Intergenic and intronic DNA hypomethylated regions as putative regulators of imprinted domains

Arundhati Bakshi¹, Corey L Bretz¹, Terri L Cain¹ & Joomyeong Kim^{*,1}

¹Department of Biological Sciences, Louisiana State University, Baton Rouge, LA 70803, USA

* Author for correspondence: Tel.: +1 225 578 7692; Fax: +1 225 578 2597; jkim@lsu.edu

Aim: To investigate the regulatory potential of intergenic/intronic hypomethylated regions (iHMRs) within imprinted domains. **Materials & methods:** Based on the preliminary results of the histone modification and conservation profiles, we conducted reporter assays on the *Peg3* and *H19* domain iHMRs. The *in vitro* results were confirmed by the *in vivo* deletion of *Peg3*-iHMR designed to test its function in the *Peg3* imprinted domain. **Results & conclusion:** Initial bioinformatic analyses suggested that some iHMRs may be noncanonical enhancers for imprinted genes. Consistent with this, *Peg3*- and *H19*-iHMRs showed context-dependent promoter and enhancer activity. Further, deletion of *Peg3*-iHMR resulted in allele- and sex-specific misregulation of several imprinted genes within the domain. Taken together, these results suggest that some iHMRs may function as domain-wide regulators for the associated imprinted domains.

Graphical abstract:

First draft submitted: 2 October 2017; Accepted for publication: 4 December 2017; Published online: 23 March 2018

Keywords: DNA methylation • iHMR • imprinted domains • imprinted gene regulation • imprinted genes • imprinting • intergenic hypomethylated region

Genomic imprinting is a unique epigenetic regulatory mechanism in placental mammals by which a subset of genes are expressed in a parental allele-specific fashion. These genes generally share roles in embryonic and neonatal growth regulation as well as brain development and behavior [1]. A fraction of imprinted genes also shows signs of tumor suppressor activity and susceptibility to epigenetic changes in many cancer states [2,3]. Only a few hundred

genes are subjected to genomic imprinting in the mammalian genome, and they are typically found clustered in relatively conserved domains of about 500 kilobases (kb) to two megabases (Mb) in length [1,4]. Imprinted genes within a domain are known to be co-regulated, based on the juxtaposition of reciprocally imprinted genes, as well as their frequently shared spatial and temporal expression patterns [1,5,6]. This co-regulation is mediated through regulatory elements called imprinting control regions (ICR). ICRs obtain DNA methylation during gametogenesis and maintain the allele-specific DNA methylation in somatic cells throughout the lifetime of eutherians. ICRs are also responsible for the imprinting and transcriptional regulation of entire imprinted domains [1,7]. However, besides ICRs, not much is currently known regarding the *cis*-regulatory elements that impact imprinted gene expression, although other enhancer-like regions have been predicted to direct the spatial and temporal expression patterns of imprinted genes [5,7–10].

In this postgenomic era, many potential regulatory elements have been revealed through high-throughput sequencing. One such class of elements is the intergenic/intronic hypomethylated region (iHMR), which are often smaller, more tissue-specific and better indicative of gene activity than hypomethylated promoter regions. The iHMRs have been characterized to resemble insulator, promoter, enhancer or bivalent elements [11]. In general, insulator-like elements are transcriptionally silent and bound by CTCF [11], with potential roles in chromatin domain organization [12]. Promoter- and enhancer-like iHMRs, on the other hand, are transcriptionally active elements bound by various transcription factors and RNA polymerase II. The enhancer-like iHMRs, however, are differentiated by their tissue-specific hypomethylation, as opposed to the constitutive hypomethylation pattern observed at many promoter- and insulator-like elements [11]. A fair share of each of the classes of iHMRs is found within imprinted domains, where they may act as enhancers, alternative promoters as well as chromatin architecture regulators. The functions of all of the aforementioned categories of *cis*-regulatory elements have been previously linked with the regulation of imprinting mechanism and/or imprinted gene expression patterns [13–17]. Moreover, many of these elements also exhibit epigenetic aberrations in cancers where imprinted genes tend to be affected [2,3]. Therefore, identifying the role of iHMRs in imprinted domains could lead to a further understanding of how imprinted genes are regulated in normal as well as clinically relevant states such as in cancer.

A few iHMRs have been previously characterized as evolutionarily conserved elements in the *Peg3*, *H19* and *Dlk1* domains. All three iHMRs are constitutively hypomethylated, and two of them bind CTCF. Yet, these iHMRs appear more enhancer-like than insulator-like in terms of their epigenetic and functional states. The list of these elements includes ECR18 (evolutionary conserved region 18) within the *Peg3* domain [8,9], *H19*-CCD (central conserved domain) within the *H19/Igf2* domain [18] and *Dlk1*-CE4 (conserved element 4) in the *Gtl2/Dlk1* domain [10]. Given the unusual co-regulation patterns of imprinted genes, we hypothesized that other such noncanonical elements may also exist within imprinted domains that are involved in properly orchestrating gene expression from opposite parental alleles. Thus, here we have surveyed five additional iHMRs within imprinted domains that share hypomethylation patterns across multiple tissues, similar to the three previously described noncanonical enhancer-like iHMRs. Analysis of the eight total elements suggest that many of these iHMRs may be transcriptional regulators within imprinted domains, despite their nontraditional presentation. Moreover, CRISPR/Cas9-based deletion of the intronic HMR, ECR18, affected multiple genes in the *Peg3* domain, suggesting that some iHMRs may even be regulators impacting the transcription of the entire imprinted domains.

Materials & methods

Bioinformatic analyses of iHMRs in imprinted domains

DNA methylation and histone modifications were visualized for eight intergenic/intronic hypomethylated regions selected for this study using the publicly available datasets on UCSC Genome Browser (Supplementary Figure 1) [19,20]. Based on placental mammal conservation data (PhyloP track) from UCSC Genome Browser, the most conserved 100–500 bp segment of the iHMR was selected for further analyses (Supplementary Material 1) [21]. Data from the MultiZ alignment tool in UCSC Genome Browser [22] were harnessed to retrieve orthologous sequence for each iHMR for all available placental mammals. Additionally, the BLAT program was used to identify orthologous iHMR sequences that had not been detected through the MultiZ alignment. Each of the sequences were then aligned against the reference mouse sequence using blastn, and the percentage identity (ID) recorded for all alignments ≥ 100 bp in length. Next, the alignments were manually surveyed for conserved DNA motifs, based on chromatin immunoprecipitation (ChIP)-seq datasets from various mouse and human tissues, available on UCSC Genome Browser [23–26]. Additionally, for five highly conserved iHMR, orthologous sequences from five species – mouse, human, cat, horse and cow – were analyzed using DCODE to identify conserved transcription

Figure 1. Intergenic/intronic hypomethylated regions in imprinted domains. (A) The various intronic hypomethylated regions of the *Gnas* domain are depicted by dark blue bars [19,20]. The iHMR selected for this study is marked by a light blue bar and labeled as 'Gnas-HMR'. Light blue bars also mark the other iHMRs selected for this study: (B) *Peg3*-ECR18 (*Peg3* domain); (C) *Ndn*-HMR (*Ndn* domain); (D) *H19*-HMR (*H19* domain). (E) The four iHMRs selected from the *Dlk1* domain are also marked and labeled as shown for the *Gnas* domain in (A).

HMR: Hypomethylated regions; ICR: Imprinting control region; iHMR: Intergenic/intronic HMR.

factor-binding motifs [27]. The phylogenetic tree in Figure 4 was downloaded in Newick format from UCSC Genome Browser and visualized using the PlyoPng program.

Combined Bisulfite Restriction Assay

Genomic DNA was isolated using phenol–chloroform (Invitrogen, Waltham, MA, USA) extraction, followed by ethanol precipitation, from 14.5-days post coitum (dpc) whole embryo, 17.5-dpc embryo head and the following adult tissues: brain, tail, kidney, spleen, lung and liver. Approximately 1 µg of genomic DNA was subjected to the sodium bisulfite treatment using the EZ DNA Methylation™ kit (Zymo Research, Irvine, CA, USA). Next, 1 µl (~20 ng) of the bisulfite-converted DNA was amplified using several primer sets (Supplementary Material 2). The PCR products were then digested with appropriate restriction enzymes to monitor the methylation level at the CpG

Figure 1. Intergenic/intronic hypomethylated regions in imprinted domains (cont.). (A) The various intronic hypomethylated regions of the *Gnas* domain are depicted by dark blue bars [19,20]. The iHMR selected for this study is marked by a light blue bar and labeled as 'Gnas-HMR'. Light blue bars also mark the other iHMRs selected for this study: (B) *Peg3*-ECR18 (*Peg3* domain); (C) *Ndn*-HMR (*Ndn* domain); (D) *H19*-HMR (*H19* domain). (E) The four iHMRs selected from the *Dlk1* domain are also marked and labeled as shown for the *Gnas* domain in (A).

HMR: Hypomethylated regions; ICR: Imprinting control region; iHMR: Intergenic/intronic HMR.

sites within their recognition sequences. In order to compare differences in DNA methylation between normal and tumor tissues, each sample was tested at least three-times (starting with the bisulfite-conversion process). The ImageJ program was subsequently used to quantify the relative band densities derived from the methylated and unmethylated DNA fractions in each sample.

Chromatin immunoprecipitation

Chromatin prepared from mouse embryonic fibroblast (MEF) cells [28] was sonicated and divided into three fractions. The first fraction was saved as input control. The second fraction was precleared with protein A/G PLUS-agarose beads (Santa Cruz Biotechnology, Dallas, TX, USA), and treated with 5 µg of either anti-Pol II (N-20: sc-899X, Santa Cruz Biotechnology) or anti-CTCF antibody (catalog number 07-729, Millipore Corporation, Darmstadt, Germany). The third fraction was precleared, but without any antibody, thus serving as the no-antibody (negative) control. The antibody/protein/DNA complex was precipitated using protein A/G PLUS-agarose beads. Subsequently, all three fractions were decrosslinked, and the DNA were recovered by phenol-chloroform (Invitrogen) extraction followed by ethanol precipitation.

One microliter DNA from all three fractions was then amplified using primers (Supplementary Material 2) that were designed to flank the predicted transcription factor-binding motif and quantified by ChIP-qPCR

Figure 2. Epigenetic profiling of intergenic/intronic hypomethylated region in imprinted domains. (A) Hypomethylation status of the iHMRs in the respective tissue is shown in blue. Although some iHMRs are constitutively hypomethylated, others are hypomethylated at various levels of tissue specificity. (B) Histone modification at iHMRs. Most of the iHMRs are marked by active histone marks (dark blue – H3K4me1; gold – H3K27ac; green – H3K4me1 and H3K27ac; magenta – H3K4me3 and H3K27ac) in a tissue-specific manner. ES: Embryonic stem; HMR: Hypomethylated regions; iHMR: Intergenic/intronic HMR.

Figure 3. DNA methylation of imprinted domain intergenic/intronic hypomethylated region in mice and humans. Methylation level of the iHMRs were analyzed with COBRA in: (A) five adult mouse tissues; (B) two embryonic stages; (C) within murine thymic lymphoma [3]; (D) at the orthologous region of the iHMRs in humans. Boxes underneath COBRA images indicate hypermethylation (red), hypomethylation (green) and no significant change (gray) in tumor tissues relative to the normal control. The band representing the methylated fraction of the DNA is marked with a red 'M' and the unmethylated fraction with a blue 'U'. Br-C: Breast cancer (unmatched); Br-mat-N: Matched breast normal; Br-mat-T: Matched breast tumor; COBRA: Combined Bisulfite Restriction Assay; E: Early stage; iHMR: Intergenic/intronic hypomethylated region; L: Late stage; Lu-C: Lung cancer (unmatched); Lu-mat-N: Matched lung normal; Lu-mat-T: Matched lung tumor; M: Middle stage.

Figure 4. Conservation profile of intergenic/intronic hypomethylated regions among placental mammals. The grayscale profile (darker the gray, higher the percent identity as compared with the reference [mouse]) shows that the iHMRs are relatively well conserved overall, and harbor some conserved DNA-binding motifs, such as CTCF ('C') and E-box ('E'/'e'). Upper case 'E' indicates that the actual sequence within the motifs are conserved in the respective species as compared with mouse. Lower case 'e' indicates the presence of a different version of the DNA-binding motif.
iHMR: Intergenic/intronic hypomethylated region.

(QuantStudio 6, Invitrogen). The Ct values for each amplicon derived from the no-antibody control (Neg) and immunoprecipitated DNA (Ab) were normalized against those from the input.

Reporter assays

A promoterless luciferase reporter construct [9] was modified by cloning *Peg3*-ECR18 (chr7:6866573–6866967) or a 1649 bp region of the *H19*-HMR (chr7:149796116–49797764) in both forward and reverse orientations upstream of the *luc2* gene. Enhancer activity for *Peg3*-ECR18 [9] and *H19*-HMR was tested by cloning the regions (as described for the promoter assay) upstream of the main promoter from each of their imprinted domains (*Peg3* and *H19*, respectively) in the *luc2* reporter construct. HEK 293T, HeLa, Neuro2A and 3T3 cells were grown in a T-75 flask with Dulbecco's modified Eagle medium Plus GlutaMAX medium (DMEM; Fisher Scientific, Pittsburgh, PA, USA) containing 10% fetal bovine serum (FBS; Fisher Scientific) and 1% antibiotic-antimycotic (A-A; Corning, Tewksbury, MA, USA). Cells were co-transfected with 2 µg of a luciferase reporter construct in a 6-well plate along with 2 µg of a β-geo reporter construct to monitor transfection efficiency. Transfections for HEK 293T, HeLa and Neuro2A cells were performed with 10 µl Lipofectamine 2000 (Invitrogen). Transfection medium (DMEM without FBS and A-A) was replaced with fresh culture medium (DMEM with 10% FBS, 1% A-A) 5.5 h post-transfection. The 3T3 cells were transfected using 6 µl GenJet *in vitro* DNA Transfection Reagent for 3T3 Cells (SignaGen Laboratories, Rockville, MD, USA), according to the manufacturer's protocol. Three independent transfections were performed for each set of reporter constructs in all four cell lines. The cells were subsequently harvested 48 h post-transfection in 250 µl of the reporter lysis buffer (0.1 M Tris-Cl, pH 7.8, 0.1% NP-40). Due to very low number of 3T3 cells, the three replicate transfections were combined when harvesting the cells (in total, 250 µl reporter lysis buffer) for the luciferase and β-galactosidase assays. The luciferase assays were conducted using a commercial reagent according to manufacturer's protocol (Promega, Fitchburg, WI, USA), and the readings were normalized against β-galactosidase activity. The graphs in Figure 6 show one representative transfection experiment

Figure 5. CTCF and RNA polymerase II binding at imprinted domain intergenic/intronic hypomethylated regions. Top: enrichment (Ab) for CTCF (dark green) and Pol II (dark blue) is shown as a percentage of the input DNA, compared with the no-antibody control (Neg; CTCF, light green; Pol II, light blue). Error bars show standard deviation between technical replicates. Bottom: significant enrichment of antibody (Ab) over control (Neg) is shown as a filled box for CTCF (green) and Pol II (blue). All normalized enrichment <1% of input DNA (gray bar) was ignored from the analysis.

* $p \leq 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$; ***** $p < 0.00001$ (Student's t-test).
N.D.: Not detected.

for HEK 293T, HeLa and Neuro2A cells and the combined readings of the three independent transfections for 3T3 cells.

Generation of the ECR18 Δ allele

A CRISPR/Cas9-based scheme was used to target a 395-bp genomic interval (mm9, chr7: 6,866,573–6,866,967) termed ECR18 for deletion. In brief, 200 fertilized eggs were isolated from time-mated C57BL/6J mice. The isolated eggs were injected with two single-stranded guide RNAs (5' guide: 5'-TGATCGATCATCACGATCACGG-3' and 3' guide: 5'-TCCACAAGCACTACTCCTCACGG-3') along with Cas9 mRNA. The injected eggs were then implanted into the uteruses of six pseudo-pregnant females, which derived 39 live pups. The pups were screened for CRISPR-mediated deletion by PCR genotyping using the following primers: P1, 5'-GCAATCTTCTCCCCGACTC-3'; P2, 5'-AGATCACATTTCAGGGGC-3'; P3, 5'-ACACCCGGAGCTATGAATGC-3'. PCR genotyping identified one female mouse with CRISPR-mediated deletion of ECR18. The genotyping PCR products were then cloned and sequenced to confirm the actual deletion of ECR18. The sequenced PCR products were aligned to the mouse genome using the UCSC genome browser, which revealed that the CRISPR-based deletion scheme removed a 273-bp region of ECR18 (mm9, chr7: 6,866,588–6,866,861). The founder female mouse containing the deletion was bred to a C57BL/6J male mouse to establish the ECR18 Δ mouse line. F1 progeny were bred to transmit the ECR18 Δ allele through both the maternal and paternal germ lines.

Gene expression analyses

Total RNA was extracted from the mouse heads of 1-day-old neonates using the Trizol Reagent (Invitrogen). Approximately 3.5 μ g total RNA was converted to cDNA using random primers and M-MuLV Reverse Transcriptase (NEB, Ipswich, MA, USA). Next, 1 μ l of 0.5 $\times$ (diluted) cDNA was used for quantitative real-time PCR (QuantStudio 6, Invitrogen) using SYBR-green (Bio-Rad, Hercules, CA, USA). Three biological replicates for each of the sexes were processed for wild-type (WT) and paternal or maternal deletion of *Peg3-ECR18*. Expression levels of two internal controls, β -actin and *Gapdh*, were measured alongside those of *Peg3*, *Usp29* and *Zim1* in neonate heads. Ct values for all genes (three technical replicates for each animal) were independently normalized, using

Figure 6. Transcriptional regulation by *Peg3*-ECR18 and *H19*-HMR. Two constructs were used to test promoter activity of the iHMR, in forward (dark colors in [B & C]) and reverse (light colors in [B & C]) orientations (A). *Peg3*-ECR18 (B) and *H19*-HMR (C) showed promoter activity in HEK293T (orange) and 3T3 (green) cells, but not in HeLa (gold) or Neuro2A (blue) cells. Luciferase activity from two constructs (D), without (hatched bars in [E & F]) and with (solid bars in [E & F]) the iHMR cloned upstream of *luc2* that was driven by the corresponding promoter in each domain. *Peg3*-ECR18 showed enhancer activity in HEK 293T, HeLa and NIH 3T3 cells (E), whereas *H19*-HMR (F) upregulated *H19* in HEK293T but downregulated *H19* in 3T3 cells. All luciferase activity is shown relative to that of the promoterless *luc2* (dotted line at 1). Error bars show standard deviation between technical replicates. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$; ***** $p < 0.00001$ (Student's t-test). iHMR: Intergenic/intronic hypomethylated region.

the $2^{\Delta\Delta C_t}$ method [29], against both internal controls to determine the accuracy of fold changes observed. Relative expression levels of *Peg3*, *Usp29* and *Zim1* were pooled for all replicates (technical and biological) of a particular sex and genotype; statistical significance of expression level differences between WT and mutant was established using the Mann–Whitney–Wilcoxon test ($p < 0.05$; RStudio). All mouse experiments were performed in accordance with NIH guidelines for care and use of animals and also approved by the Louisiana State University Institutional Animal Care and Use Committee, protocol #16-060.

Results

Identification of iHMR as potential enhancers for imprinted domains

Whole-genome bisulfite sequencing has revealed several iHMRs that do not correspond to a known promoter or enhancer element. We visualized these regions located within five imprinted domains with the UCSC Genome Browser showing DNA methylation profiles of adult mouse tissues, placenta [20] and embryonic stem (ES) cells (Figure 1) [19]. We detected the largest number of iHMR [23] in the approximately 50 kb intronic region spanning *Gnas* and *Nespas* (*Gnas* domain), 22 in the ~200 kb intron of *Usp29* (*Peg3* domain) and 21 in the ~100 kb intergenic region upstream and downstream of *Dlk1* (*Gtl2/Dlk1* domain). Five iHMRs were noted in the ~50 kb intergenic spaces between *Ndn*, *Magel2* and *Mkrn3* (*Ndn* domain), and eight in the ~40 kb region between *H19* and *Igf2* (*H19* domain) (Supplementary Figure 1). A large majority of the iHMR showed tissue-specific hypomethylation patterns; however, a few of them showed hypomethylation in a majority of the tissues studied. Although the former category has been previously shown to contain enhancer-like characteristics [11], the latter has remained largely uncharacterized. Therefore, here we largely focused on those iHMR, which showed shared hypomethylation patterns across multiple mouse tissues (*Gnas*-HMR, as an example in Figure 1A).

Eight representative iHMRs were selected from the following imprinted domains – *Gnas*, *Peg3*, *Ndn*, *H19* and *Dlk1* – and named based on the imprinted domain they were located in. The genomic location of each iHMR, along with the nomenclature, is shown in Figure 1A–E. Since the *Peg3*-HMR was initially identified as ECR18, we retained the established terminology. According to whole-genome bisulfite sequencing (Figure 2A), four iHMRs (*Peg3*-ECR18, *Ndn*-HMR, *H19*-HMR and *Dlk1*-HMR2) were hypomethylated in all adult tissues as well as placenta. The remaining four iHMRs (*Gnas*-HMR, *Dlk1*-HMR1, -HMR3 and -HMR4) were hypomethylated in a majority of the tissues. Three of the iHMRs selected showed a constitutive hypomethylation pattern (hypomethylated in ES cells as well as all adult tissues): *Ndn*-HMR, *H19*-HMR and *Dlk1*-HMR2 (Figure 2A).

In order to understand the potential functional significance of these iHMRs, we surveyed the histone modifications at these regions [23]. The iHMRs were largely marked by H3K4me1 (monomethylation of lysine 4 on histone H3) and/or H3K27ac (acetylation of lysine 27 on histone H3) (Figure 2B), indicating a potential enhancer function for these regions [30]. We noted a unique case of H3K4 trimethylation on *Peg3*-ECR18 in the embryonic heart tissue, suggesting tissue-specific promoter activity for this region. The greatest number of active histone marks was observed in E14.5 tissues (heart, limb, brain and liver), placenta and adult olfactory bulb. This is consistent with the fact that imprinted genes are mainly expressed during the embryonic and neonatal stages [7]. Therefore, the overall histone profile suggested that the iHMRs may be tissue- and stage-specific enhancers for imprinted genes in their respective domains.

DNA methylation status of iHMR in normal & tumor tissues

In order to validate the hypomethylation status of the iHMRs, we performed independent COmbined Bisulfite Restriction Assay (COBRA) [31] analyses using the DNA from the six adult tissues representing the three germ layers. According to the results (Figure 3A & Supplementary Figure 2A), five iHMRs appeared completely unmethylated in most adult tissues (*Peg3*-ECR18, *Ndn*-HMR, *Gnas*-HMR, *Dlk1*-HMR1 and -HMR3), whereas the remainder showed partial methylation patterns. The iHMRs also showed some tissue-specific methylation. Despite being generally hypomethylated, *Ndn*-HMR (Figure 3A) and *Peg3*-ECR18 (Supplementary Figure 2A) showed minor levels of methylation in the brain; *H19*-HMR, *Dlk1*-HMR2 and -HMR3, on the other hand, showed higher methylation levels in the liver, compared with the other tissues (Figure 3A). In order to test whether the hypomethylation status of the loci in the adult tissues is set up earlier during development, we performed COBRA analyses on embryonic tissues at 14.5 and 17.5 dpc. The iHMR loci appeared to have either none or very low levels of DNA methylation at both embryonic stages (Figure 3B & Supplementary Figure 2B); thus indicating that the hypomethylation status at the iHMRs is likely established during embryonic development and maintained in the adult tissues. In sum, this set of DNA methylation analyses largely agreed with the next-generation sequencing (NGS)-based results (Figure 2A), and confirmed the hypomethylation of the eight imprinted domain iHMRs selected for this study. Furthermore, the COBRA data revealed subtle tissue-specific differences in methylation levels at the iHMRs, and a tendency for greater methylation at some loci in the brain and liver.

Previously, it had been suggested that potential enhancer regions in imprinted domains (such as *Peg3*-ECR18) may be epigenetically unstable in mouse thymic lymphoma [3]. Thus, we tested the methylation levels of the remaining iHMRs at three stages of murine thymic lymphoma driven by a *Kras*^{G12D} mutation [3]. Besides *Peg3*-ECR18 (Supplementary Figure 2C), five other iHMRs also showed aberrant DNA methylation levels in the thymic neoplasm (Figure 3C). For the most part, the iHMRs tended toward hypermethylation in the cancer tissue; however, two iHMRs, *H19*-HMR and *Dlk1*-HMR1, which were partially methylated in the normal thymus, became hypomethylated in the lymphoma. Given that the open chromatin at hypomethylated regions would be susceptible to global epigenetic changes in cancer, it is interesting to note that *Ndn*-HMR and *Gnas*-HMR showed no change and remained completely unmethylated in all tumor samples (Figure 3C). This suggests that the epigenetic instability at imprinted domain iHMRs may not be simply an outcome of global changes to the epigenome, but rather a functional response to the carcinogenesis process.

Imprinted genes have also been implicated in various human cancers [2]. In particular, various epigenetic changes have also been noted at the associated ICRs and ECRs, including at the human ortholog of *Peg3*-ECR18 [2]. Thus, we assessed the methylation levels of the human orthologous regions of the iHMRs in six human samples: two pairs of breast and lung tumors along with their matched normal tissues and two unmatched samples of breast and lung cancer. The orthologous regions for the iHMRs in humans appeared largely hypomethylated in the normal human breast and lung tissue, albeit with partial methylation at all loci except h*Ndn*-HMR, h*H19*-HMR and h*Dlk1*-HMR2. Besides ECR18 (Supplementary Figure 2D), five other iHMRs showed aberrant DNA

methylation in at least one of the human tumor samples (Figure 3D). The h*Gnas*-HMR locus appeared the most sensitive, responding to all four breast and lung tumor samples. At least three loci, h*Gnas*-HMR, h*Dlk1*-HMR1 and -HMR3, were deemed capable of responding specifically to carcinogenesis in the breast and/or lung, since they showed aberrant methylation levels in at least one of the matched tumors. Two loci, h*Ndn*-HMR and h*H19*-HMR, showed no change in DNA methylation in any of the tumors, remaining completely unmethylated in all normal and tumor samples (Figure 3D). Taken together, the seemingly targeted, nonglobal nature of epigenetic changes observed at the human orthologs of the iHMRs, suggest that these elements may be functionally responsive to cancer states in humans as well as in mice.

Conserved putative regulatory motifs found within iHMR

In order to further characterize potential functions of the iHMRs, we assessed the overall sequence conservation levels at the hypomethylated loci in 40 placental mammals. Since many of the iHMRs spanned several kb, the most conserved 100–500 bp segment was selected based on the conservation data available from the UCSC Genome Browser [21]. The mouse iHMR sequence (Supplementary Material 1) was used as a reference to harness the orthologous sequences in the other mammals, using the MultiZ alignment tool [22] and BLAT. The sequences were then aligned against the mouse sequence, using the blastn program. The sequence identity (% ID) for each alignment ≥ 100 bp was recorded, and has been represented as a grayscale in Figure 4. Except for *Ndn*-HMR and *Dlk1*-HMR4, the iHMRs appeared conserved in a large number of mammalian species (Figure 4). Among the conserved iHMRs, *H19*-HMR and *Dlk1*-HMR3 showed highest level of sequence identity in humans ($\sim 88\%$) and *Dlk1*-HMR2 showed the lowest ($\sim 68\%$). The percentage ID between mouse and human for the remaining three conserved iHMRs ranged from approximately 73–79% (Supplementary Figure 3A). Overall, six out of eight imprinted domain iHMRs showed approximately 70–90% sequence identity between mice and humans, indicating high degrees of sequence conservation at these regions.

Next, we overlapped publicly available ENCODE ChIP-seq data to identify any DNA-binding factors shared among the iHMRs. We found enrichment for CTCF at *Gnas*-HMR, *Ndn*-HMR, *H19*-HMR and *Dlk1*-HMR2 in a majority of cell types, suggesting that they may be constitutive CTCF-binding sites [12]. Examination of the ChIP-seq data from the human ENCODE project also revealed CTCF, SMC3 and RAD 21 (cohesin complex factors) enrichment at the orthologous regions in humans. Furthermore, the identical CTCF-binding motif shared by h*Gnas*-HMR and m*Gnas*-HMR was found in 26 of the 40 (65%) placental mammals (Supplementary Figure 3B). Similarly, the CTCF-binding motifs at *H19*-HMR and *Dlk1*-HMR2 were also conserved, with approximately 63 and 45% of the placental mammals, respectively, also harboring the identical sequence ('C' in Figure 4). We further tested potential binding of CTCF to the iHMR through an independent series of ChIP experiments using MEF cells (Figure 5). The results confirmed the binding at the four loci, whereas the remaining iHMRs showed no significant enrichment (Figure 5). Together, these data suggest that CTCF may be the most obvious DNA-binding protein that is involved in some unknown functions of the iHMR.

We also used the following bioinformatic approach to identify additional transcription factor motifs. We selected five mammal sequences for each iHMR representing five mammalian orders, which were then used for predicting potential binding motifs for known DNA-binding proteins with the DCODE program [27]. One of the motifs conserved across multiple iHMRs turns out to be the E-box (enhancer-box); all except *Ndn*-HMR displayed some form of the generic motif (CANNTG; 'E'/'e' in Figure 4 & Supplementary Figure 3). Therefore, we conducted manual inspection of the 40 mammalian sequences to detect its presence within each iHMR. The E-box was particularly conserved at *Gnas*-HMR and *Dlk1*-HMR3. The E-box motif is detected at these loci in approximately 85% of the species surveyed, with most species harboring the identical sequence as mouse ('E' in Figure 4 & Supplementary Figure 3B). In some species, however, the exact sequence of the E-box varied from that in mouse, but a generic motif for the E-box was still present ('e' in Figure 4). Interestingly, despite the poor overall conservation status of *Dlk1*-HMR4, several placental mammals nonetheless harbor some form of the E-box motif in that region. Consistent with a conserved E-box, we discovered putative-binding sites for E2A, MYOD and AP4 at *Peg3*-ECR18 [8], *Dlk1*-HMR1, and *Gnas*-HMR (Table 1 & Supplementary Material 3), which are known to bind to the E-box motif. Overall, these surveys identify the E-box motif as another frequent motif within the imprinted domain iHMR.

Besides CTCF and E-box motifs, we also discovered other conserved motifs within the iHMRs (Table 1 & Supplementary Material 3). Both *Dlk1*-HMR1 and -HMR2 showed putative binding domains for GATA-binding factors, which are known to be over-represented at enhancer elements [32,33]. Putative motifs for hormone receptors

Table 1. Conserved putative transcription factor-binding motifs in imprinted domain intergenic/intronic hypomethylated region.

Locus	Conserved putative transcription factor motifs
Gnas-HMR	YY1, E12, MYOD, TAL1
Peg3-ECR18 [†]	MYOD, AP4, E2A, PITX2
H19-HMR	ATF1, ATF3, ATF6, ATF2/CREBP1, T3R, CREB, CREBP1CJUN
Dlk1-HMR1	E2A, MYOD, AP4, MYOGENIN, MEF2, MEF2A/RSRFC4, GATA1, GATA2, GATA3, NF1
Dlk1-HMR2	GATA1, GATA2, GATA3, GATA6, WT1
Dlk1-HMR3	AR, PR, GR, POU3F2, POU1F1, PAX6, PAX8, HNF1, IPF1/PDX1, LHX3, IRF1, NKX2-5

A detailed summary of the results from DCODE analysis is provided as Supplementary Material 3.
[†]From reference [8] (not part of DCODE analysis in this study).

(e.g., AR, GR) were found on *Dlk1*-HMR3, along with motifs for neuronal proteins such as POU1F1, PAX6 and PAX8. Motifs for sex hormone binding in a potential regulatory region for an imprinted gene is intriguing, since many of them show a sexually dimorphic expression pattern [34]. Moreover, the putative-binding motifs for the neuronal proteins are also interesting, given that *Dlk1* is functionally important for neurogenesis [35,36]. Finally, *H19*-HMR showed various cAMP response element-associated factors (e.g., CREB, ATF1 and ATF3), along with putative-binding sites for T3R and the enhancer-binding protein JUN (an AP1 transcription factor subunit). Altogether, the iHMRs showed several conserved putative binding domains for enhancer-associated transcription factors [32]. Furthermore, many of the transcription factors, whose putative DNA-binding motifs are conserved at the iHMRs, appear to be functionally relevant for the regulation of imprinted genes in the domain. These findings, along with their epigenetic profile (Figure 2), support the hypothesis that some of these iHMRs may be potential enhancer elements within imprinted domains.

Transcriptional regulation by Peg3-ECR18 & H19-HMR

To further test whether the iHMRs act as regulatory elements in imprinted domains, we conducted ChIP-qPCR for RNA polymerase II (Pol II) in MEF cells (Figure 5). According to the results, five iHMRs were significantly enriched for Pol II: *Peg3*-ECR18, *H19*-HMR, *Dlk1*-HMR2, HMR3 and HMR4. Of those, we could successfully clone two loci, *H19*-HMR (also bound by CTCF) and *Peg3*-ECR18 (not bound by CTCF), in order to conduct reporter assays *in vitro*.

Enhancer elements have been previously shown to harbor transcriptional potential [37,38]. Therefore, we first conducted a series of promoter assays for the two iHMRs, using a set of constructs designed to compare promoterless-*luc2* basal expression with the two iHMRs cloned upstream of *luc2* in both orientations (Figure 6A). According to the results, both *Peg3*-ECR18 and *H19*-HMR showed the ability to drive luciferase expression in HEK 293T and 3T3 cells. Although ECR18 appeared to have promoter functionality in an orientation-independent manner in both cell lines, *H19*-HMR was only able to drive transcription in the forward orientation in HEK293T cells (Figure 6B & C). However, neither of the iHMR showed any significant promoter activity in HeLa or Neuro2A cells in either orientation. Thus, the promoter assays concluded that the two iHMRs possess the ability to drive transcription, albeit in a tissue-specific manner. Moreover, the orientation-independent capacity of *Peg3*-ECR18 and *H19*-HMR to initiate transcription in HEK293T and/or 3T3 cells may be reminiscent of the bidirectional transcription observed at many enhancers [37,38].

Next, we directly tested the potential enhancer/repressor activity of the loci by comparing luciferase expression without and with the iHMRs cloned upstream of *luc2*, driven by the main promoter from each of the imprinted domains (Figure 6D). The results suggested that ECR18 and *H19*-HMR may possess regulatory capacity, but act in a context-dependent manner. We observed significant upregulation of *Peg3* in HEK 293T, HeLa and 3T3 cells with the upstream ECR18 (Figure 6E). *H19*-HMR, on the other hand, influenced *H19* expression levels differently based on the cell line. In HEK293T, the upstream HMR acted as a transcriptional enhancer for *H19*, whereas it had a repressive effect in 3T3 cells (Figure 6F). We were unable to determine the regulatory potential of *H19*-HMR in Neuro2A cells because of the lack of *H19* promoter activity in this cell line. Overall, the reporter assays demonstrated the *cis*-regulatory potential of *Peg3*-ECR18 and *H19*-HMR, and highlighted the context-dependent nature of their activity.

Figure 7. *In vivo* regulatory function of *Peg3*-ECR18 in *Peg3* imprinted domain. Relative expression levels of *Peg3*, *Usp29* and *Zim1* (normalized to β -actin) for: (A & C) WT (dark blue) versus *Peg3*-ECR18 deletion on the paternal allele (light blue); and (B & D) WT (dark red) versus *Peg3*-ECR18 deletion on the maternal allele (light red). Deletion of *Peg3*-ECR18 affected multiple genes in the *Peg3* domain in an allele- and sex-specific manner. Three biological replicates (individual bars) were tested for males (A & B) and females (C & D) of each genotype. All expression levels are shown relative to that of the three WT animals (dotted line at 1). Error bars show standard deviation between technical replicates for each biological replicate.

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$; ***** $p < 0.00001$.

DELm: Maternal transmission of the deletion; DELp: Paternal transmission of the deletion; WT: Wild-type.

Sex-specific regulation of the *Peg3* domain by ECR18 *in vivo*

Potential *in vivo* functions of the iHMR were further tested using a mutant mouse line, in which ECR18 was deleted with the CRISPR/Cas9 genome editing protocol. For this series of analyses, we derived two sets of mutant mice in the following manner. Male heterozygotes for the deletion of ECR18 were bred with female WT littermates to derive the pups with the paternal transmission of the deletion (DELp), while the reciprocal cross yielded the pups with maternal transmission of the deletion (DELm). These two sets of pups were used to test the potential functions of ECR18 in an allele-specific manner. Since ECR18 was previously predicted to be active in the neonate brain [9], total RNA was isolated from neonate heads for subsequent cDNA synthesis. The expression levels of the three imprinted genes within the *Peg3* domain, which are highly expressed in neonatal heads [9,39–42], were measured as a proxy for ECR18 activity (Figure 7).

According to qRT-PCR analyses, deletion of ECR18 on the paternal allele in males resulted in significant downregulation of the paternally expressed genes, *Peg3* (~50%) and *Usp29* (~60%) (Figure 7A). Downregulation of the maternally expressed *Zim1* was inconsistent across the three DELp male biological replicates, indicating that it may be affected by extraneous factors, not related directly to the ECR18 deletion. However, the maternal deletion of ECR18 resulted in consistent and significant downregulation of *Zim1*, by about 60%, in all three male biological replicates (Figure 7B). In contrast, the paternally expressed genes were less affected in the DELm samples; *Peg3* showed no difference in gene expression, and *Usp29* was downregulated by approximately 30% relative to WT. These results strongly suggest that ECR18 functionally interacts with the promoters of *Peg3*/*Usp29* and *Zim1* regulating their transcriptional activity in an allele-specific manner. However, since no effect of ECR18 deletion on either allele was observed in females (Figure 7C & D), we concluded that the modulation of imprinted genes by

ECR18 is functional only in males. Taken together, the *in vivo* regulatory activity of *Peg3*-ECR18 suggest that the iHMR may be involved in domain-wide regulation of imprinted genes in an allele- and sex-specific manner

Discussion

Recent whole-genomic bisulfite sequencing [19,20] have revealed several intergenic hypomethylated regions, whose functional significance remain largely unknown. Here we have focused on a subset of intergenic and intronic hypomethylated regions in imprinted domains (Figures 1 & 2A), which show hallmarks of regulatory elements. A majority of them bear enhancer-associated histone marks in embryonic tissues known to express imprinted genes (Figure 2B) [39,40,43–47] and are occupied by RNA polymerase II in MEF cells (Figure 5). Furthermore, most of the iHMRs harbor various conserved putative transcription factor-binding motifs that are associated with enhancers (Figure 4 & Table 1), and half of them appear to be constitutive CTCF-binding sites. A series of functional assays further suggested that these elements have transcriptional-regulatory potential (Figure 6) and may be domain-wide modulators of imprinted gene expression (Figure 7).

Imprinted genes are typically clustered within domains, where reciprocally imprinted genes are frequently co-regulated. However, much of the mechanisms through which these genes are regulated remain uncharacterized. Data from the ENCODE project have yielded many useful hints, and a few intergenic regions within imprinted domains, such as *H19*-CCD (overlapping with *H19*-HMR) and *Dlk1*-CE4 (overlapping with *Dlk1*-HMR2), have been shown to regulate *Igf2* [18] and *Dlk1* [10], respectively. Similarly, a survey on a number of ECRs in the *Peg3* domain concluded that they might harbor some regulatory potential as well. However, apart from ICRs, few other classes of regulatory elements have been characterized that have domain-wide regulatory functions. Here we show that the deletion of a conserved intronic HMR in the *Peg3* domain, ECR18, resulted in allele-specific downregulation of multiple genes in the domain (Figure 7). The results suggest that in neonate brain, ECR18 modulates the expression levels of *Peg3* and *Usp29* on the paternal allele and *Zim1* on the maternal allele *in cis*. Thus, this is the first indication that ECR18 may serve as a domain-wide enhancer element for the *Peg3* domain. Given the similarities in the epigenetic profiles between ECR18 and many other iHMRs, we propose that other intergenic hypomethylated regions may act as domain-wide regulators of imprinted genes as well. Thus, further studies aimed at elucidating the roles of individual iHMRs are likely to yield important information regarding the co-regulation of imprinted genes within a domain.

The various conserved putative transcription factor-binding sites at the iHMRs may offer a mechanistic clue toward their regulatory potential. Of note, we found shared E-box-binding motifs across multiple imprinted domain iHMRs (Figure 4), which is reminiscent of a similar finding across multiple ECRs in the *Peg3* domain [8]. A large number of imprinted genes are associated with growth regulation, myogenic and neuronal cell lineages and various mouse and human cancers [1–3,43,48]. A wide array of beta-Helix Loop Helix transcription factors that are important for the same, such as MYOD, NEUROD and MYC, are known to bind the E-box [49]. Thus, the shared E-box motif may offer a clue regarding how multiple imprinted genes might be regulated in varying contexts. Many imprinted genes also share the unique feature that their expression levels are sexually dimorphic between the two murine sexes [34]. Some iHMRs that show putative-binding motifs for sex hormone receptors (Table 1 & Supplementary Material 3) may mediate such sex-specific expression patterns of imprinted genes. In that respect, it is interesting to speculate the contribution of unknown sex-specific factors that mediate ECR18 activity only in males, but not in females. Furthermore, many of the CTCF-binding iHMRs could be involved in organizing parental allele-specific chromatin architecture, which may explain how nearby reciprocally imprinted genes are co-regulated, such as in the case of *Peg3* and *Zim1* (Figures 4 & 7) [6,50]. Moreover, the conserved CTCF-binding iHMRs (Figure 4) may be the key to understanding how the imprinting status and co-regulation of genes have remained conserved throughout all placental mammals. Many of localized transcription factor-binding sites may also help to understand associated imprinted gene function. For example, the conserved putative-binding sites for cAMP response elements at *H19*-HMR is particularly interesting given that *Igf2r* is known to be regulated by CREB in rats [51]. Similarly, the putative-binding sites for PAX6 at *Dlk1*-HMR3 could shed light on how the nearby gene is regulated in the context of neurogenesis [36]. Therefore, future molecular and mutational analyses of the various motifs within the iHMRs could be critical for understanding the regulatory importance of these regions within imprinted domains.

Detailed functional characterization of iHMRs as domain-wide gene regulators could offer important clues regarding the role of imprinted genes in cancer as well. Recent studies have implicated epigenetic instability at ICRs, along with misregulation of imprinted genes, in the context of both mouse and human cancers [2,3].

Methylation levels at putative enhancer regions (such as the iHMRs) also appeared sensitive to cancer states in both species (Figure 3C & D) [2,3]. Currently, it is unknown whether the methylation status of these regulatory elements could affect the expression level of imprinted genes; however, it is noteworthy that complete unmethylation at the *Peg3*-ICR leads to misexpression of several imprinted genes in the domain [52], many of which are also affected by ECR18 deletion (Figure 7). Based on that we propose that the methylation level at some of the ECR18-like iHMRs may have a domain-wide effect on gene regulation, when altered in the cancer environment. It is most likely that such an effect would be mediated through the activity of various transcription factors. In that respect, it is interesting to note that several transcription factors whose putative-binding sites exist on iHMRs, such as CTCF, MYOD, MYC and JUN, have been associated with various cancer states [53–56]. If the methylation status at iHMRs and the subsequent misregulation of imprinted genes are causally linked to tumor progression, then these regions could potentially be developed as novel targets of cancer therapy. Therefore, further studies aimed at understanding the functional aspects of iHMRs in imprinted domains could not only reveal how imprinted genes are regulated under normal circumstances, but the information could also have important applications in the context of cancer formation and treatment.

Conclusion & future perspective

In this study, we have focused on a set of eight iHMRs, most of which are hypomethylated in a large number of tissues and bear CTCF-binding sites. Previous studies have categorized iHMRs, which show shared hypomethylation patterns in multiple tissues largely as promoter regions and those which bind CTCF as insulators. However, a subset of those regions, such *Peg3*-HMR and *H19*-HMR, appear to function as enhancers. These noncanonical enhancer iHMRs are generally well conserved and show histone marks and DNA-binding motifs, which are consistent with their enhancer function. We confirmed their ability to both initiate transcription as well as affect the transcriptional level of a nearby gene through reporter assays in multiple cell lines. Further, deletion of the *Peg3*-HMR in mice resulted in allele- and sex-specific downregulation of several genes in the *Peg3* domain, suggesting that the iHMR may be responsible for domain-wide regulation of imprinted genes. Based on these evidences, we propose the existence of noncanonical enhancer iHMRs within imprinted domains. These iHMRs are likely to hold the mechanistic key to the co-regulation of oppositely imprinted genes within a domain by allowing for allele-specific looping interactions. Future studies aimed at mutating or deleting these regions could yield important information for imprinted domain regulation both in normal as well as disease states, such as cancer.

Summary points

- Most intergenic/intronic hypomethylated regions (iHMRs) can be categorized based on their methylation pattern as follows: promoter, hypomethylated in multiple tissue types; enhancer, tissue-specific hypomethylation; or insulator, CTCF-bound and hypomethylated in multiple tissues.
- Here we assayed eight iHMRs in five imprinted domains which comprise a subset that show a combination of these patterns: hypomethylation in multiple tissues while also being bound by CTCF and harboring epigenetic signatures of enhancer regions.
- Since many imprinting regulators remain unknown and the finding that the iHMRs tend to be epigenetically unstable in multiple cancers, we further investigated their role in regulating imprinted genes, which are themselves frequently misregulated during tumorigenesis.
- Analysis of 40 mammalian sequences indicated that most iHMRs are well conserved and harbor conserved E-box and CTCF-binding motifs.
- Besides CTCF, chromatin immunoprecipitation assays also revealed RNA polymerase II occupancy at many of the iHMRs, suggesting their potential role as regulatory elements.
- Two iHMRs, associated with the *Peg3* and *H19* imprinted domains, showed the context-dependent ability to drive low-level transcription as well as regulate the transcriptional activity of *Peg3* and *H19*, respectively.
- *In vivo* deletion of the *Peg3*-HMR resulted in allele- and sex-specific downregulation of multiple imprinted genes within the *Peg3* domain indicating the region may be a shared enhancer in the domain.
- Based on this evidence, we propose that the noncanonical iHMRs may be involved in allele-specific chromatin looping in imprinted domains, thereby mediating the co-regulation of oppositely imprinted genes.
- Future mutational analysis of these regions is likely to reveal key information regarding imprinted gene regulation, which may bear some clinical significance as well.

Supplementary data

To view the supplementary data that accompany this paper please visit the journal website at www.futuremedicine.com/doi/suppl/10.2217/epi-2017-0125

Acknowledgements

The authors thank H Kim for her assistance with the cell culture and reporter assay experiments, and C Pesson for his help with the DNA methylation analyses of the human tumor panel. The authors also thank H Kim and W Frey for their thoughtful feedback and discussion over the manuscript.

Financial & competing interests disclosure

This work was supported by NIH (R01-GM066225 and R01-GM097074 to J Kim). The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Ethical conduct of research

The authors state that they have obtained appropriate institutional review board approval or have followed the principles outlined in the Declaration of Helsinki for all human or animal experimental investigations. In addition, for investigations involving human subjects, informed consent has been obtained from the participants involved.

References

Papers of special note have been highlighted as: • of interest; •• of considerable interest

- Barlow DP, Bartolomei MS. Genomic imprinting in mammals. *Cold Spring Harb. Perspect. Biol.* 6(2), pii:a018382 (2014).
- **An excellent review summarizing our current knowledge about genomic imprinting and future directions for the field.**
- Kim J, Bretz CL, Lee S. Epigenetic instability of imprinted genes in human cancers. *Nucleic Acids Res.* 43(22), 10689–10699 (2015).
- Bretz CL, Langohr IM, Lee S, Kim J. Epigenetic instability at imprinting control regions in a Kras(G12D)-induced T-cell neoplasm. *Epigenetics* 10(12), 1111–1120 (2015).
- Kelsey G, Bartolomei MS. Imprinted genes ... and the number is? *PLoS Genet.* 8(3), e1002601 (2012).
- He H, Kim J. Regulation and function of the Peg3 imprinted domain. *Genomics Inform.* 12(3), 105–113 (2014).
- Su AI, Wiltshire T, Batalov S *et al.* A gene atlas of the mouse and human protein-encoding transcriptomes. *Proc. Natl Acad. Sci. USA* 101(16), 6062–6067 (2004).
- Barlow DP. Genomic imprinting: a mammalian epigenetic discovery model. *Annu. Rev. Genet.* 45, 379–403 (2011).
- Kim J, Ye A. Phylogenetic and epigenetic footprinting of the putative enhancers of the Peg3 domain. *PLoS ONE* 11(4), e0154216 (2016).
- **Bioinformatic analyses and MYOD-chromatin immunoprecipitation for ECR18 (and other ECRs) in Peg3 domain.**
- Thiaville MM, Kim H, Frey WD, Kim J. Identification of an evolutionarily conserved cis-regulatory element controlling the Peg3 imprinted domain. *PLoS ONE* 8(9), e75417 (2013).
- **Initial description of Peg3-ECR18 as a putative enhancer for Peg3 domain.**
- Rogers ED, Ramalie JR, McMurray EN, Schmidt JV. Localizing transcriptional regulatory elements at the mouse Dlk1 locus. *PLoS ONE* 7(5), e36483 (2012).
- **Initial description of the Dlk1-HMR2 locus and its transcriptional regulatory potential.**
- Schlesinger F, Smith AD, Gingeras TR, Hannon GJ, Hodges E. *De novo* DNA demethylation and noncoding transcription define active intergenic regulatory elements. *Genome Res.* 23(10), 1601–1614 (2013).
- **Initial description of intergenic/intronic hypomethylated regions from genome-wide analyses.**
- Li Y, Huang W, Niu L, Umbach DM, Covo S, Li L. Characterization of constitutive CTCF/cohesin loci: a possible role in establishing topological domains in mammalian genomes. *BMC Genomics* 14, 553 (2013).
- Perera BP, Kim J. Alternative promoters of Peg3 with maternal specificity. *Sci. Rep.* 6, 24438 (2016).
- Luo Z, Lin C, Woodfin AR *et al.* Regulation of the imprinted Dlk1-Dio3 locus by allele-specific enhancer activity. *Genes Dev.* 30(1), 92–101 (2016).
- Nordin M, Bergman D, Halje M, Engstrom W, Ward A. Epigenetic regulation of the Igf2/H19 gene cluster. *Cell Prolif.* 47(3), 189–199 (2014).
- Court F, Baniol M, Hagege H *et al.* Long-range chromatin interactions at the mouse Igf2/H19 locus reveal a novel paternally expressed long non-coding RNA. *Nucleic Acids Res.* 39(14), 5893–5906 (2011).

17. Nativio R, Sparago A, Ito Y, Weksberg R, Riccio A, Murrell A. Disruption of genomic neighbourhood at the imprinted IGF2-H19 locus in Beckwith–Wiedemann syndrome and Silver–Russell syndrome. *Hum. Mol. Genet.* 20(7), 1363–1374 (2011).
18. Charalambous M, Menheniott TR, Bennett WR *et al.* An enhancer element at the Igf2/H19 locus drives gene expression in both imprinted and non-imprinted tissues. *Dev. Biol.* 271(2), 488–497 (2004).
- **Initial description of the H19-HMR locus and its transcriptional regulatory potential.**
19. Stadler MB, Murr R, Burger L *et al.* DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature* 480(7378), 490–495 (2011).
20. Hon GC, Rajagopal N, Shen Y *et al.* Epigenetic memory at embryonic enhancers identified in DNA methylation maps from adult mouse tissues. *Nat. Genet.* 45(10), 1198–1206 (2013).
21. Pollard KS, Hubisz MJ, Rosenbloom KR, Siepel A. Detection of nonneutral substitution rates on mammalian phylogenies. *Genome Res.* 20(1), 110–121 (2010).
22. Blanchette M, Kent WJ, Riemer C *et al.* Aligning multiple genomic sequences with the threaded blockset aligner. *Genome Res.* 14(4), 708–715 (2004).
23. Stamatoiyannopoulos JA, Snyder M, Hardison R *et al.* An encyclopedia of mouse DNA elements (mouse ENCODE). *Genome Biol.* 13(8), 418 (2012).
24. Gerstein MB, Kundaje A, Hariharan M *et al.* Architecture of the human regulatory network derived from ENCODE data. *Nature* 489(7414), 91–100 (2012).
25. Wang J, Zhuang J, Iyer S *et al.* Sequence features and chromatin structure around the genomic regions bound by 119 human transcription factors. *Genome Res.* 22(9), 1798–1812 (2012).
26. Wang J, Zhuang J, Iyer S *et al.* Factorbook.org: a Wiki-based database for transcription factor-binding data generated by the ENCODE consortium. *Nucleic Acids Res.* 41, D171–D176 (2013).
27. Ovcharenko I, Loots GG, Giardine BM *et al.* Mulan: multiple-sequence local alignment and visualization for studying function and evolution. *Genome Res.* 15(1), 184–194 (2005).
28. Ye A, He H, Kim J. Paternally expressed Peg3 controls maternally expressed Zim1 as a trans factor. *PLoS ONE* 9(9), e108596 (2014).
29. Winer J, Jung CK, Shackel I, Williams PM. Development and validation of real-time quantitative reverse transcriptase-polymerase chain reaction for monitoring gene expression in cardiac myocytes *in vitro*. *Anal. Biochem.* 270(1), 41–49 (1999).
30. Creighton MP, Cheng AW, Welstead GG *et al.* Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl Acad. Sci. USA* 107(50), 21931–21936 (2010).
31. Clark SJ, Harrison J, Paul CL, Frommer M. High sensitivity mapping of methylated cytosines. *Nucleic Acids Res.* 22(15), 2990–2997 (1994).
32. Andersson R, Gebhard C, Miguel-Escalada I *et al.* An atlas of active enhancers across human cell types and tissues. *Nature* 507(7493), 455–461 (2014).
33. Huang J, Liu X, Li D *et al.* Dynamic control of enhancer repertoires drives lineage and stage-specific transcription during hematopoiesis. *Dev. Cell* 36(1), 9–23 (2016).
34. Faisal M, Kim H, Kim J. Sexual differences of imprinted genes' expression levels. *Gene* 533(1), 434–438 (2014).
35. Ferron SR, Charalambous M, Radford E *et al.* Postnatal loss of Dlk1 imprinting in stem cells and niche astrocytes regulates neurogenesis. *Nature* 475(7356), 381–385 (2011).
36. Surmacz B, Noisa P, Risner-Janiczek JR *et al.* DLK1 promotes neurogenesis of human and mouse pluripotent stem cell-derived neural progenitors via modulating Notch and BMP signalling. *Stem Cell Rev.* 8(2), 459–471 (2012).
37. Andersson R. Promoter or enhancer, what's the difference? Deconstruction of established distinctions and presentation of a unifying model. *BioEssays* 37(3), 314–323 (2015).
38. Kim TK, Shiekhhattar R. Architectural and functional commonalities between enhancers and promoters. *Cell* 162(5), 948–959 (2015).
39. Kuroiwa Y, Kaneko-Ishino T, Kagitani F *et al.* Peg3 imprinted gene on proximal chromosome 7 encodes for a zinc finger protein. *Nat. Genet.* 12(2), 186–190 (1996).
40. Kim J, Lu X, Stubbs L. Zim1, a maternally expressed mouse Kruppel-type zinc-finger gene located in proximal chromosome 7. *Hum. Mol. Genet.* 8(5), 847–854 (1999).
41. Kim J, Bergmann A, Lucas S, Stone R, Stubbs L. Lineage-specific imprinting and evolution of the zinc-finger gene ZIM2. *Genomics* 84(1), 47–58 (2004).
42. Kim J, Bergmann A, Wehri E, Lu X, Stubbs L. Imprinting and evolution of two Kruppel-type zinc-finger genes, ZIM3 and ZNF264, located in the PEG3/USP29 imprinted domain. *Genomics* 77(1–2), 91–98 (2001).
43. Relaix F, Weng X, Marazzi G *et al.* Pw1, a novel zinc finger gene implicated in the myogenic and neuronal lineages. *Dev. Biol.* 177(2), 383–396 (1996).
44. Bartolomei MS, Zemel S, Tilghman SM. Parental imprinting of the mouse H19 gene. *Nature* 351(6322), 153–155 (1991).

45. Dechiara TM, Robertson EJ, Efstratiadis A. Parental imprinting of the mouse insulin-like growth factor II gene. *Cell* 64(4), 849–859 (1991).
46. Yevtodiyenko A, Schmidt JV. Dlk1 expression marks developing endothelium and sites of branching morphogenesis in the mouse embryo and placenta. *Dev. Dyn.* 235(4), 1115–1123 (2006).
47. Qu C, Jiang T, Li Y *et al.* Gene expression and IG-DMR hypomethylation of maternally expressed gene 3 in developing corticospinal neurons. *Gene Expr. Patterns* 13(1–2), 51–56 (2013).
48. Perera BP, Kim J. Sex and tissue specificity of Peg3 promoters. *PLoS ONE* 11(10), e0164158 (2016).
49. Jones S. An overview of the basic helix-loop-helix proteins. *Genome Biol.* 5(6), 226 (2004).
50. Vietri Rudan M, Barrington C, Henderson S *et al.* Comparative Hi-C reveals that CTCF underlies evolution of chromosomal domain architecture. *Cell Rep.* 10(8), 1297–1309 (2015).
51. Chen WK, Kuo WW, Hsieh DJ *et al.* CREB negatively regulates IGF2R gene expression and downstream pathways to inhibit hypoxia-induced H9c2 cardiomyoblast cell death. *Int. J. Mol. Sci.* 16(11), 27921–27930 (2015).
52. Bretz CL, Kim J. Transcription-driven DNA methylation setting on the mouse Peg3 locus. *Epigenetics* 12(11), 945–952 (2017).
53. Kemp CJ, Moore JM, Moser R *et al.* CTCF haploinsufficiency destabilizes DNA methylation and predisposes to cancer. *Cell Rep.* 7(4), 1020–1029 (2014).
54. Cai C, Qin X, Wu Z *et al.* Inhibitory effect of MyoD on the proliferation of breast cancer cells. *Oncol. Lett.* 11(6), 3589–3596 (2016).
55. Gabay M, Li Y, Felsner DW. MYC activation is a hallmark of cancer initiation and maintenance. *Cold Spring Harb. Perspect. Med.* 4(6), pii:a014241 (2014).
56. Kappelmann M, Bosserhoff A, Kuphal S. AP-1/c-Jun transcription factors: regulation and function in malignant melanoma. *Eur. J. Cell Biol.* 93(1–2), 76–81 (2014).