Identification of CpG methylation and non-CpG methylation is associate with Lymphocyte specific helicase

CURRENT STATUS: POSTED

Weiwei Lai
Central South University

Lianlian Ouyang
Central South University Xiangya School of Medicine

Na Liu
Central South University

Shouping Liu
Central South University

Ying Shi
Central South University

Ling Chen
Central South University

Xiang Wang
Second Xiangya Hospital Central South University

Desheng Xiao
Xiangya Hospital Central South University

Shuang Liu
Xiangya Hospital Central South University

Yongguang Tao
Central South University

Corresponding Author
taoyong@csu.edu.cn

ORCiD: https://orcid.org/0000-0003-2354-5321

DOI:
SUBJECT AREAS
Epigenetics & Genomics

KEYWORDS
Chromatin remodeling factor, Lymphocyte specific helicase, CpG methylation, Non-CpG methation, Transcription factor, CpG islands, Differentially methylated regions
Abstract

Background: CpG methylation is crucial for normal cell development and differentiation in mammalian genomes, whereas non-CpG methylation presents in stem cells and appears to be enriched in pluripotent cell types. Lymphocyte-specific helicase (LSH) is critical for normal cell development as a regulator of CpG methylation, but whether there is potential link between LSH and non-CpG methylation remains unclear.

Results: In this study, we using reduced representation bisulfite sequencing (RRBS), we found that LSH affected DNA methylation patterns in lung adenocarcinoma (LUAD) cells. We analyzed 12,239 differentially methylated regions (DMRs) for CpG methylation and 1605 DMRs for non-CpG methylation. LSH did not change CpG methylation at a global level, but partially change the methylation of CpG islands (CGIs). Interestingly, LSH reduced approximately 30% of CHH methylations at CGIs and consistently decreased CHG and CHH methylation patterns. However, LSH increased the CHH methylation level in satellite and simple-repeat region. The KEGG analysis showed LSH to be linked with a markedly enrichment in focal adhesion and the PI3K-Akt pathway. Further analysis indicated that LSH might epigenetically regulate the binding of transcription factors, including SP2, etc.

Conclusions: These findings suggest that LSH have no remarkable influence on CpG methylation but it partially alter non-CpG methylation and repeat sequences in LUAD. LSH inhibits non-CpG methylation in cancer and may function as a regulator to promote tumorigenesis through inhibiting non-CpG methylation.

Background

Gene expression is regulated epigenetically by DNA methylation, histone variants and modifications and nucleosome positioning. The methylation of cytosines at the carbon 5 position (5-methyl-cytosine, mC) constitutes the most common covalent modification of vertebrate genomic DNA [1–3]. In mammalian genomes, CG methylation (mCG) has been shown to play critical roles in genomic imprinting, X-chromosome inactivation, cellular differentiation, and tumor development [4–7]. In addition, high CG density (also known as CpG islands, or CGIs) is scattered on mammalian genomes
and overlaps the promoter regions of ~ 70% of human genes [8]. DNA methyltransferase 1 (DNMT1) is primarily associated with the replication fork and is crucial for the maintenance of methylation; loss of DNMT1 in mouse embryonic stem cells produces a dramatic loss of CpG methylation, while non-CpG methylation levels are not affected [9]. It is well accepted that DNA methylation occurring on CpG islands acts as a silencing mechanism during normal development and tumorigenesis.

DNA methylation has also been described in non-CpG contexts, such as in CHH and CHG (H indicates A, C or T). The methylation frequency at the dinucleotides occurs as follows: CpG > CpA > CpT > CpC, which has been described in dynamic DNA methylation studies [10–13]. Non-CpG methylation was first identified in a plant genome and was found to be independent of CpG methylation in various contexts within mammalian genomes [14]. The presence of non-CpG methylation is tightly linked to the de novo methyltransferases, DNMT3A, DNMT3B, and DNMT3L; it has been proposed that non-CpG methylation seems to be a result of the sustained expression of these proteins, particularly in nondividing cells such as neurons and oocytes [9, 15–17]. Studies have shown that non-CpG methylation is enriched in human embryonic stem cells (hESCs) compared with differentiated cells [18–20]. However, the depletion of DNMT3A in hESCs exerts no significant change on the expression of key pluripotency genes (OCT4, NANOG, and SOX2), suggesting that non-CpG methylation patterns are independent of pluripotency-associated gene expression [21]. It seems to be like that non-CpG methylation enriched in hESCs but do not accompanied with the change of pluripotency. Moreover, it is unclear which factors control non-CpG methylation and whether a dynamic regulation of non-CpG methylation could alter transcription.

Lymphocyte-specific helicase (LSH), an SNF2 family member, is critical for normal cell development and cancer progression and is a regulator of CpG methylation at repeat sequences and stem cell genes [22–32]. LSH might regulate the association of DNMT3b and contribute to the CpG methylation of developmental genes that results in RNA Pol II stalling and gene silencing [33, 34]. It has been reported that LSH epigenetic silencing of Oct4 expression mediates CpG methylation [35]. Here, we examine DNA methylation in LSH-overexpressing lung adenocarcinoma cells. We used reduced representation bisulfite sequencing (RRBS) techniques to explore the patterns of DNA methylation in
the LUAD cells and found that LSH inhibited non-CpG methylation and consistently decreased DNA methylation of CHG and CHH.

Results

Efficient detection of Global Methylation Patterns by Reduced Representation Bisulfite Sequencing (RRBS)

Our previous work established stable expression of LSH in H358 cells[32], we then performed western blotting to confirm the successful establishment of LSH expression in H358 cells (supplementary Figure S1A). To identify regions that were differentially methylated among the lung adenocarcinoma cells, we carried out representation bisulfite sequencing on cellular DNA (H358-vector and H358-O/E LSH) using Illumina sequencing by 10-fold coverage synthesis technology [36–38]. We first observed the average ratio of DNA methylation patterns in the vector and LSH groups; no obvious changes were found in these two groups (supplementary Figure S1B).

There was no apparent difference in the bisulfite conversion (BC) ratio in these two groups (supplementary Figure S1C). We then investigated whether the CpG or non-CpG methylation levels changed on a genome-wide scale over the course of LSH expression. For an in-depth analysis, we merged data from the vector and LSH to compare them.

By merging all CpG or non-CpG sites from these two groups, we found that the global CpG methylation level remained the same for < 10% of the reads (pink column) and differed for 90–100% of the reads (purple column) in these two groups, whereas the CHG and CHH methylation levels (< 10% of reads, pink column) were decreased in the LSH group (Fig. 1A). Thus, the percentages of methylated CHG and CHH continued to decrease upon the expression of LSH. Next, we plotted the average methylation levels of nonoverlapping 1 Mb sliding windows along each chromosome. The global changes in non-CpG methylation patterns showed visibly decreased in multiple windows between the vector and overexpressed LSH cells, such as at chromosome 8, 16, 20 (Fig. 1B). These results indicated that CpG methylation levels remained at the same global level, and non-CpG sites were consistently not methylated in the LSH group compared to the vector.

LSH affects DNA methylation levels of different functional regions

We divided all mC patterns into specific gene features: upstream 2 kb (prepromoter 2000 bp), 5’UTR (untranslated region), exons, introns, 3’UTR, downstream 2 kb, CGIs (CpG islands), CGI-shore and
repeats (repeat sequence). The methylation levels were evaluated in these functional regions. A violin graph was plotted with dots representing different methylation levels at gene element. We found that the methylation levels had no marked differences in regions of the violin plot for CG methylation types, while, the non-CpG methylation levels of LSH group were lower in some regions compared to vector, such as intron, upstream 2 k and CGI. An approximately 30% reduction of CHH types at CGIs was found in the LSH group (Fig. 2A). Our previous study showed that LSH promoted genome stability by silencing satellite expression through affecting the 5-hmC levels in pericentromeric satellite repeats [26]. However, the levels of CpG methylation and non-CpG methylation at repeat regions showed no evidently differences between the vector and LSH group. We then subdivided the repeat region into specific genes: short interspersed nuclear elements (SINEs), long interspersed nuclear elements (LINEs), long terminal repeats (LTRs), satellites, DNA transposons, simple repeats, and low complexity. The results showed an increased level of CHH methylation at satellites and simple repeat regions (Fig. 2B). We also measured the DMR-related genes (DMGs) methylation level. The methylation level in the specified regions of these two groups (vector and LSH) were similar at the CpG methylation pattern, while the methylation level for the CG type were higher than CHG and CHH types. In addition, high levels of non-CpG methylation were detected in upstream, exon and intron region with LSH expression. Interestingly, the CHG and CHH methylation level of DMGs in LSH group showed hypomethylation and stable in each functional element compared to vector (Fig. 2C).

Overall, LSH did not affect CpG methylation but reduced non-CpG methylation, except for at satellite and simple repeat elements, in the human genome. Thus, LSH may play a unique role in the distribution of non-CpG methylation during lung adenocarcinoma development.

Characterization of DMRs; enrichment through GO and KEGG analysis

DMRs were detected and annotated into gene functional elements according to different methylation types. In total, 12,239 CG DMRs, 342 CHG DMRs and 1263 CHH DMRs were identified, most of which were in distal intragenic regions, with only 1255 (CG), 28 (CHG) and 86 DMRs in promoter regions (Fig. 3A). The length of the DMRs with different methylation patterns were measured (Fig. 3B). We further analyzed the correlation between LSH, hypo-DMRs, and hyper-DMRs for different DNA
methylation types. We found that LSH inhibited hyper-DMRs non-CpG methylation and have an opposite effect on hyper-DMRs, however, there was no apparently change on the CpG methylation pattern (Fig. 3C). DMR cluster analysis showed that LSH mainly inhibited non-CpG methylation (CHG and CHH), and few changes were found in CpG methylation (Fig. 3D). For all methylation types, the ratio of DMRs located in introns was the highest, except for that in distal intergenic regions (Fig. 4A). The evidence showed that DNA methylation modified the DNA binding specificity of transcription factors (TFs) [39]. We then observed the ability of TFs binding to DNA, a heat map was generated using a cluster analysis of the DMRs (TF binding sites per 100 DMRs) in the promoter and gene body with high level of DNA methylation or low level of DNA methylation in the LSH group. We found that the promoter region showed a high level of TFs binding sites compared to the gene body. Furthermore, hypomethylation enhanced the ability of TFs bind to DNA both in promoter and gene body region which was found to be in the LSH group with low level of DNA methylation (Fig. 4B), eg, SP2, SP3, etc. These results suggest that LSH may epigenetically regulate the ability of these TFs bind to DNA through affect DNA methylation at the gene promoter region.

To probe changes in the methylation status of gene functions under prolificacy conditions, the GO and KEGG pathway databases were analyzed to characterize the DMR-related genes (DMGs) for different methylation types. The GO analysis revealed that for CG, CHG, and CHH methylation, DMGs were markedly enriched in the categories of protein binding and plasma membrane for CG methylation and CHG methylation, and the cellular component of CHH methylation was related to the cytoplasm. The molecular function category of CG methylation was transcription DNA-template, while the non-CpG methylations (CHG, CHH) were clustered at homophillic cell adhesion (Fig. 4C). The KEGG analysis revealed that for different methylation types, the DMGs were enriched in the metabolic pathways of CG methylation. CHG methylation was enriched in metabolic pathways and the cGMP-PKG signaling pathway. The most enriched molecular functions of CHH were focal adhesion and the PI3K-Akt pathway (Fig. 4D). Importantly, we also found that some DMGs were involved in regulating the actin cytoskeleton. Taken together, our results indicate that these specific genes, which are influenced by CG methylation, mainly affect metabolism in cancer. The influence of LSH in non-CpG methylation
mostly associated with cancer development and the cellular cytoskeleton, subsequently impacting cancer cell proliferation.

Verification of sequencing results
To validate the sequencing results, we performed Venn analysis to identify the correlation between RRBS data and our previous transcriptome data in these two groups [27]. A total of 12 genes were identified to be upregulated and hypomethylated by LSH expression (Fig. 5A). Next, we observed the relationship between gene expression and methylation type in a systematic way. The transcription site (TSS) regions had a lower level of DNA methylation and higher gene expression compared with other regions, and this was not related to the DNA methylation type (CG, CHG or CHH) or LSH expression (Fig. 5B). The DNA methylation levels in the upstream 2 kb and downstream 2 kb regions were moderate. The non-CpG methylation patterns were relatively stable in the gene body region; however, the CpG methylation was disordered when integrated with the transcriptome and methylation data. This phenomenon may be involved when non-CpG methylation patterns mostly occur in the gene body region. However, there was no difference between transcriptome and methylation sequencing in our system.

We next investigated whether the differences were specifically reflected in the individual differences of certain genes. Four genes were randomly selected from the Venn analysis results in which DMRs were located in different regions. Our results showed that the CG methylation levels of transient receptor potential melastatin 5 (TRPM5), growth arrest-specific 2-like 1 (GAS2L1), intercellular adhesion molecule 5 (ICAM5) and solute carrier family 10 member 7 (SLC10A7) were not clearly changed compared to vector. However, the non-CpG methylation levels were apparently decreased under LSH expression (Fig. 6). To predict the DNA methylation level of these genes in normal and cancer samples we used MethHC, which is a DNA methylation and gene expression database for human cancer (http://methhc.mbc.nctu.edu.tw/php/index.php). GAS2L1 showed decreased DNA methylation levels in cancer despite its different regions (promoter, CpG islands and gene body) (supplementary Figure S2A). The DNA methylation levels of CpG islands and the gene body in TRPM5 were significantly reduced in cancer, while the promoter region methylation level had no significant
change (supplementary Figure S2B). Regions of CpG islands and the 5’UTR in ICAM5 displayed increased DNA methylation levels in cancer, but there was no significant change in the promoter region (supplementary Figure S2C). SLC10A7 also showed significantly decreased levels of methylation in CpG islands and the 3’UTR (supplementary Figure S2D). The TCGA database results show that these genes were significantly upregulated in cancer (except GAS2L1 was downregulated) (supplementary Figure S3A). We then carried out a correlation analysis to identify the relationship between LSH and these genes. Our results revealed that TRPM5 was positively correlated with LSH, while the other genes were negatively correlated with LSH expression (supplementary Figure S3B). We further used the KM-Plotter database (http://kmplot.com/analysis/index.php?p=background) to determine their prognosis in LUAD. High expression of TRPM5 and GAS2L1 indicated a poor prognosis in LUAD; SLC10A7 contributed to the LUAD prognosis, and ICAM5 had no significant correlation with prognosis in LUAD (supplementary Figure S3C). Taken together, these results indicate that LSH affects non-CpG methylation in certain genes and may epigenetically regulate their expression in LUAD.

Discussion

DNA methylation is the main feature of the epigenetic regulatory mechanism that plays an important role in the regulation of gene expression. Previous studies have demonstrated a role of LSH in supporting CpG methylation in normal cell developmental genes [23, 40], but there are few reports regarding cancerous non-CpG methylation. RRBS, which allows unbiased genome-wide DNA methylation profiling, has allowed us to investigate proliferation-related DNA methylation in unprecedented detail [38]. In this study, we used RRBS to investigate the DNA methylation profile of H358 cells to discover the relationship between different methylation types and LSH. We found that LSH inhibits non-CpG genome methylation, including at repeat sequence in cancer, but it does not clearly affect CG methylation. Further correlation analyses indicated that several DMR-related genes were most likely involved in cancer development and cell proliferation.

For the final analyses, we merged samples from the two groups, and this merging not only increased the genomic coverage significantly but also minimized the impact of genome complexity. However,
individual characteristics, such as cell growth status, and differences in stress responses from stimulated cell proliferation could be confounding factors. Nevertheless, our findings that LSH affects non-CpG methylation during cancer echo an earlier study on the effect of LSH on genome instability; this study found that LSH promotes genome stability through silencing satellite expression [26]. Despite the alteration of DNA methylation being considered as a classical hallmark of cancer [41, 42], the role of non-CpG methylation remains controversial, and its relevance in normal and cancer processes is poorly understood. The higher methylation levels of non-CpG sites in human pluripotent stem cells and brain cells is consistent with previous reports that found a high potential for maturation or further differentiation [13, 20, 43]. Non-CpG methylation has been reported in other important cancer-related genes, such as glutathione S-transferase pi 124(GSTP124), HIF1-α and TP53, in endometrial, breast and lung cancer specimens, respectively [44-46]. However, the potential roles of non-CpG methylation in the biological functions of cancer are not clear. Could non-CpG methylation indicate an epigenetic state for cancer cell survival and continued development? These questions require further mechanistic studies. Future investigations regarding single cell transcriptomic changes and their association with cancer may help to elucidate the physiological roles of non-CpG methylation during cancer development.

The cancer cells also have characteristics of stem cells, but they differ from embryonic stem cells. Cancer stem cells (CSCs) have stem cell abilities, such as indefinite proliferation, extensive self-renewal, and multilinkage differentiation, which are the sources of tumor initiation, metastasis and relapse [47, 48]. A recent study showed a reduction of non-CpG methylation in cutaneous squamous cell carcinoma (SCC) and actinic keratosis [41]. Although evidence has shown that certain body-methylated genes gain non-CpG methylation in the absence of IBM1 [19], methyl-CpG-binding protein 2(MeCP2) regulates non-CpG methylation in the mammalian brain [49], and domains rearranged methyltransferase (DRM) and chromomethylase 3(CMT3) maintain non-CpG methylation [50]. However, the establishment of non-CpG methylation in cancer is poorly studied. LSH functions as an oncogene by promoting metastasis and genome instability and is involved in lipid metabolism in cancer cells and epigenetic regulates stem cell fate [27, 50, 51]. Our study found that LSH inhibits
non-CpG methylation except for in repeat regions (CHH methylation). This may be connected with its ability to promote genome stability. However, there is no evidently change in CpG methylation when LSH is overexpressed. The KEGG results showed that CpG methylation is associated with metabolic and non-CpG methylation and is mostly involved in the processes of cancer survival, migration and development, such as focal adhesion, regulation of the actin cytoskeleton and PI3K-Akt signaling. TRPM5, a sodium-selective TRP channel, functions as a regulator of calcium uptake and nonselective cationic channels and was found to be upregulated in Wilms’ tumor and rhabdomyosarcomas, which may be associated with tumorigenesis [52, 53]. GAS2L1 is a GAS2-like family protein that interacts with actin filaments and microtubules. Evidence has shown mediated centrosome disjunction and movement [54] and epigenetically silenced genes in acute myeloid leukemia [55]. ICAM5 is a member of the intercellular adhesion molecule (ICAM) family that supports breast and prostate cancer susceptibility loci and promotes tumorigenesis and perineural invasion in head and neck cancer [56, 57]. SLC10A7 is involved in glycosaminoglycan synthesis, and its splicing is associated with colorectal cancer prognosis [58, 59]. These genes are associated with tumor development, and our results indicate that LSH regulates the non-CpG methylation of these genes but not their CpG methylation. To exclude the possibility that inhibition of non-CpG methylation is induced by the exogenous overexpression of LSH, we performed western blotting to clarify this in lung cancer cell lines (A549 and H358). LSH seems to work as a driver to transform tumors into malignant tumors; in short, it may function as a regulator to promote tumorigenesis through inhibiting non-CpG methylation. Transcription factors can serve as readers of DNA methylation or induce changes in DNA methylation states upon binding to target sequences [60]. We found that hypermethylation reduced TFs bind to DNA in promoters and gene body which were induced by LSH. These results indicate that LSH may epigenetically regulate the ability of TFs to bind to DNA. We also hold that LSH affects non-CpG methylation throughout the gene regions in either the coding (exon) or noncoding regions (intron, 5’UTR, 3’UTR) in certain genes. However, additional experiments and mechanistic studies are needed to clarify this.

Conclusion
In conclusion, our study utilized the RRBS technology to investigate changes in DNA methylation types upon the expression of LSH. Our findings serve as a basis for further exploring the relationship between cancer and non-CpG methylation. The results support the hypothesis that LSH inhibits non-CpG methylation in cancer and that this reduction in non-CpG methylation may affect TFs and is reflected in certain genes. The biological role that non-CpG methylation may play in cancer development and the mechanism of LSH epigenetic regulation of TFs requires further investigation.

Material And Methods

Cell culture

H358 cells were maintained in RPMI 1640 (Gibco), and A549 cells were cultured in DMEM/F12 (Gibco). All cells were supplemented with 10% newborn bovine serum (HyClone) and 1% penicillin/streptomycin/gentamicin. 293T cells were maintained in DMEM (Gibco) supplemented with 10% fetal bovine serum (HyClone) and 1% penicillin/streptomycin/gentamicin. All cells were incubated at 37°C in a humidified chamber containing 5% CO₂. All cell lines yielded negative results for mycoplasma contamination.

Plasmids and lentivirus transfection

The LSH lentiviral construct was generated by inserting the LSH cDNA into plvx-EF1a-puro vector, and an empty vector was used as negative control (Clontech, Mountain View, CA). All plasmid vectors were verified by sequencing. Plasmid transfection was performed using LipoMax (Sudgen Biotech, Bellevue, WA, USA), in accordance with the manufacturer's protocol. Cell colonies were selected using puromycin (1 μg/ml). The overexpression of LSH was confirmed by western blot.

Western blot analysis

Cells were harvested, washed twice with ice-cold phosphate-buffered saline (PBS), lysed in lysis buffer [10mM Tris–HCl, pH 8.0, 1mM EDTA, 1% SDS, 5mM dithiothreitol, 10mM phenylmethylsulfonylfluoride, 1mM Na₃VO₄, 1mM NaF, 10% (vol/vol) glycerol, protease inhibitor cocktail tablet (Roche)] and centrifuged at 15,000 × g for 10min after sonication. The supernatants were collected as whole cell lysates. A quantity of 50 μg of total protein was used for western blot analysis. Primary antibody against β-actin was purchased from Sigma-Aldrich (A5441, Sigma). Primary antibody against
LSH was purchased from Santa Cruz Biotechnology (sc-46665, Santa cruz).

**DNA, bisulfite treatment preparation**

The cells were collected and the genomic DNA was extracted using a genomic DNA kit (Sangon Biotech, Cat.#B518201-0100); bisulfite conversion was then performed using the EZ DNA Methylation Direct kit (Zymo Research Corporation, Cat#D5020). The concentration and quantity of DNA were measured by a NanoDrop instrument (NanoDrop Technologies, Wilmington, DE, USA). All operations were conducted following the manufacturers’ recommended instructions.

**RRBS library preparation and data analysis**

The RRBS libraries were prepared according to previously published protocols [61]. Briefly, genomic DNA was digested with theMspI enzyme, followed by end-repair and ligation of sequencing adaptors. The fragments were then size-selected (40–220 bp) and bisulfite-converted prior to a PCR amplification step. The quality of the libraries was checked using a bioanalyzer, and two libraries were sequenced on an Illumina HiSeq X Ten machine (100 bp, single-ended run). The peak signals produced by the Illumina HiSeq were transformed into a base sequence using base calling of the raw data or raw reads. The raw reads were then filtered for subsequent information analysis to ensure the quality of the information analysis, including the removal of reads that had adapters and filtering reads with more than 10% N content or more than 50% low-quality bases. The final filtered data were regarded as clean reads.

**Mapping reads to known genome**

Sequencing reads must be aligned with a reference genome before conducting methylation analysis. Bismark software was used to perform a comparison of the alignments of bisulfite-treated reads to a reference genome using the default parameters. Reads that aligned with the same region of the genome were regarded as duplicates. The number of duplicates was used to summarize the sequencing depth and coverage. The conversion rate of bisulfite was calculated as the percentage of the methylated clean reads as a percentage of the total number of clean reads in the lambda genome using the Bismark software. Unmethylated cytosine from the genome was converted into T after bisulfite treatment and PCR amplification, but the methylated cytosine remained unchanged. Bismark
was able to extract information about genomic cytosine sites by comparing the clean reads with the reference genome, thereby gaining cytosine site coverage statistics and the number of different types (CG as CpG, CHG and CHH) of methylated cytosine reads. As the methylation single C sites cannot be discriminated by Bismark, we used the binomial distribution test for each C site to confirm the methylated C site by screening conditions for coverage ≥4× and a false discovery rate (FDR) < 0.05.

**Estimating methylation levels and the identification of DMRs**

All cytosine sites with read coverages >10× were used for DMR analysis with MOABS [62]. First, to detect the methylated C sites in a region, we defined as the number of methylated reads at a single C site, as the number of unmethylated reads at a single C site, as the position of C, and as the total number of C positions. The methylation level of a C site was counted as follows [63]:

[Due to technical limitations, the formula could not be displayed here. Please see the supplementary files section to access the formula.]

The binomial distribution test was used to determine whether the C site was methylated. Subsequently, DMRs were defined as those with at least three different methylation sites in the region in which the difference in methylation level was greater than 0.2 (0.3 for CG type) with a p value from Fisher’s exact test of less than 0.05. The detailed DMRs were listed in supplementary file table1 to 3.

The methylation level of a region was calculated as follows [63].

**Bioinformatics analysis of DMGs**

The DMGs were compared with functional databases such as GO, COG (Cluster of Orthologous Groups of proteins) and KEGG (Kyoto Encyclopedia of Genes and Genomes) by BLAST to obtain the annotation of these genes for analyzing gene function. The GO enrichment analysis was implemented by Wallenius noncentral hypergeometric distribution in the GOseq R package[64]. KOBAS software was used to assess the statistically enrichment of differentially expressed genes in the KEGG pathways [65].

**Abbreviations**

LSH
Lymphocyte specific helicase
RRBS
Reduced Representation Bisulfite Sequencing
DMR
Differentially methylated regions
DMGs
DMR-related genes
TFs
transcription factors
CGIs
CpG islands
SINEs
short interspersed nuclear elements
LINEs
long interspersed nuclear elements
LTRs
long terminal repeats
TRPM5
transient receptor potential melastatin 5
GAS2L1
growth arrest-specific 2-like 1
ICAM5
intercellular adhesion molecule 5
SLC10A7
solute carrier family 10 member 7
Declarations

Ethics approval and consent to participate:

Not applicable.

Consent for publication:

All authors have read and consent to the publication of this research article.

Acknowledgments:

We thank all the members of the laboratory for their resourceful comments on the manuscript. We appreciate the Shenzhen Acegen Technology Co. Ltd, for guidance in data analysis and statistics, Hanlin Lu for technical guidance in RRBS library preparation.
Funding:
This work was supported by the National Natural Science Foundation of China [81872285 (Y.Shi), 81672787 (Y.Tao), 81672991 and 81874139 (S.Liu), 81672307 (X.Wang), 81772496 (D. Xiao)]; the National Basic Research Program of China [2015CB553903 (Y.Tao)], the Overseas Expertise Introduction Project for Discipline Innovation (111 Project, No. 111-2-12).

Author Contributions:
WW Lai, LL. Ouyang, Na Liu, Sp Liu, Y Shi, L Chen, X Wang, DS Xiao, S. Liu and Y. Tao made substantial contributions to the conception and design, acquisition of data, analysis and interpretation of data. Y. Tao, S. Liu and WW Lai were involved in drafting the manuscript critically for important intellectual content. All of the authors have given final approval of the version to be published.

Competing interests:
The authors declare no conflicts of interest. This manuscript has been read and approved by all authors and is not submitted or under consideration for publication elsewhere.

Availability of data and material:
The raw sequencing data and processed files are available at the SRA under the accession numbers SRX7046862 (ChIP-seq) with links to BioProject accession number PRJNA578952.

References
1. Luo C, Hajkova P, Ecker JR: Dynamic DNA methylation: In the right place at the right time. Science 2018, 361(6409):1336-1340.
2. Jiang Y, Liu S, Chen X, Cao Y, Tao Y: Genome-wide distribution of DNA methylation and DNA demethylation and related chromatin regulators in cancer. Biochim Biophys Acta 2013, 1835(2):155-163.
3. Liu S, Tao Y: Interplay between chromatin modifications and paused RNA polymerase II in dynamic transition between stalled and activated genes. Biol Rev Camb Philos Soc 2013, 88(1):40-48.
4. Bar S, Benvenisty N: Epigenetic aberrations in human pluripotent stem cells.
5. Cotton AM, Price EM, Jones MJ, Balaton BP, Kobor MS, Brown CJ: Landscape of DNA methylation on the X chromosome reflects CpG density, functional chromatin state and X-chromosome inactivation. Human molecular genetics 2015, 24(6):1528-1539.

6. Xu RH, Wei W, Krawczyk M, Wang W, Luo H, Flagg K, Yi S, Shi W, Quan Q, Li K et al: Circulating tumour DNA methylation markers for diagnosis and prognosis of hepatocellular carcinoma. Nature materials 2017, 16(11):1155-1161.

7. Xu Q, Xiang Y, Wang Q, Wang L, Brind’Amour J, Bogutz AB, Zhang Y, Zhang B, Yu G, Xia W et al: SETD2 regulates the maternal epigenome, genomic imprinting and embryonic development. Nature genetics 2019, 51(5):844-856.

8. Saxonov S, Berg P, Brutlag DL: A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. Proceedings of the National Academy of Sciences of the United States of America 2006, 103(5):1412-1417.

9. Ramsahoye BH, Biniszkiewicz D, Lyko F, Clark V, Bird AP, Jaenisch R: Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. Proceedings of the National Academy of Sciences of the United States of America 2000, 97(10):5237-5242.

10. Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo QM et al: Human DNA methylomes at base resolution show widespread epigenomic differences. Nature 2009, 462(7271):315-322.

11. Kulis M, Merkel A, Heath S, Queiros AC, Schuyler RP, Castellano G, Beekman R, Raineri E, Esteve A, Clot G et al: Whole-genome fingerprint of the DNA methylome during human B cell differentiation. Nature genetics 2015,
12. Xie W, Barr CL, Kim A, Yue F, Lee AY, Eubanks J, Dempster EL, Ren B: Base-resolution analyses of sequence and parent-of-origin dependent DNA methylation in the mouse genome. Cell 2012, 148(4):816-831.

13. Varley KE, Gertz J, Bowling KM, Parker SL, Reddy TE, Pauli-Behn F, Cross MK, Williams BA, Stamatoyannopoulos JA, Crawford GE et al: Dynamic DNA methylation across diverse human cell lines and tissues. Genome research 2013, 23(3):555-567.

14. Lindroth AM, Cao X, Jackson JP, Zilberman D, McCallum CM, Henikoff S, Jacobsen SE: Requirement of CHROMOMETHYLASE3 for maintenance of CpXpG methylation. Science 2001, 292(5524):2077-2080.

15. Sendzikaite G, Hanna CW, Stewart-Morgan KR, Ivanova E, Kelsey G: A DNMT3A PWWP mutation leads to methylation of bivalent chromatin and growth retardation in mice. Nature communications 2019, 10(1):1884.

16. Shirane K, Toh H, Kobayashi H, Miura F, Chiba H, Ito T, Kono T, Sasaki H: Mouse oocyte methylomes at base resolution reveal genome-wide accumulation of non-CpG methylation and role of DNA methyltransferases. PLoS genetics 2013, 9(4):e1003439.

17. Huang YH, Su J, Lei Y, Brunetti L, Gundry MC, Zhang X, Jeong M, Li W, Goodell MA: DNA epigenome editing using CRISPR-Cas SunTag-directed DNMT3A. Genome biology 2017, 18(1):176.

18. Lee JH, Park SJ, Nakai K: Differential landscape of non-CpG methylation in embryonic stem cells and neurons caused by DNMT3s. Scientific reports 2017, 7(1):11295.

19. Zabet NR, Catoni M, Prischi F, Paszkowski J: Cytosine methylation at CpCpG sites triggers accumulation of non-CpG methylation in gene bodies. Nucleic acids
Lister R, Pelizzola M, Kida YS, Hawkins RD, Nery JR, Hon G, Antosiewicz-Bourget J, O'Malley R, Castanon R, Klugman S et al: **Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells.** *Nature* 2011, 471(7336):68-73.

Ziller MJ, Muller F, Liao J, Zhang Y, Gu H, Bock C, Boyle P, Epstein CB, Bernstein BE, Lengauer T et al: **Genomic distribution and inter-sample variation of non-CpG methylation across human cell types.** *PLoS genetics* 2011, 7(12):e1002389.

Myant K, Termanis A, Sundaram AY, Boe T, Li C, Merusi C, Burrage J, de Las Heras JI, Stancheva I: **LSH and G9a/GLP complex are required for developmentally programmed DNA methylation.** *Genome research* 2011, 21(1):83-94.

Xi S, Geiman TM, Briones V, Guang Tao Y, Xu H, Muegge K: **Lsh participates in DNA methylation and silencing of stem cell genes.** *Stem cells* 2009, 27(11):2691-2702.

De La Fuente R, Baumann C, Fan T, Schmidtmann A, Dobrinski I, Muegge K: **Lsh is required for meiotic chromosome synapsis and retrotransposon silencing in female germ cells.** *Nature cell biology* 2006, 8(12):1448-1454.

He X, Yan B, Liu S, Jia J, Lai W, Xin X, Tang CE, Luo D, Tan T, Jiang Y et al: **Chromatin Remodeling Factor LSH Drives Cancer Progression by Suppressing the Activity of Fumarate Hydratase.** *Cancer research* 2016, 76(19):5743-5755.

Jia J, Shi Y, Chen L, Lai W, Yan B, Jiang Y, Xiao D, Xi S, Cao Y, Liu S et al: **Decrease in Lymphoid Specific Helicase and 5-hydroxymethylcytosine Is Associated with Metastasis and Genome Instability.** *Theranostics* 2017, 7(16):3920-3932.

Jiang Y, Mao C, Yang R, Yan B, Shi Y, Liu X, Lai W, Liu Y, Wang X, Xiao D et al: **EGLN1/c-Myc Induced Lymphoid-Specific Helicase Inhibits Ferroptosis**
through Lipid Metabolic Gene Expression Changes. *Theranostics* 2017, 7(13):3293-3305.

28. Mao C, Wang X, Liu Y, Wang M, Yan B, Jiang Y, Shi Y, Shen Y, Liu X, Lai W *et al*: A G3BP1-Interacting IncRNA Promotes Ferroptosis and Apoptosis in Cancer via Nuclear Sequestration of p53. *Cancer Res* 2018, 78(13):3484-3496.

29. Wang M, Mao C, Ouyang L, Liu Y, Lai W, Liu N, Shi Y, Chen L, Xiao D, Yu F *et al*: Long noncoding RNA LINC00336 inhibits ferroptosis in lung cancer by functioning as a competing endogenous RNA. *Cell Death Differ* 2019.

30. Xiao D, Huang J, Pan Y, Li H, Fu C, Mao C, Cheng Y, Shi Y, Chen L, Jiang Y *et al*: Chromatin Remodeling Factor LSH is Upregulated by the LRP6-GSK3beta-E2F1 Axis Linking Reversely with Survival in Gliomas. *Theranostics* 2017, 7(1):132-143.

31. Yang R, Liu N, Chen L, Jiang Y, Shi Y, Mao C, Liu Y, Wang M, Lai W, Tang H *et al*: LSH interacts with and stabilizes GINS4 transcript that promotes tumourigenesis in non-small cell lung cancer. *Journal of experimental & clinical cancer research: CR* 2019, 38(1):280.

32. Yang R, Liu N, Chen L, Jiang Y, Shi Y, Mao C, Liu Y, Wang M, Lai W, Tang H *et al*: GIAT4RA functions as a tumor suppressor in non-small cell lung cancer by counteracting Uchl3-mediated deubiquitination of LSH. *Oncogene* 2019.

33. Tao Y, Xi S, Briones V, Muegge K: Lsh mediated RNA polymerase II stalling at HoxC6 and HoxC8 involves DNA methylation. *PloS one* 2010, 5(2):e9163.

34. Dunican DS, Cruickshanks HA, Suzuki M, Semple CA, Davey T, Arceci RJ, Greally J, Adams IR, Meehan RR: Lsh regulates LTR retrotransposon repression independently of Dnmt3b function. *Genome biology* 2013, 14(12):R146.

35. Ren J, Hathaway NA, Crabtree GR, Muegge K: Tethering of Lsh at the Oct4 locus
promotes gene repression associated with epigenetic changes. *Epigenetics* 2018, **13**(2):173-181.

36. Lim WJ, Kim KH, Kim JY, Jeong S, Kim N: **Identification of DNA-Methylated CpG Islands Associated With Gene Silencing in the Adult Body Tissues of the Ogye Chicken Using RNA-Seq and Reduced Representation Bisulfite Sequencing.** *Frontiers in genetics* 2019, **10**:346.

37. Garrett-Bakelman FE, Sheridan CK, Kacmarczyk TJ, Ishii J, Betel D, Alonso A, Mason CE, Figueroa ME, Melnick AM: **Enhanced reduced representation bisulfite sequencing for assessment of DNA methylation at base pair resolution.** *Journal of visualized experiments: JoVE* 2015(96):e52246.

38. Chatterjee A, Macaulay EC, Ahn A, Ludgate JL, Stockwell PA, Weeks RJ, Parry MF, Foster TJ, Knarston IM, Eccles MR et al: **Comparative assessment of DNA methylation patterns between reduced representation bisulfite sequencing and Sequenom EpiTyper methylation analysis.** *Epigenomics* 2017, **9**(6):823-832.

39. Yin Y, Morgunova E, Jolma A, Kaasinen E, Sahu B, Khund-Sayeed S, Das PK, Kivioja T, Dave K, Zhong F et al: **Impact of cytosine methylation on DNA binding specificities of human transcription factors.** *Science* 2017, **356**(6337).

40. Xi S, Zhu H, Xu H, Schmidtmann A, Geiman TM, Muegge K: **Lsh controls Hox gene silencing during development.** *Proc Natl Acad Sci U S A* 2007, **104**(36):14366-14371.

41. Rodriguez-Paredes M, Bormann F, Raddatz G, Gutekunst J, Lucena-Porcel C, Kohler F, Wurzer E, Schmidt K, Gallinat S, Wenck H et al: **Methylation profiling identifies two subclasses of squamous cell carcinoma related to distinct cells of origin.** *Nature communications* 2018, **9**(1):577.

42. Mikhailenko DS, Nematsova MV: **[Point somatic mutations in bladder cancer: key
carcinogenesis events, diagnostic markers and therapeutic targets. *Urologia* 2016(1):100-105.

43. Rizzardi LF, Hickey PF, Rodriguez DiBlasi V, Tryggvadottir R, Callahan CM, Idrizi A, Hansen KD, Feinberg AP: Neuronal brain-region-specific DNA methylation and chromatin accessibility are associated with neuropsychiatric trait heritability. *Nature neuroscience* 2019, 22(2):307-316.

44. Chan QK, Khoo US, Chan KY, Ngan HY, Li SS, Chiu PM, Man LS, Ip PP, Xue WC, Cheung AN: Promoter methylation and differential expression of pi-class glutathione S-transferase in endometrial carcinoma. *The Journal of molecular diagnostics: JMD* 2005, 7(1):8-16.

45. Li C, Xiong W, Liu X, Xiao W, Guo Y, Tan J, Li Y: Hypomethylation at non-CpG/CpG sites in the promoter of HIF-1alpha gene combined with enhanced H3K9Ac modification contribute to maintain higher HIF-1alpha expression in breast cancer. *Oncogenesis* 2019, 8(4):26.

46. Kouidou S, Agidou T, Kyrkou A, Andreou A, Katopodi T, Georgiou E, Krikelis D, Dimitriadou A, Spanos P, Tsilikas C et al: Non-CpG cytosine methylation of p53 exon 5 in non-small cell lung carcinoma. *Lung cancer* 2005, 50(3):299-307.

47. Dawood S, Austin L, Cristofanilli M: Cancer stem cells: implications for cancer therapy. *Oncology* 2014, 28(12):1101-1107, 1110.

48. Munro MJ, Wickremesekera SK, Peng L, Tan ST, Itinteang T: Cancer stem cells in colorectal cancer: a review. *Journal of clinical pathology* 2018, 71(2):110-116.

49. Chen L, Chen K, Lavery LA, Baker SA, Shaw CA, Li W, Zoghbi HY: MeCP2 binds to non-CG methylated DNA as neurons mature, influencing transcription and the timing of onset for Rett syndrome. *Proceedings of the National Academy of Sciences of the United States of America* 2015, 112(17):5509-5514.
50. Cao X, Aufsatz W, Zilberman D, Mette MF, Huang MS, Matzke M, Jacobsen SE: **Role of the DRM and CMT3 methyltransferases in RNA-directed DNA methylation.** *Current biology : CB* 2003, 13(24):2212-2217.

51. Han Y, Ren J, Lee E, Xu X, Yu W, Muegge K: **Lsh/HELLS regulates self-renewal/proliferation of neural stem/progenitor cells.** *Scientific reports* 2017, 7(1):1136.

52. Mitrovic S, Nogueira C, Cantero-Recasens G, Kiefer K, Fernandez-Fernandez JM, Popoff JF, Casano L, Bard FA, Gomez R, Valverde MA et al: **TRPM5-mediated calcium uptake regulates mucin secretion from human colon goblet cells.** *eLife* 2013, 2:e00658.

53. Prawitt D, Enklaar T, Klemm G, Gartner B, Spangenberg C, Winterpacht A, Higgins M, Pelletier J, Zabel B: **Identification and characterization of MTR1, a novel gene with homology to melastatin (MLSN1) and the trp gene family located in the BWS-WT2 critical region on chromosome 11p15.5 and showing allele-specific expression.** *Human molecular genetics* 2000, 9(2):203-216.

54. Au FK, Jia Y, Jiang K, Grigoriev I, Hau BK, Shen Y, Du S, Akhmanova A, Qi RZ: **GAS2L1 Is a Centriole-Associated Protein Required for Centrosome Dynamics and Disjunction.** *Developmental cell* 2017, 40(1):81-94.

55. Desmond JC, Raynaud S, Tung E, Hofmann WK, Haferlach T, Koeffler HP: **Discovery of epigenetically silenced genes in acute myeloid leukemias.** *Leukemia* 2007, 21(5):1026-1034.

56. Kammerer S, Roth RB, Reneland R, Marnellos G, Hoyal CR, Markward NJ, Ebner F, Kiechle M, Schwarz-Boeger U, Griffiths LR et al: **Large-scale association study identifies ICAM gene region as breast and prostate cancer susceptibility locus.** *Cancer research* 2004, 64(24):8906-8910.
57. Maruya SI, Myers JN, Weber RS, Rosenthal DI, Lotan R, El-Naggar AK: **ICAM-5 (telencephalin) gene expression in head and neck squamous carcinoma tumorigenesis and perineural invasion.** *Oral oncology* 2005, **41**(6):580-588.

58. Dubail J, Huber C, Chantepie S, Sonntag S, Tuysuz B, Mihci E, Gordon CT, Steichen-Gersdorf E, Amiel J, Nur B et al: **SLC10A7 mutations cause a skeletal dysplasia with amelogenesis imperfecta mediated by GAG biosynthesis defects.** *Nature communications* 2018, **9**(1):3087.

59. Liu J, Li H, Shen S, Sun L, Yuan Y, Xing C: **Alternative splicing events implicated in carcinogenesis and prognosis of colorectal cancer.** *Journal of Cancer* 2018, **9**(10):1754-1764.

60. Zhu H, Wang G, Qian J: **Transcription factors as readers and effectors of DNA methylation.** *Nature reviews Genetics* 2016, **17**(9):551-565.

61. Chatterjee A, Rodger EJ, Stockwell PA, Le Mee G, Morison IM: **Generating Multiple Base-Resolution DNA Methylation Using Reduced Representation Bisulfite Sequencing.** *Methods in molecular biology* 2017, **1537**:279-298.

62. Sun D, Xi Y, Rodriguez B, Park HJ, Tong P, Meong M, Goodell MA, Li W: **MOABS: model based analysis of bisulfite sequencing data.** *Genome biology* 2014, **15**(2):R38.

63. Schultz MD, Schmitz RJ, Ecker JR: **'Leveling' the playing field for analyses of single-base resolution DNA methylomes.** *Trends in genetics : TIG* 2012, **28**(12):583-585.

64. Young MD, Wakefield MJ, Smyth GK, Oshlack A: **Gene ontology analysis for RNA-seq: accounting for selection bias.** *Genome biology* 2010, **11**(2):R14.

65. Mao X, Cai T, Olyarchuk JG, Wei L: **Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary.**
Figures
Figure 1

Genome-wide CpG and Non-CpG Methylation Patterns. A. The color bars indicate the read depth, and the Y-axis shows the percent of global CpGs, CHGs, and CHHs in each group. B. Circular genomic plot of average DNA methylation level in 1-Mb non-overlapping bins. Circles from outside to the center: cytochromosome, CHH methylation level (green, vector; red, LSH), CHG methylation level, CpG methylation level. Each peak represents a 1-Mb bin.
Figure 1

Genome-wide CpG and Non-CpG Methylation Patterns. A. The color bars indicate the read depth, and the Y-axis shows the percent of global CpGs, CHGs, and CHHs in each group. B. Circular genomic plot of average DNA methylation level in 1-Mb non-overlapping bins. Circles from outside to the center: cytobchromosome, CHH methylation level (green, vector; red, LSH), CHG methylation level, CpG methylation level. Each peak represents a 1-Mb bin.
Figure 2

A

B

C

Figure 2
Composite Plot of DNA Methylation Levels across Genomic Features. A. Violin plot for the overall distribution of methylation levels for different methylation types in the regions (CpGs, CHGs, CHHs) for each group. B. Violin plot for the distribution of methylation levels at repeat-specific genes. C. The methylation level of CpG, CHG and CHH in different functional elements; blue, vector; red, LSH. Each line consists of the mean and 95% confidence interval. CGI (CpG island): a region with at least 200bp, a GC percentage greater than 50%, and an observed-to-expected CpG ratio greater than 60%, as defined in the UCSC genome browser. CGI-shore (CpG island shore): 2-kb regions flanking a CpG island.
Composite Plot of DNA Methylation Levels across Genomic Features. A. Violin plot for the overall distribution of methylation levels for different methylation types in the regions (CpGs, CHGs, CHHs) for each group. B. Violin plot for the distribution of methylation levels at repeat-specific genes. C. The methylation level of CpG, CHG and CHH in different functional elements; blue, vector; red, LSH. Each line consists of the mean and 95% confidence interval. CGI (CpG island): a region with at least 200bp, a GC percentage greater than 50%, and an observed-to-expected CpG ratio greater than 60%, as defined in the UCSC genome browser. CGI-shore (CpG island shore): 2-kb regions flanking a CpG island.
DMR length and methylation levels in different methylation patterns. A. DMRs in promoters for the different methylation types. B. The length of DMRs. C. Hyper-DMRs and hypo-DMRs in different methylation types. D. Heat map of differential DMRs in the vector and LSH groups: CG methylation; CHG methylation; CHH methylation.
DMR length and methylation levels in different methylation patterns. A. DMRs in promoters for the different methylation types. B. The length of DMRs. C. Hyper-DMRs and hypo-DMRs in different methylation types. D. Heat map of differential DMRs in the vector and LSH groups: CG methylation; CHG methylation; CHH methylation.
Figure 4

Top transcription binding sites, top GO and top KEGG pathway analyses of CG, CHG, CHH type DMRs. A. Ratio of DMRs with different methylation types in different gene functional regions, CG, CHG, CHH. H = A, C or T. The colors represent the promoter, (≤1 kb), upstream promoter (2 kb), exon, intron, downstream 2 kb, 5’UTR, 3’UTR, CpG islands and CpG shore regions, respectively. B. A Heat map of the top 20 TFs binding ability to DNA at promoter and gene body region, high represents of higher DNA methylation in LSH group(LSH>Vector); low stands for lower DNA methylation in LSH group(LSH<Vector). C. Top GO analysis. The ordinate represents the GO terms that were the most enriched; the abscissa represents the P-value that was calculated using -log10 transformed values. The blue, green and red colors indicate biological process, cellular component and molecular function, respectively. The circle size represents the number of genes contained in the particular class in the graph. D. Top KEGG analysis. Larger circles indicate more genes. Differently colored circles represent the enrichment degree of false positives, where redder
circles represent lower false positive rates.

Figure 4

Top transcription binding sites, top GO and top KEGG pathway analyses of CG, CHG, CHH type DMRs. A. Ratio of DMRs with different methylation types in different gene functional regions, CG, CHG, CHH. H = A, C or T. The colors represent the promoter, (≤1 kb), upstream promoter (2 kb), exon, intron, downstream 2 kb, 5’UTR, 3’UTR, CpG islands and CpG shore regions, respectively. B. A Heat map of the top 20 TFs binding ability to DNA at promoter and gene body region, high represents of higher DNA methylation in LSH group (LSH>Vector); low stands for lower DNA methylation in LSH group (LSH<Vector). C. Top GO analysis. The ordinate represents the GO terms that were the most enriched; the abscissa represents the P-value that was calculated using -log10 transformed values. The blue, green and red colors indicate biological process, cellular component and molecular function, respectively. The circle size represents the number of genes contained in the particular class in the graph. D. Top KEGG analysis. Larger circles indicate more genes.
Differently colored circles represent the enrichment degree of false positives, where redder circles represent lower false positive rates.

Figure 5

Correlation of methylome and transcriptome. A. Venn analysis of methylome and transcriptome; 158 differentially expressed genes with a threshold of FDR < 0.01, [log2-fold] ≥1. B. Correlation of CpG, CHG and CHH methylation levels in different functional elements and gene expression levels. Purple, high gene expression; blue, medium expression; green, low expression; red, undetectable expression.
Figure 5

Correlation of methylome and transcriptome. A. Venn analysis of methylome and transcriptome; 158 differentially expressed genes with a threshold of FDR < 0.01, [log2-fold] ≥1. B. Correlation of CpG, CHG and CHH methylation levels in different functional elements and gene expression levels. Purple, high gene expression; blue, medium expression; green, low expression; red, undetectable expression.
Figure 6

IGV analysis of example genes with different methylation types in different regions.

Integrative Genomics Viewer (IGV) software analysis of gene methylation levels. CpG, CHG, CHH methylation levels (scale: 0%-100%). Blue, vector; red, LSH; yellow, DMR; green, CpG islands (UCSC browser), light blue, exon. Different genes DMRs are located in different regions. A. TRPM5 (chr11:2,423,746-2,446,275) DMR located in intron. B. GAS2L1 (chr22:29,700,985-29,710,778) DMR located in exon. C. ICAM5 (chr19:10,398,655-10,409,454) DMR located near 3’UTR region. D. DMR of SLC10A7 (chr4:147,441,459-147,444,978) located near the 5’UTR region.
IGV analysis of example genes with different methylation types in different regions.

Integrative Genomics Viewer (IGV) software analysis of gene methylation levels. CpG, CHG, CHH methylation levels (scale: 0%–100%). Blue, vector; red, LSH; yellow, DMR; green, CpG islands (UCSC browser), light blue, exon. Different genes DMRs are located in different regions. A. TRPM5 (chr11:2,423,746-2,446,275) DMR located in intron. B. GAS2L1 (chr22:29,700,985-29,710,778) DMR located in exon. C. ICAM5 (chr19:10,398,655-10,409,454) DMR located near 3’UTR region. D. DMR of SLC10A7 (chr4:147,441,459-147,444,978) located near the 5’UTR region.

Supplementary Files
This is a list of supplementary files associated with this preprint. Click to download.

Formula.docx
Supplementary Figure and figure legends.docx
Formula.docx
Supplementary Figure and figure legends.docx