Epigenetic modification of retinoic acid-treated human embryonic stem cells

Hyun Sub Cheong1, Han Chul Lee2, Byung Lae Park1, Hyemin Kim3, Mi Jin Jang3, Yong Mahn Han3, Seun-Young Kim2, Yong Sung Kim2,* & Hyoung Doo Shin1,4,*

1Department of Genetic Epidemiology, SNP Genetics, Inc., Seoul 153-803, 2Medical Genomics Research Center, Korea Research Institute of Bioscience and Biotechnology, Daejeon 305-806, 3Department of Biological Sciences and Center for Stem Cell Differentiation, Korean Advanced Institute of Science and Technology, Daejeon 305-701, 4Department of Life Science, Sogang University, Seoul 121-742, Korea

INTRODUCTION

Human embryonic stem cells (hESCs) are unique in their ability to maintain pluripotence. This property makes hESCs leading candidates for use in cell therapy and in studies on early human development. Retinoic acid (RA), the most potent natural form of vitamin A, plays an important role in mediating the growth and differentiation of both normal and transformed cells (1, 2). It is essential for many diverse biological functions including growth, vision, reproduction, embryonic development, differentiation of epithelial tissues, and immune responses (2). In vitro, RA induces differentiation of hESCs into a number of specific cell types.

Differentiation of a specific cell type involves the establishment of a precise epigenetic profile composed of genome-wide epigenetic modifications such as DNA methylation and histone modification. Since epigenetic modifications in gene areas regulate transcriptional activity, the epigenetic profile of the cell reflects the transcriptome, at least partially (3-5).

hESCs have been investigated using multiple techniques, including gene expression profiling, mitochondrial sequencing, immunocytochemistry, genotyping, functional assays, and DNA methylation assay (6-10). DNA methylation of the genome is the key to maintaining the differentiated state of hESCs (11, 12), and it must be reset during differentiation by RA treatment.

Differences between hESC lines with respect to gene expression profiles have been investigated before (13), and it has also been demonstrated that hESCs have unique DNA methylation profiles compared to other cell types, including embryonic germ cells, trophoblast stem cells, and several adult stem cell populations (8, 14). Key regulators of development such as Oct4 and NANOG are also controlled by epigenetic mechanisms (15, 16). However, a whole-genomic correlation study on DNA methylation and gene expression has not been reported.

The present study utilized DNA methylation and gene expression assays to generate whole-genomic methylation and gene expression profiles for both undifferentiated hESCs and RA-treated hESCs. These results provide valuable information that can be used to identify differentially methylated CpG sites and differentially expressed genes.

RESULTS

We applied a comprehensive DNA methylation profiling approach to assess the epigenetic states of three hESC lines (CHA3-hES, CHA4-hES, and SNUhiES3) as well as their epigenetic modifications after RA treatment. A whole-genome DNA methylation assay method was used to analyze the methylation status of 27,578 CpG sites selected from more than

*Corresponding authors. Hyoung Doo Shin, Tel: 82-2-705-8615; Fax: 82-2-2026-4299; E-mail: hds@SORGANG.ac.kr; Yong Sung Kim, Tel: 82-42-879-8110; Fax: 82-42-879-8119; E-mail: yongsung@kribb.re.kr
DOI 10.5483/BMBRep.2010.43.12.830
Received 18 June 2010, Accepted 25 October 2010

Keywords: DNA methylation, Epigenetic modification, Gene expression, Human embryonic stem cell, Retinoic acid
14,000 well-annotated genes. We measured the overall methylation levels after RA treatment. We found that the average methylation level in the RA-treated hESCs (29.5%) was greater than in the undifferentiated hESCs (27.1%). The lower methylation level we obtained for the undifferentiated hESCs was expected since global hypomethylation has been reported often in embryonic stem cells (17).

To discover which CpG sites contribute the most to the epigenetic modification of hESCs by RA treatment, we compared the DNA methylation patterns between undifferentiated hESCs and RA-treated hESCs. This analysis produced a list of 166 CpG sites from 151 genes that significantly contribute to the separation of the two groups. Among them, the top 100 CpG sites, based on $|\Delta \beta|$, are shown in Fig. 1A. We then clustered all of the samples based on their relative methylation levels at these 166 CpG sites (data not shown). Three hESC lines were correctly aggregated into the two other major clusters, which comprise undifferentiated hESC lines and RA-treated cells.

We also investigated three hESC lines for differential expression of genes upon RA treatment. A total of 9,736 distinct genes (23% of the RefList) passed the expression criteria of a Detection Score $\geq 0.99$. Among them, 2,013 genes were differentially expressed. We observed that 1,003 genes were up-regulated (>1.5-fold) and 1,010 genes were downregulated (<0.66-fold). The extreme differences observed between undifferentiated hESCs and RA-treated hESCs are shown in Fig. 1B.

Combined analysis of methylation and expression data revealed that 19 genes (STAP2, VAMP8, C10orf26, WFIKKN1, ELF3, C10Tnf6, C10orf10, MRCPRF, ARSE, LSAMP, CENTD3, LDB2, POUSF1, GSPT2, THY1, ZNF574, MSX1, SCMH1, and RARB) were highly correlated with each other (Pearson correlation coefficient $\geq 0.8$) (Supplementary Table 1).

To validate the methylation status of the highly correlated genes, we selected two genes (CENTD3 and MSX1) and performed bisulfate sequencing. Bisulfate sequencing of 400-500 bp including Illumina probe position revealed hypermethylation (36.7% and 19.6%) in SNUhES3 cells after RA treatment that was consistent with the genome-wide DNA methylation (Supplementary Fig. 1).

**DISCUSSION**

Despite their differences in origin, different sample preparation methods, and karyotypes, three hESC lines were correctly aggregated into two other major clusters, which comprise undifferentiated hESC lines and RA-treated cells. This suggests that the three hESC lines share a common epigenetic signature, which is likely linked to embryonic stem (ES) cell-specific properties such as self-renewal and pluripotency.

CpGs on the C10orf10, FAM12B, VAMP8, CLDN15, and FLJ20273 genes were the most hypomethylated, whereas C7orf29, CHFR, GSPT2, HDCMA18P, and MSX1 were the most hypermethylated after RA treatment. These genes, methylation of CHFR is known to be associated with silencing of CHFR expression in various types of cancer (18), and CHFR is also known as a tumor suppressor (19). This means that cells in which CHFR was epigenetically inactivated constituted differentiated hESCs.

In order to define the relationship between methylation and expression of genes, we performed gene expression profiling to compare both methylation status and gene expression levels. Among the differentially methylated genes, HOXB5, INS-IGF2, HOX5, LCP1, and ANKRD38 were the most highly upregulated, whereas PRDM14, ZIC2, C9orf135, MIAT, and SFRP2 were downregulated after RA treatment. Among these genes, HOX5 was found to be rapidly induced within mouse ES cells as a result of RA treatment (20). In addition, knock-