Tissue-specific demethylation in CpG-poor promoters during cellular differentiation

Genta Nagae¹, Takayuki Isagawa¹, Nobuaki Shiraki³, Takanori Fujita¹, Shogo Yamamoto¹, Shuichi Tsutsumi¹, Aya Nonaka¹, Sayaka Yoshida¹, Kelsuke Matussaka¹,², Yutaka Midorikawa¹, Shumpei Ishikawa¹,², Hidenobu Soejima⁴, Masashi Fukayama², Hiroyumi Suemori⁵, Norio Nakatsuji⁶, Shoen Kume³ and Hiroyuki Aburatani¹,*

¹Genome Science Division, Research Center for Advanced Science and Technology and ²Department of Pathology, University of Tokyo, Tokyo, Japan, ³Department of Stem Cell Biology, Institute of Molecular Embryology and Genetics, University of Kumamoto, Kumamoto, Japan, ⁴Division of Molecular Genetics and Epigenetics, Department of Biomolecular Sciences, Faculty of Medicine, Saga University, Saga, Japan, ⁵Laboratory of Embryonic Stem Cell Research, Stem Cell Research Center, Institute for Frontier Medical Sciences and ⁶Institute for Integrated Cell-Material Sciences, Kyoto University, Kyoto, Japan

Received December 13, 2010; Revised and Accepted April 14, 2011

Epigenetic regulation is essential in determining cellular phenotypes during differentiation. Although tissue-specific DNA methylation has been studied, the significance of methylation variance for tissue phenotypes remains unresolved, especially for CpG-poor promoters. Here, we comprehensively studied methylation levels of 27,578 CpG sites among 21 human normal tissues from 12 anatomically different regions using an epigenotyping beadarray system. Remarkable changes in tissue-specific DNA methylation were observed within CpG-poor promoters but not CpG-rich promoters. Of note, tissue-specific hypomethylation is accompanied by an increase in gene expression, which gives rise to specialized cellular functions. The hypomethylated regions were significantly enriched with recognition motifs for transcription factors that regulate cell-type-specific differentiation. To investigate the dynamics of hypomethylation events, we analyzed methylation levels of the entire APOA1 gene locus during in vitro differentiation of embryonic stem cells toward the hepatic lineage. A decrease in methylation was observed after day 13, coinciding with alpha-fetoprotein detection, in the vicinity of its transcription start sites (TSSs), and extends up to ~200 bp region encompassing the TSS at day 21, equivalent to the hepatoblastic stage. This decrease is even more pronounced in the adult liver, where the entire APOA1 gene locus is hypomethylated. Furthermore, when we compared the methylation status of induced pluripotent stem (iPS) cells with their parental cell, IMR-90, we found that fibroblast-specific hypomethylation is restored to a fully methylated state in iPS cells after reprogramming. These results illuminate tissue-specific methylation dynamics in CpG-poor promoters and provide more comprehensive views on spatiotemporal gene regulation in terminal differentiation.

INTRODUCTION

In a series of differentiation processes during embryogenesis, a wide variety of cells are generated and organized in a spatiotemporal manner. They acquire distinctive patterns of gene expression to execute specialized cellular functions. In mammals, this gene specification is tightly regulated by multiple levels of epigenetic systems such as DNA methylation, histone modification, chromatin remodeling and non-coding RNA guidance (1,2). Mammalian cells coordinate the complex transcriptional networks, which are essential for establishment of cellular programming and maintenance of given cellular phenotypes.

*To whom correspondence should be addressed at: Genome Science Division, Research Center for Advanced Science and Technology, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8904, Japan. Tel: +81 354525352; Fax: +81 354525355; Email: haburata-tky@umin.ac.jp

© The Author 2011. Published by Oxford University Press. All rights reserved. For Permissions, please email: journals.permissions@oup.com
DNA methylation has a strong impact on transcriptional repression. Because covalent modification of DNA itself is chemically stable when compared with other epigenetic marks, methylation-mediated repression is thought to be an effective mechanism to maintain long-lasting cell memories. Indeed, it plays pivotal roles in fundamental biological processes, including genome imprinting, retrotransposon silencing, X chromosome inactivation and tissue-specific gene expression (1). The lethality due to selective ablation of DNA methyltransferase with global loss of 5-methylcytosine also provides solid evidence for its significance in mammalian embryogenesis (3,4). Embryonic stem (ES) cells deficient in maintenance methyltransferase, Dnmt1, are viable, but die when induced to differentiate (5), demonstrating Dnmt1 is essential for the dynamic epigenetic changes in cellular differentiation.

For many years, tissue-specific differentially methylated regions (tDMRs) have been of great interest (6–9). In somatic tissues, which include terminally differentiated cells, significant methylation variance between cells have been reported (6,8–11). Although recent technological advances in methylation profiling have broadened our understanding of the human methylome, we are still far from a comprehensive map required for deeper understanding of developmental epigenomics. That is partly because, due to technological limitations, most of earlier studies on human tDMRs have focused on CpG island promoters (12,13). In general, housekeeping genes, which constitutively express across many tissues, have such CpG-rich promoters. However, more than half the genes which have a tissue-specific pattern of expression have CpG-poor promoters (14). Therefore, it is important to analyze CpG-poor promoters in addition to CpG-rich promoters to elucidate regulatory changes of methylation during terminal differentiation.

Recent large-scaled methylation analyses of human normal tissues have revealed that methylation variance can be identified outside CpG islands and at CpG-poor promoters (6). The significance of methylation in the marginal regions of CpG islands (so-called CpG shore methylation) has also been proposed (15). In addition, tissue-specific binding of RNA polymerase II is often observed in CpG-poor promoters (16). These observations point to a significant role for epigenetic dynamics in CpG-poor promoters for terminal differentiation.

There are some difficulties in analyzing methylation levels in human tissue samples with accuracy. Cell populations in human tissues are not homogenous but rather are composed of a heterogeneous cell population which originates from different lineages. Because measurements of methylation are derived from these different methylomes of component cells, large differences in methylation between different cell types can be obscured. It is necessary to evaluate the methylation status quantitatively, rather than just qualitatively, to allow any comparison of methylation profiles between different samples. This requires good assay reproducibility to detect the more subtle methylation differences. In this study, we performed genome-wide promoter methylation analysis of human normal tissues using an epigenotyping beadarray system, which allows methylated CpG quantification in CpG-poor promoters as well as in CpG-rich promoters (17). We utilized inclusive probe sets for tissue-specific hypermethylation and hypomethylation, which occur mainly in CpG-poor promoters.

Of note, we found that tissue-specific hypomethylation is well correlated with gene expression profiles that underlie tissue phenotypes. Around these cell-type-specific hypomethylated regions, binding motifs of particular transcription factors are remarkably enriched. These results suggest that a combination of tissue-specific promoter hypomethylation and selective binding of transcription factors is deeply involved in targeting specific genes during terminal differentiation. In addition, we demonstrated spreading of hypomethylation in CpG-poor promoters by in vitro cellular differentiation. The restoration of the fibroblast-specific hypomethylation was also observed during cellular reprogramming into induced pluripotent stem (iPS) cells. These results emphasize the importance of methylation dynamics in CpG-poor regions for multilayered epigenetic regulation in mammalian embryogenesis.

RESULTS

Genome-wide methylation analysis of human normal tissue reveals methylation variances in CpG-poor promoters

To develop a better understanding of methylation diversity among human normal tissues, we performed promoter methylation analysis of 21 human normal tissue samples from 12 anatomically different regions (Supplementary Material, Table S1). A HumanMethylation27 BeadChip® (Illumina, Inc) was used to quantify the methylation level of 27 578 CpG sites harboring 14 475 Refseq promoter regions (17).

First, we evaluated the accuracy and sensitivity of the assay using the modified DNA samples as methylation controls (0, 25, 50, 75 and 100% of methylation). The observed values of the methylated CpG ratio for the control samples were well correlated with the expected ratio of methylated CpGs (Supplementary Material, Fig. S1). Thus, methylation changes are quantitatively detectable using this system.

Next, we analyzed inter-individual methylation differences. The comparison plots of autosomal probes using biological duplicates of nine human tissues (brain, oral mucosa, lung, stomach, colon, liver, peripheral blood, kidney and skeletal muscle) showed good correlation between each pair (Pearson correlation coefficient; r > 0.97) (Supplementary Material, Fig. S2). For X-linked genes, most promoters on the inactivated allele are methylated in female cells. As expected, 0% methylation in male cells and ~50% methylation in female cells are accurately reported by the system (Supplementary Material, Fig. S2).

In this epigenotyping beadarray, most probes are designed to bind at and around the promoter regions, which are from 1.5 kb upstream to 1 kb downstream of transcription start sites (TSSs) of Refseq genes (Supplementary Material, Fig. S3). On the basis of the classification by the local CpG observed to expected ratio (CpG o/e) and the GC content ratio (GCR) around the probe, probes are divided into three groups: high-CpG density probes (HCG; CpG o/e > 0.75, GCR > 0.55), low-CpG density probes (LCG; CpG o/e < 0.48) and intermediate-CpG density probes (ICG; neither HCG nor LCG) (Supplementary Material, Fig. S4). The promoter methylation status is strongly affected by local CpG density. Most probes in relatively CpG-rich regions (HCG
Identification of tissue-specific hypermethylated and hypomethylated regions

To identify the tissue-specific differential gene methylation, we compared the methylation profiles of seven representative tissues. These were the brain and oral mucosa from the ectodermal lineage, the colon and liver from the endodermal lineage, the peripheral blood and skeletal muscle from the mesodermal lineage, and the testis. First, we ranked the 26,486 autosomal probes in order of difference of the methylation level between the one tissue and the average of the other tissues. In case that tissue-specific hypomethylation or hypermethylation are sorted by the absolute values of the difference of the methylation level (more than 0.25 or less than −0.25), the number of distinctive gene sets varies widely (Supplementary Material, Fig. S6). There are more tissue-specific hypermethylated genes in the brain, liver, blood and testis than in other tissues. As for the hypomethylation, a large number of genes are selected in the testis and oral mucosa by this criterion. To evaluate the specific differential methylation equally among human tissues, we selected the top 250 probes of tissue-specific hypomethylation and hypermethylation for each tissue (Supplementary Material, Table S2). The methylation panel clearly shows specific hypomethylation as well as hypermethylation among seven tissues (Fig. 2A and B). We validated the methylation levels of the distinctive genes using the MassARRAY system. These methylation levels were consistent with the microarray data (Supplementary Material, Fig. S7). With respect to CpG density, most tissue-specific hypomethylated sites (80–90%) are associated with CpG-poor promoters (Fig. 2A and B). A notable exception was the testis, as testis-specific hypomethylated genes are associated with CpG-rich promoters. In CpG-rich promoter regions (HCP and ICP, n = 18,481), ~900 regions (5.10 ± 0.48%) were found to be densely hypermethylated (mCpG > 70%) in somatic tissues. In the testis, only 286 regions (1.55%) were methylated (Supplementary Material, Fig. S5). These results are in agreement with earlier systematic screens that found the major fraction of tDMRs corresponding to CpG island methylation are either sperm- or testis-specific (6,9,10).

Variable hypomethylation patterns are associated with tissue-specific gene functions, gene expression patterns and selective binding of transcription factors

To characterize the gene function related to tissue-specific hypomethylation, we examined the enrichment of the specific gene ontology (GO) biological process categories in the top 250 hypomethylated gene sets. As shown in Table 1, the tissue-specific hypomethylated genes are closely related to cell-type-specific functions. For example, oral mucosa-specific hypomethylated genes show over-representation of genes related to ectoderm or epidermis development. In the gene set of liver-specific hypomethylation, we found several protein families synthesized by hepatocytes, such as serpin peptidase inhibitors and complement factors. Thus, these gene sets show over-representation of genes associated with acute inflammatory response. In the blood set, we found genes related to immune response, a key role for white blood cells. Genes related to the reproductive process are the major targets for CpG methylation in somatic cells besides sperm and its progenitor cells in the testis.

In contrast, we could not find any meaningful functional association between gene sets which undergo tissue-specific hypermethylation (Supplementary Material, Table S3). Although previous reports have identified a substantial number of confirmed sets of tissue-specific hypermethylation, it has been difficult to associate these with the tissue phenotype. De novo hypermethylation in differentiated cells might be often induced independently of functional specification.

While dense methylation of the CpG island promoter deeply contributes to gene silencing in pathological conditions such as cancer (18,19), the influence of sparse methylation in CpG-poor promoters on gene expression still remains controversial. CpG-poor promoters preferentially display the TATA box and numerous transcription factor-binding motifs around the TSS (20). Combinations of transcription factor binding in regulatory elements are involved in targeting gene expression. Thus, we analyzed expression levels of representative gene sets of tissue-specific hypomethylation and hypermethylation across seven human tissues (Fig. 3). The average expression level of hypomethylated genes is significantly higher than that of hypermethylated genes in a tissue-specific manner.
specific manner. In contrast, tissue-specific hypermethylated genes are suppressed among all tissues. These results indicate that hypomethylation in the CpG-poor promoters identified here underlie tissue-specific expression in a given cell type.

Local epigenetic modification and recruitment of transcription factors are a fundamental part of the system for appropriate transcriptional regulation (21). We performed enrichment analysis of 746 recognition motifs for transcription factors to examine the relationship between cis-regulatory elements of promoters and tissue-specific hypomethylation. As shown in Figure 4, some matrices are significantly enriched (Z-score > 8.0) in the hypomethylated regions. In oral-mucosa-specific hypomethylated regions, the binding motifs of p53 family genes are highly enriched. p63, the master regulator of keratinocyte differentiation, has similar DNA-binding domains to p53 and half of p63-bound regions in the squamous cell carcinoma cell line have p53 consensus motifs (22). In liver-specific hypomethylated regions, the matrices for the C4 zinc finger domain of the PPAR family (PPARA, PPARG and RXRs) and the NR2F family (HNF4A) are enriched compared with the background sequences. In the blood set, the matrices for the ETS domain of ETS factors (ETS1, ETS2, ELF2, ELK1) and the Runt domain of AML factors (RUNX1) are enriched. Similarly, MyoD-binding motifs are enriched in skeletal muscle-specific hypomethylated regions. In contrast, we could not find significant enrichment of transcription factor-binding motifs in tissue-specific hypermethylated regions (data not shown). Although the molecular mechanism of de novo hypermethylation and hypomethylation remains unknown, it is suggested that selective binding of transcription factors are at least significantly associated with regional hypomethylation during terminal differentiation.

Dynamic changes of CpG-poor promoter methylation during in vitro differentiation and cellular reprogramming

Although tissue-specific hypomethylation in CpG-poor promoters are closely related to gene specification for the tissue phenotype, when and how these variable methylation statuses are established remain unknown. To elucidate the methylation changes during cellular differentiation, we performed clustering analysis of human somatic tissue and normal cells.
including human ES cells, iPS cells and primary fibroblast cells using tissue-specific hypomethylation sites (Fig. 5). The heatmap shows distinct methylation patterns between the pluripotent cells and somatic tissues composed of the terminally differentiated cells. Seven human ES cell lines and two iPS cell lines show similar methylation patterns. Intriguingly, most genes representing specific hypomethylation in differentiated cells are densely methylated in both ES cells and iPS cells, raising the possibility that the default state of low CpG promoters in the embryonic stage is totally methylated and erasure of methylation may occur during terminal differentiation in a cell-type-specific manner.

Next we compared the methylation status of the adult human liver and the fetal liver. Liver-specific hypomethylated genes are heavily methylated in KhES3, a human ES cell line, but are hypomethylated in the adult liver tissue (Fig. 6B and C). In the fetal liver, the methylation level of these genes shows a mild decrease in these regions. Bisulfite sequencing also revealed the partial hypomethylation of ITIH3 and APOA1 promoters (Supplementary Material, Fig S8A and B).

To further analyze the demethylation dynamics during hepatic differentiation, we analyzed methylation during in vitro differentiation toward hepatic lineages (23). On day 7, the cells began to express an endoderm marker, SOX17 (Fig. 6A). AFP expression was detected on day 13 and ALB expression was detected on day 21. The methylation status of liver-specific hypomethylated genes showed a slight decrease during hepatic differentiation (Fig. 6B and C). Indeed, bisulfite sequencing of the APOA1 promoter region demonstrated that CpG sites in this promoter region are fully hypermethylated in KhES3 and gradually become demethylated during in vitro differentiation (Supplementary Material, Fig. S8B). Demethylated regions are observed only in the vicinity of APOA1 TSSs at day 21 of differentiation, and spread over 1 kb beyond the APOA1 TSS in adult liver tissues. Sparse non-CpG methylation is observed in KhES3 and lost at day 21 of differentiation and also in adult liver tissues. This demethylation in non-CpG sites in KhES3 is also observed in the promoter region of CD6 in adult blood and of STMN4 in the adult brain (Supplementary Material, Fig. S9).

We then analyzed further the methylation status over the entire APOA1 gene locus to determine the extent of demethylation events (Fig. 6D). Demethylation starts from the vicinity of APOA1 TSSs at day 13 and extends to 200 bp around the TSS on day 21. Hypomethylated regions in human liver tissues spread over the APOA1 region, from TSSs to the CpG island of the 3’ end and the further downstream region, suggesting the correlation of extensive demethylation with the stable expression of specific gene sets and cell fate determination.

Epigenetic reprogramming using defined factors enables terminally differentiated cells to gain pluripotency (24). Re-expression of pluripotency genes associated with these promoters, which are methylated in differentiated somatic cells, is important for iPS cell generation (25). The heatmap shows that the four human primary fibroblast cell lines (IMR90, MRC-9, KMS-6 and TIG-103) share specific hypomethylation. After cellular reprogramming into iPS cells, the IMR90 cells show restoration of methylation in these tissues spread over the APOA1 region, from TSSs to the CpG island of the 3’ end and the further downstream region, suggesting this type of tDMRs might be a consequence of repression of the unnecessary genes (28). Although

**DISCUSSION**

In this study, we analyzed inclusive gene sets for tissue-specific hypomethylation and hypermethylation among human normal tissues. Of note, the former gene subsets are remarkably associated with cellular functions characterizing the tissue phenotypes. Although we have examined the limited sites of promoter regions, we reveal here that these hypomethylated genes display tissue-specific patterns of gene expression and specific enrichment of transcription factor recognition motifs in their promoters. This indicates the methylation changes in these regulatory regions might have functional roles in spatiotemporal transcriptional control. Furthermore, the hypomethylation panel showed an unexpected dense methylation pattern in pluripotent stem cells and regional hypomethylation in differentiated cells, suggesting this type of tDMRs might be a consequence of methylation erasure or a dilution process.

To date, the exploration of tDMRs was performed on the premise that stepwise addition of promoter methylation contributes to cell fate determination during early embryogenesis (26,27). It has been widely accepted that the genomic DNA of the embryo, which has pluripotency to differentiate into multiple lineages, is initially unmethylated and subsequent accumulations of hypermethylation in CpG island promoters are important for lineage restriction by reinforcing transcriptional repression of the unnecessary genes (28). Although
this concept was true for some validated examples, it cannot adequately explain the global control of gene expression. In fact, consistent with the previous studies (6,10), we observed that most CpG island promoters are invariably unmethylated among normal tissues. In contrast with tissue-specific hypermethylation in CpG island promoters, tissue-specific hypomethylation in CpG-poor promoters has been underestimated so far and is significantly associated with the tissue phenotype.

These observations raise a new question about the molecular mechanism of tissue-specific hypomethylation established during terminal differentiation. Promoter demethylation in the differentiated cells is an old concept (29,30), but it has been forgotten while mammalian DNA demethylase was yet to be discovered. Now, two types of mechanisms for DNA demethylation, namely active demethylation and passive demethylation, are widely accepted for mammals (31,32).

Figure 3. The gene expression level of tissue-specific differentially methylated genes. Shown box plots (from 25th percentile to the 75th percentile with heavy lines at the median) represent average gene expression levels (the log scale of the GeneChip score) of tissue-specific hypomethylated genes (A) and tissue-specific hypermethylated genes (B), for each tissue. The dotted lines extend above and below the box to show the first and ninth deciles. Black and white boxes below the bar graphs represent hypermethylation and hypomethylation of the given tissue, respectively.
Active demethylation is observed in the paternal genome of an embryo during the first few days (33,34). In this process, demethylation occurs globally except for the limited foci such as imprinting control regions and centromeric and pericentromeric heterochromatin (35). Although recent reports suggested the ten-eleven translocation (TET) family proteins,

Figure 4. Enrichment of transcription factor recognition motifs in the tissue-specific hypomethylated regions. Each row represents a cis-regulatory module family with significant over-representation relative to a random set of mammalian promoters (Z-score > 8.0). Each column represents a tissue type. Four tissues (oral mucosa, liver, blood and skeletal muscle) show some specific enrichment of their master regulators binding motifs, respectively.

Figure 5. Hierarchical clustering analysis of human somatic tissues and normal cells. The dendrogram in the upper panel was obtained on the basis of the representative gene sets of tissue-specific hypomethylation using average linkage correlation. Each row represents a CpG locus (250 tissue-specific hypomethylation for each) and each column represents a sample. The colored boxes above the dendrogram indicate the nature of the samples; human somatic tissues (blue), human ES cells (red), human iPS cells (orange) and human primary fibroblast (green). The color scale bar at the right side shows the percentage of the methylation level (0–100%).
TET1, TET2 and TET3, are candidate proteins responsible for the erasure process through an oxidative demethylation pathway (32,36), further investigations are needed. The unexpected dynamics of DNA methylation during cellular differentiation might give us an important clue to elucidate the mechanism of cell fate determination during embryogenesis.

An alternative explanation for the tissue-specific demethylation seen in CpG-poor promoters is passive demethylation, which is usually observed in asymmetric cell division or highly proliferating cells like cancer cells. Inhibiting maintenance of cytosine methylation of the template strand could result in dilution of methylation in differentiated daughter cells. According to this scenario, transcription factor-related inhibition of DNA methyltransferase at the timing of cell division might be necessary because the developmental hypomethylation we observed here occurs not in a genome-wide manner but in a regional manner. Indeed, the enrichment of transcription factor-binding motifs is seen at the demethylated regions in a tissue-specific manner. Recently, it was shown that mitotically retained transcription factors are associated with the asymmetric cell division in some contexts (37,38). If sustained binding of transcription factors inhibits propagation of DNA methylation into the newly synthesized strand, transcription factor-driven demethylation will be inherited in proliferating cells. In our study, we examined in vitro differentiation in a series of promoters and found that a wave of demethylation develops from the TSS of \( APOA1 \) and \( ITIH3 \) promoters. Once the binding of transcription factors at demethylated regions induces gene expression in the tissue progenitor cells, sustained induction in response to appropriate extrinsic stimuli may result in loss of propagation of DNA methylation marks in the promoter regions for...
Figure 6. (Continued).
long-lasting maintenance of a transcriptionally active state. Subsequently, in this model, chromatin conformation changes in terminally differentiated cells would expand the demethylated regions and contribute to the establishment of stable and highly efficient expression of specific gene subsets.

Growing evidence suggests that forced induction of master regulator genes has the potential to change the fate of lineage-restricted cells, even in terminally differentiated cells (39–41). We identified restoration of methylation during reprogramming into iPS cells. The feasibility of cell reprogramming suggests that differentiated cells still have much more plasticity in the epigenetic status including DNA methylation than we had expected. Further analysis of methylation changes might provide novel insight into mechanisms that will generate a transcriptional repertoire for variable cell lineages and give us useful clues to control cell fate fixation, which might be applicable for regenerative medicine.

MATERIALS AND METHODS

Genomic DNA from human normal tissues

Frozen tissues of the brain, lung, liver and kidney were obtained from surgical specimens. Patients undergoing surgical resection at the Tokyo University General Hospital provided tissue after obtaining informed consent. Buccal swabs of oral mucosa, peripheral blood and placental tissue were from healthy volunteers. This study was certified by the Ethics Committee of Tokyo University. Genomic DNA from these clinical samples was extracted using the QIAamp DNA Mini Kit (QIAGEN). Genomic DNA of further individuals was purchased from Bio-Chain (details are listed in Supplementary Material, Table S1). For the methylation-negative control, totally unmethylated genomic DNA was synthesized by a whole-genome amplification system, GenomiPhi (GE healthcare). For a positive control, fully methylated genomic DNA was generated by SssI CpG methylase (New England Biolabs) treatment of lymphocyte DNA.

Human ES cell lines

Human ES cell lines, KhES1, KhES2, KhES3, KhES4, were established and maintained as described previously (42). Human ES cell lines (H1, H9) and human iPS cell lines [iPS(IMR90)-1 and iPS(IMR90)-4] were obtained from WiCell Research Institute. HES3 cell line was obtained from ES Cell International.

Briefly, undifferentiated human ES cells were maintained on a feeder layer of MEF in DMEM/F12 (Sigma) supplemented with 20% KSR, L-Glu, NEAA and β-ME under 3% CO2. To passage ES cells, ES cell colonies were detached from the feeder layer by treatment with 0.25% trypsin and 0.1 mg/ml of collagenase IV in PBS containing 20% KSR and 1 mm of CaCl2 at 37°C for 5 min, followed by the addition of culture medium. ES cell clumps were disaggregated into smaller pieces by gentle pipetting.

An in vitro differentiation experiment was performed following the reported method, with some modification (43). Briefly, KhES3 cells were cultured in differentiation medium [RPMI supplemented with human recombinant activin A (100 ng/ml) and defined FBS]. FBS concentrations were 0% for the first 24 h, 0.2% for the second 48 h and 2.0% for subsequent days of differentiation. Media were replaced every 2 days with fresh differentiation medium supplemented with growth factors. ES cells were cultured in differentiation medium (DMEM supplemented with 10% KSR, Dex and HGF) for up to 30 days.

Methylation profiling

Methylation status was analyzed using HumanMethylation27 BeadChip (Illumina). Genomic DNA for methylation profiling was quantified using the Quant-iT dsDNA BR Assay Kit (Invitrogen). Five hundred nanograms of genomic DNA was bisulfite-converted using an EZ DNA Methylation Kit (Zymo Research). The converted DNA was amplified, fragmented and hybridized to a BeadChip according to the manufacturer’s instructions. The raw signal intensity for both methylated (M) and unmethylated (U) DNA was measured using a BeadArray scanner (Illumina). The methylation level of the each individual CpG is obtained using the formula (M)/(M)+U+100 by the GenomeStudio (Illumina).

Quantitative methylation analysis using the MassARRAY system

Bisulfite treatment of genomic DNA was performed using an EZ Methylation Kit (Zymo Research). Primer sequences are given in Supplementary Material, Table S4. This system utilizes MALDI-TOF mass spectrometry in combination with RNA base-specific cleavage (MassCLEAVE). A detectable pattern is analyzed for the methylation status. Mass spectra were acquired using a MassARRAY Compact MALDI-TOF (Sequenom) and spectra’s methylation ratios were generated using Epityper software v1.0 (Sequenom).

Bisulfite sequencing

Bisulfite sequencing analysis was performed as described previously (44). Bisulfite treatment of genomic DNA was performed using an EZ Methylation Kit (Zymo Research). All primer sequences and melting temperatures for the polymerase chain reaction (PCR) are given in Supplementary Material, Table S4. PCR amplicons were subcloned into the pGEM-T vector (Promega). Clones were sequenced using PRISM3100 Sequencer (Applied Biosystems).

RNA extraction and gene expression microarray analysis

Genome-wide analysis of mRNA expression levels using U133plus2.0 human expression array® (Affymetrix) was done essentially as described previously (45). Briefly, total RNA was isolated using TRIzol reagent (Invitrogen), according to the manufacturer’s instructions. One microgram of RNA was used for the generation of double-stranded cDNA with the SuperScript Double-Stranded cDNA Synthesis Kit (Invitrogen) according to the manufacturer’s protocol. Double-stranded cDNAs were hybridized to the microarray.
Reverse transcription–polymerase chain reaction analysis

RNA extraction and reverse transcription–polymerase chain reaction (RT–PCR) were done as described (46). Total RNA was extracted using TRI Reagent (Sigma-Aldrich) or the RNeasy micro-kit (Qiagen) and then treated with DNase (Sigma-Aldrich). Three micrograms of RNA was reverse-transcribed using Moloney Murine Leukemia Virus reverse transcriptase (Toyobo, Japan) and oligo(dT) primers (Toyobo). The primer sequences are shown in Supplementary Material, Table S4. The PCR conditions for each cycle were as follows: denaturation at 96°C for 30 s, annealing at 60°C for 2 s and extension at 72°C for 45 s. RT–PCR products were separated by 5% non-denaturing polyacrylamide gel electrophoresis, stained with SYBR Green I (Molecular Probes), and visualized using a Gel Logic 200 Imaging System (Kodak).

Definition of probe classes and promoter classes

We classified 27,578 probes into three categories: HCG, ICG and LCG. Each probe position was defined with respect to the position of a given CpG site. We determined the GC content and the ratio of observed versus expected CpG dinucleotides in a surrounding 500 bp window. The CpG ratio was calculated using the following formula: (number of CpGs × number of bp) (number of Cs × number of Gs). Three categories of probes were determined as follows: (i) HCGs (8098 probes) covering a 500 bp area with a CpG ratio above 0.75 and GC content above 55%; (ii) LCGs (8374 probes) excluded from a 500 bp area with a CpG ratio above 0.48; and (iii) ICGs (11,106 probes) that could not be categorized as either HCGs or LCGs.

Clustering analysis

To analyze the similarity of the methylation levels among human somatic tissues, ES cells and iPS cells, we used the data set of tissue-specific hypomethylation selected in Figure 2A for the cluster analysis. We applied a hierarchical clustering algorithm using the uncentered correlation coefficient as the measure of similarity and average linkage clustering (47) and visualized the dendrogram and the heatmap using TreeView (48).

GO functional annotation analysis

GO functional annotations for differentially hypomethylated and hypermethylated gene sets were performed using the Database for Annotation, Visualization and Integrated Discovery (DAVID) Bioinformatic Resources v6.7 (http://niaid.abbcc.ncifcrf.gov/home.jsp). The lists of 250 gene symbols that show specific hypermethylation or hypomethylation for each tissue were submitted and DAVID default population background (Homo sapiens) was chosen to detect significantly over-represented GO biological processes (GOTERM BP-FAT). P-values were calculated by a modified Fisher’s exact test and adjusted for multiple hypotheses testing using Bonferroni correction. The three GO terms with the most significant P-value and the number of genes involved in the term were listed for each tissue.

Enrichment analysis of transcription factor-binding motifs

To determine over-represented transcription factor-binding sites in tissue-specific hypomethylated and hypermethylated regions, sequences around the probe within a 500 bp window were screened for the presence of binding sites using Genomatix RegionMiner (http://www.genomatix.de, matrix library version 7.1). The number of binding site motifs was determined and over-representation over the background of random mammalian promoter sequences was calculated as the Z-score. Transcription factor families with a Z-score greater than 8.0 were considered highly significant. The Z-scores of these representative TF modules are visualized in the heatmap using TreeView (48).

SUPPLEMENTARY MATERIAL

Supplementary Material is available at HMG online.

ACKNOWLEDGEMENTS

We are grateful to Hiroko Meguro for microarray experiment, Kaoru Nakano for MassARRAY analysis, Elodie Lebre donchel for bisulfite sequencing experiment and Michael Jones for critical reading of the manuscript.

Conflict of Interest statement. None declared.

FUNDING

This work was mainly supported by a Grant-in-Aid for Scientific Research (S) 20221009 (H.A.) from the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan, and the Program of Fundamental Studies in Health Sciences of the National Institute of Biomedical Innovation (NIBIO), Japan.

REFERENCES


