



Systemic analysis of osteoblast-specific DNA methylation marks reveals novel epigenetic basis of osteoblast differentiation



Fangtang Yu^{a,1}, Hui Shen^{a,*,1}, Hong-Wen Deng^{a,b,c}

^a Center for Bioinformatics and Genomics, Department of Global Biostatistics and Data Science, School of Public Health and Tropical Medicine, Tulane University, New Orleans, LA 70112, USA

^b College of Life Sciences, Hunan Normal University, Changsha, Hunan 410081, China

^c College of Life Sciences and Bioengineering, Beijing Jiaotong University, Beijing 100044, China

ARTICLE INFO

Article history:

Received 17 October 2016

Received in revised form 20 March 2017

Accepted 1 April 2017

Available online 3 April 2017

Keywords:

Osteoblast

Methylation

Cell-specific

Differential methylation analysis

Transcription

Alternative splicing

ABSTRACT

DNA methylation is an important epigenetic modification that contributes to the lineage commitment and specific functions of different cell types. In this study, we compared ENCODE-generated genome-wide DNA methylation profiles of human osteoblast with 21 other types of human cells in order to identify osteoblast-specific methylation events. For most of the cell strains, data from two isogenic replicates were included, resulting in a total of 51 DNA methylation datasets. We identified 852 significant osteoblast-specific differentially methylated CpGs (DMCs) and 295 significant differentially methylated regions (DMRs). Significant DMCs/DMRs were not enriched in CpG islands (CGIs) and promoters, but more strongly enriched in CGI shores/shelves and in gene body and intergenic regions. The genes associated with significant DMRs were highly enriched in biological processes related to transcriptional regulation and critical for regulating bone metabolism and skeletal development under physiologic and pathologic conditions. By integrating the DMR data with the extensive gene expression and chromatin epigenomics data, we observed complex, context-dependent relationships between DNA methylation, chromatin states, and gene expression, suggesting diverse DNA methylation-mediated regulatory mechanisms. Our results also highlighted a number of novel osteoblast-relevant genes. For example, the integrated evidences from DMR analysis, histone modification and RNA-seq data strongly support that there is a novel isoform of *neurexin-2* (*NRXN2*) gene specifically expressed in osteoblast. *NRXN2* was known to function as a cell adhesion molecule in the vertebrate nervous system, but its functional role in bone is completely unknown and thus worth further investigation. In summary, we reported a comprehensive analysis of osteoblast-specific DNA methylation profiles and revealed novel insights into the epigenetic basis of osteoblast differentiation and activity.

© 2017 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

DNA methylation of cytosine is a crucial epigenetic mechanism for transcriptional regulation and has profound impacts on embryonic development, genomic imprinting, X-chromosome inactivation, and the pathogenesis of various human disorders (Weber et al., 2007). Though the regulatory function of DNA methylation is generally thought to be associated with transcriptional repression when occurring in gene promoter regions and with transcriptional activation when occurring in gene bodies (Jones, 2012; Ball et al., 2009; Rauch et al., 2009), recent studies revealed a far more complicated relationship between DNA methylation and gene expression. Both negative and positive

correlations between methylation and expression have been observed across all gene structural regions, and DNA methylation can also regulate alternative splicing through effects on RNA Pol II elongation (Jones, 2012; Shukla et al., 2011; Chandra et al., 2014; Ehrlich and Lacey, 2013; Liu et al., 2013; Deaton et al., 2011; Lee et al., 2014; Varley et al., 2013), indicating that DNA methylation can have diverse, chromatin context- and cell type-dependent regulatory functions on transcription.

With recent advance in high-throughput technology for DNA methylation analysis (Sun et al., 2015), a number of studies have demonstrated that DNA methylation profiles vary in diverse human tissues and cell types (Jones, 2012; Løkk et al., 2014; Yang et al., 2015), which contribute to the regulation of cell type-specific gene expression and determine the differentiation and specific function of different cell types (Futscher et al., 2002; de la Rica et al., 2013; Tsumagari et al., 2013). For example, Ziller et al. (2013) found that 21.8% of autosomal CpGs showed dynamic DNA methylation changes in a range of human cell and tissue types and these dynamic CpGs co-localized with gene regulatory elements,

* Corresponding author at: Center for Bioinformatics and Genomics, Department of Global Biostatistics and Data Science, School of Public Health and Tropical Medicine, Tulane University, 1440 Canal St., Suite 2001, New Orleans, LA 70112, USA.

E-mail address: hshen3@tulane.edu (H. Shen).

¹ These authors contributed equally to this work.

particularly enhancers and transcription-factor-binding sites, allowing identification of key lineage-specific regulators. In addition, Rica et al. (2013) identified hyper-/hypo-methylation changes in several thousand genes during in vitro monocyte-to-osteoclast differentiation, including all relevant osteoclast differentiation and function categories. DNA methylation has also been implicated in the regulation of differentiation and function of osteoblasts, the bone-forming cell with main function of mineralizing the bone matrix (Eslaminejad et al., 2013). For example, the promoter of *osteocalcin* gene, a gene solely expressed by osteoblasts, is highly methylated in cells not expressing osteocalcin, including the mesenchymal stem cells (MSCs) (Villagra et al., 2002). Interestingly, during in vitro MSC-to-osteoblast differentiation, as the osteocalcin gene becomes increasingly expressed, CpG methylation of the osteocalcin promoter is significantly reduced (Villagra et al., 2002). Similarly, reduced DNA methylation along with transcriptional upregulation were also observed for two additional osteogenic genes, namely, *alpha 1 type I collagen (COL1A1)* and *osteopontin* (Arnsdorf et al., 2010). In addition to hypomethylation mediated gene activation, hypermethylation induced silencing of specific genes were also crucial in osteoblast differentiation. For instance, Hsiao et al. (2010) found that *Trip10* (thyroid hormone receptor interactor 10), an adaptor protein involved in diverse cellular functions, shows significant alterations in promoter methylation and mRNA levels during lineage-specific induction of human bone marrow-derived MSCs. Remarkably, targeted induction of *Trip10* promoter methylation resulted in reduced *Trip10* expression and accelerated MSC differentiation towards osteogenic lineage at the expense of MSC-to-adipocyte differentiation. Taken together, these results strongly supported that DNA methylation plays a significant role in mediating cell-specific gene transcription and interfering with cell fate determination, including osteoblast differentiation.

In this study, we compared the genome-wide DNA methylation profiles between human osteoblasts and a wide range of other types of human cells in order to identify and characterize osteoblast-specific methylation patterns on a global scale. The purpose is to identify those genes and regulatory mechanisms underlying specific functions of osteoblasts. Our results revealed many osteoblastic hyper-/hypo-methylated CpGs across the genome. By integrating the DNA methylation patterns with transcriptomic and other epigenomic profiles, we further showed that these osteoblastic-specific methylation events were enriched in regulatory regions beyond the promoter areas and may influence gene expression and the use of alternative promoters in a cell-type specific manner. Collectively, these data may provide novel insight into the regulatory role of DNA methylation in osteoblast differentiation and functioning.

2. Results and discussion

2.1. Identification and characterization of osteoblast-specific DMCs/DMRs

We compared ENCODE-generated DNA methylation profiles of osteoblasts with those of 20 different types of non-transformed human cell strains plus Epstein-Barr virus-transformed lymphoblastoid cell lines (LCLs) (Supplementary Table 1). For most of the cell strains, DNA methylation data generated by reduced representation bisulfite sequencing (RRBS) from two isogenic replicates were included, resulting in a total of 51 DNA methylation datasets. The number of CpGs assessed per sample ranged from 960,300 to 1,489,630, including ~31.6–43.7% of CpGs with sequence coverage $\geq 10\times$ (Supplementary Table 1). We compared a total of 182,518 CpGs with coverage $\geq 10\times$ across all 51 samples and identified 852 significant differentially methylated CpGs (DMCs) with stringent criteria ($q < 0.01$, difference in methylation $\geq 50\%$), which were distributed across the entire genome (Supplementary Fig. 1). Hierarchical clustering analysis using the significant DMCs correctly grouped cells from similar tissues and of similar biological functions (Supplementary Fig. 2). Interestingly, we observed high similarity of the DNA methylation patterns between osteoblast and skeletal muscle

myoblast. This is not completely unexpected, because both osteoblast and myoblast are mesodermal descendent of the bone-marrow mesenchymal stem cells (BMSCs) (Gimble et al., 2008). Moreover, it has been shown that myoblastic cells can be differentiated into osteoblastic cells (Tanaka et al., 2012), and the muscle-derived MSCs were more effective in differentiation into osteoblastic cells than BMSCs (Glass et al., 2011). In fact, a high similarity of chromatin states between osteoblast and skeletal muscle myoblast has also been observed in the NIH Roadmap Epigenomics project (C. Roadmap Epigenomics et al., 2015).

Of the total 852 DMCs, 685 (80.40%) were hypermethylated and 167 (19.60%) were hypomethylated, in osteoblasts vs. other cell types. While the majority of the DMCs was mapped to CpG islands (CGIs), DMCs were more strongly enriched in non-CGI regions, including CGI shore ($p = 5.73 \times 10^{-7}$, Fisher's exact test), CGI shelf ($p = 7.54 \times 10^{-11}$) and open sea ($p = 4.54 \times 10^{-11}$), when taking into account of the number of CpGs tested in each CpG annotation class (Fig. 1A). We observed a marked difference in the distributions with respect to CGIs between hyper- and hypo-methylated DMCs (Fig. 1A, $p = 1.34 \times 10^{-64}$), with the over majority (78%) of hypermethylated DMCs associated with CGIs, in contrast to hypomethylated DMCs, which were mainly mapped to open sea (~56%) and relatively infrequent in CGIs (~13%). Interestingly, the enrichment of cell lineage-/tissue-specific DNA methylation events in non-CGI regions but depletion in CGIs have also been observed by others (Lokk et al., 2014; Yang et al., 2015; Byun et al., 2009; Sliker et al., 2013), highlighting the importance of exploring the functional significance of non-CGI methylation.

We considered the location of DMCs across different parts of individual genes. We observed a significant depletion of DMCs in 5'-untranslated regions (UTRs) ($p = 4.2 \times 10^{-34}$) and promoters ($p = 7.87 \times 10^{-21}$) but a significant enrichment of DMCs in exons ($p = 1.64 \times 10^{-07}$), 3'UTR ($p = 3.89 \times 10^{-09}$) as well as intergenic regions (p

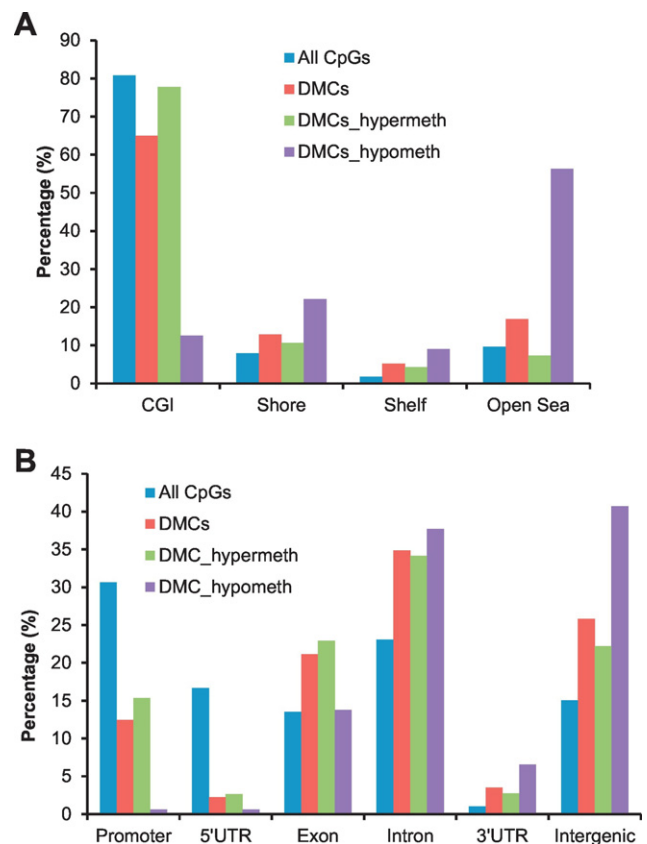


Fig. 1. Distribution of all 182,518 tested CpGs and 852 significant osteoblast-specific DMCs, including 685 hypermethylated and 167 hypomethylated DMCs, across (A) different regions related to CGIs and (B) different genic regions.

$= 1.02 \times 10^{-31}$) (Fig. 1B), when comparing to the overall distribution of all the tested CpGs. These results are in line with previous findings that cell type-/tissue-specific epigenetic marks were more often found in gene body regions than in promoter areas (Lokk et al., 2014; Yang et al., 2015; Tsumagari et al., 2013; Ernst et al., 2011), suggesting that the intragenic DNA methylation may play significant roles in the differentiation of diverse cell/tissue types. There is also a significant difference in the genic distribution between hyper- and hypo-methylated DMCs (Fig. 1B, $p = 5.89 \times 10^{-11}$). Although over 50% of both hyper- and hypo-methylated DMCs were mapped to gene body (introns and exons), the hypermethylated DMCs were observed much more frequently at promoters and 5-UTRs than hypomethylated DMCs, and the latter were more frequently associated with 3'-UTRs and intergenic regions.

While methylation at individual CpG sites may possibly be linked to gene expression regulation (Venza et al., 2012; Chen, 1983), methylation levels at neighboring CpGs are often highly correlated and methylation-mediated regulatory elements often extend across genomic regions (Davies et al., 2012; Baylin and Jones, 2011; Song et al., 2009). Therefore, we further conducted differentially methylated region (DMR) analysis based on the computed DMC data. In total, we identified 295 significant DMRs, including 247 (83.73%) hypermethylated and 48 (16.27%) hypomethylated DMRs in osteoblasts vs. other cell types. The full list of significant DMRs was provided in the Supplementary Table 2. The hypomethylated DMRs (mean size = 72 bp) were significantly longer ($p = 0.008$) but contained comparable numbers of individual DMCs (mean number of DMCs = 5, $p = 0.10$) than the hypermethylated DMRs (mean size = 61 bp), suggesting more spreading of demethylation than of de novo methylation. When mapping the 295 DMRs to different CGI regions, we found 197 (66.8%) DMRs were mapped to CGIs, 49 (16.6%) to CGI shores, 16 (5.4%) to CGI shelves, and 33 (11.2%) to open sea (Supplementary Fig. 3A). In addition, most of the DMRs were observed in gene bodies and intergenic regions rather than in promoters of RefSeq genes (Supplementary Fig. 3B). Similar to what we observed when comparing the CGI/genic distribution of hyper- vs. hypo-methylated DMCs, we observed a highly significant difference in the distribution of hyper- vs. hypo-methylated DMRs with respect to CGIs (Supplementary Fig. 3A, $p = 3.22 \times 10^{-25}$) and RefSeq genes (Supplementary Fig. 3B, $p = 3.22 \times 10^{-25}$). Specifically, the hypermethylated DMRs were much more frequently associated with CGIs and promoters than hypomethylated DMRs, and the latter were mainly in open sea and intergenic regions. This difference is also reflected by the distinct distribution of hyper- and hypo-methylated DMRs relative to the transcription start sites (TSSs) of associated genes (Supplementary Fig. 3C), with hypo-methylated DMRs showing a marked depletion within 5 kb of the TSSs but not for hyper-methylated DMRs.

To better understand the biological context of the osteoblast-specific DMRs, we examined the distribution of the DMRs with respect to the different chromatin states in osteoblasts, which were characterized by the NIH Roadmap Epigenomic Consortium (C. Roadmap Epigenomics et al., 2015). Overall, the distribution of hyper- and hypo-methylated DMRs in chromatin states were not significantly ($p = 0.68$) different. Hyper-/hypo-methylated DMRs were often associated with elements in weak transcription, repressed/quiescent chromatin states polycomb, but not with active/flanking TSS regions (Supplementary Fig. 3D), suggesting that these osteoblastic-specific DMRs mainly affect weakly/low-level transcribed elements rather than active promoters/TSS flanking regions.

To further explore the potential functional significance of the osteoblastic-specific DMRs, we next tested whether the nearby genes of DMRs were enriched for certain functional terms by using the GREAT program (McLean et al., 2010). The gene ontology (GO) enrichment analysis revealed that genes associated with the hypermethylated DMRs were highly enriched in a number of biological process terms that are relevant to embryo and skeletal development (Fig. 2A), such

as embryonic development (p value = 3.91×10^{-24} , fold enrichment = 3.02) and skeletal system development (p value = 2.24×10^{-18} , fold enrichment = 3.86). Remarkably, the top 10 mouse and human phenotypes that were most significantly enriched for genes associated with the hypermethylated DMRs were almost all related to skeletal abnormalities (Fig. 2B–C), such as abnormal axial skeleton morphology (p -value = 1.03×10^{-18} , fold enrichment = 2.78), abnormal cartilage morphology (p -value = 1.70×10^{-17} , fold enrichment = 3.88), and abnormality of the mouth/hand/teeth (p -value = 2.58×10^{-7} – 1.18×10^{-6} , fold enrichment = 2.07–2.52). In addition, by integrating with results from a large meta-analysis of genome-wide association studies (GWASs) for osteoporosis risk (Estrada et al., 2012), we demonstrated significant enrichment of osteoporosis-associated genes in both hyper- and hypo-methylated DMRs (p value = 2.00×10^{-4} and 0.0187 respectively). Specifically, of the 178 genes annotated to hypermethylated DMRs, 6 genes (*ESR1*, *FOXL1*, *HOXC4*, *HOXC5*, *HOXC6*, and *WNT3*) showed significant genetic association with osteoporosis in the GWAS meta-analysis. Similarly, one (*PTPRN2*) of the 31 genes annotated to the hypomethylated DMRs are associated with osteoporosis risks. These are strongly contrasted with the background gene set, for which of the 11,329 genes annotated to all tested CpGs, only 77 genes were associated with osteoporosis. These results strongly suggested the identified DMRs and their associated genes may play functionally significant roles in bone metabolism and skeletal development in physiologic and pathologic conditions.

By integrative analysis of the DMR data with the extensive gene expression and chromatin epigenetics data in ENCODE (E.P. Consortium, 2012), we observed complex, context-dependent relationships between DNA methylation, chromatin states, and gene expression, which are illustrated below with some representative genes.

2.2. Hypermethylated DMRs at promoters/5'-end regions repress gene expression in osteoblasts: *SIM2* and *GLIS1*

Promoters and 5'end regions are usually constitutively unmethylated, especially when they overlap with CGIs, even in genes with cell type-specific expression (Meissner et al., 2008). Nonetheless, there are notable exceptions as illustrated by gene *SIM2*. Specifically, we detected multiple osteoblast-specific hypermethylated DMRs (q -value = 8.57×10^{-19} – 1.25×10^{-38} , DM% = 32.3–82.4%) at *SIM2* promoter and an immediate downstream region of its TSS (Fig. 3). These DMRs were all distributed within a large CGI.

Gene-repressive DNA hypermethylation in promoter regions normally localize to chromatin with repressive histone modification markers, such as H3K27me3 and H3K9me3 (Hagarman et al., 2013). However, this was not the case for the promoter region of *SIM2*. Instead, the promoter and 5'end region of *SIM2* display strong signal for active promoter (H3K4me3) and transcriptional elongation (H3K79me2) in osteoblast (Fig. 3 and Supplementary Fig. 4). Consistently, we also observed low but detectable levels of *SIM2* expression in osteoblast (RNA-seq track in Fig. 3). Similar active transcription chromatin states were also observed at this region in human skeletal muscle myoblasts (HSMM), but the expression levels of *SIM2* in HSMM were considerably higher. This suggested the osteoblast-specific DNA hypermethylation at this region imposes a repressive effect on the transcription of *SIM2*, even when separated from the typical promoter-inhibiting chromatin marks. We speculate that the co-existence of cell-specific DNA hypermethylation and active transcription/elongation chromatin marks, but lacking repressive histone modifications, at *SIM2* promoter/5'end regions allows a tight control of repressed but not completely abolished expression of this genes in osteoblasts. In contrast, the *SIM2* gene is completely silenced in lymphocyte B-cells (GM12891) and Human Mammary Epithelial Cells (HMEC), which are likely to be mediated by the strong and wide-spread repressive histone modification of H3K27me3 (Fig. 3 and Supplementary Fig. 4). *SIM2* gene encodes a transcription factor that is generally known as a mast regulator of

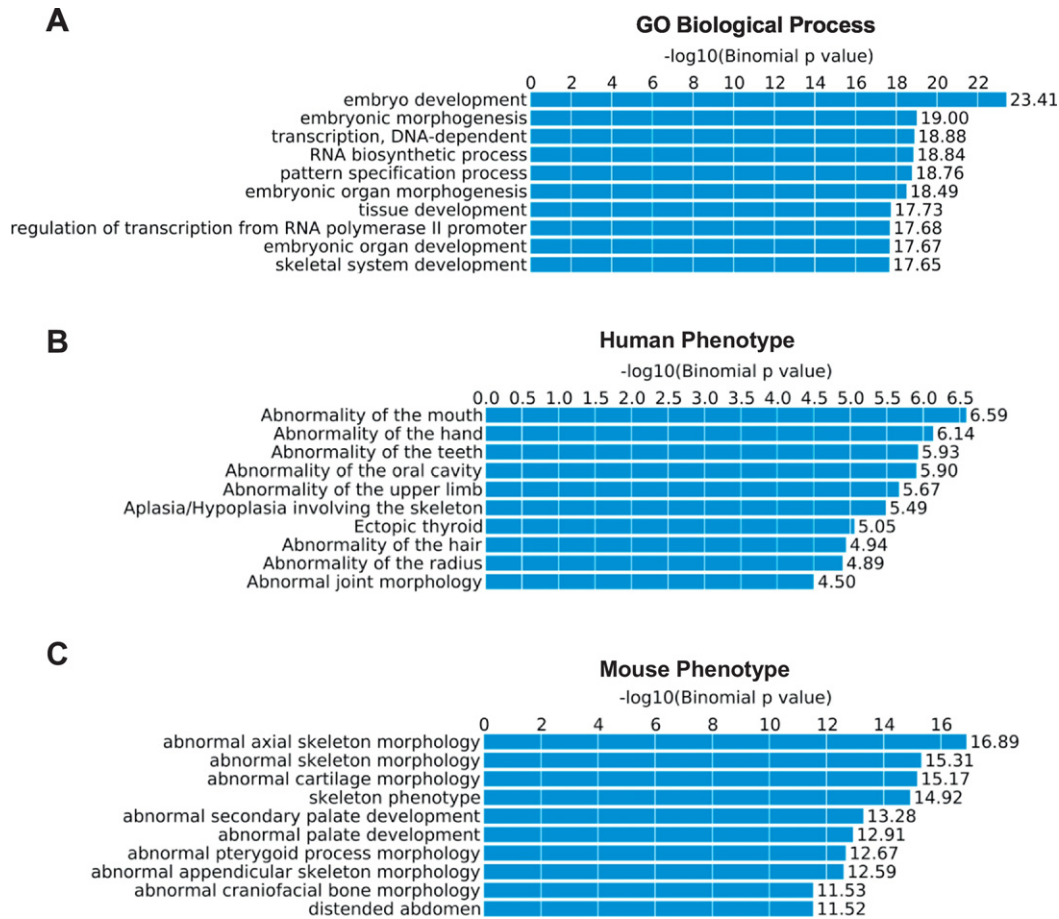


Fig. 2. Top 10 results from functional annotation and GO enrichment analysis of osteoblast-specific hypermethylated DMRs by using the GREAT package (McLean et al., 2010).

neurogenesis. However, several studies indicated that *SIM2* also plays a critical role in the regulation of osteogenesis and skeletal development (Shamblott et al., 2002; Goshu et al., 2002; Kubo et al., 2009). Specifically, siRNA knockdown of *SIM2* in MSCs suppressed osteogenesis potential and delayed matrix calcification (Kubo et al., 2009), and *SIM2* knockout mice exhibited prominent craniofacial and vertebrae abnormalities (Shamblott et al., 2002; Goshu et al., 2002). On the other hand, over-expression of *SIM2* has been implicated in the pathogenesis of Down syndrome (Meng et al., 2006). Therefore, the temporal and spatial expression of *SIM2* may have to be tightly regulated to prevent pathological consequences in a cell type-specific manner, and DNA methylation may be a critical mechanism for this fine-tuning of expression. Interestingly, osteoblast-specific hypermethylated signals also extended to multiple CpGs further deep in the *SIM2* gene body (in introns 1–2 and exon 2), which precisely bound a potential non-coding *SIM2* isoform (Fig. 3). Therefore, hypermethylated DMRs may also regulate the *SIM2* isoform expression in osteoblast.

Similar to *SIM2*, we also detected highly significant osteoblast-specific hypermethylated DMRs ($q\text{-value} = 1.06 \times 10^{-45}$ – 1.08×10^{-49} , $\text{DM}\% = 64.8$ – 69.5%) at the promoter of *GLIS1* gene, which overlap with a single CGI (Fig. 4). The RNA-seq and ExonArray data indicate *GLIS1* gene is preferentially expressed in osteoblasts and to a less extent, in HSMM, among the cells used for DMR detection (Fig. 4). Consistent with the gene expression data, histone modification marks indicate the existence of strong enhancers and active promoter (H3K4me3 and H3K27ac) at the *GLIS1* promoter regions specifically in osteoblast and HSMM, and poised promoter (H3K4me3 and H3K27me3 bivalent marks) in *GLIS1* non-expressing cells e.g., LBL, HMEC (Fig. 4 and Supplementary Fig. 5). Interestingly, the two osteoblast-specific hypermethylated DMRs precisely bound a segment exhibiting active

promoter- and strong enhancer-like histone modification marks (H3K4me3 and H3K27ac) specifically in HSMM but barely detectable in osteoblasts (Supplementary Fig. 5). Therefore, these osteoblast-specific hypermethylated DMRs might repress a myogenic-specific promoter/enhancer in osteoblasts, allowing for precise regulation of *GLIS1* expression in a cell type-specific manner. *GLIS1* encodes for a GLI-related Kruppel-like zinc finger transcription factor and can effectively promote the reprogramming of somatic cells during induced pluripotent stem cells (iPSC) generation (Maekawa and Yamanaka, 2011; Maekawa et al., 2011). Importantly, *GLIS1* is upregulated during the osteoblastic differentiation (Bustos-Valenzuela et al., 2011) and has been linked to coronary artery calcified plaque (Divers et al., 2013), which is closely related to osteoblastic differentiation and activity (Doherty et al., 2003). Together, these results suggest that *GLIS1* expression is likely to be tightly regulated in osteoblasts and cell type-specific DNA methylation may help to achieve this fine-tuning of expression.

2.3. Hypermethylation at alternative promoters contributes to cell type-specific isoform expression: *MEST* and *NRXN2*

MEST (mesoderm specific transcript) is a member of the α/β hydrolase fold family and may play a role in development, including bone growth (Andrade et al., 2010). *MEST* gene has multiple, complicated mRNA isoforms, including 6 RefSeq annotates and at least 17 alternative mRNA variants identified by AceView program (Thierry-Mieg and Thierry-Mieg, 2006). The 6 RefSeq *MEST* annotates resulted from the usage of 4 alternative promoters/TSSs and 2 alternatively spliced exons (Fig. 5). Interestingly, we detected a significant osteoblast-associated hypermethylated DMR ($q\text{-value} = 1.47 \times 10^{-39}$, $\text{DM}\% = 55.9\%$) overlapping one of the alternative promoters/TSSs that encode RefSeq

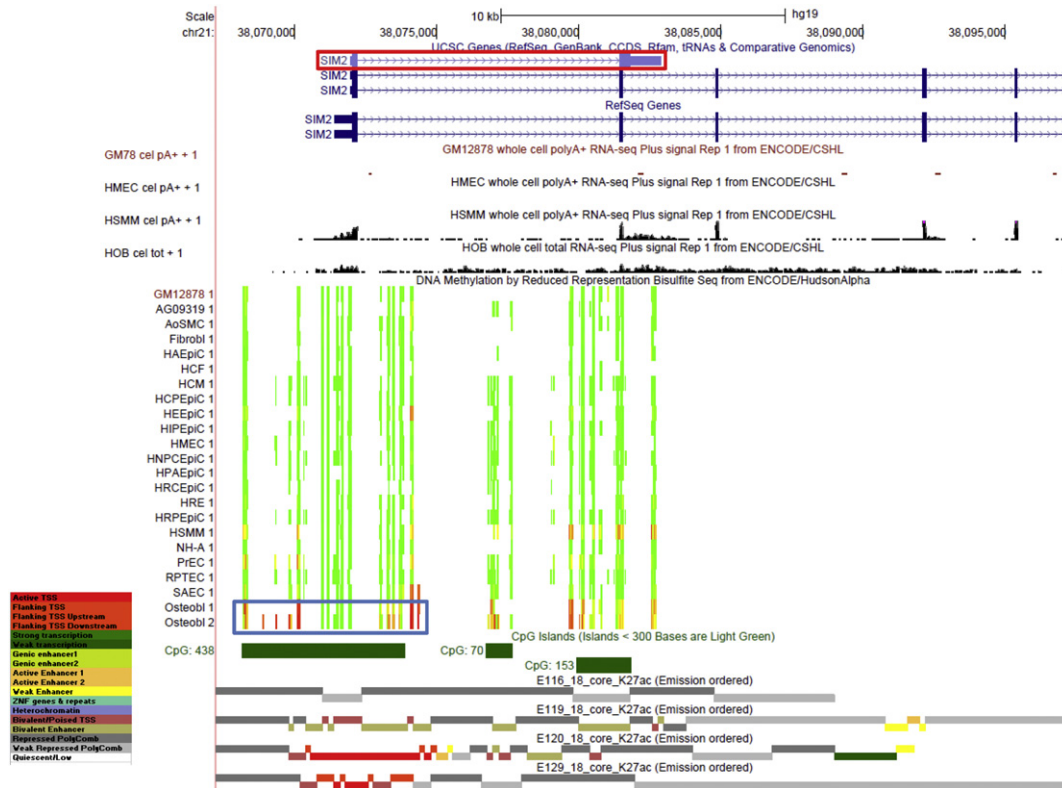


Fig. 3. Osteoblastic hypermethylation at *SIM2* TSS surrounding region. The following profiles are shown using the UCSC Genome Browser (<http://genome.ucsc.edu>, version hg19) for the *SIM2* gene region: UCSC genes, RefSeq genes, RNA-seq data from ENCODE/Cold Spring Harbor, DNA methylation levels assessed by RRBS from ENCODE/HudsonAlpha, CpG islands, Chromatin state (18-state) annotation from NIH Roadmap Epigenome project. LCL (only GM12878 is shown, but the other LCL samples give similar results), HMEC, HSMM, and osteoblast (HOB) are the only studied cell types having both data in RNA-seq and histone modification tracks available from ENCODE. Methylation status is represented with an 11-color gradient for which red, yellow, and green represent that 100%, 50%, and 0% of molecules sequenced are methylated, respectively. Osteoblast-specific DMRs are indicated in the blue box. At this scale, individual differentially methylated CpG sites cannot be resolved from neighboring sites. The specific non-coding *SIM2* isoform (Ensembl transcript ID: ENST00000460783.1) is marked in the red box. The 18 chromatin states are represented with the indicated colors, with E116 representing states for LCL (GM12878), E119 for HMEC, E120 for HSMM, and E129 for osteoblast. The detailed ChIP-seq data for various histone modification marks (H3K4me1, H3K4me3, H3K9me3, H3K27ac, H3K27me3, H3K36me3, H3K79me2, and H4K20me1) are presented in the Supplementary Fig. 3.

transcript variants 2 and 5 (Fig. 5). CpGs within this DMR exhibited strong methylation in osteoblast and HSMM, but are largely unmethylated in other cell types, including LCL and HMEC (Fig. 5). The ENCODE RNA-seq data indicate that there are considerable transcription signals of *MEST* from this alternative TSS in LCL and HMEC, suggesting the expression of RefSeq transcript variants 2 and/or 5 in LCL and HMEC. In contrast, these two *MEST* alternative transcripts were barely detectable in osteoblast and HSMM (Fig. 5), for which *MEST* RefSeq transcript variants 1 and/or 4 are dominant. Consistently, there are strong active promoter/enhancer (H3K4me1 and H3K4me3) signals at this DMR-overlapped alternative TSS in LCL and HMEC, but not in osteoblast and HSMM (Fig. 5 and Supplementary Fig. 6). Interestingly, *MEST* is known to exhibit isoform-specific imprinting (Kosaki et al., 2000; Huntriss et al., 2013; Kamei et al., 2007) and the promoter switching may lead to loss of imprinting and aberrant expression of *MEST* gene, which has been linked to several types of cancers (Pedersen et al., 2002; Nestheide et al., 2013; Li et al., 2008). Particularly, aberrant expression of *MEST* gene has been detected in human osteoblast cell lines (hFOB1.19 cells) in a model of human osteosarcoma (Li et al., 2008). Taken together, DNA methylation may represent a critical epigenetic mechanism for regulation of alternative promoter usage at *MEST* gene in a cell-type specific manner, and dysregulation of this epigenetic mechanism may contribute to the pathogenesis of *MEST* loss-of-imprinting associated disorders.

Another potential connection between osteoblast-associated DMR and cell-type specific isoform expression was detected in *NRXN2* gene. *NRXN2* gene encodes a member of the neurexin gene family and has

very complex transcription architecture. Though RefSeq annotates only 3 representative transcripts (Fig. 6), the alternative transcription of this genes is likely to be far more complicated, with 31 (annotated by NCBI *Homo sapiens* Annotation Release 107, Supplementary Fig. 7) and possibly thousands of alternative isoforms generated through the usage of multiple alternative promoters and extensive alternative splicing events (Tabuchi and Sudhof, 2002; Rowen et al., 2002). Specifically, we identified a significant, osteoblast-associated hypomethylated DMR ($q\text{-value} = 1.58 \times 10^{-51}$, $DM\% = -64.8\%$) spanning 8 CpGs in the exon 10 of the *NRXN2* RefSeq transcript variant alpha-1 (Fig. 6). Similar hypomethylation was also observed in HSMM, whereas almost all the other cell types exhibited hypermethylation in this region (Fig. 6). Interestingly, the ENCODE RNA-seq data indicate that a specific *NRXN2* isoform initiating from the immediate upstream of the exon 10 is highly expressed in osteoblast and HSMM, but not detected in LCL and HMEC (Fig. 6). Though the mRNA transcript of this specific isoform delineated by RNA-seq does not match any of the three RefSeq transcripts, it is consistent with the predicted *NRXN2* transcript variant X29 (XM_011545385.1) by the NCBI annotation (Supplementary Fig. 7), strongly supporting the authentic and predominant expression of this transcript variant in osteoblast and HSMM. Moreover, there were strong signals of active promoter-like (H3K4me3 and H3K27ac) and transcriptional activity-associated (H3K79me2) chromatin states around the DMR in osteoblast and HSMM, whereas these histone modification marks were depleted around the DMR in cell types that did not express this isoform, such as LCL and HMEC (Supplementary Fig. 7). These results strongly support that the intragenic DNA methylation may have

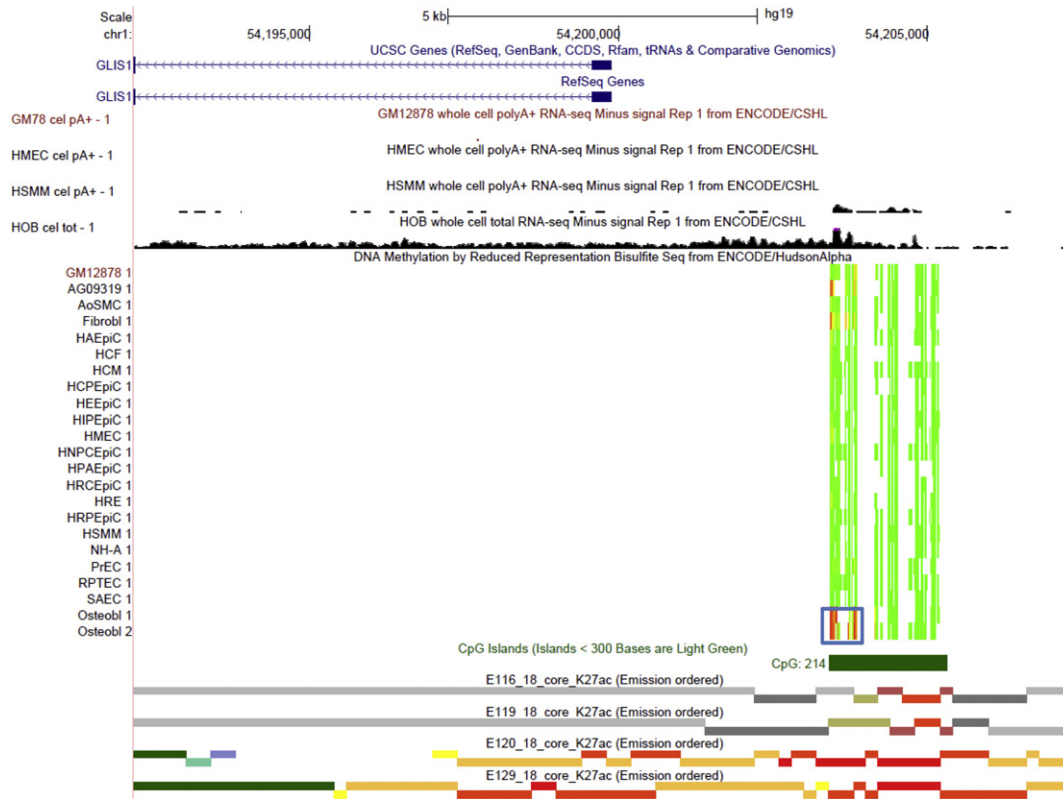


Fig. 4. Osteoblastic hypermethylation at *GLIS1* TSS surrounding region. The same tracks as illustrated in Fig. 3 are shown using the UCSC Genome Browser (<http://genome.ucsc.edu>, version hg19). Osteoblast-specific DMRs are highlighted in the blue box.

a crucial role in regulating cell context-specific alternative promoters in gene bodies (Maunakea et al., 2010).

NRXN2 is known as a cell surface protein involved in cell recognition and cell adhesion in the vertebrate nervous system. It plays an essential role in synapse function and its alterations have been linked to autistic spectrum disorder (Gauthier et al., 2011; Dachtler et al., 2014). The majority of *NRXN2* transcripts in the nervous system are produced from the upstream promoter and encode alpha-neurexin isoforms while a smaller number of transcripts are produced from the downstream promoter and encode beta-neurexin isoforms. The alpha-neurexins contain one epidermal growth factor-like (EGF-like) sequence and six laminin G domains, and have been shown to interact with neurexophilins. The beta-neurexins lack EGF-like sequences and contain only one laminin G domain, and bind to alpha-dystroglycan. The *NRXN2* variant X7 is also predicted to lack the EGF-like sequences but contain three laminin G domains. The functional roles of this *NRXN2* variant in osteoblast warrant further exploration.

2.4. Hypomethylated DMRs at promoters contribute to active expression of primary osteoblastic genes: *BGLAP*

We specifically examined the DNA methylation patterns around a number of genes that were known to play key roles in osteoblastic differentiation (Hojo et al., 2015; Kirkham and Cartmell, 2007; Cawthorn et al., 2012), including *Runx2* (*Cbfa1*), *Sp7* (*osterix*), *Dlx5*, *Msx2* (*HOX8*), *BGLAP* (*osteocalcin*), *COL1A1*, *MEF2C*, *BMPs* (*BMP-2*, *-4*, *-6*, *-7*, and *-9*), and *WNTs* (*WNT-6*, *-8*, *-10a* and *-10b*). We identified a highly significant osteoblast-specific hypomethylated DMRs (q -value = 7.15×10^{-45} , $DM\% = -44.89\%$) at the promoter region of *BGLAP* gene (Fig. 7 and Supplementary Table 2). This significant DMR overlapped with a region showing strong active promoter-/enhancer-related chromatin states (Fig. 7) and histone modification marks (H3K4me1, H3K4me3, and H3K27ac) (Supplementary Fig. 8). These findings were consistent with

the evident *BGLAP* expression in osteoblast (Fig. 7). In contrast, *BGLAP* promoter showed hypermethylation and/or weak enhancer-related chromatin marks in HMEC and LCL (Fig. 7 and Supplementary Fig. 8). Therefore, our findings provided direct evidence that promoter methylation may interactively work with other epigenomic mechanisms to regulate the cell-type specific expression of some key osteogenic genes. Interestingly, promoter hypomethylation and active transcription of *BGLAP* were also observed in HSMM (Fig. 7 and Supplementary Fig. 8). This again reflected the close connections between osteoblast and HSMM. In fact, a recent study has demonstrated that *BGLAP* expression in myofibers is necessary and sufficient to maintain muscle mass in mice (Mera et al., 2016).

In contrast, no significant DMRs were detected at other selected osteoblastic genes (including ± 5 kb upstream/downstream regions) (Supplementary Table 3). On one hand, this may reflect the inadequate coverage of genome-wide CpGs by RRBS, which was known to be biased towards regions rich in CpG sites (e.g., CGIs). For instance, no DNA methylation data were available for CpGs located within the ± 5 kb surrounding regions of *Runx2* gene from the ENCODE RRBS dataset. Future studies using more comprehensive DNA methylation techniques, such as whole-genome bisulfite sequencing, are needed to investigate the DNA methylation mediated regulations for these genes. On the other hand, the lack of significant DMRs in these selected osteoblastic genes may imply that various other mechanisms (e.g., histone modification) may regulate the expression of these genes independent of the effects of DNA methylation (Weber et al., 2007; Jones, 2012).

3. Summary

In this study, we identified and characterized human osteoblast-specific DNA methylation profiles by comparing the genome-wide DNA methylation profiles between human osteoblasts and 21 other types of human cells and by integrating the DNA methylation

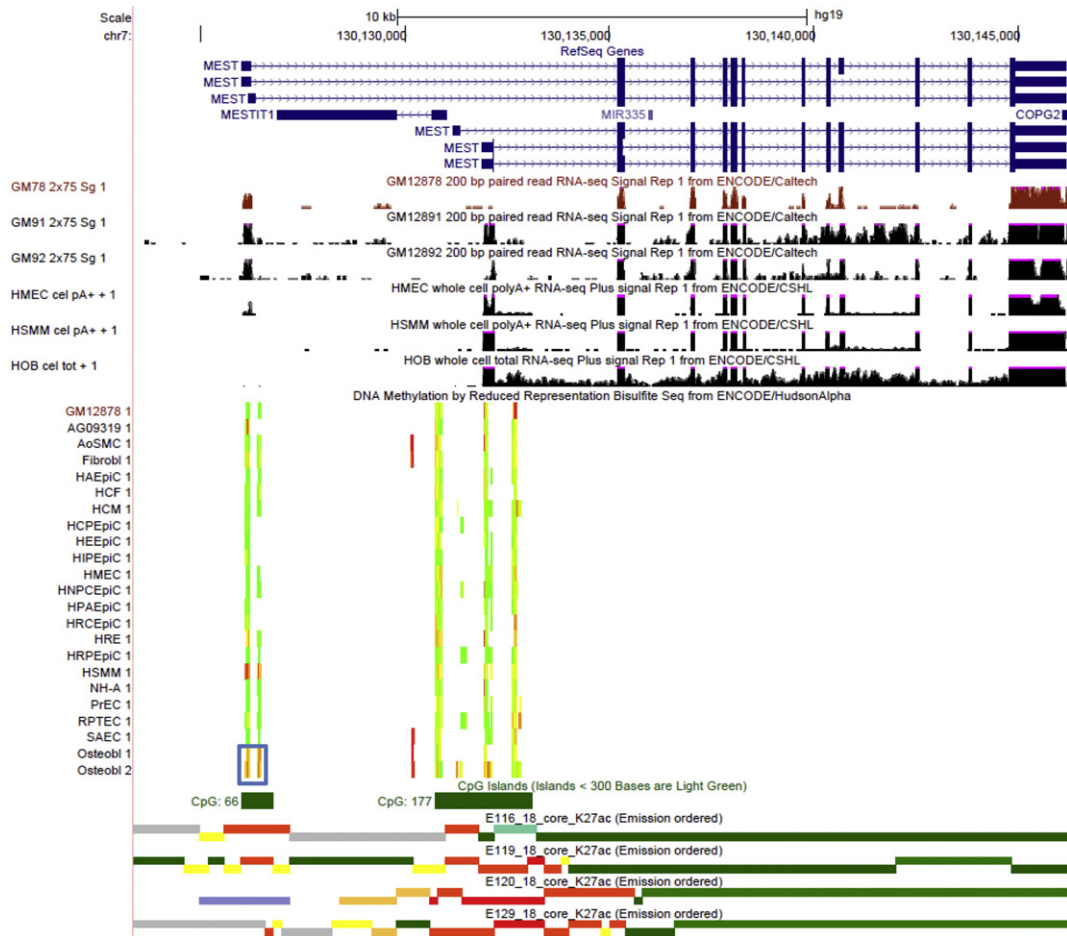


Fig. 5. Osteoblastic hypermethylation at *MEST* gene region. The same tracks as illustrated in Fig. 3 are shown using the UCSC Genome Browser (<http://genome.ucsc.edu>, version hg19), with the additions of RNA-seq tracks for LCLs (GM12891 and GM12892) from ENCODE/Caltech. Osteoblast-specific DMR is highlighted in the blue box.

patterns with transcriptomic and other epigenomic profiles. This study has a few notable limitations. First, most of the analyzed cell types only have two isogenic replicates and thus the potential inter-individual variability of DNA methylation patterns within each cell type has not been taken into account. In addition, all the epigenomic and expression data were generated from cells expanded *in vitro*, which may exhibit distorted profiles from their *in vivo* status (Caliskan et al., 2011; Saferali et al., 2010). Despite these limitations, several evidences provided strong support for the general reliability of our findings. For instance, the identified osteoblastic-specific DNA methylation sites were distributed across different genomic regions in a pattern that was largely in agreement with the patterns previously observed by other tissue-/cell-type specific DNA methylation profiling studies. More importantly, the identified osteoblastic-specific DMRs were significantly enriched for genes that are critical for bone metabolism and skeletal development in physiologic and pathologic conditions, providing compelling evidence that DNA methylation may regulate transcription including cell-type specific isoform expression of many genes that are important for osteoblast differentiation and activities. Our results provided a framework for development of more specific hypotheses concerning epigenetic regulation of osteogenesis and highlighted several interesting targets for further evaluation. Future studies with multiple biological replicates and *in vitro* as well as *in vivo* functional assays are required to further replicate our findings and elucidate the molecular mechanisms underlying the DMR-mediated regulation of osteoblast differentiation and function, particularly for the numerous DMRs found in gene body areas.

4. Materials and methods

4.1. Samples and DNA methylation profiling

Genome-wide DNA methylation profiles of osteoblast and 20 additional different types of non-transformed human cell strains plus 4 Epstein-Barr virus-transformed LCLs were downloaded from the ENCODE website (<http://genome.ucsc.edu/cgi-bin/hgFileUi?db=hg19&g=wgEncodeHaibMethylRbbs>). There are two isogenic replicates for each cell line (except for myoblasts), which were replicates derived from the same human donor but have been treated separately, i.e., two growths of the same cell line, two separate library preparations, and two separate sequencing runs. A total of 51 DNA methylation datasets (BED files) were obtained (Supplementary Table 1).

4.2. Differentially methylation analysis

All statistical analyses were performed using R version 3.0.2. The identification of DMCs was performed by using the methylKit package (Akalin et al., 2012). Specifically, at each tested CpG site, we fitted a logistic regression model for the proportion of methylated cytosines in osteoblasts vs. all other samples. To be conservative, only CpGs with sequence coverage $\geq 10\times$ across all the cell lines were included in the analysis, and the significant DMCs were defined as CpGs showing absolute difference in methylation level of $\geq 50\%$ between osteoblasts and other cells at a significance level of $q\text{-value} \leq 0.01$. The $q\text{-values}$ correspond to multiple testing adjusted $p\text{-values}$ using the sliding linear

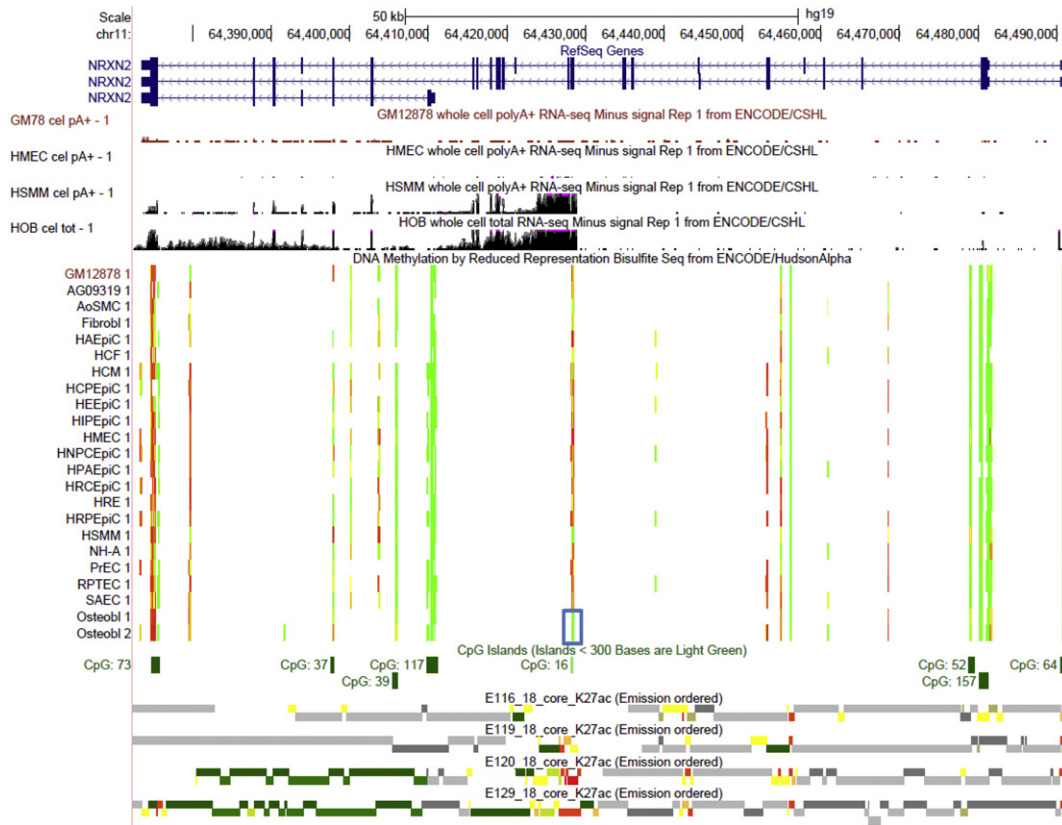


Fig. 6. Osteoblastic hypomethylation at *NRXN2* gene region. The same tracks as illustrated in Fig. 3 are shown using the UCSC Genome Browser (<http://genome.ucsc.edu>, version hg19). Osteoblast-specific DMR is highlighted in the blue box.

model (SLIM) (Wang et al., 2011). Hierarchical clustering analysis using the significant DMCs was also carried out in the methylKit.

For the identification of DMRs, we used the eDMR package (Li et al., 2013), which can directly take objects from methylKit and perform regional optimization calling and DMR statistical analysis and filtering. Specifically, the program uses a bimodal normal distribution to identify the optimum cutoff for calling a gap between two DMRs (Li et al., 2013). The DMR identification were restricted to those regions that contain ≥ 5 CpGs including ≥ 3 DMCs and have absolute mean methylation difference $> 20\%$ between the osteoblasts and the other cell types. The statistical significance of DMRs was calculated by combining the p-values of DMCs within the refined regions through the Stouffer-Liptak test (Pedersen et al., 2012). A FDR (False Discovery Rate) correction was also applied to correct for multiple hypothesis testing for the combined p-values. The significant DMRs are those with $q < 0.001$.

4.3. Annotation analysis

DMCs and DMRs were characterized with respect to different genic regions (promoters, exons, introns, 5'UTRs, 3'UTRs and intergenic regions) and different regions relative to CGIs, including CGIs, CGI shores (2 kb regions flanking CGIs), CGI shelf (2 kb regions flanking CGI shores), and open sea (> 4 kb to the nearest CGIs). The annotation files of RefSeq genes and CGIs were downloaded from the UCSC genome browser (<http://genome.ucsc.edu/cgi-bin/hgTables>). In the event that a DMC was mapped to multiple different CGI regions, we assigned the DMC to a single CGI region based on the priority order: CGI > CGI shore > CGI shelf > open sea.

To assist the functional annotation of the identified DMCs/DMRs, a variety of chromatin epigenomic (histone modification marks and DNase I hypersensitivity) and transcriptomic (RNA-seq) profiles were obtained from the ENCODE project (E.P. Consortium, 2012) via the UCSC genome browser. All these transcriptomic and chromatin-related epigenomic

data were generated from the same set of cell lines as those for the DNA methylation data. Additionally, we obtained combinatorial chromatin states (the 18-state model) in several cell lines (osteoblast, HSMM, GM12878, and HMEC) from the NIH Roadmap Epigenomics Consortium (Bernstein et al., 2010), which predicted the chromatin states by using the ChromHMM (Ernst and Kellis, 2012) approach on the ENCODE chromatin-related data. For the DMRs overlapping with regions of different chromatin states, we assigned them to the one state with larger proportion of overlap. Functional annotation and GO enrichment analysis of DMRs were carried out by using the GREAT package (McLean et al., 2010) with human reference genome GRCh37 (UCSC hg19, Feb/2009) as background. GREAT assigns biological meaning to a set of potential regulatory genomic regions (e.g., DMRs) by associating genomic regions with nearby genes and applying the gene annotations to the regions. Association is a two-step process. First, every gene is assigned a regulatory domain consisting of a basal domain that extends 5 kb upstream and 1 kb downstream from its TSS (regardless of other nearby genes), and an extension domain that extends in both directions up to the basal regulatory domain of the nearest upstream and downstream genes within 1 Mb (McLean et al., 2010). Then, each potential regulatory genomic region is associated with all genes whose regulatory domain it overlaps.

4.4. Enrichment of osteoporosis-associated genes among DMRs

We also investigated whether genes associated with DMRs are enriched for genetic variants underlying osteoporosis. Specifically, 143 genes which contain SNPs with p-value $\leq 5 \times 10^{-6}$ for association with bone mineral density (BMD) were recognized as osteoporosis-associated genes, based on the data released from the largest to date meta-analysis in the bone field by the Genetic Factors for Osteoporosis Consortium (GEFOS), including 32,961 individuals from 17 GWASs for BMD (Estrada et al., 2012). The significance of enrichment for osteoporosis-associated genes is tested by hypergeometric test comparing the DMR-annotated

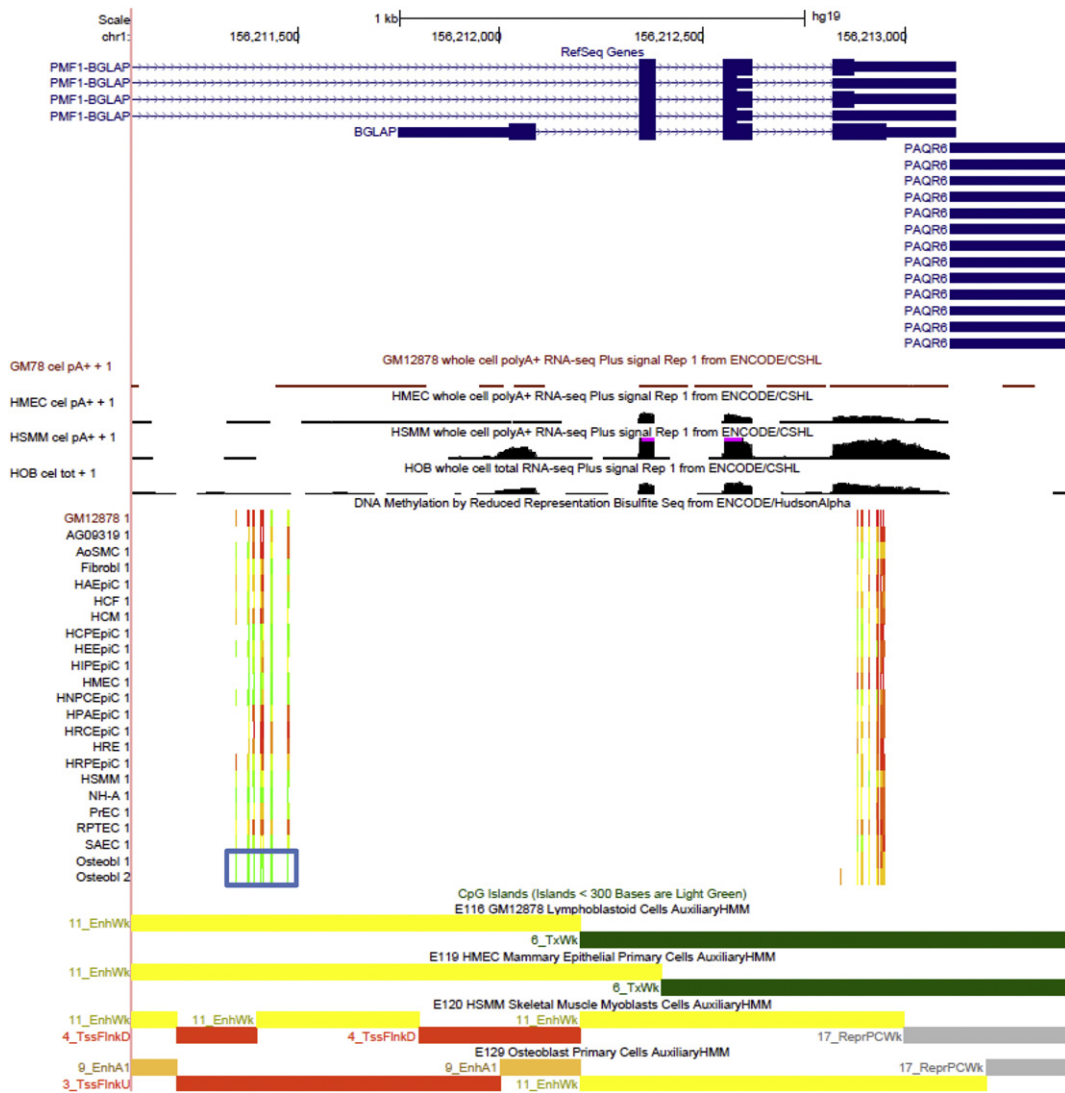


Fig. 7. Osteoblastic hypomethylation at *BGLAP* gene region. The same tracks as illustrated in Fig. 3 are shown using the UCSC Genome Browser (<http://genome.ucsc.edu>, version hg19). Osteoblast-specific DMR is highlighted in the blue box.

genes and all the 11,329 genes (as a background set) annotated to the 182,518 tested CpGs.

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.bonr.2017.04.001>.

Disclosure of interest

All authors declare that they have no conflicts of interest.

Acknowledgements

The investigators of this work were partially supported by grants from the NIH (R01AG026564, R01AR050496, R01AR057049, R01AR059781 and P20GM109036), and the Edward G. Schlieder Endowment as well as the Drs. W. C. Tsai and P. T. Kung Professorship in Biostatistics from Tulane University. The authors declare no conflicts of interest.

References

Akalın, A., Kormaksson, M., Li, S., Garrett-Bakelman, F.E., Figueroa, M.E., Melnick, A., Mason, C.E., 2012. methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biol.* 13 (10), R87.

- Andrade, A.C., Lui, J.C., Nilsson, O., 2010. Temporal and spatial expression of a growth-regulated network of imprinted genes in growth plate. *Pediatr. Nephrol.* 25 (4), 617–623.
- Arnsdorf, E.J., Tummala, P., Castillo, A.B., Zhang, F., Jacobs, C.R., 2010. The epigenetic mechanism of mechanically induced osteogenic differentiation. *J. Biomech.* 43 (15), 2881–2886.
- Ball, M.P., Li, J.B., Gao, Y., Lee, J.H., LeProust, E.M., Park, I.H., Xie, B., Daley, G.Q., Church, G.M., 2009. Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. *Nat. Biotechnol.* 27 (4), 361–368.
- Baylin, S.B., Jones, P.A., 2011. A decade of exploring the cancer epigenome - biological and translational implications. *Nat. Rev. Cancer* 11 (10), 726–734.
- Bernstein, B.E., Stamatoyannopoulos, J.A., Costello, J.F., Ren, B., Milosavljevic, A., Meissner, A., Kellis, M., Marra, M.A., Beaudet, A.L., Ecker, J.R., Farnham, P.J., Hirst, M., Lander, E.S., Mikkelsen, T.S., Thomson, J.A., 2010. The NIH roadmap epigenomics mapping consortium. *Nat. Biotechnol.* 28 (10), 1045–1048.
- Bustos-Valenzuela, J.C., Fujita, A., Halcsik, E., Granjeiro, J.M., Sogayar, M.C., 2011. Unveiling novel genes upregulated by both rhBMP2 and rhBMP7 during early osteoblastic transdifferentiation of C2C12 cells. *BMC Res. Notes* 4, 370.
- Byun, H.M., Siegmund, K.D., Pan, F., Weisenberger, D.J., Kanel, G., Laird, P.W., Yang, A.S., 2009. Epigenetic profiling of somatic tissues from human autopsy specimens identifies tissue- and individual-specific DNA methylation patterns. *Hum. Mol. Genet.* 18 (24), 4808–4817.
- C. Roadmap Epigenomics, Kundaje, A., Meuleman, W., Ernst, J., Bilenyk, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J., Ziller, M.J., Amin, V., Whitaker, J.W., Schultz, M.D., Ward, L.D., Sarkar, A., Quon, G., Sandstrom, R.S., Eaton, M.L., Wu, Y.C., Pfennig, A.R., Wang, X., Clausnitzer, M., Liu, Y., Coarfa, C., Harris, R.A., Shores, N., Epstein, C.B., Gjonneska, E., Leung, D., Xie, W., Hawkins, R.D., Lister, R., Hong, C., Gascard, P., Mungall, A.J., Moore, R., Chuah, E., Tam, A., Canfield, T.K., Hansen, R.S., Kaul, R., Sabo, P.J., Bansal, M.S., Carles, A., Dixon, J.R., Farh, K.H., Feizi, S., Karlic, R., Kim, A.R., Kulkarni, A., Li, D., Lowdon, R., Elliott, G., Mercer, T.R., Neph,

- S.J., Onuchic, V., Polak, P., Rajagopal, N., Ray, P., Sallari, R.C., Siebenthal, K.T., Sinnott-Armstrong, N.A., Stevens, M., Thurman, R.E., Wu, J., Zhang, B., Zhou, X., Beaudet, A.E., Boyer, L.A., De Jager, P.L., Farnham, P.J., Fisher, S.J., Haussler, D., Jones, S.J., Li, W., Marra, M.A., McManus, M.T., Sunyaev, S., Thomson, J.A., Tlsty, T.D., Tsai, L.H., Wang, W., Waterland, R.A., Zhang, M.Q., Chadwick, L.H., Bernstein, B.E., Costello, J.F., Ecker, J.R., Hirst, M., Meissner, A., Milosavljevic, A., Ren, B., Stamatoyannopoulos, J.A., Wang, T., Kellis, M., 2015. Integrative analysis of 111 reference human epigenomes. *Nature* 518 (7539), 317–330.
- Caliskan, M., Cusanovich, D.A., Ober, C., Gilad, Y., 2011. The effects of EBV transformation on gene expression levels and methylation profiles. *Hum. Mol. Genet.* 20 (8), 1643–1652.
- Cawthorn, W.P., Bree, A.J., Yao, Y., Du, B., Hemati, N., Martinez-Santibanez, G., MacDougald, O.A., 2012. Wnt6, Wnt10a and Wnt10b inhibit adipogenesis and stimulate osteoblastogenesis through a beta-catenin-dependent mechanism. *Bone* 50 (2), 477–489.
- Chandra, S., Baribault, C., Lacey, M., Ehrlich, M., 2014. Myogenic differential methylation: diverse associations with chromatin structure. *Biology (Basel)* 3 (2), 426–451.
- Chen, M.F., 1983. Endocrinologic and immunologic function in uremic patients with as-thenia of the kidney. *Zhong Xi Yi Jie He Za Zhi* 3 (6), 328–330.
- E.P. Consortium, 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature* 489 (7414), 57–74.
- Dachtler, J., Glasper, J., Cohen, R.N., Ivorra, J.L., Swiffen, D.J., Jackson, A.J., Harte, M.K., Rodgers, R.J., Clapcote, S.J., 2014. Deletion of alpha-neurexin II results in autism-related behaviors in mice. *Transl. Psychiatry* 4, e484.
- Davies, M.N., Volta, M., Pidsley, R., Lunnon, K., Dixit, A., Lovestone, S., Coarfa, C., Harris, R.A., Milosavljevic, A., Troakes, C., Al-Sarraj, S., Dobson, R., Schalkwyk, L.C., Mill, J., 2012. Functional annotation of the human brain methylome identifies tissue-specific epigenetic variation across brain and blood. *Genome Biol.* 13 (6), R43.
- Deaton, A.M., Webb, S., Kerr, A.R., Illingworth, R.S., Guy, J., Andrews, R., Bird, A., 2011. Cell type-specific DNA methylation at intragenic CpG islands in the immune system. *Genome Res.* 21 (7), 1074–1086.
- Divers, J., Palmer, N.D., Lu, L., Register, T.C., Carr, J.J., Hicks, P.J., Hightower, R.C., Smith, S.C., Xu, J., Cox, A.J., Hruska, K.A., Bowden, D.W., Lewis, C.E., Heiss, G., Province, M.A., Borecki, I.B., Kerr, K.F., Chen, Y.D., Palmas, W., Rotter, J.L., Wassel, C.L., Bertoni, A.G., Herrington, D.M., Wagenknecht, L.E., Langefeld, C.D., Freedman, B.I., 2013. Admixture mapping of coronary artery calcified plaque in African Americans with type 2 diabetes mellitus. *Circ. Cardiovasc. Genet.* 6 (1), 97–105.
- Doherty, T.M., Asotra, K., Fitzpatrick, L.A., Qiao, J.H., Wilkin, D.J., Detrano, R.C., Dunstan, C.R., Shah, P.K., Rajavashisth, T.B., 2003. Calcification in atherosclerosis: bone biology and chronic inflammation at the arterial crossroads. *Proc. Natl. Acad. Sci. U. S. A.* 100 (20), 11201–11206.
- Ehrlich, M., Lacey, M., 2013. DNA methylation and differentiation: silencing, upregulation and modulation of gene expression. *Epigenomics* 5 (5), 553–568.
- Ernst, J., Kellis, M., 2012. ChromHMM: automating chromatin-state discovery and characterization. *Nat. Methods* 9 (3), 215–216.
- Ernst, J., Kheradpour, P., Mikkelsen, T.S., Shores, N., Ward, L.D., Epstein, C.B., Zhang, X., Wang, L., Issner, R., Coyne, M., Ku, M., Durham, T., Kellis, M., Bernstein, B.E., 2011. Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* 473 (7345), 43–49.
- Eslaminejad, M.B., Fani, N., Shahhoseini, M., 2013. Epigenetic regulation of osteogenic and chondrogenic differentiation of mesenchymal stem cells in culture. *Cell J.* 15 (1), 1–10.
- Estrada, K., Styrkarsdottir, U., Evangelou, E., Hsu, Y.H., Duncan, E.L., Ntzani, E.E., Oei, L., Albagha, O.M., Amin, N., Kemp, J.P., Koller, D.L., Li, G., Liu, C.T., Minster, R.L., Moayyeri, A., Vandenput, L., Willner, D., Xiao, S.M., Yerges-Armstrong, L.M., Zheng, H.F., Alonso, N., Eriksson, J., Kammerer, C.M., Kaptoge, S.K., Leo, P.J., Thorleifsson, G., Wilson, S.G., Wilson, J.F., Aalto, V., Alen, M., Aragaki, A.K., Aspelund, T., Center, J.R., Dailiana, Z., Duggan, D.J., Garcia, M., Garcia-Giralt, N., Giroux, S., Hallmans, G., Hocking, L.J., Husted, L.B., Jameson, K.A., Khusainova, R., Kim, G.S., Kooperberg, C., Koromila, T., Kruk, M., Laaksonen, M., Lacroix, A.Z., Lee, S.H., Leung, P.C., Lewis, J.R., Masi, L., Mencej-Bedrac, S., Nguyen, T.V., Nogueux, X., Patel, M.S., Prezelj, J., Rose, L.M., Scollen, S., Siggeirsdottir, K., Smith, A.V., Svensson, O., Trompet, S., Trummer, O., van Schoor, N.M., Woo, J., Zhu, K., Balcells, S., Brandi, M.L., Buckley, B.M., Cheng, S., Christiansen, C., Cooper, C., Dedoussis, G., Ford, I., Frost, M., Goltzman, D., Gonzalez-Macias, J., Kahonen, M., Karlsson, M., Khusnutdinova, E., Koh, J.M., Kollia, P., Langdahl, B.L., Leslie, W.D., Lips, P., Ljunggren, O., Lorenc, R.S., Marc, J., Mellstrom, D., Obermayer-Pietsch, B., Olmos, J.M., Pettersson-Kymmer, U., Reid, D.M., Riancho, J.A., Ridker, P.M., Rousseau, F., Sulgebom, P.E., Tang, N.L., Urreitzti, R., Van Hul, W., Viikari, J., Zarrabeitia, M.T., Aulchenko, Y.S., Castano-Betancourt, M., Grundberg, E., Herrera, L., Ingvarsson, T., Johannsdottir, H., Kwan, T., Li, R., Luben, R., Medina-Gomez, C., Palsson, S.T., Reppe, S., Rotter, J.L., Sigurdsson, G., van Meurs, J.B., Verlaan, D., Williams, F.M., Wood, A.R., Zhou, Y., Gautvik, K.M., Pastinen, T., Raychaudhuri, S., Cauley, J.A., Chasman, D.I., Clark, G.R., Cummings, S.R., Danoy, P., Dennis, E.M., Eastell, R., Eisman, J.A., Gudnason, V., Hofman, A., Jackson, R.D., Jones, G., Jukema, J.W., Khaw, K.T., Lehtimaki, T., Liu, Y., Lorentzon, M., McCloskey, E., Mitchell, B.D., Nandakumar, K., Nicholson, G.C., Oostra, B.A., Peacock, M., Pols, H.A., Prince, R.L., Raitakari, O., Reid, I.R., Robbins, J., Sambrook, P.N., Sham, P.C., Shuldiner, A.R., Tylavsky, F.A., van Duijn, C.M., Wareham, N.J., Cupples, L.A., Econs, M.J., Evans, D.M., Harris, T.B., Kung, A.W., Psaty, B.M., Reeve, J., Spector, T.D., Streeten, E.A., Zillikens, M.C., Thorsteinsdottir, U., Ohlsson, C., Karasik, D., Richards, J.B., Brown, M.A., Stefansson, K., Uitterlinden, A.G., Ralston, S.H., Ioannidis, J.P., Kiel, D.P., Rivadeneira, F., 2012. Genome-wide meta-analysis identifies 56 bone mineral density loci and reveals 14 loci associated with risk of fracture. *Nat. Genet.* 44 (5), 491–501.
- Futscher, B.W., Oshiro, M.M., Wozniak, R.J., Holtan, N., Hanigan, C.L., Duan, H., Domann, F.E., 2002. Role for DNA methylation in the control of cell type specific maspin expression. *Nat. Genet.* 31 (2), 175–179.
- Gauthier, J., Siddiqui, T.J., Huashan, P., Yokomaku, D., Hamdan, F.F., Champagne, N., Lapointe, M., Spiegelman, D., Noreau, A., Lafreniere, R.G., Fathalli, F., Joobar, R., Krebs, M.O., DeLisi, L.E., Mottron, L., Fombonne, E., Michaud, J.L., Drapeau, P., Carbonetto, S., Craig, A.M., Rouleau, G.A., 2011. Truncating mutations in NRXN2 and NRXN1 in autism spectrum disorders and schizophrenia. *Hum. Genet.* 130 (4), 563–573.
- Gimble, J.M., Guilak, F., Nuttall, M.E., Sathishkumar, S., Vidal, M., Bunnell, B.A., 2008. In vitro differentiation potential of mesenchymal stem cells. *Transfus. Med. Hemother.* 35 (3), 228–238.
- Glass, G.E., Chan, J.K., Freidin, A., Feldmann, M., Horwood, N.J., Nanchahal, J., 2011. TNF-alpha promotes fracture repair by augmenting the recruitment and differentiation of muscle-derived stromal cells. *Proc. Natl. Acad. Sci. U. S. A.* 108 (4), 1585–1590.
- Goshu, E., Jin, H., Fasnacht, R., Sepenski, M., Michaud, J.L., Fan, C.M., 2002. Sim2 mutants have developmental defects not overlapping with those of Sim1 mutants. *Mol. Cell Biol.* 22 (12), 4147–4157.
- Hagarman, J.A., Motley, M.P., Kristjansdottir, K., Soloway, P.D., 2013. Coordinate regulation of DNA methylation and H3K27me3 in mouse embryonic stem cells. *PLoS One* 8 (1), e53880.
- Hojo, H., Ohba, S., Chung, U., 2015. Signaling pathways regulating the specification and differentiation of the osteoblast lineage. *Regen. Ther.* 1, 57–62.
- Hsiao, S.H., Lee, K.D., Hsu, C.C., Tseng, M.J., Jin, V.X., Sun, W.S., Hung, Y.C., Yeh, K.T., Yan, P.S., Lai, Y.Y., Sun, H.S., Lu, Y.J., Chang, Y.S., Tsai, S.J., Huang, T.H., Leu, Y.W., 2010. DNA methylation of the Trip10 promoter accelerates mesenchymal stem cell lineage determination. *Biochem. Biophys. Res. Commun.* 400 (3), 305–312.
- Huntriss, J.D., Hemmings, K.E., Hinkins, M., Rutherford, A.J., Sturme, R.G., Elder, K., Picton, H.M., 2013. Variable imprinting of the MEST gene in human preimplantation embryos. *Eur. J. Hum. Genet.* 21 (1), 40–47.
- Jones, P.A., 2012. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat. Rev. Genet.* 13 (7), 484–492.
- Kamei, Y., Suganami, T., Kohda, T., Ishino, F., Yasuda, K., Miura, S., Ezaki, O., Ogawa, Y., 2007. Peg1/Mest in obese adipose tissue is expressed from the paternal allele in an isoform-specific manner. *FEBS Lett.* 581 (1), 91–96.
- Kirkham, G.R., Cartmell, S.H., 2007. Genes and proteins involved in the regulation of osteogenesis. In: Ashammakhi, N., Reis, R., Chiellini, E. (Eds.), *Topics in Tissue Engineering*.
- Kosaki, K., Kosaki, R., Craigen, W.J., Matsuo, N., 2000. Isoform-specific imprinting of the human PEG1/MEST gene. *Am. J. Hum. Genet.* 66 (1), 309–312.
- Kubo, H., Shimizu, M., Taya, Y., Kawamoto, T., Michida, M., Kaneko, E., Igarashi, A., Nishimura, M., Segoshi, K., Shimizu, Y., Tsuji, K., Aoba, T., Kato, Y., 2009. Identification of mesenchymal stem cell (MSC)-transcription factors by microarray and knockdown analyses, and signature molecule-marked MSC in bone marrow by immunohistochemistry. *Genes Cells* 14 (3), 407–424.
- Lee, S.M., Lee, Y.G., Bae, J.B., Choi, J.K., Tayama, C., Hata, K., Yun, Y., Seong, J.K., Kim, Y.J., 2014. Hbx induces hypomethylation of distal intragenic CpG islands required for active expression of developmental regulators. *Proc. Natl. Acad. Sci. U. S. A.* 111 (26), 9555–9560.
- Li, Y., Meng, G., Guo, Q.N., 2008. Changes in genomic imprinting and gene expression associated with transformation in a model of human osteosarcoma. *Exp. Mol. Pathol.* 84 (3), 234–239.
- Li, S., Garrett-Bakelman, F.E., Akalin, A., Zumbo, P., Levine, R., B.L. To, Lewis, I.D., Brown, A.L., D'Andrea, R.J., Melnick, A., Mason, C.E., 2013. An optimized algorithm for detecting and annotating regional differential methylation. *BMC Bioinf.* 14 (Suppl. 5), S10.
- Liu, Y., Ding, J., Reynolds, L.M., Lohman, K., Register, T.C., De La Fuente, A., Howard, T.D., Hawkins, G.A., Cui, W., Morris, J., Smith, S.G., Barr, R.G., Kaufman, J.D., Burke, G.L., Post, W., Shea, S., McCall, C.E., Siscovick, D., Jacobs Jr., D.R., Tracy, R.P., Herrington, D.M., Hoeschele, I., 2013. Methylocines of gene expression in human monocytes. *Hum. Mol. Genet.* 22 (24), 5065–5074.
- Lokk, K., Modhukur, V., Rajashekar, B., Martens, K., Magi, R., Kolde, R., Koltšina, M., Nilsson, T.K., Vilo, J., Salumets, A., Tonisson, N., 2014. DNA methylome profiling of human tissues identifies global and tissue-specific methylation patterns. *Genome Biol.* 15 (4), r54.
- Maekawa, M., Yamanaka, S., 2011. Glis1, a unique pro-reprogramming factor, may facilitate clinical applications of iPSC technology. *Cell Cycle* 10 (21), 3613–3614.
- Maekawa, M., Yamaguchi, K., Nakamura, T., Shibukawa, R., Kodanaka, I., Ichisaka, T., Kawamura, Y., Mochizuki, H., Goshima, N., Yamanaka, S., 2011. Direct reprogramming of somatic cells is promoted by maternal transcription factor Glis1. *Nature* 474 (7350), 225–229.
- Maunakea, A.K., Nagarajan, R.P., Bilenyk, M., Ballinger, T.J., D'Souza, C., Fouse, S.D., Johnson, B.E., Hong, C., Nielsen, C., Zhao, Y., Turecki, G., Delaney, A., Varhol, R., Thiessen, N., Shchors, K., Heine, V.M., Rowitch, D.H., Xing, X., Fiore, C., Schillebeckx, M., Jones, S.J., Haussler, D., Marra, M.A., Hirst, M., Wang, T., Costello, J.F., 2010. Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature* 466 (7303), 253–257.
- McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., Bejerano, G., 2010. GREAT improves functional interpretation of cis-regulatory regions. *Nat. Biotechnol.* 28 (5), 495–501.
- Meissner, A., Mikkelsen, T.S., Gu, H., Wernig, M., Hanna, J., Sivachenko, A., Zhang, X., Bernstein, B.E., Nusbaum, C., Jaffe, D.B., Gnirke, A., Jaenisch, R., Lander, E.S., 2008. Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* 454 (7205), 766–770.
- Meng, X., Peng, B., Shi, J., Zheng, Y., Chen, H., Zhang, J., Li, L., Zhang, C., 2006. Effects of overexpression of Sim2 on spatial memory and expression of synapsin I in rat hippocampus. *Cell Biol. Int.* 30 (10), 841–847.
- Mera, P., Laue, K., Wei, J., Berger, J.M., Karsenty, G., 2016. Osteocalcin is necessary and sufficient to maintain muscle mass in older mice. *Mol. Metab.* 5 (10), 1042–1047.

- Nestheide, S., Bridge, J.A., Barnes, M., Frayer, R., Sumegi, J., 2013. Pharmacologic inhibition of epigenetic modification reveals targets of aberrant promoter methylation in Ewing sarcoma. *Pediatr. Blood Cancer* 60 (9), 1437–1446.
- Pedersen, I.S., Dervan, P., McGoldrick, A., Harrison, M., Ponchel, F., Speirs, V., Isaacs, J.D., Gorey, T., McCann, A., 2002. Promoter switch: a novel mechanism causing biallelic PEG1/MEST expression in invasive breast cancer. *Hum. Mol. Genet.* 11 (12), 1449–1453.
- Pedersen, B.S., Schwartz, D.A., Yang, I.V., Kechris, K.J., 2012. Comb-p: software for combining, analyzing, grouping and correcting spatially correlated P-values. *Bioinformatics* 28 (22), 2986–2988.
- Rauch, T.A., Wu, X., Zhong, X., Riggs, A.D., Pfeifer, G.P., 2009. A human B cell methylome at 100-base pair resolution. *Proc. Natl. Acad. Sci. U. S. A.* 106 (3), 671–678.
- de la Rica, L., Rodriguez-Ubrea, J., Garcia, M., Islam, A.B., Urquiza, J.M., Hernando, H., Christensen, J., Helin, K., Gomez-Vaquero, C., Ballestar, E., 2013. PU.1 target genes undergo Tet2-coupled demethylation and DNMT3b-mediated methylation in monocyte-to-osteoclast differentiation. *Genome Biol.* 14 (9), R99.
- Rowen, L., Young, J., Birditt, B., Kaur, A., Madan, A., Philipps, D.L., Qin, S., Minx, P., Wilson, R.K., Hood, L., Graveley, B.R., 2002. Analysis of the human neurexin genes: alternative splicing and the generation of protein diversity. *Genomics* 79 (4), 587–597.
- Saferali, A., Grundberg, E., Berlivet, S., Beauchemin, H., Morcos, L., Polychronakos, C., Pastinen, T., Graham, J., McNeney, B., Naumova, A.K., 2010. Cell culture-induced aberrant methylation of the imprinted IG DMR in human lymphoblastoid cell lines. *Epigenetics* 5 (1), 50–60.
- Shamblott, M.J., Bugg, E.M., Lawler, A.M., Gearhart, J.D., 2002. Craniofacial abnormalities resulting from targeted disruption of the murine Sim2 gene. *Dev. Dyn.* 224 (4), 373–380.
- Shukla, S., Kavak, E., Gregory, M., Imashimizu, M., Shutinoski, B., Kashlev, M., Oberdoerffer, P., Sandberg, R., Oberdoerffer, S., 2011. CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. *Nature* 479 (7371), 74–79.
- Slieker, R.C., Bos, S.D., Goeman, J.J., Bovee, J.V., Talens, R.P., van der Breggen, R., Suchiman, H.E., Lameijer, E.W., Putter, H., van den Akker, E.B., Zhang, Y., Jukema, J.W., Slagboom, P.E., Meulenbelt, I., Heijmans, B.T., 2013. Identification and systematic annotation of tissue-specific differentially methylated regions using the Illumina 450k array. *Epigenetics Chromatin* 6 (1), 26.
- Song, F., Mahmood, S., Ghosh, S., Liang, P., Smiraglia, D.J., Nagase, H., Held, W.A., 2009. Tissue specific differentially methylated regions (TDMR): changes in DNA methylation during development. *Genomics* 93 (2), 130–139.
- Sun, Z., Cunningham, J., Slager, S., Kocher, J.P., 2015. Base resolution methylome profiling: considerations in platform selection, data preprocessing and analysis. *Epigenomics* 7 (5), 813–828.
- Tabuchi, K., Sudhof, T.C., 2002. Structure and evolution of neurexin genes: insight into the mechanism of alternative splicing. *Genomics* 79 (6), 849–859.
- Tanaka, K., Inoue, Y., Hendy, G.N., Canaff, L., Katagiri, T., Kitazawa, R., Komori, T., Sugimoto, T., Seino, S., Kaji, H., 2012. Interaction of Tmem119 and the bone morphogenetic protein pathway in the commitment of myoblastic into osteoblastic cells. *Bone* 51 (1), 158–167.
- Thierry-Mieg, D., Thierry-Mieg, J., 2006. AceView: a comprehensive cDNA-supported gene and transcripts annotation. *Genome Biol.* 7 (Suppl. 1), S12 1–S1214.
- Tsumagari, K., Baribault, C., Terragni, J., Varley, K.E., Gertz, J., Pradhan, S., Badoo, M., Crain, C.M., Song, L., Crawford, G.E., Myers, R.M., Lacey, M., Ehrlich, M., 2013. Early de novo DNA methylation and prolonged demethylation in the muscle lineage. *Epigenetics* 8 (3), 317–332.
- Varley, K.E., Gertz, J., Bowling, K.M., Parker, S.L., Reddy, T.E., Pauli-Behn, F., Cross, M.K., Williams, B.A., Stamatoyannopoulos, J.A., Crawford, G.E., Absher, D.M., Wold, B.J., Myers, R.M., 2013. Dynamic DNA methylation across diverse human cell lines and tissues. *Genome Res.* 23 (3), 555–567.
- Venza, I., Visalli, M., Fortunato, C., Ruggeri, M., Ratone, S., Caffo, M., Caruso, G., Alafaci, C., Tomasello, F., Teti, D., Venza, M., 2012. PGE2 induces interleukin-8 derepression in human astrocytoma through coordinated DNA demethylation and histone hyperacetylation. *Epigenetics* 7 (11), 1315–1330.
- Villagra, A., Gutierrez, J., Paredes, R., Sierra, J., Puchi, M., Imschenetzky, M., Wijnen Av, A., Lian, J., Stein, G., Stein, J., Montecino, M., 2002. Reduced CpG methylation is associated with transcriptional activation of the bone-specific rat osteocalcin gene in osteoblasts. *J. Cell. Biochem.* 85 (1), 112–122.
- Wang, H.Q., Tuominen, L.K., Tsai, C.J., 2011. SLIM: a sliding linear model for estimating the proportion of true null hypotheses in datasets with dependence structures. *Bioinformatics* 27 (2), 225–231.
- Weber, M., Hellmann, I., Stadler, M.B., Ramos, L., Paabo, S., Rebhan, M., Schubeler, D., 2007. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat. Genet.* 39 (4), 457–466.
- Yang, X., Shao, X., Gao, L., Zhang, S., 2015. Systematic DNA methylation analysis of multiple cell lines reveals common and specific patterns within and across tissues of origin. *Hum. Mol. Genet.* 24 (15), 4374–4384.
- Ziller, M.J., Gu, H., Muller, F., Donaghey, J., Tsai, L.T., Kohlbacher, O., De Jager, P.L., Rosen, E.D., Bennett, D.A., Bernstein, B.E., Gnirke, A., Meissner, A., 2013. Charting a dynamic DNA methylation landscape of the human genome. *Nature* 500 (7463), 477–481.