



OPEN

Transcription factor-mediated direct cellular reprogramming yields cell-type specific DNA methylation signature

Kenichi Horisawa¹, Shizuka Miura¹, Hiromitsu Araki², Fumihito Miura³, Takashi Ito³ & Atsushi Suzuki¹✉

Direct reprogramming, inducing the conversion of one type of somatic cell into another by the forced expression of defined transcription factors, is a technology with anticipated medical applications. However, due to the many unresolved aspects of the induction mechanisms, it is essential to thoroughly analyze the epigenomic state of the generated cells. Here, we performed comparative genome-wide DNA methylation analyses of mouse embryonic fibroblasts (MEFs) and cells composing organoids formed by intestinal stem cells (ISCs) or induced ISCs (iISCs) that were directly induced from MEFs. We found that the CpG methylation state was similar between cells forming ISC organoids and iISC organoids, while they differed widely from those in MEFs. Moreover, genomic regions that were differentially methylated between ISC organoid- and iISC organoid-forming cells did not significantly affect gene expression. These results demonstrate the accuracy and safety of iISC induction, leading to the medical applications of this technology.

A complex molecular mechanism known as epigenetics exists behind the diversity of cells with the same genetic information and enables the readout of specific genetic information on demand. DNA methylation is an epigenetic process in which a methyl group is added at the fifth position of the cytosine pyrimidine ring. Cytosine methylation is classified by sequence context into CG (commonly called CpG), CHH (H = C, A, or T), and CHG sequences, which are regulated by different molecular mechanisms¹. CpG sequences are dominant in mammals and are considered major targets of epigenetic regulation². In general, CpG methylation near the transcription start site (TSS) has a negative control over gene expression³, whereas CpG methylation can also contribute to the enhancement of transcription^{4,5}. The DNA methylation patterns of mother cells are accurately transmitted to daughter cells during cell division by a mechanism called maintenance methylation in multicellular organisms, whereas de novo methylation, the addition of new methyl groups to specific DNA regions, is another important mechanism for changing the transcriptional state of cells⁶. Maintenance and de novo methylation are mainly controlled by DNMT1 and DNMT3A/3B, respectively⁷.

De novo methylation is involved in cell differentiation, and its disruption leads to cancer development⁸. Cellular reprogramming can be artificially induced by using knowledge of cell differentiation and often occurs during cancer development^{9,10}. Thus, the regulatory system of de novo methylation may be an important subject in the study of cellular reprogramming, cell differentiation and cancer development. In fact, the state of DNA methylation dramatically changes not only during cell differentiation but also during cellular reprogramming. Somatic cell nuclear transplantation (SCNT) can reset the state of DNA methylation in transplanted nuclei¹¹, and the DNA methylation state of induced pluripotent stem cells (iPSCs) transition to a state similar to that of embryonic stem cells (ESCs) during reprogramming¹². However, an inadequate transition of the state of DNA methylation occasionally occurs in SCNT, resulting in a significant reduction in the production ratios of cloned animals^{13,14}. Moreover, during iPSC reprogramming, differentially methylated regions (DMRs) inherited from the original cells¹⁵ and appearing ectopically¹⁶ are often abnormally generated. Thus, to develop safe and stable cellular reprogramming methods, it is necessary to elucidate, understand, and control the molecular basis of the DNA methylation transition.

¹Division of Organogenesis and Regeneration, Medical Institute of Bioregulation, Kyushu University, Fukuoka 812-8582, Japan. ²Insect Science and Creative Entomology Center, Kyushu University Graduate School of Agriculture, Fukuoka 819-0395, Japan. ³Department of Biochemistry, Kyushu University Graduate School of Medical Sciences, Fukuoka 812-8582, Japan. ✉email: suzukicks@bioreg.kyushu-u.ac.jp

As with the technology for inducing iPSCs, direct reprogramming, which can induce a type of somatic cell from another type of somatic cell by the forced expression of defined transcription factors, is considered a promising therapeutic strategy for the treatment of diseases⁹. Various types of somatic cells, e.g. hepatocytes^{17,18}, cardiomyocytes¹⁹, and neuronal cells²⁰, have been generated by inducing direct reprogramming using a combination of transcription factors involved in cell differentiation^{21–23}. A previous study showed that a set of transcription factors used to directly induce neuronal cells synergistically affected the levels of DNA methylation in mouse fibroblasts to establish a neuronal non-CpG methylation pattern²⁴. However, further investigation is required for a deeper understanding of the genome-wide DNA methylation changes and its functional roles in direct reprogramming. DNA methylation and demethylation may play essential roles in transcriptional regulation during the progression of direct reprogramming. In fact, it is possible to efficiently induce direct reprogramming by adding 5-aza-2-deoxycytidine, a DNA methyltransferase inhibitor, to the culture medium in combination with the introduction of defined transcription factors into the cells²⁵. Therefore, it is important to comprehensively understand the role of DNA methylation in direct reprogramming.

In our previous study, we established the method for inducing fetal intestinal progenitor cells (FIPCs) by expressing a set of four genes encoding Hnf4a, Foxa3, Gata6, and Cdx2 in mouse embryonic fibroblasts (MEFs)²⁶. Under three-dimensional (3D) culture conditions, these induced FIPCs (iFIPCs) formed spherical organoids (SOs) and subsequently gave rise to adult-type induced intestinal stem cells (iISCs) that formed budding organoids (BOs)²⁶. The morphology and gene expression signature of iISC-derived BOs (iISC-BOs) closely resemble those of intestinal stem cell (ISC)-derived BOs (ISC-BOs), and both iISCs and ISCs can differentiate into functional intestinal epithelial cells, i.e. as enterocytes, Paneth cells, goblet cells, and enteroendocrine cells. Moreover, iISCs and ISCs undergo self-renewing cell division and are thus stably maintained in long-term cultures by serial passaging. Transplantation of iISC-BOs and ISC-BOs into a chemically induced colonic injury model results in the reconstitution of intestinal epithelial tissue in damaged colons^{26,27}. Based on these findings, we performed a genome-wide and base-resolution methylome analysis using a post-bisulfite adaptor-tagging (PBAT) method^{28,29} for MEFs, iISC-BOs, and ISC-BOs, and investigated the transition of the DNA methylation state during direct reprogramming of MEFs to iISCs and the similarities and differences between iISC-BOs and ISC-BOs. Moreover, by combining our present methylome data with our previous transcriptome data²⁶, we examined the correlation between the abnormal DNA methylation state found in iISCs and the levels of transcription in iISCs.

Results

iISC-BOs exhibit a CpG methylation signature that closely resembles ISC-BOs

Genomic DNA was extracted from MEFs, iISC-BOs, and ISC-BOs, and high-resolution methylome analysis was performed using the PBAT method (Fig. 1A). The genome-wide methylation states of cytosine-containing sequences such as CpG, CHH, and CHG were compared among the samples. The data showed that CpG methylation signatures were similar among replicates and between iISC-BOs and ISC-BOs (Pearson's correlation coefficient [R] = 0.76–0.84), whereas substantial differences were found between iISC-BOs and MEFs, and between ISC-BOs and MEFs (R = 0.45–0.65; Fig. 1B,C, and Supplementary Fig. 1). The methylation states of CHH and CHG were highly correlated in comparison to iISC-BOs with their replicates and with ISC-BOs, and MEFs (R = 0.97–1.00; Fig. 1C and Supplementary Fig. 1).

In a genome-wide view, the total CpG methylation levels of MEFs were 69.6–70.1%, whereas those of ISC-BOs and iISC-BOs were 64.3–66.3% and 61.5–63.8%, respectively (Fig. 1D, the left panel). The CpG methylation levels of the introns and intergenic regions that occupy most of the genome showed the same tendencies as those of the total regions (Fig. 1D, right bottom panels). In contrast, the exons and CpG islands (CGIs) that may have an impact on transcriptional regulation exhibited slightly higher and higher levels of CpG methylation, respectively, in both iISC-BOs and ISC-BOs in comparison to that of MEFs (Fig. 1D, right upper panels). Moreover, the percentage of DMRs between iISC-BOs and ISC-BOs was one-tenth that between MEFs and iISC-BOs or ISC-BOs (Fig. 1E). Taken together, our data demonstrate that direct reprogramming from MEFs to iISCs is associated with variations in genome-wide CpG methylation and allows the generation of iISC-BOs that have a CpG methylation state similar to ISC-BOs. Additionally, we confirmed the gene expression of writers, readers, and erasers of the DNA methylation from previous microarray data²⁶. The expression of some of the genes was changed during the direct reprogramming from MEFs to iISC-BOs, however, no genes showed significant difference of expression between ISC-BOs and iISC-BOs (Supplementary Fig. 2).

Genome-wide similar distribution patterns of DMRs between MEFs and iISC-BOs or ISC-BOs result in a similar gene expression pattern of iISC-BOs and ISC-BOs

We sought to identify the genomic loci of DMRs between MEFs and iISC-BOs or ISC-BOs to investigate the transition of the CpG methylation state from MEFs to iISCs in more detail. The numbers of DMRs with hypermethylation in iISC-BOs (iISC-high DMRs) or ISC-BOs (ISC-high DMRs) and those with hypomethylation in iISC-BOs (iISC-low DMRs) or ISC-BOs (ISC-low DMRs) were not significantly different, whereas the numbers of iISC/ISC-low DMRs were slightly higher than those of iISC/ISC-high DMRs (Fig. 2A). Genome-wide distribution analysis revealed that the genomic loci of iISC-low and ISC-low DMRs and those of iISC-high and ISC-high DMRs were similarly distributed in the intergenic regions and regions of introns and exons, respectively (Fig. 2B). To examine how similar distribution patterns of iISC/ISC-low and iISC/ISC-high DMRs affected gene expression, we investigated the expression of DMR-associated genes by reanalyzing our previous transcriptome data obtained from MEFs, iISC-BOs, and ISC-BOs²⁶. To identify DMR-associated genes, we focused on iISC/ISC-low DMRs and iISC/ISC-high DMRs located within 10 kbp upstream and downstream of the TSS (Fig. 2C) because CpG methylation near the TSS affects transcription³⁰. We found that genes associated with iISC-low DMRs and ISC-low DMRs and those associated with iISC-high DMRs and ISC-high DMRs were similarly expressed

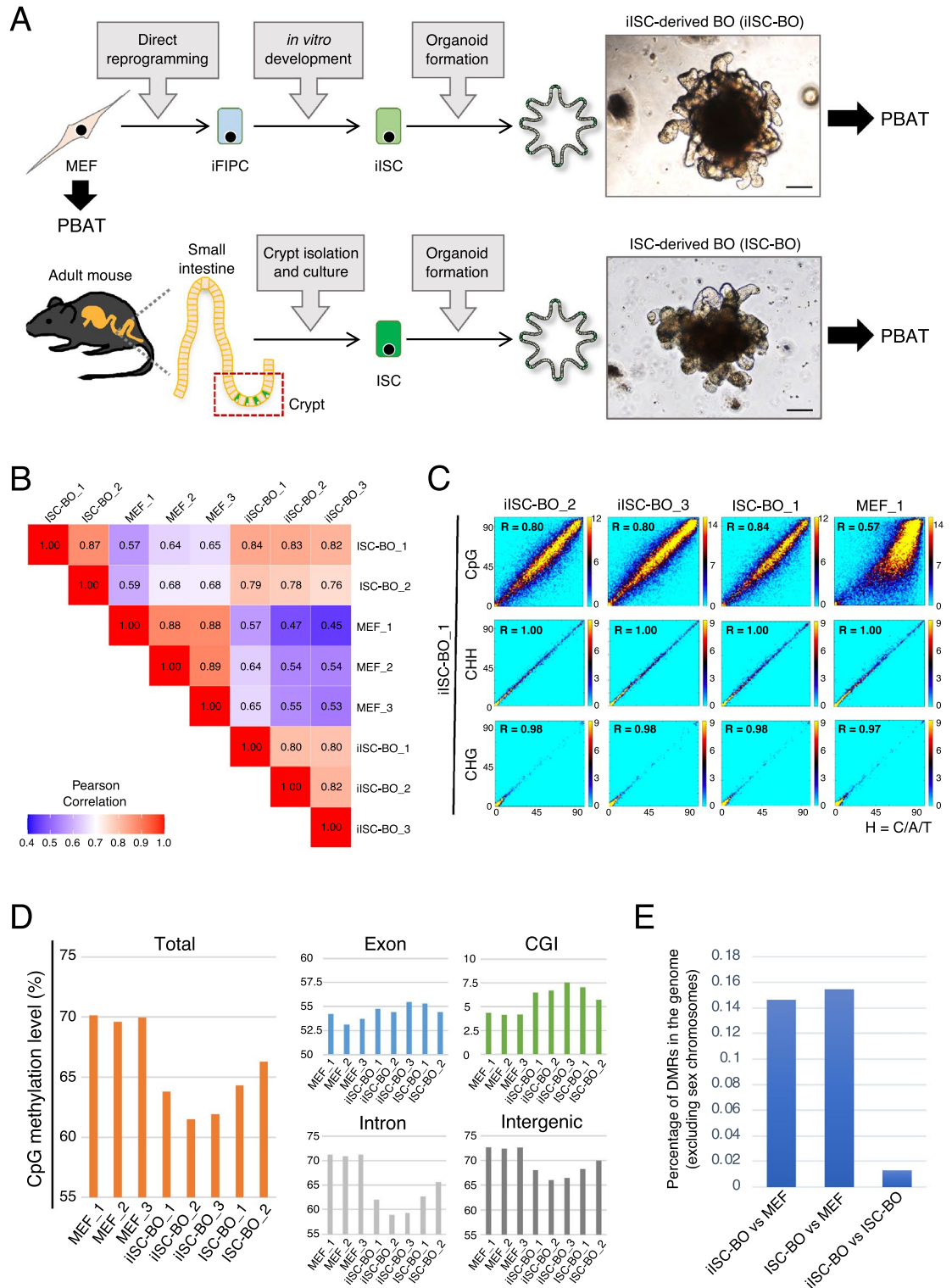


Figure 1. Genome-wide DNA methylation analysis of MEFs, iISC-BOs, and ISC-BOs. **(A)** Scheme of the overall experiment. Photos show bright field images of BOs. Scale bars, 100 μ m. **(B)** Pearson's correlation analysis of genome-wide CpG methylation between samples. **(C)** Density plots comparing genome-wide CpG, CHH, and CHG methylation status between samples. The window and step sizes were set to be 1 kbp and 500 bp, respectively. R scores indicate Pearson's correlation coefficient. **(D)** Averaged CpG methylation levels in each region, which was calculated when the number of reads mapped within a given region was five or more and there were five or more cytosines. **(E)** Genomic occupancy of the DMRs between cells. As a pretreatment of the methylation data analysis, only cytosines assigned with 10 or more reads and located on autosomes were selected.

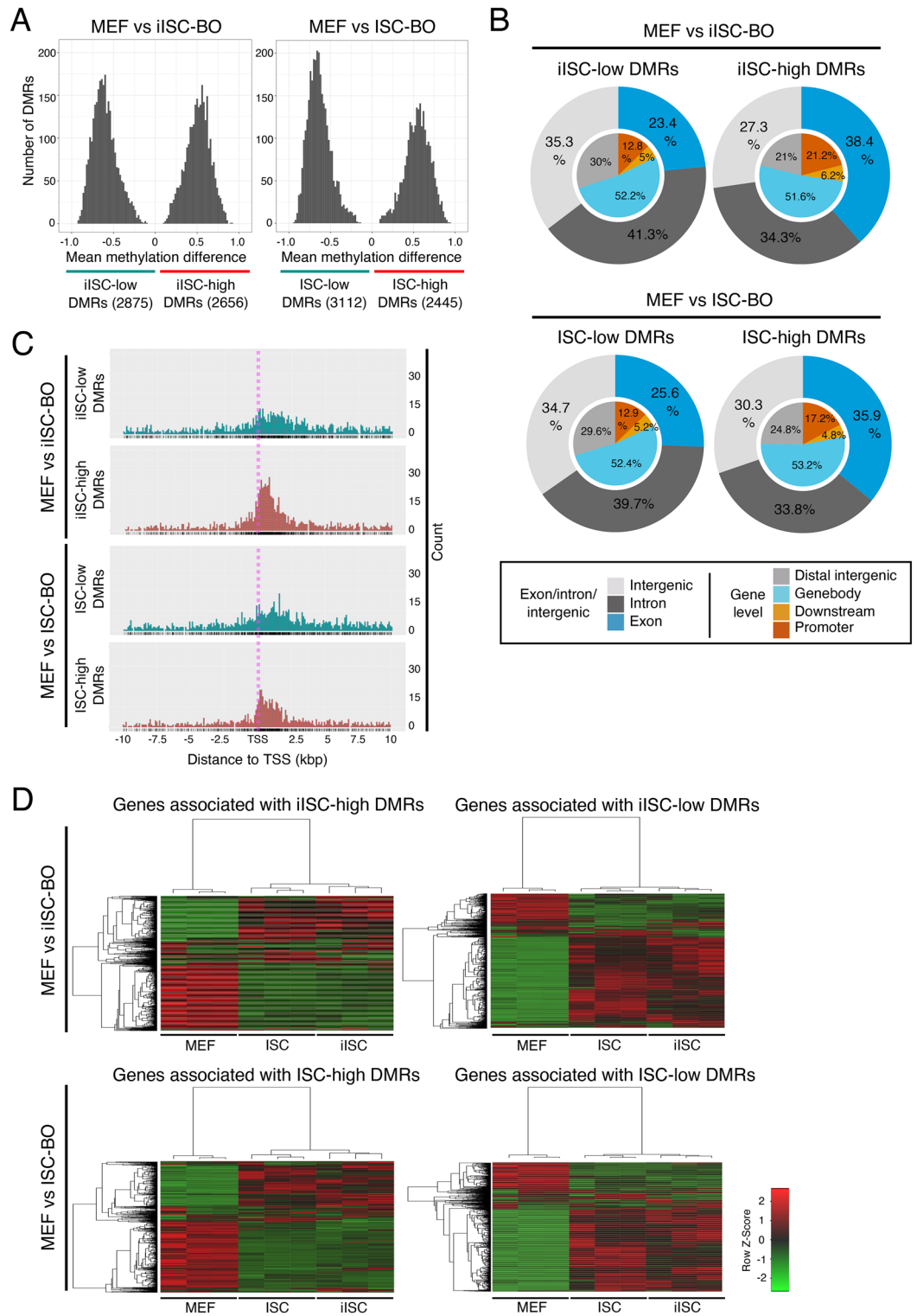


Figure 2. Analysis of DMRs between MEFs and iISC-BOs or ISC-BOs. **(A)** Histograms showing mean methylation difference of DMRs between MEFs and iISC-BOs or ISC-BOs. **(B)** Pie charts showing genomic position of detected DMRs. Outer and inner circles indicate genomic position with different definitions. **(C)** Dot plots showing distribution of DMRs around TSS of DMR-associated genes. **(D)** Heatmaps showing expression level of the DMR-associated genes. As a pretreatment of the methylation data analysis, only cytosines assigned with 10 or more reads and located on autosomes were selected.

between iISCs and ISCs compared with MEF (Fig. 2D and Supplementary Table 1). Thus, the similarity in the genome-wide distribution patterns of iISC-low and ISC-low DMRs and those of iISC-high and ISC-high DMRs may contribute to the similar gene expression patterns of iISC-BOs and ISC-BOs.

iISC-BOs have hyper- and hypo-methylated DNA, compared with ISC-BOs

As shown above, the CpG methylation state and the resulting gene expression pattern in iISC-BOs were similar to those in ISC-BOs. However, they were not the same. Thus, we examined the genomic loci of DMRs between iISC-BOs and ISC-BOs to explore the differences in CpG methylation states between these two samples. Our data showed that hypermethylation of DMRs in iISC-BOs (hyper DMRs) was more than twice as high as hypomethylation in iISC-BOs (hypo DMRs), suggesting that DNA methylation is induced more frequently than DNA demethylation during direct reprogramming from MEFs to iISCs (Fig. 3A). Heatmaps showing the average CpG methylation rates in the hyper-/hypo-DMRs among MEFs, iISC-BOs, and ISC-BOs revealed that these hyper-/hypo-DMRs could be mainly divided into two groups: those that maintain the levels of CpG methylation in the process of direct reprogramming from MEFs to iISCs, and those that represent insufficient or excessive CpG methylation in only iISC-BOs (Fig. 3B).

It has been reported that CpG methylation yielded around the TSS affects their transcription³⁰. Thus, we analyzed the genomic loci of the hyper-/hypo-DMRs around the TSS to examine the possibility that these hyper-/hypo-DMRs affect gene transcription. Our data demonstrated that the hyper-DMRs were located in regions distal and proximal to the TSS, whereas the hypo-DMRs were mainly located in regions distal to the TSS (Fig. 3C). Thus, hyper DMRs contribute more frequently to the misregulation of gene expression in iISCs than hypo DMRs.

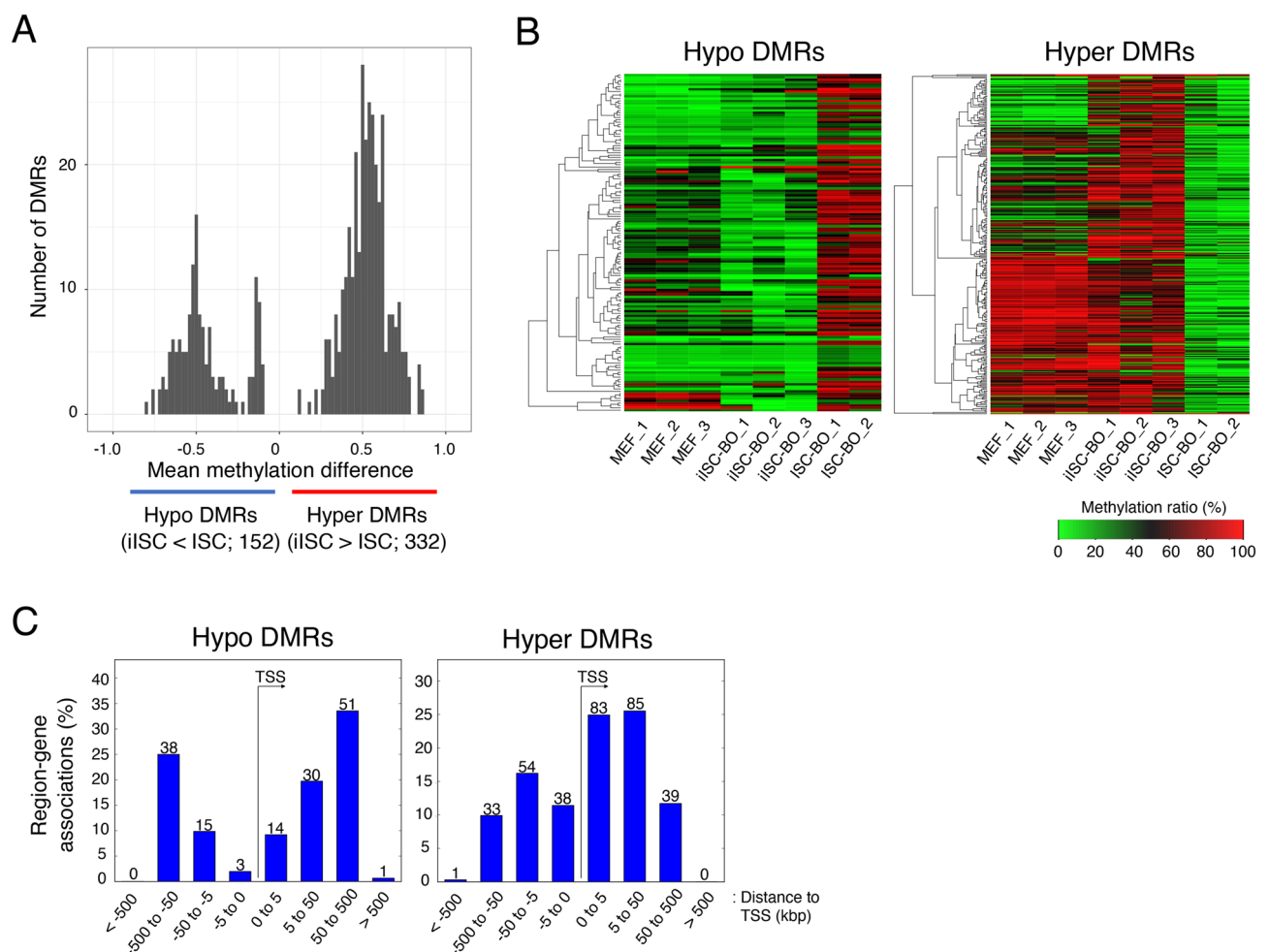


Figure 3. Analysis of DMRs between iISC-BOs and ISC-BOs. **(A)** Histogram showing mean methylation difference of DMRs between iISC-BOs and ISC-BOs. **(B)** Heatmaps showing averaged methylation ratio of detected DMRs in each sample. **(C)** Relative distance of detected DMRs from TSS of DMR-associated genes. The numbers above the bar plots indicate the counts of DMRs. As a pretreatment of the methylation data analysis, only cytosines assigned with 10 or more reads and located on autosomes were selected.

iISC-BOs have a cellular reprogramming-associated aberrant DNA methylation signature

As shown in Fig. 3, an abnormal DNA methylation state may be induced during direct reprogramming. Reprogramming-associated aberrant DNA methylation signatures can be classified into at least two groups. One is an original cell type-specific DNA methylation signature that should be changed but are abnormally maintained in reprogrammed cells (maintained DNA methylation: mDNA methylation), and the other is a DNA methylation signature that is abnormally acquired or erased in only reprogrammed cells (specific DNA methylation: sDNA methylation). To identify regions with mDNA or sDNA methylation in iISC-BOs, DMRs between iISC-BOs and ISC-BOs were divided into two groups based on the overlap with DMRs between MEF and ISC-BOs (Fig. 4A). When there was an overlap, iISCs still had an MEF-specific DNA methylation signature, indicating an mDNA methylation signature in iISCs. When there was no overlap, iISCs had a DNA methylation signature that was distinct from that of MEFs and iISCs, indicating an sDNA methylation signature in iISCs. Our data demonstrated that both hyper- and hypo-DMRs could be divided into regions with mDNA or sDNA methylation signatures, and that DMRs with sDNA methylation signatures were found more frequently than those with mDNA methylation signatures (Fig. 4B). Genome-wide distribution analyses revealed that both mDNA and sDNA methylation was detected at higher frequencies in exons with hyper-DMRs than in those with hypo-DMRs (Fig. 4C). Because hypermethylation of gene body CGIs is considered a result of enhanced transcription³¹, our data suggest that the mDNA and sDNA methylation signatures found in exons with hypermethylated DMRs reflect errors in transcriptional activation during direct reprogramming from MEFs to iISCs.

Abnormal DNA methylation in iISC-BOs has little effect on gene expression

Our present data suggest that hyper-/hypo-DMRs affect the transcriptional levels of genes and lead to the misregulation of gene expression in iISC-BOs. Our previous study demonstrated that the transcriptional signature of

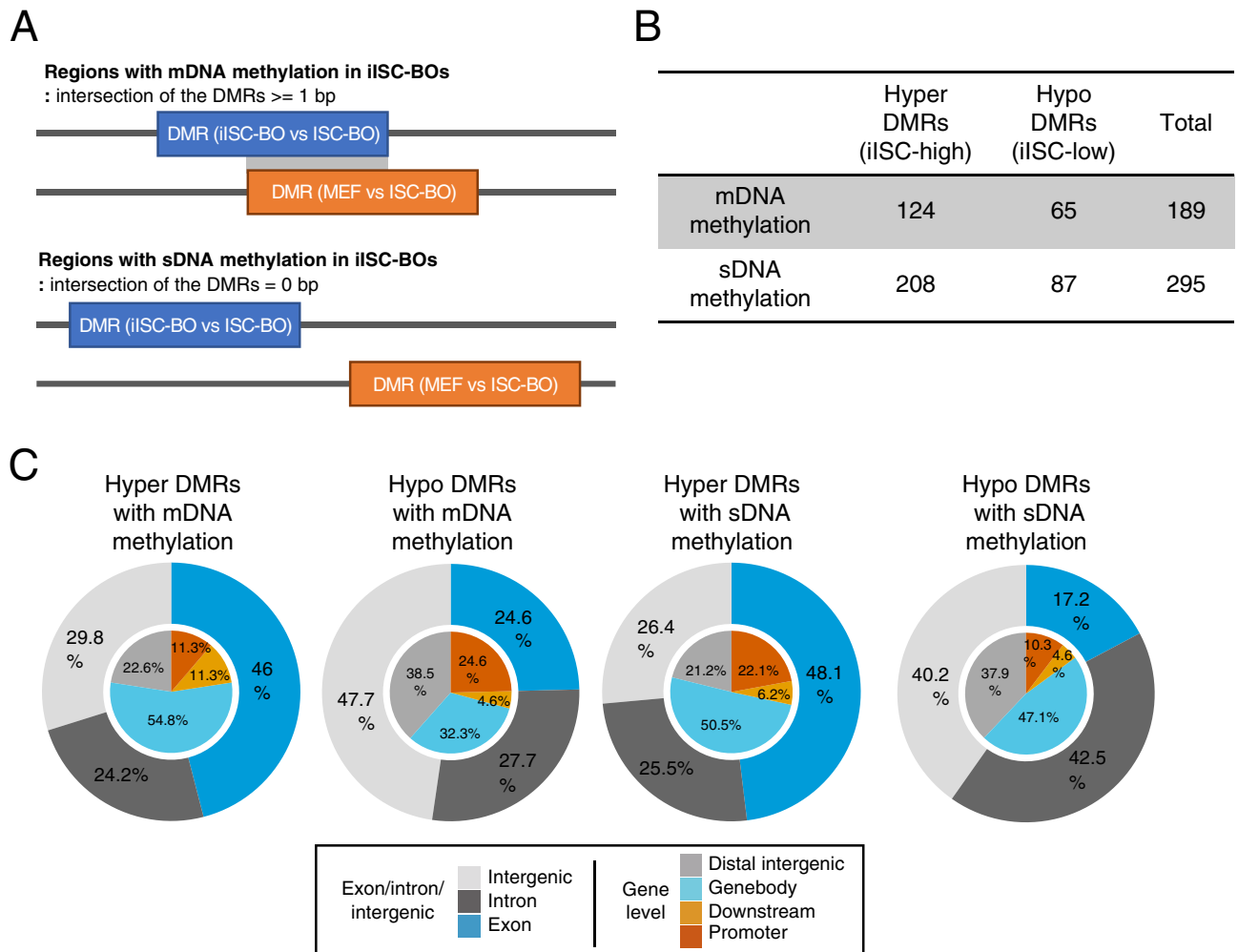


Figure 4. Classification and genomic positioning of abnormally methylated/unmethylated regions in iISC-BOs. (A) Definition of DMRs with mDNA and sDNA methylation in iISC-BOs (upper and bottom figures, respectively). (B) Number of DMRs in each class: hyper-DMRs with mDNA methylation, hyper-DMRs with sDNA methylation, hypo-DMRs with mDNA methylation, and hypo-DMRs with sDNA methylation. (C) Pie charts showing genomic position of DMRs in each class. Outer and inner circles indicate genomic position with different definition.

iISC-BOs closely resembles that of ISC-BOs, although a small number of differentially expressed genes (DEGs) were found between iISC-BOs and ISC-BOs²⁶. Thus, in this study, we examined the relationship between hyper-/hypo-DMRs and DEGs, both of which were identified in iISC-BOs and ISC-BOs. Genes that contained hyper-/hypo-DMRs within 10 kb upstream and downstream of the TSS were extracted and divided into DMRs with mDNA or sDNA methylation signatures (Fig. 5A). Among these hyper-/hypo-DMR-associated genes, hyper-DMRs were found at high frequencies downstream of the TSS, including the gene body, whereas only a few hypo-DMRs were found within 10 kbp upstream and downstream of the TSS (Fig. 5A). In addition, sDNA methylation was found more frequently than mDNA methylation in hyper-DMR-associated genes (Fig. 5A). These data suggest that hyper-DMRs are involved in the regulation of DEGs between iISC-BOs and ISC-BOs.

To examine this possibility, we classified DEGs into genes expressed at higher or lower levels in iISC-BOs than in ISC-BOs, designated iISC-high DEGs or iISC-low DEGs, respectively, and compared each of them with hyper-DMR-associated genes (Fig. 5B). Our data demonstrated that the number of genes contained in the hyper-DMR-associated genes and iISC-high DEGs or iISC-low DEGs were 9 and 12, respectively (Fig. 5B). These genes comprised only 6.5% and 8.6% of the hyper-DMR-associated genes, respectively. Scatter plots showing fold-differences in the expression of hyper-DMR-associated genes revealed that nine genes have hyper-DMRs with an sDNA methylation signature, while the 12 genes are divided into four and eight genes that have DMRs with an mDNA and sDNA methylation signature, respectively (Fig. 5C). Nevertheless, the expression levels of the majority of hyper-DMR-associated iISC-high and iISC-low DEGs (*Foxc1*, *Gprc5b*, *Hoxa5*, *Satb2*, *H2-T3*, *Gata4*, *Lefty1*, *Foxa1*, *Tubb2b*, *Pbld1*, and *Proz*) were significantly upregulated or downregulated in iISC-BOs compared to MEF, in the same manner as in ISC-BOs, while those of other DEGs (*Hoxa13*, *Gsc*, *Hoxd9*, *Mkx*, and *Adprhl1*) did not change significantly (Fig. 6).

In addition to the hyper-DMR-associated genes, the hypo-DMR-associated genes also overlapped with iISC-high and iISC-low DEGs, although the number of overlapping genes was only three and five, respectively (Fig. 7A). Five genes had hypo-DMRs with an sDNA methylation signature, while three genes were divided into one and two genes that had DMRs with mDNA and sDNA methylation signatures, respectively (Fig. 7B). Similar to the case of hyper-DMR-associated iISC-high and iISC-low DEGs, the hypo-DMR-associated iISC-high and iISC-low DEGs (*Foxa1*, *Onecut1*, *Onecut2* and *Ddr2*), except for *Hoxa9*, were expressed at much higher or lower levels than in MEFs and ISC-BOs (Fig. 7C). Taken together, our data demonstrated that almost all abnormal DNA methylation states in iISC-BOs may not be involved in the dysregulation of genes that are normally expressed or silenced in ISC-BOs.

Discussion

Our genome-wide methylome analysis revealed that the methylation state of CpG, but not that of CHH or CHG, changed during the direct reprogramming process of iISC-BOs. CpG methylation, the most frequently observed form of DNA methylation in almost all mammalian somatic cells, is important for transcription³². Previous studies have shown that CHH and CHG methylation may also play important roles in plant and mammalian pluripotency and neuronal cells^{33,34}. In the reprogramming of MEFs to iPSCs and neuronal cells, not only the methylation state of CpG, but also that of CHH and CHG changed significantly^{24,35}. Thus, it is suggested that variation in CHH and CHG methylation is found in specific cell types and should be corrected during reprogramming from other cell types.

Abnormal CpG methylation signature that affects transcription is often observed during the reprogramming process of iPSCs^{15,36}. These abnormally methylated regions found in iPSCs are similarly observed at high frequencies in the original cells, which are known as memory DMRs^{15,36}. These memory DMRs are involved in the differentiation potential of iPSCs^{15,16,37–40}. Meanwhile, 39% of the abnormally methylated regions found in iISC-BOs were identified as memory DMRs, suggesting that these memory DMRs negatively affect the differentiation potential of iISCs. However, our previous study demonstrated that iISCs can differentiate into functional intestinal epithelial cells in the same manner as ISCs²⁶. Moreover, among the memory DMR-associated genes, only *Hoxa9* was abnormally expressed in iISC-BOs. Thus, it is suggested that cellular properties and gene expression signatures are less influenced by memory DMRs in reprogramming-induced cells than in iPSCs.

While the molecular mechanisms underlying the transition of DNA methylation during the direct reprogramming from MEFs to iISC-BOs remain unclear, it is plausible that the alteration in expression of DNA methylation-related genes plays a role, because some of these genes were up- or down-regulated significantly during the cell-fate conversion. Meanwhile, the expression level of the genes in iISC-BOs was almost equal to that in ISC-BOs. This suggests that the abnormal methylation sites in iISC-BO are not attributable to abnormal expression of the DNA methylation-related genes.

Our present data demonstrated that hyper-/hypo-DMRs, at least, within 10 kbp upstream and downstream of the TSS had no significant impact on transcription. However, it is possible that hyper-/hypo-DMRs at a distance of more than 10 kbp from the TSS affect transcription. In particular, hypo-DMRs were predominantly found more than 50 kbp away from the TSS, suggesting that putative enhancers are regulated by DNA demethylation in iISC-BOs. To overcome the limitation of our analyzing strategies, a combined analysis of DMRs with three-dimensional genome structures will allow us to investigate the effects of DMRs located in regions distal to the TSS on transcription⁴¹. Moreover, to elucidate the impact of specific aberrant DNA methylation sites on cellular function and phenotype, it is imperative to employ new technologies like epigenome editing⁴².

The genome-wide methylation state of iISC-BOs was similar to that of ISC-BOs, although a small number of aberrant methylation sites were observed in iISC-BOs. In iPSCs, the number of aberrant methylation sites gradually decreases because of serial cell division in long-term culture, and the DNA methylation signature of iPSCs approaches that of ESCs⁴³. Thus, it is suggested that the DNA methylation signature of iISC-BOs gradually resembles that of ISC-BOs by continuous self-renewal cell divisions of iISCs under suitable cell culture

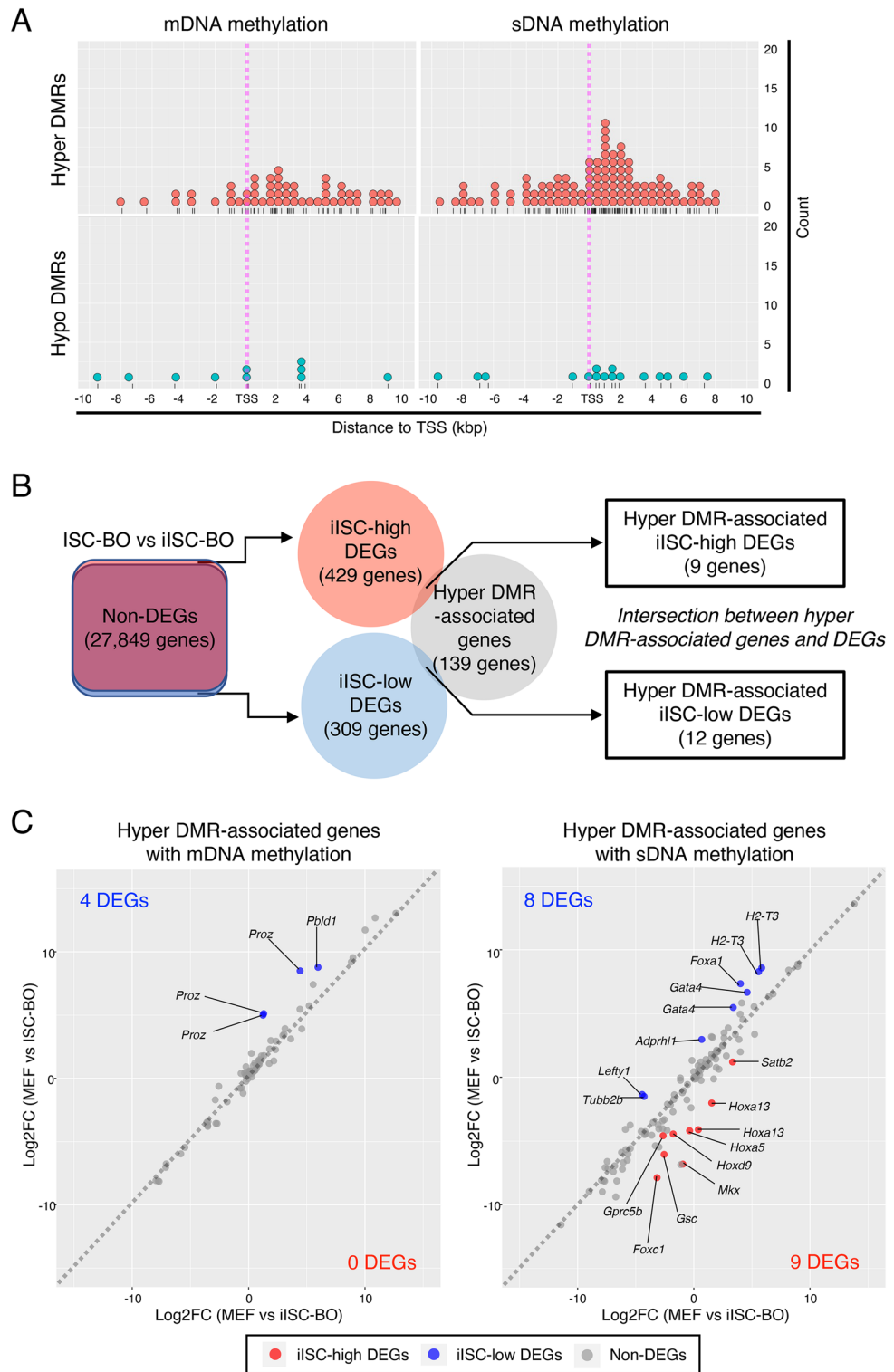


Figure 5. Trans-omic analysis for hyper-DMR-associated genes. **(A)** Dot plots showing distribution of hyper- (upper plots) and hypo- (lower plots) DMRs with mDNA (left plots) and sDNA (right plots) methylation around TSS of DMR-associated genes. **(B)** Right Venn diagram indicates intersection between genes associated with hyper-DMRs and iSC-high or -low DEGs which were detected from a comparative analysis between iSC-BO and iSC-BOs (left Venn diagram). **(C)** Scatter plots showing log-fold change of gene expression (Log_2FC) of the hyper-DMR-associated genes between MEFs and iSC-BOs (vertical axis), and these between MEFs and iSC-BOs (horizontal axis). Left and right plots indicate hyper-DMR-associated genes with mDNA and sDNA methylation, respectively.

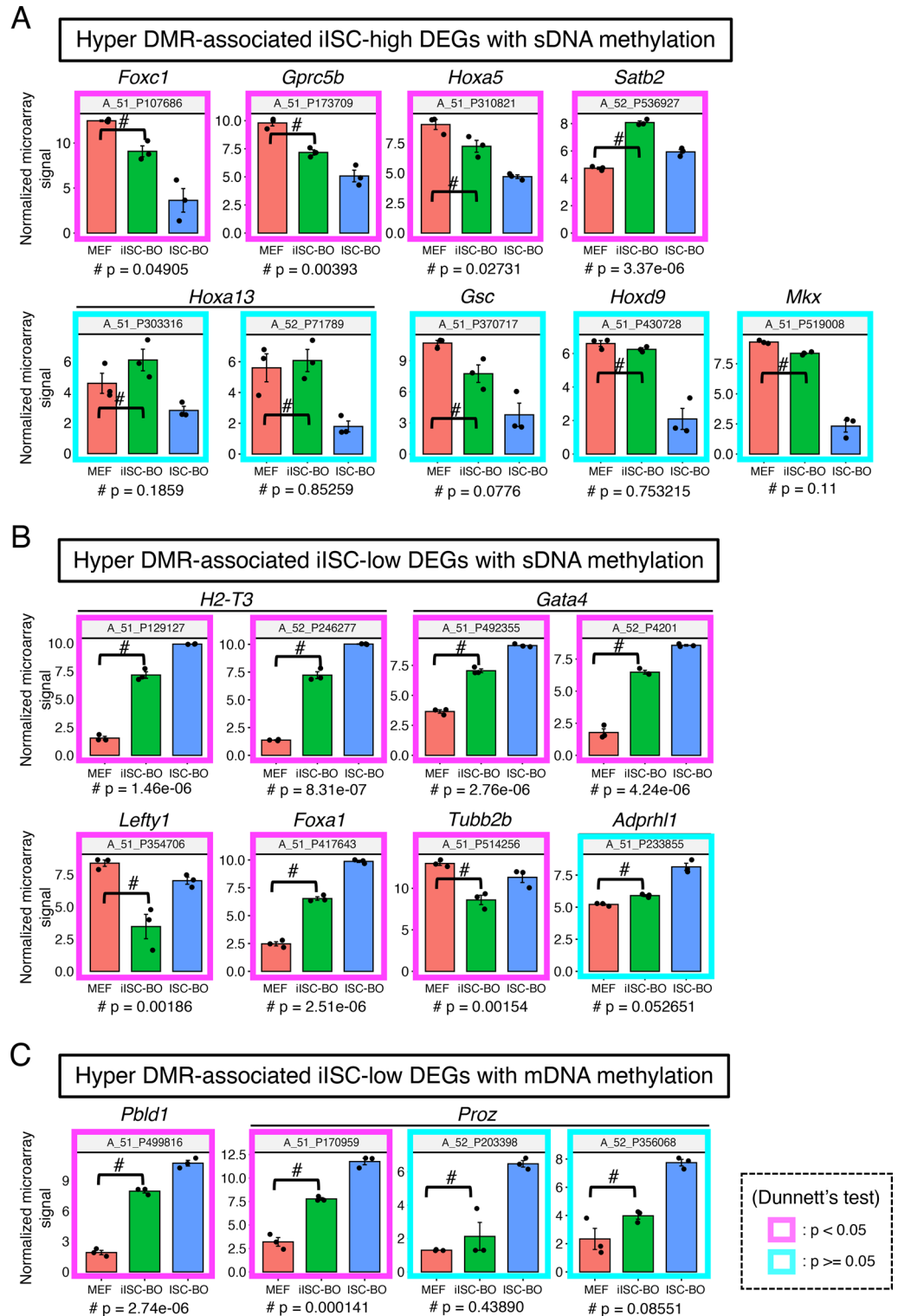


Figure 6. Gene expression analysis for hyper-DMR-associated DEGs. **(A)** Gene expression levels of hyper-DMR-associated iISC-high DEGs with sDNA methylation. **(B)** Expression levels of hyper-DMR-associated iISC-low DEGs with sDNA methylation. **(C)** Expression levels of hyper-DMR-associated iISC-low DEGs with mDNA methylation. Vertical axes of all plots indicate normalized and log-transformed microarray signals. Probe IDs of the microarray corresponding to the transcripts are indicated at the top of each plot. Transcripts showing significant differences in expression between MEFs and iISC-BOs are shown in the magenta box ($p < 0.05$) and those with no difference are shown in cyan boxes ($p \geq 0.05$). Dunnett's test was used for statistical analysis.

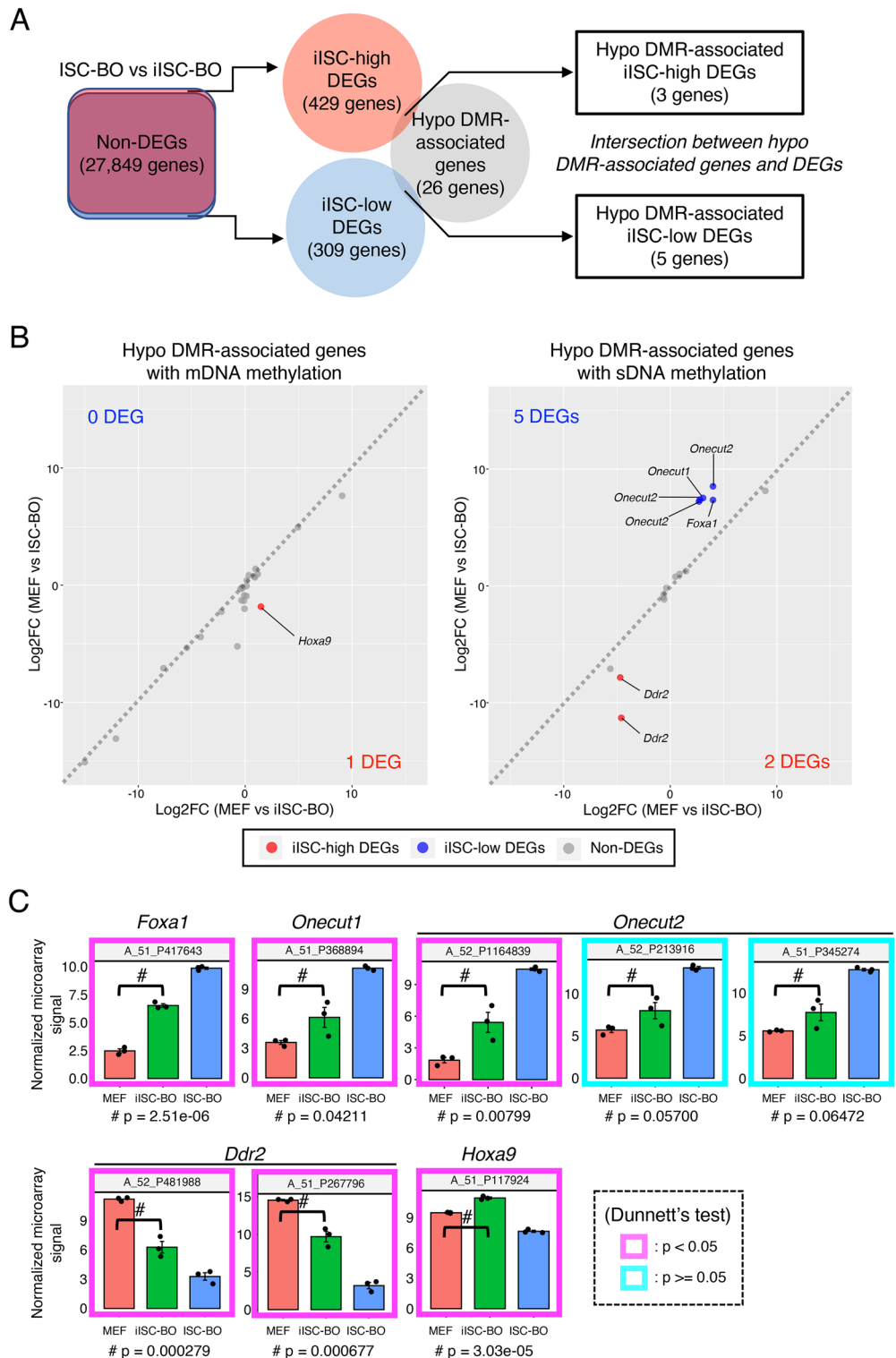


Figure 7. Trans-omic analysis for hypo-DMR-associated genes. **(A)** The right Venn diagram indicates the intersection between genes associated with hypo-DMRs and iISC-high or -low DEGs, which were detected from a comparative analysis between iISC-BO and ISC-BOs (left Venn diagram). **(B)** Scatter plots showing log-fold changes in gene expression (Log₂FC) of the hypo-DMR-associated genes between MEFs and ISC-BOs (vertical axis) and between MEFs and iISC-BOs (horizontal axis). Left and right plots indicate hypo-DMR-associated genes associated with mDNA and sDNA methylation, respectively. **(C)** Expression levels of hypo-DMR-associated DEGs. Vertical axes of all plots indicate normalized and log-transformed microarray signals. Probe IDs of the microarray corresponding to the transcripts are indicated at the top of each plot. Transcripts showing significant differences in expression between MEFs and iISC-BOs are shown in the magenta box ($p < 0.05$) and those with no difference are shown in cyan boxes ($p \geq 0.05$). Dunnnett's test was used for statistical analysis.

conditions. Tissue stem/progenitor cells induced using direct reprogramming technology can obtain a stable DNA methylation state similar to that of tissue-derived allogeneic cells during propagation in culture, which is more suitable for clinical applications than non-proliferative differentiated cells.

Methods

Cell source and generating procedure of iISC-BOs and ISC-BOs

Intestinal organoids (iISC-BOs and ISC-BOs) used in this study were prepared as described previously²⁶. The procedure is briefly described as follows: MEFs were prepared using embryos from E13.5 male and female C57BL/6 mice (Clea, Tokyo, Japan) and cultured in MEF medium¹⁷. To induce direct reprogramming, MEFs were infected with retroviruses expressing *Hnf4a*, *Foxa3*, *Gata6*, and *Cdx2*. Adult intestinal crypts were isolated from 10-week-old C57BL/6 male mice (Clea). The reprogrammed cells and the cells obtained from adult intestinal crypts were embedded in Matrigel (BD Biosciences, Franklin Lakes, NJ, USA) and cultured in mouse intestinal basal medium [advanced DMEM/F-12 (Thermo Fisher Scientific, MA, USA), 2 mM L-Alanyl-L-glutamine (Nacalai Tesque, Kyoto, Japan), 10 mM HEPES (Nacalai Tesque), N-2 supplement (1 ×) (Thermo Fisher Scientific), B-27 supplement (1 ×) (Thermo Fisher Scientific), 1 mM N-acetylcysteine (Sigma-Aldrich, MO, USA), and penicillin/streptomycin (Nacalai Tesque)], supplemented with 50 ng/mL human recombinant EGF (Sigma-Aldrich), 100 ng/mL murine recombinant Noggin (PeproTech, NJ, USA), and 500 ng/mL human recombinant R-spondin1 (PeproTech). The iISC-BOs and ISC-BOs were used after more than 10 passages in 3D culture.

Whole-genome bisulfite sequencing (WGBS) with a PBAT

Three, two, and three independent replicates were prepared for iISC-BOs, ISC-BOs, and MEFs, respectively, and these samples were individually converted into WGBS libraries. The genomic DNA used for WGBS was prepared using a NucleoSpin Tissue Kit (Macherey–Nagel, Dueren, Germany) according to the manufacturer's instructions. For sequencing library preparation, 100 ng of purified genomic DNA was spiked with 1 ng of unmethylated lambda DNA (Promega, Madison, WI, USA) and subjected to bisulfite treatment using the EZ DNA methylation gold kit (Zymo Research, Irvine, CA, USA). Library preparation was based on the tPBAT protocol, an improved version of the PBAT strategy²⁸. After library preparation, sequencing was performed using the Illumina HiSeq X Ten system (Macrogen, Tokyo, Japan), assigning one-third of the lanes per sample. The sequenced reads were mapped to the mouse reference genome mm9 combined with the genome sequence of Escherichia phage lambda using BMap as previously described²⁸. The mapped reads were summarized using in-house software, and the basic statistics of the methylome data are provided in Supplementary Table 2. The reads obtained from the independent replicates were treated individually and as groups, depending on the purpose of the analysis.

Transcriptome data analysis

DEGs were detected in the previous study with the GeneSpring software (Agilent, Pal Alto, CA, USA)²⁶. The log-fold change and normalized signal of each probe were calculated from the raw data.

Identification of DMRs

The DMRs between MEFs, ISC-BOs, and iISC-BOs were identified using the Metilene program (version 0.2.8) with default parameters⁴⁴. DMRs were filtered to retain those containing at least 20 CpGs with a q-value less than 0.05, and a methylation difference larger than 10%.

Analysis of the genomic location of DMRs

The genomic locations of the DMRs were analyzed using the ChIPpeakAnno program (version 3.28.1)^{45,46} in R (version 4.1.2). The distance between the DMRs and the TSS was determined using the GREAT program (version 4.0.4)⁴⁷.

Statistics and reproducibility

Two ISC-BO, three iISC-BO, and three MEF samples were employed as biological replicates of methylome data. Dunnett's test was used for statistical analysis of transcriptome data.

Ethics statement

This study was approved by the Kyushu University Animal Experiment Committee and were performed according to the ethical guidelines of Kyushu University (<https://ura.kyushu-u.ac.jp/animal/en/>). All animal experiments were carried out in accordance with the ARRIVE guidelines (<https://arriveguidelines.org/>).

Data availability

Raw WGBS data have been deposited in the NCBI for Biotechnology Information Gene Expression Omnibus (GEO) database (GSE222669). Public datasets of microarray analysis for gene expression in MEFs, ISC-BOs, and iISC-BOs are available in the NCBI GEO database (GSE85232).

Code availability

The source codes for the programs used in this study can be downloaded from GitHub (<https://github.com/FumihitoMiura/Project-2>).

Received: 27 October 2023; Accepted: 9 December 2023

Published online: 15 December 2023

References

- Sarkies, P. Encyclopaedia of eukaryotic DNA methylation: From patterns to mechanisms and functions. *Biochem. Soc. Trans.* **50**, 1179–1190 (2022).
- Suzuki, M. M. & Bird, A. DNA methylation landscapes: Provocative insights from epigenomics. *Nat. Rev. Genet.* **9**, 465–476 (2008).
- Deaton, A. M. & Bird, A. CpG islands and the regulation of transcription. *Genes Dev.* **25**, 1010–1022 (2011).
- Wan, J. *et al.* Characterization of tissue-specific differential DNA methylation suggests distinct modes of positive and negative gene expression regulation. *BMC Genomics* **16**, 49 (2015).
- Rauluseviute, I., Drablos, F. & Rye, M. B. DNA hypermethylation associated with upregulated gene expression in prostate cancer demonstrates the diversity of epigenetic regulation. *BMC Med. Genom.* **13**, 6 (2020).
- Greenberg, M. V. C. & Bourc'his, D. The diverse roles of DNA methylation in mammalian development and disease. *Nat. Rev. Mol. Cell Biol.* **20**, 590–607 (2019).
- Chen, Z. & Zhang, Y. Role of mammalian DNA methyltransferases in development. *Annu. Rev. Biochem.* **89**, 135–158 (2020).
- Dor, Y. & Cedar, H. Principles of DNA methylation and their implications for biology and medicine. *The Lancet* **392**, 777–786 (2018).
- Horisawa, K. & Suzuki, A. Direct cell-fate conversion of somatic cells: Toward regenerative medicine and industries. *Proc. Jpn. Acad. Ser. B* **96**, 131–158 (2020).
- Shekhani, M. T., Jayanthi, A.-S., Maddodi, N. & Setaluri, V. Cancer stem cells and tumor transdifferentiation: Implications for novel therapeutic strategies. *Am. J. Stem Cells* **2**, 52–61 (2013).
- Niemann, H., Tian, X. C., King, W. A. & Lee, R. S. F. Epigenetic reprogramming in embryonic and foetal development upon somatic cell nuclear transfer cloning. *Reproduction* **135**, 151–163 (2008).
- Mikkelsen, T. S. *et al.* Dissecting direct reprogramming through integrative genomic analysis. *Nature* **454**, 49–55 (2008).
- Chan, M. M., Smith, Z. D., Egli, D., Regev, A. & Meissner, A. Mouse ooplasm confers context-specific reprogramming capacity. *Nat. Genet.* **44**, 978–980 (2012).
- Zhang, S. *et al.* Aberrant DNA methylation reprogramming in bovine SCNT preimplantation embryos. *Sci. Rep.* **6**, 30345 (2016).
- Kim, K. *et al.* Epigenetic memory in induced pluripotent stem cells. *Nature* **467**, 285–290 (2010).
- Lister, R. *et al.* Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature* **471**, 68–73 (2011).
- Sekiya, S. & Suzuki, A. Direct conversion of mouse fibroblasts to hepatocyte-like cells by defined factors. *Nature* **475**, 390–393 (2011).
- Huang, P. *et al.* Induction of functional hepatocyte-like cells from mouse fibroblasts by defined factors. *Nature* **475**, 386–389 (2011).
- Ieda, M. *et al.* Direct reprogramming of fibroblasts into functional cardiomyocytes by defined factors. *Cell* **142**, 375–386 (2010).
- Vierbuchen, T. *et al.* Direct conversion of fibroblasts to functional neurons by defined factors. *Nature* **463**, 1035–1041 (2010).
- Lee, C. S., Friedman, J. R., Fulmer, J. T. & Kaestner, K. H. The initiation of liver development is dependent on Foxa transcription factors. *Nature* **435**, 944–947 (2005).
- Castro, D. S. *et al.* A novel function of the proneural factor *Ascl1* in progenitor proliferation identified by genome-wide characterization of its targets. *Genes Dev.* **25**, 930–945 (2011).
- Hong, J. H. & Zhang, H. G. transcription factors involved in the development and prognosis of cardiac remodeling. *Front. Pharmacol.* **13**, 828549 (2022).
- Luo, C. *et al.* Global DNA methylation remodeling during direct reprogramming of fibroblasts to neurons. *eLife* **8**, e40197 (2019).
- Koblas, T., Leontovyc, I., Loukotova, S., Kosinova, L. & Saudek, F. Reprogramming of pancreatic exocrine cells AR42J into insulin-producing cells using mRNAs for Pdx1, Ngn3, and MafA transcription factors. *Mol. Ther. Nucleic Acids* **5**, e320 (2016).
- Miura, S. & Suzuki, A. Generation of mouse and human organoid-forming intestinal progenitor cells by direct lineage reprogramming. *Cell Stem Cell* **21**, 456–471.e5 (2017).
- Fukuda, M. *et al.* Small intestinal stem cell identity is maintained with functional Paneth cells in heterotopically grafted epithelium onto the colon. *Genes Dev.* **28**, 1752–1757 (2014).
- Miura, F., Enomoto, Y., Dairiki, R. & Ito, T. Amplification-free whole-genome bisulfite sequencing by post-bisulfite adaptor tagging. *Nucleic Acids Res.* **40**, e136 (2012).
- Miura, F. & Ito, T. Post-bisulfite adaptor tagging for PCR-free whole-genome bisulfite sequencing. In *DNA Methylation Protocols* (ed. Tost, J.) 123–136 (Springer, 2018). https://doi.org/10.1007/978-1-4939-7481-8_7.
- Smith, Z. D. & Meissner, A. DNA methylation: Roles in mammalian development. *Nat. Rev. Genet.* **14**, 204–220 (2013).
- Jeziorska, D. M. *et al.* DNA methylation of intragenic CpG islands depends on their transcriptional activity during differentiation and disease. *PNAS* **114**, E7526–E7535 (2017).
- Pelizzola, M. & Ecker, J. R. The DNA methylome. *FEBS Lett.* **585**, 1994–2000 (2011).
- Lister, R. *et al.* Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462**, 315–322 (2009).
- Lister, R. *et al.* Global epigenomic reconfiguration during mammalian brain development. *Science* **341**, 1237905 (2013).
- Araki, H. *et al.* Base-resolution methylome of retinal pigment epithelial cells used in the first trial of human induced pluripotent stem cell-based autologous transplantation. *Stem Cell Rep.* **13**, 761–774 (2019).
- Bar, S. & Benvenisty, N. Epigenetic aberrations in human pluripotent stem cells. *EMBO J.* **38**, e101033 (2019).
- Ohji, Y. *et al.* Incomplete DNA methylation underlies a transcriptional memory of somatic cells in human iPS cells. *Nat. Cell Biol.* **13**, 541–549 (2011).
- Kim, K. *et al.* Donor cell type can influence the epigenome and differentiation potential of human induced pluripotent stem cells. *Nat. Biotechnol.* **29**, 1117–1119 (2011).
- Bar-Nur, O., Russ, H. A., Efrat, S. & Benvenisty, N. Epigenetic memory and preferential lineage-specific differentiation in induced pluripotent stem cells derived from human pancreatic islet beta cells. *Cell Stem Cell* **9**, 17–23 (2011).
- Roost, M. S. *et al.* DNA methylation and transcriptional trajectories during human development and reprogramming of isogenic pluripotent stem cells. *Nat. Commun.* **8**, 908 (2017).
- Gong, H., Yang, Y., Zhang, S., Li, M. & Zhang, X. Application of Hi-C and other omics data analysis in human cancer and cell differentiation research. *Comput. Struct. Biotechnol. J.* **19**, 2070–2083 (2021).
- Goell, J. H. & Hilton, I. B. CRISPR/Cas-based epigenome editing: Advances, applications, and clinical utility. *Trends Biotechnol.* **39**, 678–691 (2021).
- Nishino, K. *et al.* DNA methylation dynamics in human induced pluripotent stem cells over time. *PLOS Genet.* **7**, e1002085 (2011).
- Jühling, F. *et al.* metilene: Fast and sensitive calling of differentially methylated regions from bisulfite sequencing data. *Genome Res.* **26**, 256–262 (2016).
- Zhu, L. J. *et al.* ChIPpeakAnno: A bioconductor package to annotate ChIP-seq and ChIP-chip data. *BMC Bioinform.* **11**, 237 (2010).
- Zhu, L. J. Integrative analysis of ChIP-chip and ChIP-seq dataset. In *Tiling Arrays: Methods and Protocols* (eds Lee, T.-L. & Shui Luk, A. C.) 105–124 (Humana Press, 2013). https://doi.org/10.1007/978-1-62703-607-8_8.
- McLean, C. Y. *et al.* GREAT improves functional interpretation of cis-regulatory regions. *Nat. Biotechnol.* **28**, 495–501 (2010).

Acknowledgements

We thank Drs. Toshio Kitamura, Masafumi Onodera, and Hiroyuki Miyoshi for sharing the Plat-E cells, pGCDNsam and pGCDNsam-ires-EGFP vectors, and the pCMV-VSV-G plasmid, respectively, and Yuuki Honda, Mariko Tasai, Mitsuhiro Kurata, Yoshimi Iwasaki, Emiko Koba, Ryo Ugawa, and Kanako Ichikawa for their excellent technical assistance. This work was supported in part by the JSPS KAKENHI (Grant Numbers: JP16K08592 and JP23K11851 to K.H.; JP18H06069 and JP21K18039 to S.M.; JP18H05102, JP19H01177, JP19H05267, JP20H05040, JP21K19916, JP22H05634, JP22H04698, and JP22H00592 to A.S.), the Program for Basic and Clinical Research on Hepatitis of the Japan Agency for Medical Research and Development (AMED) (JP23fk0210116 to A.S.), the Research Center Network for Realization of Regenerative Medicine of AMED (JP23bm1123005 to A.S.), the Platform Project for Supporting Drug Discovery and Life Science Research (Basis for Supporting Innovative Drug Discovery and Life Science Research (BINDS)) from AMED under Grant Number JP21am0101103 (to T.I.), the Medical Research Center Initiative for High Depth Omics (to K.H., S.M., and A.S.), the Takeda Science Foundation (to S.M. and A.S.), the Uehara Memorial Foundation (to K.H., S.M., and A.S.), the Kato Memorial Trust for Nambyo Research (to A.S.), the Suzuken Memorial Foundation (to A.S.), the Naito Foundation (to S.M. and A.S.), and the Shinnihon Foundation of Advanced Medical Treatment Research (to S.M.).

Author contributions

K.H. and S.M. performed the experiments, collected samples, and conducted data analyses. H.A., F.M., and T.I. performed methylome analysis using the PBAT method. A.S. contributed to the conception, design, and overall project management. K. H. and A. S. wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-49546-8>.

Correspondence and requests for materials should be addressed to A.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023