are known to be regulated through similar repression mechanisms involving H3K9me3 (Kim et al., 2013). Therefore, it is most likely that PEG3’s regulation on Zim1 may be also through H3K9me3.

2.5 Reference


CHAPTER 3
PEG3 BINDS TO H19-ICR AS A TRANSCRIPTIONAL REPRESSOR

3.1 Introduction

Genomic imprinting is an epigenetic process by which one allele of autosomal genes is repressed based on their parental origin (Barlow, 2011). At present, less than 200 genes have been identified as imprinted genes, making up about 1% of the mammalian gene catalogue (Barlow and Bartolomei, 2014). Imprinted genes are usually clustered in specific chromosomal regions, and these imprinted domains are regulated by cis regulatory elements, such as imprinting control regions (ICRs) (Edwards and Ferguson-Smith, 2007; Wan and Bartolomei, 2008). Imprinted genes encode either proteins or non-coding RNA (ncRNA), the majority of which have been shown to play critical roles in controlling embryonic growth and development (Ivanova and Kelsey, 2011; Miyoshi et al., 2006). Some imprinted genes are also known to encode DNA-binding proteins, such as Zac1 and Peg3 (Thiaville et al., 2013; Varrault et al., 2006). According to the results, the DNA-binding protein encoded by Zac1 binds to the 3’-side enhancer of H19 as a transcriptional activator, suggesting a functional connection between the two imprinted genes, paternally expressed Zac1 and maternally expressed H19 (Varrault et al., 2006). Detailed surveys further revealed that many imprinted genes behave similarly in response to environmental and developmental cues (Varrault et al., 2006). Thus, the similar expression responses shared among individual imprinted genes have been a basis for the imprinting network model, in which imprinted genes are connected to each other and co-regulated for the common biological outcomes (Varrault et al., 2006).
Peg3 (paternally expressed gene 3) is another imprinted gene that encodes a DNA-binding protein (Thiaville et al., 2013). Peg3 is also a member of an evolutionarily conserved imprinted domain located in human chromosome 19q13.4/proximal mouse chromosome 7 (Kim et al., 1997; Kuroiwa et al., 1996). This domain is located in the middle of C2H2 Kruppel-type zinc finger gene families (Kim et al., 1997). In fact, Peg3 itself is predicted to encode a protein with twelve C2H2 Kruppel-type zinc fingers (Kim et al., 1997; Kuroiwa et al., 1996). Recent studies confirmed that the protein PEG3 indeed binds to a large number of genomic targets as a DNA-binding protein (Thiaville et al., 2013). According to the results, PEG3 functions as a transcriptional repressor for these downstream genes (Thiaville et al., 2013). In particular, PEG3 binds to its adjacent imprinted gene, maternally expressed Zim1. Detailed analyses indicated that PEG3 functions as a transcriptional repressor for Zim1 possibly through H3K9me3 (trimethylation on lysine 9 of histone H3)-mediated mechanisms (Ye et al., 2014). This possibility has been further supported by the observation that many placenta-specific gene families associated with H3K9me3 modification are de-repressed in the mutant embryos lacking PEG3 (Kim et al., 2013). Although premature at the moment, a series of these recent studies suggest that PEG3 is a transcriptional repressor for its downstream genes through H3K9me3 mediated mechanisms (Kim et al., 2013; Ye et al., 2014).

As an ongoing effort, the current study performed another series of ChIP-seq experiments using mouse embryonic fibroblast (MEF) cells. This series of analyses identified a set of 16 potential downstream genes, and the majority of these genes appear to share oocyte-specific expression patterns. Interestingly, one particular target happens to be located within the imprinting control region of H19. Detailed analyses confirmed the binding of PEG3 to the H19-ICR, and also expression and in vitro analyses further suggest that PEG3 may function as a
transcriptional repressor for the maternal allele of the H19-ICR. More detailed results have been described in the following section.

3.2 Materials and Methods

Ethics Statement

All the experiments related to mice were performed in accordance with National Institutes of Health guidelines for care and use of animals, and also approved by the Louisiana State University Institutional Animal Care and Use Committee (IACUC), protocol #13-061.

Derivation of MEF (Mouse Embryonic Fibroblast) cells

Two litters of 14.5-dpc embryos were harvested through timed mating of the male mice heterozygous for the KO allele with the female wild-type littermates. The mutant allele of Peg3 used for the current study has been previously reported and characterized in detail (Kim et al., 2013). The head portion and the red tissues were removed from the embryos, and the remaining portions were minced with razor blades. These minced tissues were transferred to a 15 mL conical tubes containing 1 mL trypsin (Invitrogen, Cat. No. 25300062). After 5 min incubation at 37°C, the cells were harvested with centrifugation, and later resuspended in 15 mL media (Life technologies, Cat. No.10566024). Finally, the resuspended cells were plated onto a T-75 flask. The MEF from each embryo was first genotyped using the following primer set: Peg3-for (5’-ATGAGTCTCGATCCAGGTATGCC-3’) and LoxR (5’-TGAACTGATGGCGAGCTCAGACC-3’). Gender of each MEF was also determined using the following primer set: mSry-F (5’-GTCCCGTGGTGAGGCACAAG-3’) and mSry-R (5’-GCAGCTCTACTCCAGTCTTGCC-3’).

Chromatin ImmunoPrecipitation (ChIP) and ChIP-seq analyses
Chromatins were prepared from two different types of samples, MEF and neonatal brains, according to the method previously described (Thiaville et al., 2013). In brief, the homogenized samples were first cross-linked with 1% formaldehyde for 20 mins, and then lysed with the buffer containing protease inhibitor cocktail (Millipore, Cat. No. 539131). The released nuclei were fractionated with sonication to derive a pool of DNA fragments size-ranging from 300 to 500 bp in length. The prepared chromatin was immunoprecipitated with a commercial anti-PEG3 antibody (Abcam, Cat. No. ab99252). The immunoprecipitated DNA was dissolved in 100 µl of TE for PCR analyses. For ChIP-seq analysis, two pools of MEF cells (WT and KO) were individually immunoprecipitated with the anti-PEG3 antibody. The two immunoprecipitated DNA along with the two corresponding input DNA were used for constructing libraries for ChIP-seq experiments according to the manufacturers’ protocol (Illumina FC4014003). The raw sequence reads derived from these four libraries, on average 35 millions read per sample, were mapped to the mouse reference genome sequence (mm9) using Bowtie2 (Langmead et al., 2009). The sam files from the mapping were first converted into bed files using Samtools, and later the bed files were used for predicting peaks using MACS2 (Zhang et al., 2008). The final outputs from MACS2 describing ChIP-seq peaks are available (Table 3.1 and Table 3.2).

Transfection experiments

MEF cells were transfected with the following two constructs, GFP (pIRES-puro-GFP) and FLP (pIRES-puro-FLP), using the GenJet transfection reagent (Cat. No. SL100489-MEF). Transfection efficiency was also monitored through GFP expression under a fluorescence microscope at 24-hour post transfection. The transfected cells were harvested at 72-hour post transfection for RNA and DNA isolation. The reverted allele of Peg3 by FLP was detected through PCR with the following primer set: Flpko-F and Flpko-R.
Table 3.1 Final output of MACS2 describ

<table>
<thead>
<tr>
<th>chr</th>
<th>start</th>
<th>end</th>
<th>length</th>
<th>abs_summi</th>
<th>piled</th>
<th>#NAME?</th>
<th>fold_en</th>
<th>#NAME?</th>
</tr>
</thead>
<tbody>
<tr>
<td>chr1</td>
<td>4E+07</td>
<td>4E+07</td>
<td>214</td>
<td>40332825</td>
<td>24</td>
<td>19.98059</td>
<td>9.169</td>
<td>14.56673</td>
</tr>
<tr>
<td>chr10</td>
<td>2E+07</td>
<td>2E+07</td>
<td>130</td>
<td>22168150</td>
<td>71</td>
<td>16.95287</td>
<td>3.188</td>
<td>11.6175</td>
</tr>
<tr>
<td>chr10</td>
<td>6E+07</td>
<td>6E+07</td>
<td>139</td>
<td>61065544</td>
<td>28</td>
<td>22.57681</td>
<td>9.406</td>
<td>17.10032</td>
</tr>
<tr>
<td>chr10</td>
<td>7E+07</td>
<td>7E+07</td>
<td>159</td>
<td>69182041</td>
<td>19</td>
<td>11.1882</td>
<td>5.571</td>
<td>6.07643</td>
</tr>
<tr>
<td>chr11</td>
<td>9E+07</td>
<td>9E+07</td>
<td>692</td>
<td>94902750</td>
<td>39</td>
<td>19.21628</td>
<td>5.679</td>
<td>13.81907</td>
</tr>
<tr>
<td>chr11</td>
<td>9E+07</td>
<td>9E+07</td>
<td>204</td>
<td>94902953</td>
<td>24</td>
<td>13.17297</td>
<td>5.614</td>
<td>7.96399</td>
</tr>
<tr>
<td>chr11</td>
<td>1E+08</td>
<td>1E+08</td>
<td>866</td>
<td>98657854</td>
<td>34</td>
<td>26.64127</td>
<td>9.749</td>
<td>21.08695</td>
</tr>
<tr>
<td>chr11</td>
<td>1E+08</td>
<td>1E+08</td>
<td>202</td>
<td>98797465</td>
<td>31</td>
<td>28.34715</td>
<td>11.62</td>
<td>22.76277</td>
</tr>
<tr>
<td>chr11</td>
<td>1E+08</td>
<td>1E+08</td>
<td>458</td>
<td>101033294</td>
<td>51</td>
<td>27.20384</td>
<td>6.577</td>
<td>21.64065</td>
</tr>
<tr>
<td>chr11</td>
<td>1E+08</td>
<td>1E+08</td>
<td>976</td>
<td>102977885</td>
<td>114</td>
<td>117.7367</td>
<td>21.8</td>
<td>108.31483</td>
</tr>
<tr>
<td>chr12</td>
<td>9E+07</td>
<td>9E+07</td>
<td>256</td>
<td>88235875</td>
<td>54</td>
<td>34.79762</td>
<td>8.485</td>
<td>29.13021</td>
</tr>
<tr>
<td>chr12</td>
<td>1E+08</td>
<td>1E+08</td>
<td>172</td>
<td>116610545</td>
<td>52</td>
<td>75.90919</td>
<td>29.91</td>
<td>69.61341</td>
</tr>
<tr>
<td>chr13</td>
<td>3E+07</td>
<td>3E+07</td>
<td>253</td>
<td>34629123</td>
<td>34</td>
<td>42.61245</td>
<td>18.78</td>
<td>36.88287</td>
</tr>
<tr>
<td>chr13</td>
<td>1E+08</td>
<td>1E+08</td>
<td>403</td>
<td>108621779</td>
<td>52</td>
<td>42.57547</td>
<td>11.9</td>
<td>36.84896</td>
</tr>
<tr>
<td>chr13</td>
<td>1E+08</td>
<td>1E+08</td>
<td>521</td>
<td>113836658</td>
<td>115</td>
<td>115.5785</td>
<td>20.7</td>
<td>106.93469</td>
</tr>
</tbody>
</table>
(Table 3.1 continued)

<table>
<thead>
<tr>
<th>chr</th>
<th>2E+07</th>
<th>2E+07</th>
<th>417</th>
<th>21294543</th>
<th>35</th>
<th>27.78518</th>
<th>10.03</th>
<th>22.21164</th>
</tr>
</thead>
<tbody>
<tr>
<td>chr15</td>
<td>8E+07</td>
<td>8E+07</td>
<td>269</td>
<td>81688817</td>
<td>32</td>
<td>12.14457</td>
<td>4.174</td>
<td>6.98778</td>
</tr>
<tr>
<td>chr15</td>
<td>1E+08</td>
<td>1E+08</td>
<td>199</td>
<td>98702755</td>
<td>30</td>
<td>19.6769</td>
<td>7.417</td>
<td>14.27236</td>
</tr>
<tr>
<td>chr15</td>
<td>1E+08</td>
<td>1E+08</td>
<td>457</td>
<td>100446605</td>
<td>65</td>
<td>55.06908</td>
<td>13.34</td>
<td>49.24123</td>
</tr>
<tr>
<td>chr17</td>
<td>2E+07</td>
<td>2E+07</td>
<td>135</td>
<td>17964860</td>
<td>23</td>
<td>12.31064</td>
<td>5.389</td>
<td>7.13832</td>
</tr>
<tr>
<td>chr17</td>
<td>4E+07</td>
<td>4E+07</td>
<td>473</td>
<td>35196645</td>
<td>72</td>
<td>51.20489</td>
<td>10.36</td>
<td>45.41695</td>
</tr>
<tr>
<td>chr17</td>
<td>8E+07</td>
<td>8E+07</td>
<td>227</td>
<td>84105682</td>
<td>18</td>
<td>13.16938</td>
<td>6.898</td>
<td>7.96383</td>
</tr>
<tr>
<td>chr18</td>
<td>7E+07</td>
<td>7E+07</td>
<td>297</td>
<td>70728886</td>
<td>35</td>
<td>15.9922</td>
<td>5.111</td>
<td>10.6842</td>
</tr>
<tr>
<td>chr19</td>
<td>5E+06</td>
<td>5E+06</td>
<td>139</td>
<td>5050679</td>
<td>36</td>
<td>18.67471</td>
<td>5.908</td>
<td>13.29805</td>
</tr>
<tr>
<td>chr19</td>
<td>5E+06</td>
<td>5E+06</td>
<td>297</td>
<td>5059356</td>
<td>30</td>
<td>16.03755</td>
<td>5.831</td>
<td>10.72671</td>
</tr>
<tr>
<td>chr19</td>
<td>5E+07</td>
<td>5E+07</td>
<td>342</td>
<td>46471902</td>
<td>41</td>
<td>38.2826</td>
<td>13.43</td>
<td>32.59413</td>
</tr>
<tr>
<td>chr2</td>
<td>5E+06</td>
<td>5E+06</td>
<td>313</td>
<td>4541720</td>
<td>41</td>
<td>26.30039</td>
<td>7.9</td>
<td>20.75279</td>
</tr>
<tr>
<td>chr2</td>
<td>2E+07</td>
<td>2E+07</td>
<td>355</td>
<td>16277581</td>
<td>20</td>
<td>12.09984</td>
<td>5.85</td>
<td>6.94564</td>
</tr>
<tr>
<td>chr2</td>
<td>9E+07</td>
<td>9E+07</td>
<td>315</td>
<td>92886671</td>
<td>44</td>
<td>33.32309</td>
<td>10.1</td>
<td>27.67196</td>
</tr>
<tr>
<td>chr3</td>
<td>1E+08</td>
<td>1E+08</td>
<td>177</td>
<td>143869201</td>
<td>31</td>
<td>12.95673</td>
<td>4.543</td>
<td>7.7613</td>
</tr>
<tr>
<td>chr4</td>
<td>5E+07</td>
<td>5E+07</td>
<td>244</td>
<td>47069810</td>
<td>25</td>
<td>24.9939</td>
<td>11.79</td>
<td>19.46714</td>
</tr>
<tr>
<td>chr5</td>
<td>7E+07</td>
<td>7E+07</td>
<td>164</td>
<td>66089746</td>
<td>21</td>
<td>19.76101</td>
<td>9.972</td>
<td>14.35459</td>
</tr>
<tr>
<td>chr6</td>
<td>9E+07</td>
<td>9E+07</td>
<td>194</td>
<td>85283650</td>
<td>38</td>
<td>25.15462</td>
<td>8.062</td>
<td>19.6273</td>
</tr>
<tr>
<td>chr6</td>
<td>1E+08</td>
<td>1E+08</td>
<td>177</td>
<td>137703225</td>
<td>22</td>
<td>13.84072</td>
<td>6.334</td>
<td>8.61041</td>
</tr>
<tr>
<td>chr7</td>
<td>7E+06</td>
<td>7E+06</td>
<td>4261</td>
<td>6662311</td>
<td>67</td>
<td>104.3578</td>
<td>38.37</td>
<td>96.62609</td>
</tr>
<tr>
<td>chr7</td>
<td>7E+06</td>
<td>7E+06</td>
<td>151</td>
<td>6665441</td>
<td>32</td>
<td>40.96827</td>
<td>18.62</td>
<td>35.25648</td>
</tr>
<tr>
<td>chr7</td>
<td>7E+06</td>
<td>7E+06</td>
<td>141</td>
<td>6668799</td>
<td>27</td>
<td>32.95075</td>
<td>15.8</td>
<td>27.30277</td>
</tr>
<tr>
<td>chr7</td>
<td>7E+06</td>
<td>7E+06</td>
<td>127</td>
<td>6678086</td>
<td>16</td>
<td>12.0638</td>
<td>6.735</td>
<td>6.91605</td>
</tr>
<tr>
<td>chr7</td>
<td>1E+08</td>
<td>1E+08</td>
<td>255</td>
<td>134758760</td>
<td>23</td>
<td>10.34227</td>
<td>4.514</td>
<td>5.27608</td>
</tr>
<tr>
<td>chr7</td>
<td>1E+08</td>
<td>1E+08</td>
<td>127</td>
<td>134780575</td>
<td>25</td>
<td>16.94107</td>
<td>7.242</td>
<td>11.6063</td>
</tr>
<tr>
<td>chr7</td>
<td>1E+08</td>
<td>1E+08</td>
<td>420</td>
<td>149766405</td>
<td>30</td>
<td>22.19105</td>
<td>8.635</td>
<td>16.7255</td>
</tr>
<tr>
<td>chr8</td>
<td>2E+07</td>
<td>2E+07</td>
<td>141</td>
<td>19926784</td>
<td>95</td>
<td>12.99106</td>
<td>2.309</td>
<td>7.79544</td>
</tr>
<tr>
<td>chr8</td>
<td>6E+07</td>
<td>6E+07</td>
<td>133</td>
<td>56199328</td>
<td>26</td>
<td>9.46264</td>
<td>3.833</td>
<td>4.4665</td>
</tr>
<tr>
<td>chr8</td>
<td>9E+07</td>
<td>9E+07</td>
<td>652</td>
<td>93478789</td>
<td>110</td>
<td>102.4554</td>
<td>17.72</td>
<td>94.85301</td>
</tr>
<tr>
<td>chr8</td>
<td>1E+08</td>
<td>1E+08</td>
<td>149</td>
<td>114060235</td>
<td>31</td>
<td>23.08086</td>
<td>8.812</td>
<td>17.59557</td>
</tr>
<tr>
<td>chr9</td>
<td>3E+07</td>
<td>3E+07</td>
<td>170</td>
<td>30983979</td>
<td>23</td>
<td>18.07835</td>
<td>8.38</td>
<td>12.71156</td>
</tr>
<tr>
<td>chr9</td>
<td>4E+07</td>
<td>4E+07</td>
<td>133</td>
<td>43551823</td>
<td>19</td>
<td>14.35476</td>
<td>7.335</td>
<td>9.09958</td>
</tr>
<tr>
<td>chr9</td>
<td>5E+07</td>
<td>5E+07</td>
<td>158</td>
<td>50693365</td>
<td>20</td>
<td>11.03346</td>
<td>5.303</td>
<td>5.93413</td>
</tr>
<tr>
<td>chr9</td>
<td>5E+07</td>
<td>5E+07</td>
<td>158</td>
<td>53676597</td>
<td>17</td>
<td>17.19154</td>
<td>9.589</td>
<td>11.85277</td>
</tr>
<tr>
<td>chr9</td>
<td>5E+07</td>
<td>5E+07</td>
<td>291</td>
<td>53677330</td>
<td>24</td>
<td>13.17297</td>
<td>5.614</td>
<td>7.96399</td>
</tr>
<tr>
<td>chr9</td>
<td>8E+07</td>
<td>8E+07</td>
<td>170</td>
<td>83470446</td>
<td>26</td>
<td>16.16732</td>
<td>6.634</td>
<td>10.85411</td>
</tr>
<tr>
<td>chr9</td>
<td>1E+08</td>
<td>1E+08</td>
<td>2115</td>
<td>122670298</td>
<td>40</td>
<td>42.54447</td>
<td>16.17</td>
<td>36.81857</td>
</tr>
<tr>
<td>chrX</td>
<td>1E+08</td>
<td>1E+08</td>
<td>261</td>
<td>103382197</td>
<td>36</td>
<td>21.46082</td>
<td>6.959</td>
<td>16.01181</td>
</tr>
<tr>
<td>chrX</td>
<td>2E+08</td>
<td>2E+08</td>
<td>168</td>
<td>165112444</td>
<td>26</td>
<td>17.96076</td>
<td>7.521</td>
<td>12.59682</td>
</tr>
<tr>
<td>chrX</td>
<td>2E+08</td>
<td>2E+08</td>
<td>138</td>
<td>166430854</td>
<td>97</td>
<td>23.641</td>
<td>3.323</td>
<td>18.14262</td>
</tr>
</tbody>
</table>
Table 3.2 Final output of MACS2 describing ChIP-seq peaks derived from KO MEF cells

<table>
<thead>
<tr>
<th>chr</th>
<th>start</th>
<th>end</th>
<th>length</th>
<th>abs_summ</th>
<th>pileup</th>
<th>#NAME?</th>
<th>fold_enric</th>
<th>#NAME?</th>
</tr>
</thead>
<tbody>
<tr>
<td>chr10</td>
<td>7E+06</td>
<td>7E+06</td>
<td>189</td>
<td>7232759</td>
<td>45</td>
<td>15.73724</td>
<td>4.12341</td>
<td>10.29044</td>
</tr>
<tr>
<td>chr11</td>
<td>1E+08</td>
<td>1E+08</td>
<td>637</td>
<td>102977887</td>
<td>54</td>
<td>64.83344</td>
<td>22.44136</td>
<td>58.95597</td>
</tr>
<tr>
<td>chr11</td>
<td>1E+08</td>
<td>1E+08</td>
<td>132</td>
<td>104396112</td>
<td>40</td>
<td>55.54784</td>
<td>23.76239</td>
<td>49.73617</td>
</tr>
<tr>
<td>chr12</td>
<td>9E+07</td>
<td>9E+07</td>
<td>180</td>
<td>88235935</td>
<td>21</td>
<td>14.52334</td>
<td>6.92641</td>
<td>9.12068</td>
</tr>
<tr>
<td>chr12</td>
<td>1E+08</td>
<td>1E+08</td>
<td>110</td>
<td>116610524</td>
<td>27</td>
<td>37.30627</td>
<td>18.23666</td>
<td>31.54494</td>
</tr>
<tr>
<td>chr13</td>
<td>1E+08</td>
<td>1E+08</td>
<td>181</td>
<td>108621755</td>
<td>21</td>
<td>18.55709</td>
<td>9.25034</td>
<td>13.01896</td>
</tr>
<tr>
<td>chr13</td>
<td>1E+08</td>
<td>1E+08</td>
<td>300</td>
<td>113836650</td>
<td>39</td>
<td>32.80038</td>
<td>11.30277</td>
<td>27.06999</td>
</tr>
<tr>
<td>chr15</td>
<td>1E+08</td>
<td>1E+08</td>
<td>247</td>
<td>100446588</td>
<td>40</td>
<td>25.17904</td>
<td>7.65998</td>
<td>19.52791</td>
</tr>
<tr>
<td>chr17</td>
<td>4E+07</td>
<td>4E+07</td>
<td>162</td>
<td>35196700</td>
<td>23</td>
<td>19.585</td>
<td>9.2453</td>
<td>14.0203</td>
</tr>
<tr>
<td>chr18</td>
<td>7E+07</td>
<td>7E+07</td>
<td>204</td>
<td>70728833</td>
<td>23</td>
<td>13.89606</td>
<td>6.15122</td>
<td>8.52842</td>
</tr>
<tr>
<td>chr18</td>
<td>8E+07</td>
<td>8E+07</td>
<td>100</td>
<td>82390457</td>
<td>49</td>
<td>31.35868</td>
<td>8.22651</td>
<td>25.63886</td>
</tr>
<tr>
<td>chr2</td>
<td>5E+06</td>
<td>5E+06</td>
<td>156</td>
<td>4541786</td>
<td>19</td>
<td>14.98483</td>
<td>7.70441</td>
<td>9.56219</td>
</tr>
<tr>
<td>chr2</td>
<td>9E+07</td>
<td>9E+07</td>
<td>157</td>
<td>92886581</td>
<td>19</td>
<td>10.32927</td>
<td>5.12602</td>
<td>5.21948</td>
</tr>
<tr>
<td>chr6</td>
<td>1E+08</td>
<td>1E+08</td>
<td>100</td>
<td>140582309</td>
<td>18</td>
<td>14.61188</td>
<td>7.75247</td>
<td>9.20108</td>
</tr>
<tr>
<td>chr7</td>
<td>7E+06</td>
<td>7E+06</td>
<td>4318</td>
<td>6660650</td>
<td>71</td>
<td>117.3989</td>
<td>43.56048</td>
<td>108.4094</td>
</tr>
</tbody>
</table>
Electromobility shift assay (EMSA)

EMSA was performed using a gel shift assay system kit (Promega Cat. No. E3053). This series of assays also used mouse brain nuclear extract (Active Motif Cat. No. 36053) since Peg3 is highly expressed in the brain tissue. The competition assays were performed in the following manner. Briefly, the binding buffer, 2.72 μg mouse brain nuclear extract and unlabeled competitor oligonucleotide duplexes (1.74 pico mole, 200X) were first incubated at room temperature 10 mins. Later, 1 μl of P32-labeled duplex probe (1X) was added and incubated at room temperature for additional 20 mins. The reaction mixture was subsequently separated on a
5% TBE gel (Bio-Rad Cat. No. 456-5014), and exposed to a film for 2 to 6 hours at -80. For supershift assays, the initial reaction mixture was incubated along with an antibody, either anti-PEG3 antibody or anti-YY1 antibody (SantaCruz, Cat No. 1703X), and later incubated with P32-labeled duplex probe.

RNA isolation and qRT-PCR analyses

Total RNA was isolated from either MEF or the brains of one-day-old neonates using a commercial kit (Trizol, Invitrogen). The total RNA was then reverse-transcribed using the M-MLV kit (Invitrogen), and the subsequent cDNA was used as a template for quantitative PCR. The qRT-PCR analysis was performed with SYBR Select Master Mix (Applied Biosystems, Life Technologies) using the iCycler iQTM multicolor real-time detection system (Bio-Rad). All qRT-PCR reactions were carried out for 40 cycles under standard PCR conditions with internal controls (Gapdh and β-actin). The results derived from qRT-PCR were further analyzed using the threshold (Ct) value. The ΔCt value was initially calculated by subtracting Ct value of a testing replicate of a given gene from the average Ct value of the internal control (28S and β-actin). The fold difference for each replicate was then calculated by raising the ΔΔCt value as a power of 2 (Winer et al., 1999). The average and standard deviation for each sample were then calculated by compiling the normalized values. The information regarding individual primer sequences is also available (Table 3.3).
Table 3.3 Primer sequences used in this study.

<table>
<thead>
<tr>
<th>RT-Primers</th>
<th>Oligo sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Msl1 rt a</td>
<td>GGAGATAGAGGAGCTGAAAGTC</td>
</tr>
<tr>
<td>Msl1 rt b</td>
<td>TAAGGGGAGGAGTCTGGAAGAGT</td>
</tr>
<tr>
<td>Ilr1 rt a</td>
<td>TGCTGGATTGGATGAACATC</td>
</tr>
<tr>
<td>Ilr1 rt b</td>
<td>TCTCCACTGCTGGAGGTAC</td>
</tr>
<tr>
<td>Pdk2 rt a</td>
<td>CGTTGCTCATGAACAGCAGT</td>
</tr>
<tr>
<td>Pdk2 rt b</td>
<td>CTGCGGGCTCTTGGTCCAGGA</td>
</tr>
<tr>
<td>Tob2 rt a</td>
<td>CTGCAGAGGCGCATCCCTG</td>
</tr>
<tr>
<td>Tob2 rt b</td>
<td>CTCTCTCCCCAAAACAGGCTGT</td>
</tr>
<tr>
<td>Mta3 rt a</td>
<td>CGGCAACATGTACCCTGGGT</td>
</tr>
<tr>
<td>Mta3 rt b</td>
<td>GCCGCTCTCCTCCGTCAGTC</td>
</tr>
<tr>
<td>Dazap2 rt a</td>
<td>GAGTGGGAGCGAAGCTAGAG</td>
</tr>
<tr>
<td>Dazap2 rt b</td>
<td>ATCGTACGACCTGCACTAC</td>
</tr>
<tr>
<td>Znrfl1 rt a</td>
<td>CCCAGGTGGTACGTCAACGCA</td>
</tr>
<tr>
<td>Znrfl1 rt b</td>
<td>CTCTCTGGTGGGGTGATGATC</td>
</tr>
<tr>
<td>SIC35F2 rt a</td>
<td>GAGCAGGATCCCGAGCATAAC</td>
</tr>
<tr>
<td>SIC35F2 rt b</td>
<td>ATCTCCAGGAGGTTACTACT</td>
</tr>
<tr>
<td>Prdm11 rt a</td>
<td>GGTCGGAAGCACAAGGAGGCA</td>
</tr>
<tr>
<td>Prdm11 rt b</td>
<td>GCCGAAGTGGCTGGAGTGATG</td>
</tr>
<tr>
<td>Pprf18 rt a</td>
<td>GAGGACAGAAGACCTGCTG</td>
</tr>
<tr>
<td>Pprf18 rt b</td>
<td>CCCAACACGTCTAATCCGAG</td>
</tr>
<tr>
<td>Msl3 rt a</td>
<td>AGAAGAGCGCTGTAGGTTG</td>
</tr>
<tr>
<td>Msl3 rt b</td>
<td>ACGCAGCGCACAATACCTT</td>
</tr>
<tr>
<td>Ddah2 rt a</td>
<td>AATCAACAGGCGTCAGAGAGT</td>
</tr>
<tr>
<td>Ddah2 rt b</td>
<td>GAGCTTTAGCCAGGTCCAGA</td>
</tr>
<tr>
<td>Rara rt a</td>
<td>ACGAGTCTCCCTGAGCATTG</td>
</tr>
<tr>
<td>Rara rt b</td>
<td>GTGCTCTGGTCAGGCTGTA</td>
</tr>
<tr>
<td>Tnik rt a</td>
<td>GGACCCCTAGTGGTACAGA</td>
</tr>
<tr>
<td>Tnik rt b</td>
<td>TCTGATTTCCCGGTTGAGTC</td>
</tr>
<tr>
<td>Aplp2 rt a</td>
<td>GAGCCCTCAGATGCGAGT</td>
</tr>
<tr>
<td>Aplp2 rt b</td>
<td>ATTGACCTGGCTGTTGCTT</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ChIP-Primers</th>
<th>Oligo sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>H19 ICR-F</td>
<td>GTGTTGCTACATAAGGAAGGCA</td>
</tr>
<tr>
<td>H19 ICR-R</td>
<td>CCTATGCGCGCTGTCCCGAGCAA</td>
</tr>
<tr>
<td>H19 Peak-F</td>
<td>GTACGACCCCATGACCTTTAT</td>
</tr>
<tr>
<td>H19 Peak-R</td>
<td>ATGGGCTGGTAAGACCGAGT</td>
</tr>
<tr>
<td>Pgrn211 F</td>
<td>ACCAAGCCCTGAGGCGAAGA</td>
</tr>
<tr>
<td>Pgrn211 R</td>
<td>TGCCCTTCACAGCTGTCTTTC</td>
</tr>
</tbody>
</table>
Luciferase reporter assay

This series of reporter assays used the modified version of β-geo vector (Kim et al., 2009) as a control vector monitoring transfection efficiency, and also the modified version of luciferase vector (Promega, PGL3) as a basic construct testing the promoter and the PEG3-bound regions of H19. The 610-bp-long promoter region of H19 was first inserted into the NotI site of the promoterless luciferase vector (Basic-Luc). Then, this constructed vector (H19-Pro) was further modified by inserting individually three 152-bp-long target and its mutated versions (WT-152,
Mu2a, Mu2b) into the 5'-side region of the 610-bp promoter of H19. For luciferase assay, HEK 293 cells were cultured in DMEM plus GlutaMAX medium with 10% fetal bovine serum and 1% antibiotic-antimycotic (GibcoBRL), and plated in 12-well plates for plasmid transfection. The cells on each well were transfected with 4 μl lipofectamine 2000 (Invitrogen) and 1.6 μg of total reporter construct (0.8 μg β-geo vector and 0.8 μg luciferase vector) either alone or with the expression vector containing full-length PEG3. For a series of co-transfection experiments, an additional 10 ng of the expression vector of PEG3 was regarded as 1X, 50 ng as 5X, and so forth. Fresh complete media was added 6 hours after transfection, and total cell lysates were harvested in 300 μl of cell lysis buffer 48 hours later according to our previously published protocol (Kim et al., 2009). The luciferase assay was performed in triplicate as previously published.

3.3 Results

Identification of the downstream genes of Peg3 in MEF cells

In the current study, we used a different but improved approach for PEG3 ChIP-seq experiments compared to the previous study (Thiaville et al., 2013). First, we used the chromatin isolated from homogenous cell populations, mouse embryonic fibroblast cells (MEFs). Second, we used one mutant model targeting the Peg3 locus as a negative control (Kim et al., 2013). In this KO model, the transcription of Peg3 is truncated by the two poly(A) signals that have been inserted into the intron 5 as part of the expression cassette, thus the protein PEG3 is depleted (Figure 3.1 A). The actual experiments were performed as described below. We first derived a set of MEF cells, WT and KO, from the 14.5-dpc embryos that had been prepared through timed mating of the male heterozygotes for the KO allele and female littermates. The chromatin prepared from WT and KO MEF cells were individually immunoprecipitated with polyclonal anti-PEG3 antibody (Ye et al., 2014). The two immunoprecipitated DNA along with the two
input DNA were used for library construction and subsequent sequencing, resulting in 30 to 40 millions reads per sample. These four sets of raw sequence reads have been processed for predicting the potential targets (peaks) of PEG3. This series of bioinformatic processes, described in Material and Methods, derived two sets of target regions: one set (56 peaks) from WT and the other set (36 peaks) from KO (Table 3.1 and Table 3.2). The target regions predicted only from WT, but not from KO, have been further considered as the potential target regions of PEG3 (41 peaks).

Figure 3.1 PEG3 ChIP-seq using WT and KO MEF cells. (A) Schematic representation of the genomic structure of the paternally expressed Peg3. In the mutant allele, a 7.1-kb cassette containing a promoterless β-galactosidase (β-Gal) and human β-actin promoter-driven neomycin resistant gene (NeoR) has been inserted between exon 5 and 6 of Peg3. The inserted cassette is flanked by two FRT sites (open ovals), thus can be removed through FLP-mediated recombination. (B) CTCF-binding profile and PEG3 ChIP-seq profile on the H19 locus. The CTCF-binding profiles were derived from MEF cells, brain and liver tissues (upper). PEG3 ChIP-seq profiles were derived from WT and KO MEF cells (lower).
Detailed inspection of these 41 target regions indicated that 31 targets are derived either from the 5’ enhancer, promoter, or 1st intron of 16 individual genes. The remaining 10 targets are derived from the intergenic regions without any obvious gene association.

In this study, we further consider a set of 16 genes as the potential downstream genes of PEG3 hereafter (Table 3.4). Initial inspection of this set of genes provided the following observations.

First, the majority of these genes (13 out of 16) tend to be highly expressed in mature oocytes, in which Peg3 is repressed by DNA methylation.

Second, several genes are closely associated with various cancers, in which Peg3 is also known to be repressed by DNA methylation (Dowdy et al., 2005; Feng et al., 2008). These include Il1r1, Tnik, Pdk2, Rara, Tob2 and Mta3. The expression profiles observed among the downstream genes are somewhat inversely correlated with those of Peg3. This may be an indication that Peg3 is a potential repressor for the identified downstream genes.

Third, one particular target turns out to be localized within the imprinting control region of H19, which has been shown as a representative peak in Figure 3.1 B. This peak is localized between the 3rd and 4th CTCF binding sites of the H19-ICR.

The potential binding and subsequent connection of PEG3 to the H19-ICR is very significant given the functional roles played by these two imprinted genes, thus this possibility has been further analyzed in the latter half of the current study.
Overall, this series of ChIP-seq experiments identified a set of 16 genes that may be regulated by Peg3 in MEF cells.

Table 3.4 Downstream genes of PEG3 in MEF cells (mm9)

<table>
<thead>
<tr>
<th>downstream genes</th>
<th>chr</th>
<th>start</th>
<th>end</th>
<th>log(p value)</th>
<th>position</th>
<th>expression in oocytes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Il1r1</td>
<td>chr1</td>
<td>40332678</td>
<td>40332891</td>
<td>10.98059</td>
<td>1st intron</td>
<td>+++</td>
</tr>
<tr>
<td>Prpf18</td>
<td>chr2</td>
<td>4541553</td>
<td>4541865</td>
<td>26.30039</td>
<td>5' enhancer</td>
<td>++</td>
</tr>
<tr>
<td>Prdm11</td>
<td>chr2</td>
<td>92886527</td>
<td>92886836</td>
<td>33.32309</td>
<td>5' enhancer</td>
<td>+</td>
</tr>
<tr>
<td>Tnkl</td>
<td>chr3</td>
<td>28161777</td>
<td>28162142</td>
<td>26.96248</td>
<td>5' enhancer</td>
<td>+</td>
</tr>
<tr>
<td>H19</td>
<td>chr7</td>
<td>1.5E+08</td>
<td>1.5E+08</td>
<td>22.19105</td>
<td>5' enhancer</td>
<td>+</td>
</tr>
<tr>
<td>Znrf1</td>
<td>chr8</td>
<td>1.14E+08</td>
<td>1.14E+08</td>
<td>23.08086</td>
<td>promoter</td>
<td>+</td>
</tr>
<tr>
<td>Aplp2</td>
<td>chr9</td>
<td>30983874</td>
<td>30984043</td>
<td>18.07835</td>
<td>intron</td>
<td>+</td>
</tr>
<tr>
<td>Slc35f2</td>
<td>chr9</td>
<td>53677080</td>
<td>53677370</td>
<td>13.17297</td>
<td>3' enhancer</td>
<td>++</td>
</tr>
<tr>
<td>Pdk2</td>
<td>chr11</td>
<td>94902125</td>
<td>94902816</td>
<td>19.21628</td>
<td>promoter</td>
<td>no</td>
</tr>
<tr>
<td>Msl1</td>
<td>chr11</td>
<td>98657458</td>
<td>98658323</td>
<td>26.64127</td>
<td>promoter</td>
<td>+</td>
</tr>
<tr>
<td>Rara</td>
<td>chr11</td>
<td>98797334</td>
<td>98797535</td>
<td>28.34715</td>
<td>5' enhancer</td>
<td>no</td>
</tr>
<tr>
<td>Tob2</td>
<td>chr15</td>
<td>81688742</td>
<td>81689010</td>
<td>12.14457</td>
<td>promoter</td>
<td>++</td>
</tr>
<tr>
<td>Dazap2</td>
<td>chr15</td>
<td>1E+08</td>
<td>1E+08</td>
<td>55.06908</td>
<td>1st intron</td>
<td>++</td>
</tr>
<tr>
<td>Ddaah2</td>
<td>chr17</td>
<td>35196409</td>
<td>35196881</td>
<td>51.20489</td>
<td>1st intron</td>
<td>no</td>
</tr>
<tr>
<td>Mtn3</td>
<td>chr17</td>
<td>84105659</td>
<td>84105885</td>
<td>13.16938</td>
<td>1st intron</td>
<td>+++</td>
</tr>
<tr>
<td>Msl3</td>
<td>chrX</td>
<td>1.65E+08</td>
<td>1.65E+08</td>
<td>17.90076</td>
<td>promoter</td>
<td>+++</td>
</tr>
</tbody>
</table>

Expression level changes of the downstream genes in the MEF cells and neonatal brains

The identified downstream genes of Peg3 were further analyzed in the following manner. The expression levels of these genes were compared between the WT and KO cells lacking the protein PEG3 with qRT-PCR analyses (Figure 3.2). This series of expression analyses used two sets of total RNA: the first set from MEF cells and the second set from the neonatal brains. According to the results from MEF cells, the majority of the tested genes were upregulated in the KO cells except Il1r1 and Slc35f2. The observed upregulations were statistically significant: 10 genes showed 1.3 to 2.3-fold upregulation while 4 genes displayed more than 3-fold upregulation. Among all the tested genes, H19 displayed the most dramatic upregulation, 6-fold, followed by Pdk2 (4.3-fold), Msl1 (3.6-fold) and Msl3 (3.2-fold). On the other hand, the results from the
neonatal brains displayed overall less dramatic changes. Among the 16 tested genes, 9 genes displayed statistically significant changes: 8 upregulation and 1 downregulation. Five out of the 8 upregulated genes showed less than 1.6-fold upregulation. In particular, the expression level change of H19 (1.4-fold) was much smaller than the 6-fold upregulation observed from the MEF cells. In contrast, the three following genes showed more than 2-fold upregulation, including Tnik (2.3-fold), Il1r1 (2.9-fold) and Mta3 (2.7-fold).

Figure 3.2 Expression level changes of the identified downstream genes of Peg3. A series of qRT-PCR analyses were performed to measure the expression level changes of the potential downstream genes of Peg3 using the total RNA isolated from the MEF cells (A) and the neonatal brains (B). For each gene, the expression levels were first normalized with an internal control, and the normalized values were subsequently compared between KO and WT samples. The graph summarizes the relative expression levels of each gene with standard deviation, the statistical significance of which was further tested with student t-test. This series of expression analyses were also performed using two independent sets of MEFs and neonatal brains.

It is prudent to note that the expression levels of the majority of the genes observed from the neonatal brains were much lower than those observed from the MEF cells. For instance, the average Ct (threshold cycle) values of the tested genes were around 25 for the MEF cells and 30 for the neonatal brains, while the Ct values for an internal control Gapdh were around 19 to 20 in
both cases. This is particularly the case for the two genes, Msl1 and Msl3, in which their expression levels were almost undetectable in the neonatal brains. Thus, the values for these two genes are missing in the neonatal brain set. A similar series of qRT-PCR analyses were also performed using 13-dpc embryos, and the results also indicated that less dramatic changes were observed from this stage of embryos (Figure 3.3). Overall, this series of analyses demonstrated that the expression levels of the majority of the identified downstream genes of Peg3 were upregulated in the cells lacking PEG3, suggesting that Peg3 may function as a transcriptional repressor for the identified downstream genes. The observed upregulation was also much more pronounced in the MEF cells than in the neonatal brains and embryos.

Figure 3.3 Expression level changes using 13dpc embryos. A series of qRT-PCR analyses were performed to measure the expression level changes of the potential downstream genes of Peg3 using the total RNA isolated from the 13dpc embryos (left). And expression levels of a series of H19-related genes were also tested by qRT-PCR analysis.
H19 upregulation by the depletion of PEG3

The upregulation of H19 observed from the MEF cells and the neonatal brains was further characterized (Figure 3.4). This upregulation could be either a direct or indirect outcome of the depletion of PEG3. Therefore, another series of expression analyses were performed with a set of additional genes that may be associated with the function of the H19-ICR or the transcription of H19 itself. The list of these additional genes includes Igf2, Igf2r, Ctcf and Zac1. The transcriptional levels of Igf2 and Igf2r are closely associated with those of H19 (Gabory et al., 2009). On the other hand, both Ctcf and Zac1 are known to control the transcription levels of H19 as activators (Engel et al., 2006; Kurukuti et al., 2006; Schoenherr et al., 2003; Varrault et al., 2006). This list also includes two unrelated genes, Yy1 and p53, as negative controls.

Figure 3.4 Expression level changes of H19-related genes. A series of qRT-PCR analyses were performed to measure the expression level changes of the genes associated with the H19-ICR or H19 transcription using the total RNA isolated from the MEF cells (A) and the neonatal brains (B). For each gene, the expression levels were first normalized with an internal control, and the normalized values were subsequently compared between KO and WT samples. The graph summarizes the relative expression levels of each gene with standard deviation, the statistical significance of which was further tested with student t-test. This series of expression analyses were also performed using two independent sets of MEFs and neonatal brains.
According to the results from the MEF cells, the expression levels of four genes were upregulated in the KO cells: Ctcf (1.5-fold), Zac1 (2.7-fold), Igf2r (2-fold) and Igf2 (7-fold). In contrast, the two negative control genes, Yyl and p53, did not show any major difference between the WT and KO cells, indicating that the depletion of PEG3 did not have any global impact on the transcription of the MEF cells. Interestingly, the two known activators for H19, Ctcf and Zac1, turned out to be upregulated, thus it is possible that the depletion of PEG3 might have caused the upregulation of H19 through these two transcription factors. However, given the expression level changes in these genes, 6-fold change for H19 versus 1.5-fold change for Ctcf and 2.7-fold change for Zac, it might be PEG3 that regulates the expression of H19 as a trans factor. On the other hand, the upregulation of Igf2 and Igf2r was unexpected since their transcriptional levels, at least for Igf2, are usually inversely correlated with those of H19 (Thorvaldsen et al., 1998). Nevertheless, it is interesting to point out the fact that the levels of upregulation observed from Igf2 (7-fold) were somewhat similar to those of H19 (6-fold). This might be an indication that the depletion of PEG3 affects both H19 and Igf2 through the H19-ICR. A similar survey was also performed using the total RNA from the neonatal brains (Figure 3.4 B). According to the results, no major changes were observed except the fact that H19 was still upregulated while Igf2 and Ctcf were downregulated. Overall, this series of analyses concluded that the upregulation of H19 is likely an outcome of some changes that happen on the H19/Igf2 locus, but not of global effects that are driven by the depletion of PEG3.

**PEG3 binding to the H19-ICR**

We also further characterized the potential binding of PEG3 to the H19-ICR with the following approaches. We repeated ChIP experiments using the two sets of chromatin prepared
from MEF cells and neonatal brains (Figure 3.5 A). We first tested the feasibility of the ChIP experiments using the known target region (Thiaville et al., 2013).

Figure 3.5 PEG3 binding to the H19-ICR. Individual ChIP experiments were performed to confirm the in vivo binding of PEG to the H19-ICR using two sets of chromatins prepared from MEF cells and neonatal brains (A). Each set of chromatin was also derived from two different samples, representing WT and KO (+/-p) cells. Since Peg3 is expressed mainly from the paternal allele, the heterozygotes with the paternal transmission are considered to be null. Three individual DNA were derived from each ChIP experiment, and used as templates for PCR survey: Input, Negative control (Neg), and the immunoprecipitated DNA with anti-PEG3 antibody (PEG3 IP). PCR-based surveys tested three individual regions: Pgm2l1 as a positive control that has been known to be bound by PEG3, H19 ICR to test the 420-bp-long peak region and H19 Peak to test the 152-bp-long narrower peak region of the H19-ICR. (B) Allele test of PEG3 binding. Individual ChIP experiments were repeated with the chromatin isolated from the 11.5-dpc F1 hybrid embryos that had been prepared through the crossing between male Spretus and female C57BL/6J (B6). A restriction enzyme digestion (Acul) showed two alleles in the Input as well as the immunoprecipitated DNA, but with different ratios between the two alleles, which are shown underneath the gel images. (C) Imprinting test of H19 expression. The total RNA were isolated from two sets of the F1 hybrid samples and two sets of hybrid MEF Cell samples that had been prepared through the crossing between male B6 and female PWD/PhJ. A restriction enzyme digestion (BCLI) showed the two parental alleles (lane 1-2), and also the maternal-specific expression in neonatal brains (lane 3-4) and in MEF cells (lane 5-6).
These primer sets were tested on the two sets of ChIP DNA. As shown in Figure 3.5 A, specific enrichments were indeed observed only from the WT sets, but not from KO sets, confirming the actual binding of PEG3 to the H19-ICR. Quantitative measurement of these enrichments also supported the binding of PEG3 to the H19-ICR (Figure 3.6).

Figure 3.6 Quantitative measurement of the immunoprecipitated DNA by polyclonal antibodies against PEG3. A series of qPCR experiments were performed using the two sets of Input and IP DNA from MEF cells (left) and neonatal brain (right). For each locus, the relative enrichment levels were first normalized with the values from Input, and the normalized values were subsequently compared between KO and WT samples. The graph summarizes the relative enrichment levels of each locus with standard deviation. The relative enrichment values are much higher in WT than in KO samples since the chromatin from KO does not contain the PEG3 protein.
The enrichment (or binding) was much more obvious with the second primer set, further suggesting that PEG3 may bind to some cis-regulatory motifs located within this 152-bp-long genomic interval. As shown in the H19-ICR, an independent ChIP experiments also confirmed again the binding of PEG3 to a half of the potential targets identified through ChIP-seq (Figure 3.7).

Figure 3.7 Individual ChIP experiment confirming the binding of PEG3 to the identified targets by ChIP-Seq. Besides H19-ICR, 9 individual targets have been further tested by two separate series of ChIP experiments. Two of these loci, Msl1 and Msl3, were further analyzed with qPCR as shown Figure 3.6.
Since H19 is imprinted, we also tested the allele specificity of PEG3 binding to the H19-ICR. For this test, we prepared another set of chromatin isolated from 11-dpc F1 hybrid embryos that had been prepared through the interspecific crossing between male Spretus and female C57BL/6J (B6).

As shown in Figure 3.5 B, a restriction enzyme digestion of the input DNA showed two parental alleles although the Spretus allele was over-represented relative to the B6 allele due to heteroduplex formation during PCR followed by inefficient enzyme digestion (1 to 0.29).

Nevertheless, the digestion of the immunoprecipitated DNA by anti-PEG3 antibody displayed much higher levels of the B6 allele (1 to 3.20), suggesting that PEG3 most likely binds to the maternal allele of the H19-ICR. This further implicates that PEG3 may be functionally involved in the maternal allele of the H19-ICR.

To further test this possibility, we also performed another imprinting test with the total RNA isolated from the MEF cells and neonatal brains of F1 hybrid that had been prepared through the interspecific crossing between male B6 and female PWD/PhJ.

A restriction enzyme digestion showed that H19 was still expressed mainly from the maternal allele in both cells and neonatal brain tissues. This is also the case for the imprinting status of Igf2, still showing the paternal expression (Figure 3.8). This indicated that the depletion of PEG3 did not cause any change in the imprinting status of H19, meaning no impact on the imprinted paternal allele.

Instead, the depletion of PEG3 may have an impact on the active maternal allele of the H19-ICR, resulting in the upregulation of H19 in MEF cells and neonatal brains. Taken together,
this series of analyses concluded that PEG3 most likely binds to the maternal allele of the H19-ICR as a repressor to control the transcriptional levels of H19.

![Imprinting test](image)

**Figure 3.8** Imprinting test of *Igf2* expression and other imprinted genes’ expression level changes after *Peg3* knockout. (A) Imprinting test results indicated that the *Igf2* is still a paternally expressed gene in the KO. (B) The expression levels of three imprinted gene (*Igf2, Grb10, Peg10*) elevated in the KO.

**PEG3’s binding site within the H19-ICR**

The predicted binding sites of PEG3 localized within the H19-ICR were further characterized with a series of gel shift assays. This series of gel shift assays were performed with several sets of oligonucleotide duplexes.

First, a 39-bp-long duplex from the Pgm2l1 locus was used as a positive control since this locus has been proven to be a target of PEG3 (Thiaville et al., 2013). Second, the 152-bp-long H19 Peak region, the smallest target of PEG3 within the H19-ICR (Figure 3.5 A), was further
divided into three 50-bp-long individual regions, which were subsequently used as a set of three testing duplexes, H19-1, -2, and -3.

Finally, the 50-bp-long region surrounding the 3rd CTCF site of the H19-ICR was used as a negative control. According to the results, the P32-labeled Pgm2l1 probe was bound by a protein complex, which is known to be the complex containing PEG3 (Thiaville et al., 2013).

Among the tested duplexes, one particular duplex, H19-1, was shown to compete well against the Pgm2l1 probe, whereas the other two duplexes, H19-2 and H19-3, and also the CTCF duplex did not compete at all.

This was further confirmed through a reciprocal set of gel shift assays, in which the H19-1 duplex was used as a P32-labeled probe (Figure 3.9 B). Consistent with the initial result, the Pgm2l1 duplex was the only duplex that competed well against the H19-1 probe.

As expected, the three remaining duplexes, H19-2, H19-3 and CTCF, did not compete against the H19-1 probe. Also, a set of supershift assays demonstrated that the protein complex bound by the H19-1 probe is indeed the complex containing PEG3 (Figure 3.9 B).

Taken together, this series of gel shift assays confirmed that the 50-bp-long region covered by the H19-1 duplex most likely contains the binding site for PEG3. Furthermore, this 50-bp-long region may be the only region that is responsible for the binding to PEG3 within the H19-ICR.

The actual binding sites of PEG3 within the H19-1 region were further narrowed down with another set of gel shift assays (Figure 3.10). This series of analyses used two sets of mutant oligonucleotide duplexes. For the first set of mutant duplexes, the 50-bp-long H19-1 region was
again divided into three regions, and the sequence of each region was all mutated into As: Mu L1, L2, and L3 (Figure 3.10).

Figure 3.9 Electromobility shift assay for the DNA-binding sites of PEG3 within the H19-ICR. (A) The 152-bp-long region within the H19-ICR was divided into three individual 50-bp-long regions, and used as oligonucleotide duplexes, H19-1, H19-2, H19-3. The 32-bp-long duplex from Pgm2l1 was used as a P$^{32}$-radiolabeled probe. The 50-bp-long region surrounding the 3rd CTCF site of the H19-ICR was also included as a control. Excessive amount (200X) of unlabeled Pgm2l1, H19-1, H19-2, H19-3 and CTCF were competed against the radiolabeled H19-1 probe. The H19-1 duplex competed well against the labeled Pgm2l1 probe, indicating that the DNA-binding site for PEG3 is localized within the covered region by H19-1. (B) A reciprocal set of competition assays were also performed using H19-1 as a radiolabeled probe (left). A separate supershift assay was also performed to confirm that the complex binding to H19-1 is indeed the protein PEG3 (right). (C) Shown are the sequences of the duplexes used for the EMSA.
These three mutant duplexes were tested against the original H19-1 probe. The results indicated that the two duplexes, Mu L1 and Mu L2, did not compete well, suggesting that the region covered by these two duplexes likely contains the actual targets of PEG3.

**Figure 3.10 Mutational analyses of the PEG3 binding site of the H19-ICR.** The identified 50-bp-long H19-1 region was further analyzed with two sets of mutant oligonucleotide duplexes. (A) The 50-bp-long H19-1 was divided into three 14-bp-long regions, and the bases within each region were all mutated into As. These mutants, Mu L1, L2, L3, were competed against the original H19-1 probe (left). The region covered by Mu L1 and L2 was further divided into 4 individual 7-bp-long regions, and each region was again mutated with As. These four mutants were competed against the H19-1 probe (right). (B) The results indicated that two 7-bp-long regions, Mu 1a and Mu 2b, are the most critical for the binding to PEG3, which are indicated by asterisks (top). The schematic diagram shows part of genomic structure of the H19-ICR with the relative positions of H19-1, H19-2, H19-3, AB-2, AB-3 and two CTCF sites (bottom).
Thus, we prepared the second set of mutant duplexes. In this set, the 14-bp-long A stretches of each of the two duplexes, Mu L1 and Mu L2, was further divided into two individual stretches of 7-bp-long As: Mu 1a, Mu 1b, Mu 2a, Mu 2b. We repeated another competition assay with these 4 mutant duplexes. The results indicated that the two duplexes, Mu 1a and Mu 2b, did not compete well against the original H19-1 probe, suggesting that these two small regions may be the most critical regions for the binding to PEG3. Inspection of these two small regions revealed that these two regions contain or overlap with a small motif, which resembles part of the known motif of PEG3, GTGG (Figure 3.10 B) (Thiaville et al., 2013). It is also interesting to note that the 450-bp region surrounding two CTCF sites also contains two potential binding sites for an orphan nuclear receptor family (Bowman et al., 2003). These two sites, AB-2 and AB-3, are localized just outside of the two PEG3 binding sites, thus suggesting that this 450-bp region may be a main regulatory center attracting several transcription factors besides CTCF and PEG3 for the transcription control of H19 and Igf2. Overall, this series of analyses identified two 7-bp-long small regions within the H19-ICR as potential binding sites for PEG3.

**PEG3 as a transcriptional repressor for H19**

We further tested a potential repressor role of PEG3 in the transcription of H19 using the following in vitro experiments (Figure 3.11). As mentioned before, the transcription of Peg3 is truncated by the inserted cassette, which is flanked by two FRT (Flippase Recombination Target) sites. A vector construct expressing Flippase (FLP) was transiently transfected into the KO cells to remove the cassette, subsequently restoring the transcription and translation of Peg3. Two additional sets of MEF cells were also transfected as controls: one without vector (Mock) and the other with Green Florescent Protein vector (GFP). As shown in Figure 3.11 B, the transient
expression of FLP indeed removed the inserted cassette based on the detection of a genomic fragment without the inserted cassette. Also, RT-PCR analyses indeed detected the transcription of Peg3 only from the KO cells with FLP transfection, but not from the two controls, confirming the restored expression of Peg3.

Figure 3.11 Expression level analyses of *H19* using the MEF cells with the restored expression of Peg3. (A) Genomic structure of the mutant allele of Peg3 and FLP-mediated recombination scheme to restore the expression of Peg3. The inserted cassette is flanked by two FRT sites, thus can be removed by Flippase (FLP). (B) Three pools of KO MEF cells were individually transfected with the following constructs: no vector as a mock control (lane 1), a Green Fluorescent Protein (GFP) expression vector as a negative control (lane 2), and a FLP expression vector (lane 3). The total RNA isolated from these cells were analyzed with qRT-PCR to measure the expression levels of β-actin and Peg3. The bottom panel shows genotyping results confirming FLP-mediated recombination of the mutant allele (Rev KO) and endogenous allele (WT) of Peg3. (C) The total RNA isolated from the three sets of MEF cells were also used for measuring the expression levels of *H19*. The expression level of *H19* was first normalized with an internal control (*Gapdh*), and the normalized values from the three sets of MEFs were subsequently compared. The values from the MEF cells transfected with GFP and FLP were divided by the value from a Mock control. Finally, these relative values were presented in the graph with standard deviation.
Thus, total RNA was subsequently isolated from three sets of MEF cells for the expression analyses of H19. According to the results from qRT-PCR, the expression levels of H19 in the sample with FLP transfection was indeed down-regulated by 20% relative to that from the sample without vector (Mock) (Figure 3.11 C).

This down-regulation with statistical significance was also observed only from the sample with FLP, but not from the sample with GFP, thus confirming that the observed down-regulation is likely caused by the restored expression of Peg3. This further confirmed the repressor role of PEG3 in the transcription of H19.

The repressor role of PEG3 was further tested using a series of reporter assays (Figure 3.12). A promoterless luciferase vector was used as a basic construct (Basic-Luc), modified first by including the 610-bp-long promoter region of H19 (H19-Pro), and later by including the 152-bp-long target region of PEG3 (WT-152). Transfection of these constructs into HEK293 cells derived the following outcomes.

As expected, the promoter region of H19 dramatically increased the reporter activity, more than 10 fold. On the other hand, the inclusion of the 152-bp target region of PEG3 resulted in 50% reduction in the promoter activity (Figure 3.12 A).

This suggests that the 152-bp region likely contains repressor elements for the transcription driven by the H19 promoter. As a second set of assays, the WT-152 construct were co-transfected with varying amounts of the expression vector producing the protein PEG3 7.
As shown in Figure 3.12 B, the transfection with this expression vector at 5X dosage resulted in almost 40% reduction in the reporter activity compared to that with no expression vector.

Figure 3.12 Testing the repressor role of PEG3 through reporter assays. (A) This series of reporter assays used the following three constructs. The Basic-Luc construct with a promoterless luciferase reporter was modified first by inserting the 610-bp-long promoter region of \textit{H19} (H19-Pro), and later by inserting the additional 152-bp-long target region of PEG3 (WT-152). The reporter activity from these three constructs was summarized and presented as a graph on left. (B) The WT-152 construct was co-transfected with varying amount of the expression vector producing the full-length PEG3 protein. In this series of co-transfection experiments, 10 ng of the expression vector was considered as 1X. The luciferase activity for each sample was first compared to that of the control sample with no expression vector, and this relative value was summarized and presented as a graph. (C) The WT-152 construct was further tested after mutating the potential binding sites of PEG3. The constructs, Mu2a and Mu2b, are identical to WT-152 except that each construct has a 7-bp-long mutation on the critical region for the binding to PEG3. The exact mutation spot for each construct is same as its corresponding duplex mutant used for gel shift assays (Figure 3.10).
This reduction is again consistent with the prediction that PEG3 may function as a repressor for the transcription of H19. As a final set of assays, the target region of PEG3 within the WT-152 construct was first mutated, and later co-transfected with the expression vector of PEG3 (Figure 3.12 C). The results from gel shift assays demonstrated that one 7-bp-long region is critical (Mu2b in Figure 3.10). Thus, this region along with the adjacent 7-bp-long region (Mu2a) were similarly mutated in the WT-152 construct. According to the results, co-transfection of WT-152 with the 5X dosage of the PEG3 expression vector resulted in 30% reduction in the reporter activity. However, the observed reduction was disappeared slightly in the Mu2a construct, but completely in the Mu2b construct (Figure 3.12 C). This indicated that the second 7-bp-long region (Mu2b) is very critical for the repressor role by PEG3. Since this small region is also known to be critical for the binding to PEG3 (Figure 3.10), these results strongly support an idea that the repressor role of PEG3 may be mediated through its direct binding to this particular small region of H19. Overall, the results described above confirmed that the protein PEG3 likely functions as a repressor for the transcription of H19.

3.4 Discussion

In the current study, we have identified a set of 16 genes as potential downstream genes of Peg3 by performing ChIP-seq experiment with MEF cells. The identified genes tend to be highly expressed in oocytes and also during early embryogenesis. Expression analyses further indicated that the majority of these genes were upregulated in the MEF cells lacking the protein PEG3, thus confirming the repressor role for PEG3. Interestingly, the imprinting control region of the H19/Igf2 domain turns out to be one of the downstream targets of Peg3. The target site is specifically localized between the 3rd and 4th CTCF sites of the H19-ICR. According to the results, PEG3 binds to the active maternal allele of the H19-ICR, and also the depletion of PEG3...
resulted in the upregulation of both H19 and Igf2 without any changes in their imprinting status. This confirms again a repressor role for PEG3 in the transcriptional regulation of H19 and Igf2. Overall, the current study uncovers a functional connection between the two imprinted genes, the paternally expressed Peg3 as a trans factor controlling the maternally expressed H19.

The protein PEG3 appears to control a set of 16 potential downstream genes as a DNA-binding transcription factor (Table 1). The results from ChIP-seq provide the following insights regarding the function of PEG3. First, the number of potential downstream genes identified from MEF cells is relatively small compared to those from the other known DNA-binding transcription factors, 16 versus several hundreds (Kim et al., 2007). The expression levels of PEG3 are known to be very high in early-stage tissues and neuronal cells (Kuroiwa et al., 1996). Nevertheless, PEG3 seems to bind to a very small subset of genes in MEF cells. This might be related to the following possibility. PEG3 has been detected not only in the nucleus but also in the cytoplasm, suggesting unknown cytoplasmic functions other than DNA-binding nuclear functions (Jiang et al., 2010; Relaix et al., 1998). It is possible that only a small fraction of the protein PEG3 might function as a DNA-binding protein, especially in MEF cells. As a consequence, this very limited amount of PEG3 may be available for binding and controlling a small subset of genes. Second, the identified downstream genes were shown to be all upregulated by the depletion of PEG3 (Figure 3.2 & Figure 3.4). This appears to be consistent with the observations derived from the previous studies (Thiaville et al., 2013; Ye et al., 2014). The two known downstream genes, Pgm2l1 and another imprinted gene Zim1, were also upregulated in the KO cells lacking PEG3. Therefore, it is most likely that PEG3 may also function as a repressor for the newly identified downstream genes. Third, the identified genes seem to share several features, such as high levels of expression in oocytes and also close association with
human cancers. As described earlier, it is interesting to note that the expression level of Peg3 is very low in both oocytes and some cancers, especially breast and ovarian cancers (Feng et al., 2008). This inverse correlation again supports a repressor role of PEG3 for the identified downstream genes. Furthermore, since human PEG3 has been regarded as a tumor suppressor (Kim et al., 2015), the identified downstream genes with cancer connection might provide potential mechanisms, by which PEG3 may suppresses tumor formation. For instance, MTA3 is very closely associated with breast cancer as an oncogene (Zhang et al., 2006), thus it would be interesting to further characterize a potential connection between PEG3 and MTA3, specifically to test whether PEG3 functions as a tumor suppressor by repressing oncogenic MTA3. In sum, the protein PEG3 appears to be a transcriptional repressor for a small number of genes that are closely associated with either early developmental processes or human cancers.

The results described in the current study also revealed that the imprinting control region of the H19/Igf2 domain is one of main targets of PEG3 (Table 3.4, Figure 3.1). ChIP experiment confirmed the in vivo binding of PEG3 to the maternal allele of the H19-ICR (Figure 3.5). A series of expression analyses further demonstrated the upregulation of H19 and Igf2 in the MEF cells lacking PEG3 (Figure 3.4). In the simplest scenario, this upregulation could be an outcome of the depletion of the transcriptional repressor PEG3 in the MEF cells. However, the depletion of PEG3 seems to have a more complicated and global outcome than expected since several other imprinted genes were also affected (Figure 3.4). In particular, the upregulation of Igf2 is somewhat perplexing given the fact that the expression is still from the paternal allele, which is not bound by PEG3. This observation, however, made sense at the beginning since PEG3 binds to the H19-ICR, which controls both H19 and Igf2. Nevertheless, the upregulation of Igf2 might be an indirect outcome of the depletion of PEG3. This might also be the case for several other
imprinted genes that were also affected, including Igf2r, Peg10 and Grb10. One plausible scenario would be that the observed changes might be reflecting some changes in the proposed imprinted gene network, which might be triggered by the depletion of PEG3 (Varrault et al., 2006). In that regard, it is relevant to point out the fact that Peg3 is thought to be at the center of this proposed network, meaning that Peg3 has the most connections with the other imprinted genes (Varrault et al., 2006). Thus, it is reasonable to predict that the upregulation observed from some of the imprinted genes might be caused by potential disturbance in the imprinted gene network. Overall, the depletion of the protein PEG3 has a global impact on the transcription of several other imprinted genes although it appears to controls only a few imprinted genes, such as H19.

3.5 Reference


CHAPTER 4
CONCLUSIONS AND DISCUSSIONS

According to our current knowledge, less than 200 mouse genes and a few human genes have been identified as imprinted gene (Weksberg, 2010). Most of these imprinted genes are clustered into 16 genomic regions that contain two or more genes (Barlow, 2011). Therefore, two genes with opposite imprinting are often localized right next to each other, such as \textit{H19/Igf2} and \textit{Gtl2/Dlk1}. At present, two genes with opposite imprinting are well-known regulated through shared \textit{cis} regulatory element, such as ICR. For instance, \textit{H19/Igf2} ICR localizes the upstream of \textit{H19} gene, which bound by CTCF to mediate the imprinting expression of these two genes (Singh et al., 2012). CTCF binding sites mutation in the maternal allele of \textit{H19} ICR disrupted imprinted expression and gene transcription, indicating ICR is the critical element to regulate two adjacent genes with opposite imprinting (Szabo et al., 2004).

In the Peg3 imprinted domain, \textit{Peg3} and \textit{Zim1} are two adjacent imprinted genes with opposite imprinting. We hypothesize that Peg3 ICR is responsible for the imprinted expression of these two genes. However, my current results illustrate a new regulatory mechanism that two oppositely imprinted genes can interact through their gene products rather than shared \textit{cis} regulatory element (Ye et al., 2014). Although the precise mechanism of PEG3 action remains unclear, we conclude that PEG3 may repress \textit{Zim1} expression through establishing more repressor marker-H3K9me3 (Ye et al., 2014). To gain insights into PEG3 repression mechanism of action, we further demonstrate that PEG3 might recruit KAP1 which subsequently interacts with methyltransferase (SETDB1) to achieve targeted gene repression. Meanwhile, comparing the results derived from microarray assay analysis between the WT and \textit{Peg3} mutant mice, we found that expression level of a subset of placenta-specific genes elevated in \textit{Peg3} mutant mice,
indicating the a repressor function of PEG3 (Kim et al., 2013). PEG3 is also known as a zinc finger DNA-binding protein and its DNA-binding motif is well characterized by our group (Thiaville et al., 2013). Thus, in this dissertation, we further explore whether PEG3 represses these target genes through direct or indirect interactions with regulatory factors.

Therefore, we performed a ChIP-seq analysis to identify the potential downstream genes of PEG3. Based on these results, 16 genes were identified as the downstream genes of PEG3. Our qRT-PCR analysis further demonstrated that majority of these genes had elevated expression levels in Peg3 mutant MEF cells. These results strongly suggested that PEG3 could repress its downstream genes through direct binding. However, most of the up-regulated genes found from the microarray assay had not been shown as downstream genes of PEG3, which suggested that PEG3 might mainly repress gene expression indirectly using an unknown regulatory mechanism. Interestingly, one of the downstream genes of PEG3 is H19, which is a maternally expressed gene. Moreover, the binding region is a partial region of H19 ICR which has been extensively studied in the past decades by several independent groups. The precise binding sites were further refined and characterized by ChIP and gel shift assays. According to the results, it is interesting to note that the binding region of PEG3 is surrounded by two CTCF binding sites and two potential binding sites for an orphan nuclear receptor family. These results indicate that this small binding region may be a main regulatory center that can attract several transcription factors for transcriptional control of H19/Igf2. qRT-PCR analysis and our transfection results suggested that PEG3 functioned as a repressor to down-regulate H19 expression.

According to my current results, PEG3 protein can directly repress two other imprinted genes (Zim1 and H19), indicating the existence of a proposed “imprinting network” where Peg3 has the most connections with the other imprinted genes (Varrault et al., 2006; Ye et al., 2014).
Therefore, it is reasonable to predict that the upregulation observed from these two imprinted genes might be caused by potential disturbance in the imprinted gene network.

Moreover, a majority of downstream genes of PEG3 are highly expressed in the oocyte, where Peg3 cannot be expressed due to its imprinted maternal allele. This observation suggests that these downstream genes may play a critical role during oogenesis because of their high expression in the oocyte. Moreover, PEG3 mainly functions as a repressor for these genes. Therefore, this observation may explain why Peg3 biologically functions as a paternally expressed gene as the maternal allele expression may disrupt oogenesis. In order to test this hypothesis, overexpression of Peg3 during the process of oogenesis needs to be manipulated and oocyte phenotype and genotype changes need to be further characterized.

Besides the potential enrollment in the process of oogenesis, Peg3 is also known as tumor repressor gene. The human PEG3 has been considered as a tumor repressor gene because of its low expression in cancer tissues and cells, especially in breast and ovarian cancers (Kim et al., 2015; Zhang et al., 2006). However, the molecular mechanisms governing the tumor repressor function of PEG3 remain unclear and unexplored. According to our ChIP-seq result, the identified downstream genes that are associated with cancer may provide potential mechanisms to explain PEG3 tumor repressor function. One of the identified downstream genes of PEG3 that is associated with cancer development is Mta3, which is associated with breast cancer as an oncogene (Fujita et al., 2003). Thus, PEG3’s tumor repressor function, especially during breast cancer development, can be further characterized by a potential connection between Peg3 and Mta3.
Another two downstream genes (Msl1 and Msl3) of PEG3 have been extensively studied in *Drosophila melanogaster* in the past decades, however they are rarely studied in mammals so far. They are two critical genes that encode two important components of male-specific lethal (MSL) complex (MSL1, MSL2, MSL3, MOF, MLE, roX1 and roX2 lncRNAs) in *Drosophila melanogaster* (Gelbart and Kuroda, 2009). The *Drosophila melanogaster* MSL complex binds the single male X chromosome to upregulate gene expression by twofold to achieve dosage compensation (Conrad and Akhtar, 2011). In this complex, MSL1 and MSL2 are essential elements that recognize sites on the X chromosome, whereas MSL3, MOF and MLE are required specifically for the MSL complex to spread from chromatin entry sites (CES) (Gu et al., 1998; Lyman et al., 1997; Palmer et al., 1994). One critical member in this complex (MOF) is well known as a histone acetyltransferase, which can establish activator histone modification markers (H4K16ac) to achieve gene up-regulation (Hilfiker et al., 1997). Chromatin immunoprecipitation combined with microarray (ChIP-chip) results indicated that H4K16ac is broadly associated with the male X chromosome (Gelbart et al., 2009). According to our current results, we predict that PEG3 may function as a repressor for these two important genes. According to Peg3 KO mouse phenotypes, the birth weight of male neonates are always less than that of females, and ChIP-seq reads analysis comparing WT and Peg3 KO input samples suggested 3-fold more reads located in the Peg3 KO Y chromosome than WT Y chromosome. These results indicate that male mice are more susceptible compared to female mice after Peg3 knockout as Y chromosome may lose its compressed structure because of the accumulation of H4k16ac histone marker after Peg3 KO. Even though the molecular mechanism of the potential connection between PEG3 and Y chromosome remains unclear, it is reasonable to state that male mouse susceptible phenotype
may be correlated with MSL complex dysregulation, which subsequently result in Y chromosome decompression through accumulating H4K16ce histone marker.

Taken together, my current researches identify a list of downstream genes of PEG3 and PEG3 mainly functions as repressor for these genes. Potential connections between PEG3 and several well-characterized downstream genes may help contribute to explain PEG3 tumor repressor function and some male specific phenotypes after Peg3 knockout.

4.1 References


VITA

An Ye is the son of Mr. Hegen Ye and Mrs. Meifang Ye. He was born in Pujiang, China, in 1987 and graduated from Pujiang High School in 2005. Five years later, he received M.D. from Hebei Medical University in 2010. In August 2012, he was accepted by the department of Biological Sciences at Louisiana State University and began his doctoral research at Dr. Joomeyong Kim’s lab on mouse epigenetic study. Mr. Ye will graduate with the degree of Doctor of Philosophy in Biological Sciences in August 2016.