

# Abnormal CpG island methylation occurs during *in vitro* differentiation of human embryonic stem cells

Yin Shen<sup>1</sup>, Janet Chow<sup>1</sup>, Zunde Wang<sup>2,†</sup> and Guoping Fan<sup>1,\*</sup>

<sup>1</sup>Department of Human Genetics, Institute of Stem Cell Biology and Medicine, David Geffen School of Medicine, UCLA, 695 Charles Young Drive South, Los Angeles, CA 90095, USA and <sup>2</sup>EpiGenX Pharmaceuticals Inc., 5385 Hollister Avenue, Santa Barbara, CA 93111, USA

Received June 7, 2006; Revised and Accepted July 21, 2006

Directed differentiation of human embryonic stem cells (hESCs) into specific somatic cells holds great promise for cell replacement therapies. However, it is unclear if *in vitro* hESC differentiation causes any epigenetic abnormality such as hypermethylation of CpG islands. Using a differential methylation hybridization method, we identified 65 CpG islands (out of 4608 CpG islands or 1.4%) that exhibited increased DNA methylation during the conversion of hESCs into neural progenitor/stem cells (NPCs). These methylated CpG islands belong to genes in cell metabolism, signal transduction and cell differentiation, which are distinctively different from oncogenic CpG island hypermethylation observed in cancer-related genes during tumorigenesis. We further determined that methylation in these CpG islands, which is probably triggered by *de novo* DNA methyltransferase Dnmt3a, is abnormally higher in hESC–NPCs than in primary NPCs and astrocytes. Correlating with hypermethylation in promoter CpG islands of metabolic enzyme gene *CPT1A* and axoneme apparatus gene *SPAG6*, levels of *CPT1A* and *SPAG6* mRNAs are significantly reduced in hESC–NPCs when compared with hESCs or primary neural cells. Because *CPT1A* is involved in lipid metabolism and *CPT1A* deficiency in human is associated with the hypoketotic hypoglycemia disorder, the reduced *CPT1A* expression in hESC–NPCs raises a potential concern for the suitability of these cells in cell transplantation. Collectively, our data show that abnormal CpG island methylation takes place in a subset of genes during the differentiation/expansion of hESC derivatives under current culture conditions, which may need to be monitored and corrected in future cell transplantation studies.

## INTRODUCTION

DNA cytosine methylation is one of the major epigenetic factors in vertebrate animals that is involved in gene regulation, genomic imprinting, X-chromosome inactivation and genome stability (1). During embryonic development, DNA methylation is established by the coordinated actions of a family of DNA (cytosine-5) methyltransferases (Dnmts) including the maintenance enzyme Dnmt1 and *de novo* DNA methyltransferases Dnmt3a and Dnmt3b (2). DNA methylation is required for animal development because mutant mice lacking either the maintenance enzyme

(Dnmt1) or both of the *de novo* DNA methyltransferases (Dnmt3a/3b) exhibit significant demethylation in the genome and die at embryonic day (E) 8–10 just after gastrulation (3,4). Alterations in DNA methylation have been associated with cancer and human genetic disorders (5). For example, DNMT3B deficiency in humans leads to significant demethylation in centromeric minor satellite repeats, and patients suffer a rare genetic disorder, namely ICF syndrome (Immuno-deficiency, Centromere instability and Facial anomalies) (4,6). DNA methylation deficiency in mouse neural precursor/stem cells induces a precocious astroglial phenotype (7,8), suggesting that proper control of DNA methylation is essential

\*To whom correspondence should be addressed. Tel: +1 3102670439; Fax: +1 3107945446; Email: gfan@mednet.ucla.edu

†Present address: Division of Biology, Beckman Research Institute, City of Hope, Duarte, CA 91010, USA.

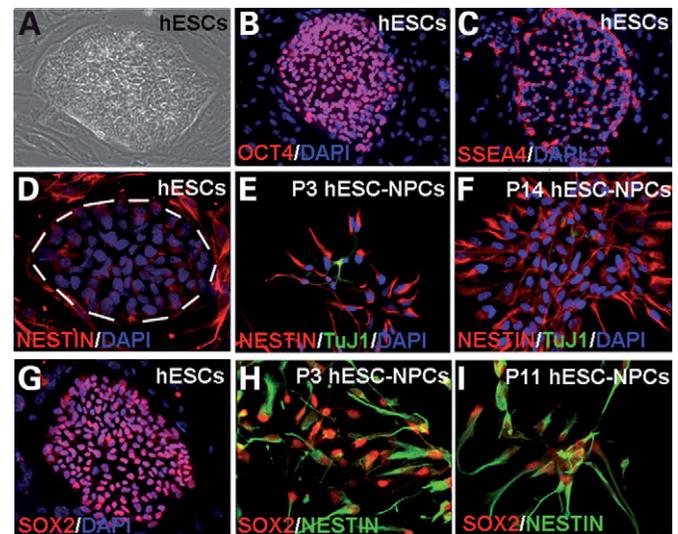
for the cell differentiation program. In somatic cell nuclear transfer experiments, reprogramming the DNA methylation pattern is one of the essential epigenetic changes and its faulty reprogramming in cloned embryos is postulated to be one of the leading causes for the very low efficiency of animal cloning (9).

Pluripotent human embryonic stem cells (hESCs) have the capacity to differentiate into multi-lineage cell types and the potential to be a major source of donor cells for cell transplantation and tissue repair (10,11). For example, hESC-derived oligodendrocyte progenitor cells have been shown to remyelinate and restore locomotion after spinal cord injury in rodents (12). However, recent work suggests that genetic lesions and epigenetic instability including abnormal DNA methylation in cancer-related genes occur during long-term passage of hESCs (13). This raises concerns about whether *in vitro* differentiation of hESCs may also lead to abnormalities in DNA methylation patterns in comparison with normal somatic cells. In this study, we used a DNA microarray-based method to first assess if changes in DNA methylation occur in a library of 4608 CpG islands during the conversion of undifferentiated hESCs into neural progenitor/stem cells (NPCs). The majority of CpG islands are associated with the gene promoter/first exon and are normally protected from DNA methylation during embryogenesis (14–16). Our results indicate that only a small fraction of CpG islands are subject to *de novo* DNA methylation upon hESC differentiation. However, this subset of CpG islands is distinctively different from CpG islands methylated in cancer cells (17). Furthermore, levels of DNA methylation in these selected CpG islands in hESC-derived NPCs (hESC–NPCs) become abnormally high when compared with levels in primary human cells, leading to the silencing of these genes in hESC–NPCs.

## RESULTS

### *In vitro* neural differentiation of hESCs

We have adopted a culture protocol for directed neural differentiation of hESCs, which has been successfully used to derive homogenous populations of NPCs from mouse and human ESCs (18,19) (Fig. 1A–I, see Materials and Methods). Through consecutive passages in serum-free DMEM/F12 medium with B27 supplement and bFGF treatment, homogenous populations of NPCs can be obtained from the third passage of NPCs (P3). Using antibodies against neuronal marker  $\beta$ III TUBULIN (TuJ1 antigen), astroglial marker glial fibrillary acidic protein (GFAP) and neural stem cell marker NESTIN (7), we found that P3 NPC cultures contained over 95% NESTIN- and SOX2-positive NPCs and a small number of spontaneously differentiated  $\beta$ III TUBULIN-positive neurons (Fig. 1E). In late passage, hESC-derived NPCs ( $\geq$ P8), while virtually all cells are NESTIN- and SOX2-positive NPCs (Fig. 1F and I), a small number of cells (5–10%) are also GFAP-positive (data not shown), suggesting the presence of a small percentage of glial progenitors in these late passage hESC–NPCs. Upon withdrawal of bFGF treatment and treatment of neurotrophins, ~50% of P3 hESC–NPCs differentiate into post-mitotic neurons within a



**Figure 1.** Characterization of hESCs and hESC-derived NPCs. (A–C) Undifferentiated hESCs in phase-contrast microscopic graph (A), or stained with antibodies against OCT4 (B) and SSEA-4 (C). (D–F) show immunostaining with NESTIN (NPC marker in red) and TuJ1 (neuronal marker in green) antibodies of hESCs (D) and P3 (E) and P14 (F) hESC–NPCs. The dotted line in (D) outlines a hESC colony that exhibits moderate NESTIN staining signals when compared with the strong staining in hESC–NPCs (E and F). Mouse MEF feeder cells also have a background staining of NESTIN, which is much weaker than hESC–NPCs. (G–I) SOX2 staining of undifferentiated hESCs and hESC–NPCs. (H) and (I) also show colocalization of SOX2 and NESTIN in hESC–NPCs. DAPI counterstaining (in blue) shows nuclear morphology.

week. In contrast, most late passage hESC–NPCs are prone to differentiate into mixed population of neurons and astrocytes (Supplementary Material, Fig. S1), suggesting a switch of neurogenic to gliogenic competence over the time (20). The hESC–NPCs after long-term passaging ( $>$ P20) maintain a normal karyotype as assayed by standard G-band staining of metaphase chromosome spreads (data not shown).

### Selective increase in CpG island methylation

We have used a modified method of differential DNA methylation hybridization (DMH) (21) to identify potential methylation changes in genomic DNA harvested from hESCs and P3 and P14 hESC–NPCs. Genomic DNA was cut with methyl-sensitive *Sma*I restriction enzyme at the site CCC<sup>↓</sup>GGG that is preferentially localized within CpG islands (22) and ligated to blunt adaptors for ligation-mediated PCR amplification. The amplified DNA fragments from hESCs, and P3 and P14 hESC–NPCs were then labeled with either Cy3 or Cy5 fluorescent dye and cross-hybridized to a custom-made DNA microarray containing 4608 CpG island fragments to identify differentially hybridized spots that resulted from different levels of DNA methylation. This CpG island library was constructed with presumably unmethylated CpG islands purified from *Sma*I-cut DNA fragments of adult human leukocyte DNA and a number of known CpG islands that become hypermethylated in cancer cells (Z.W. *et al.* in preparation, see Materials and Methods). Our first analysis of the microarray data show that there are two distinct

groups of hybridized spots on each CpG island microarray among three pairs of comparisons (hESCs versus P3 hESC–NPCs, hESCs versus P14 hESC–NPCs and P3 hESC–NPCs versus P14 hESC–NPCs). In the first group of hybridized spots, the signal intensity is biased towards probes from DNA samples of hESCs and P3 hESC–NPCs when compared with P14 hESC–NPCs, suggesting that these CpG island clones could be genomic regions with increased DNA methylation in P14 hESC–NPCs. The second group of hybridized spots, representing the majority of hybridized signals with no significant difference or biased hybridization towards P14 hESC–NPCs. We experimentally checked methylation status of several clones in this second group and confirmed that these CpG clones did not exhibit changes in DNA methylation between hESCs and P14 hESC–NPCs (data not shown).

Among the first group of hybridized clones that exhibited stronger hybridized signals in hESCs samples (thus suggesting increased DNA methylation in differentiated hESC–NPCs), we first set up a 1.5-fold cut-off (hESC–NPCs/hESCs ratio  $<0.667$ ) and identified 45 candidate clones. By sampling six microarray data sets between hESCs and P14 hESC–NPCs, or P3 and P14 hESC–NPCs, we further identified additional 20 clones that consistently showed a ratio  $<0.90$  in all the arrays. By combining these two sets of clones together, we estimated that 1.4% of clones (or 65 out of 4608 probe sets on the CpG island microarray) that presumably have higher levels of DNA methylation in P14 hESC–NPCs than in undifferentiated hESCs or early passage P3 hESC–NPCs. It is worth noting that our current result is solely based on the *SmaI* site-based DMH method, which cannot detect increased methylation in CpG islands lacking the *SmaI* restriction enzyme site (CCCGGG). Thus, our result could be potentially an under-estimate of CpG island methylation across the genome during hESC differentiation.

To identify the nature of these genomic clones that exhibited increased DNA methylation during neural differentiation of hESCs and passages of hESC–NPCs, we randomly sequenced 56 clones and blasted them against the human genome sequence database. As annotated in Table 1, we found that 91.1% of clones (51/56) are located in designated CpG island regions (designated as GC%  $>50\%$ ,  $>200$  bp and the ratio of observed/expected CpG  $>0.6$ ) (23), confirming the high quality of this CpG island library. Among these CpG island clones,  $\sim 51\%$  (26/51) are localized at promoter regions, 25% in exon and/or intron regions and the remaining 24% in intergenic regions or without any hit against the human sequence database (UCSC hg17 freeze May 2004). Gene ontology analysis of these annotated CpG islands (Table 1) suggests that these genes are involved in cell metabolism, signal transduction and cell differentiation and development. Among them are metabolic enzymes (e.g. carnitine palmitoyl-transferase 1A or CPT1A), transcription regulators (e.g. zinc finger protein 451) and signaling molecules (e.g. BMP4 and brain adenylate cyclase 1), which could be important for the survival and differentiation of hESC–NPCs. To confirm the quality of DMH microarray results, we randomly picked 11 CpG islands and experimentally verified the increase in CpG island methylation in all these CpG islands in hESC–NPCs (Fig. 2 and Supplementary Material, Fig. S2). Bisulfite sequencing analysis (24) and COBRA assay (25) were used to

quantify levels of DNA methylation of selected CpG sites in three gene promoters [*CPT1A*, *ARMADILLO REPEAT CONTAINING 7* (*ARMC7*) and *SPERM-ASSOCIATED ANTIGEN 6 ISOFORM 1* (*SPAG6*)] and in two coding exon regions [*PROTOCOLADHERIN 17* (*PCDH17*) and *STATHMIN 3* (*STMN3*)] (Fig. 2A–D). The *CPT1A* gene actually contains two alternative gene promoters: the exon1a promoter contains a 1.2 kb CpG island that was identified in our DMH array, and the exon1b promoter contains a 2.5 kb CpG island that is  $\sim 1.5$  kb downstream of exon1a. We, therefore, analyzed if an increase in methylation in the CpG island occurs in both exon1a and 1b promoters (Fig. 2A) with the aim to see if methylation is regulated in a coordinated fashion on these two alternative *CPT1A* gene promoters. *SPAG6* and *ARMC7* gene promoters contain a single CpG island in the proximal gene promoter/first exon region (Fig. 2B and D). Our results confirmed that a significant increase in DNA methylation, albeit at various degrees, occurs in all these CpG islands during the conversion of hESCs into early passages of hESC–NPCs (e.g. P3). Moreover, all CpG islands reach the highest level of DNA methylation in late passage P14 hESC–NPCs, suggesting that levels of DNA methylation are further elevated during the prolonged passaging of hESC–NPCs.

#### Selective DNA hypermethylation in hESC–NPCs in comparison to primary human NPCs, astrocytes and leukocytes

To determine if methylation patterns of these CpG islands in hESC–NPCs (Table 1) are unique to hESC–NPCs or simply reflect neural cell-type specific methylation that could occur in primary neural cells during embryogenesis, we performed extensive bisulfite methylation analysis in four representative CpG islands within *CPT1A* and *SPAG6* gene promoters and the *PCDH17* exon region in hESC–NPCs and primary NPCs and astrocytes derived from human fetal brain (Fig. 2A–C). We also analyzed adult leukocytes which represent a completely different cell lineage. If CpG island methylation in hESC–NPCs belongs to cell-type specific methylation, we would expect to see a similar degree of CpG island methylation in hESC–NPCs when compared with primary NPCs and astrocytes. Overall, these select CpG islands remain unmethylated in cultured primary NPCs and astrocytes, as well as in fetal cortical tissue and adult leukocytes. Our results suggest that higher-than-normal levels of DNA methylation in these select CpG islands in hESC–NPCs represent an accumulation of aberrant DNA hypermethylation during the conversion of hESCs into NPCs and the expansion of hESC–NPCs in culture.

#### Hypermethylated CpG islands in hESC–NPCs are distinctively different from those in the promoter of cancer-related genes

To further determine whether hypermethylation in select gene loci is unique to hESC–NPCs or shares the same characteristics as hypermethylation found in cancer cells, we compared our list of hypermethylated loci to the list of known hypermethylated genes as reported for cancer cells

**Table 1.** Candidate CpG islands that exhibit increased DNA methylation during hESC differentiation and NPC passage

ID	Mean $\pm$ SD	Refseq/representative mRNA	Gene Name	Position	Location	Annotation of DNA sequences of CpG islands
<b>7E6</b>	<b>0.589</b>	<b>0.14</b>	<b>NM_001876</b>	<b>CPT1A</b>	<b>Promoter</b>	<b>Chr11</b> <b>Carnitine palmitoyltransferase 1A (liver)</b>
<b>4E10</b>	<b>0.731</b>	<b>0.43</b>	<b>NM_024585</b>	<b>ARMC7</b>	<b>Promoter</b>	<b>Chr17</b> <b>Armadillo repeat containing 7</b>
<b>33A11</b>	<b>0.667</b>	<b>0.06</b>	<b>NM_024585</b>	<b>ARMC7</b>	<b>Promoter</b>	<b>Chr17</b> <b>Armadillo repeat containing 7</b>
7H6	0.615	0.06	NM_002342	LTBR	Promoter	Chr12 Lymphotoxin beta receptor
11A12	0.472	0.05	NM_024821	FLJ22349	Promoter	Chr22 Hypothetical protein LOC79879
26B6	0.623	0.16	NM_006113	VAV3	Promoter	Chr1 vav 3, a Rho family guanine nucleotide exchange factor
<b>23D3</b>	<b>0.429</b>	<b>0.04</b>	<b>NM_021192</b>	<b>HOXD11</b>	<b>Promoter</b>	<b>Chr2</b> <b>Homeo box D11</b>
25E3	0.611	0.15	NM_002307	LGALS7	Promoter	Chr19 galectin 7
<b>26C7</b>	<b>0.469</b>	<b>0.12</b>	<b>NM_021116</b>	<b>ADCY1</b>	<b>Promoter</b>	<b>Chr7</b> <b>Brain adenylate cyclase 1</b>
33A10	0.661	0.07	NM_173680.2	MGC33584	Promoter	Chr7 Hypothetical protein LOC285971
<b>36C1</b>	<b>0.555</b>	<b>0.04</b>	<b>NM_024722</b>	<b>ACBD4</b>	<b>Promoter</b>	<b>Chr17</b> <b>Acyl-Coenzyme A binding domain containing 4</b>
<b>40C1</b>	<b>0.616</b>	<b>0.08</b>	<b>NM_001202</b>	<b>BMP4</b>	<b>Promoter</b>	<b>Chr14</b> <b>Bone morphogenetic protein 4 preproprotein</b>
27D10	0.804	0.23	NM_025079	MCPIP	Promoter	Chr1 Hypothetical protein FLJ23231
			NM_000095	COMP	Exon and intron	Chr19 Cartilage oligomeric matrix protein precursor
3E10	0.673	0.15	NM_032930	MGC13040	Promoter	Chr11 Hypothetical protein LOC85016
			NM_018176	LG12	Promoter (not CpG island)	Chr4 Leucine-rich repeat LGI family, member 2
18G7	0.786	0.13	NM_020644	C110rf15	Promoter (not CpG island)	Chr11 Hypothetical protein LOC56674
41E1	0.744	0.18	NM_030912	TRIM8	Promoter/1st exon	Chr10 Tripartite motif-containing 8
9E11	0.863	0.2		AK000246	Promoter/1st exon	Chr13 Homo sapiens cDNA FLJ20239 fis, clone COLF5934
<b>33B1</b>	<b>0.702</b>	<b>0.12</b>	<b>NM_015555</b>	<b>ZNF451</b>	<b>Promoter/1st exon</b>	<b>Chr6</b> <b>Zinc finger protein 451</b>
25B2	0.564	0.12	NM_130900	RAET1L	Promoter/1st exon (not CpG island)	Chr8 Retinoic acid early transcript 1L
CGI4b12	0.594	0.17		AK090480	Promoter/1st exon (not CpG island)	Chr15 Homo sapiens mRNA for FLJ00402 protein.
<b>43F9</b>	<b>0.654</b>	<b>0.1</b>	<b>NM_012443</b>	<b>SPAG6</b>	<b>Promoter/1st exon and 1st intron</b>	<b>Chr10</b> <b>Sperm associated antigen 6 isoform 1</b>
3A5	0.807	0.2			Promoter/1st exon and 1st intron	Chr4 KIAA0922
11D12	0.656	0.08	NM_133261	GIPC3	Promoter/1st exon and 1st intron	Chr19 PDZ domain protein GIPC3
24H3	0.732	0.15		BG753981	Promoter/1st exon and 1st intron	Chr14 Human EST from primary b-cells from tonsils (cell line)
33F8	0.445	0.08	NM_014071	NCOA6	Promoter/1st exon and 1st intron	Chr20 Nuclear receptor coactivator 6
33H6	0.766	0.17	NM_054020	CATSPER2	Promoter/1st exon and 1st intron	Chr15 Sperm-associated cation channel 2 isoform 1
<b>34A5</b>	<b>0.677</b>	<b>0.15</b>	<b>NM_001009812</b>	<b>LBX2</b>	<b>Promoter/1st exon and 1st intron</b>	<b>Chr2</b> <b>Ladybird homeobox homolog 2</b>
36A12	0.645	0.1	NM_152377	C1orf87	Promoter/1st exon and 1st intron	Chr1 Hypothetical protein LOC127795
38D2	0.787	0.2	NM_014604	TAX1BP3	Promoter/1st exon and 1st intron	Chr17 Tax1 (human T-cell leukemia virus type 1)
			Chr2: 236,864,253-236,864,502	Intergenic	Chr2	CpG island
<b>34A4</b>	<b>0.634</b>	<b>0.09</b>	<b>NM_014459</b>	<b>PCDH17</b>	<b>Exon</b>	<b>Chr13</b> <b>Protocadherin 17</b>
34D6	0.527	0.07		AK090480	Exon	Chr15 Homo sapiens mRNA for FLJ00402 protein.
44H4	0.659	0.14		AK055250	Exon	Chr11 Homo sapiens cDNA FLJ30688 fis, clone FCBBF2000473
<b>6F5</b>	<b>0.532</b>	<b>0.13</b>	<b>NM_015894</b>	<b>STMN3</b>	<b>Exon and intron</b>	<b>Chr20</b> <b>Stathmin 3 (SCG10-like-protein)</b>
2A12	0.614	0.11	AY356349	CYP26C1	Exon and intron	Chr10 Cytochrome P450, family 26, subfamily C, polypeptide 1
14G3	0.665	0.12	NM_030572	MGC10946	Exon and intron	Chr12 Hypothetical protein LOC80763
27B5	0.456	0.1	NM_021948	BCAN	Exon and intron	Chr1 Brevican isoform 1
29H7	0.579	0.08	BC007360	BC007360	Exon and intron	ChrX MGC16121 protein
32G10	0.574	0.19	NM_001012334	MDK	Exon and intron	Chr11 Midkine
16F6	0.591	0.07	NM_006480	RGS14	Intron	Chr5 <i>RGS14</i> gene for regulator of G protein signalling 14
19D9	0.673	0.23	AF410154	MHC2TA	Intron	Chr16 MHC class II transactivator
21C3	0.436	0.14	NM_016307	PRRX2	Intron	Chr9 Paired related homeobox 2
22H12	0.319	0.08	NM_002398	MEIS1	Intron	Chr2 Meis1 homolog
13A11	0.856	0.22	Chr2: 45,088,398-45,088,572	Intergenic	Chr2	CpG island
17A5	0.687	0.13	Chr6: 21,772,601-21,772,897	Intergenic	Chr6	CpG island

23H1	0.643	0.07	Chr18: 11,937,736–11,937,908	Intergenic	Chr18	CpG island
25H5	0.796	0.21	Chr18: 11,937,591–11,937,911	Intergenic	Chr18	CpG island
30A7	0.621	0.15	Chr2: 66,720,759–66,720,954	Intergenic	Chr2	CpG island
31C9	0.657	0.11	Chr2: 66,720,854–66,720,952	Intergenic	Chr2	CpG island
32H3	0.446	0.12	Chr17: 43,959,366–43,959,604	Intergenic	Chr17	CpG island
36E6	0.551	0.07	Chr14: 99,750,365–99,750,723	Intergenic	Chr14	CpG island
37H5	0.677	0.07	Chr2: 66,720,759–66,720,954	Intergenic	Chr2	CpG island
40A4	0.546	0.15	Chr17: 43,959,366–43,959,604	Intergenic	Chr17	CpG island
44F2	0.688	0.1	Chr3: 129,809,598–129,809,918	Intergenic	Chr3	CpG island
11D11	0.546	0.11	Chr10: 73,758,730–73,759,224	Intergenic (not CpG island)	Chr10	CpG island
14H10	0.584	0.29	Chr12: 123,108,420–123,108,832	Intergenic (not CpG island)	Chr12	CpG island
8C1	0.526	0.16				No hit

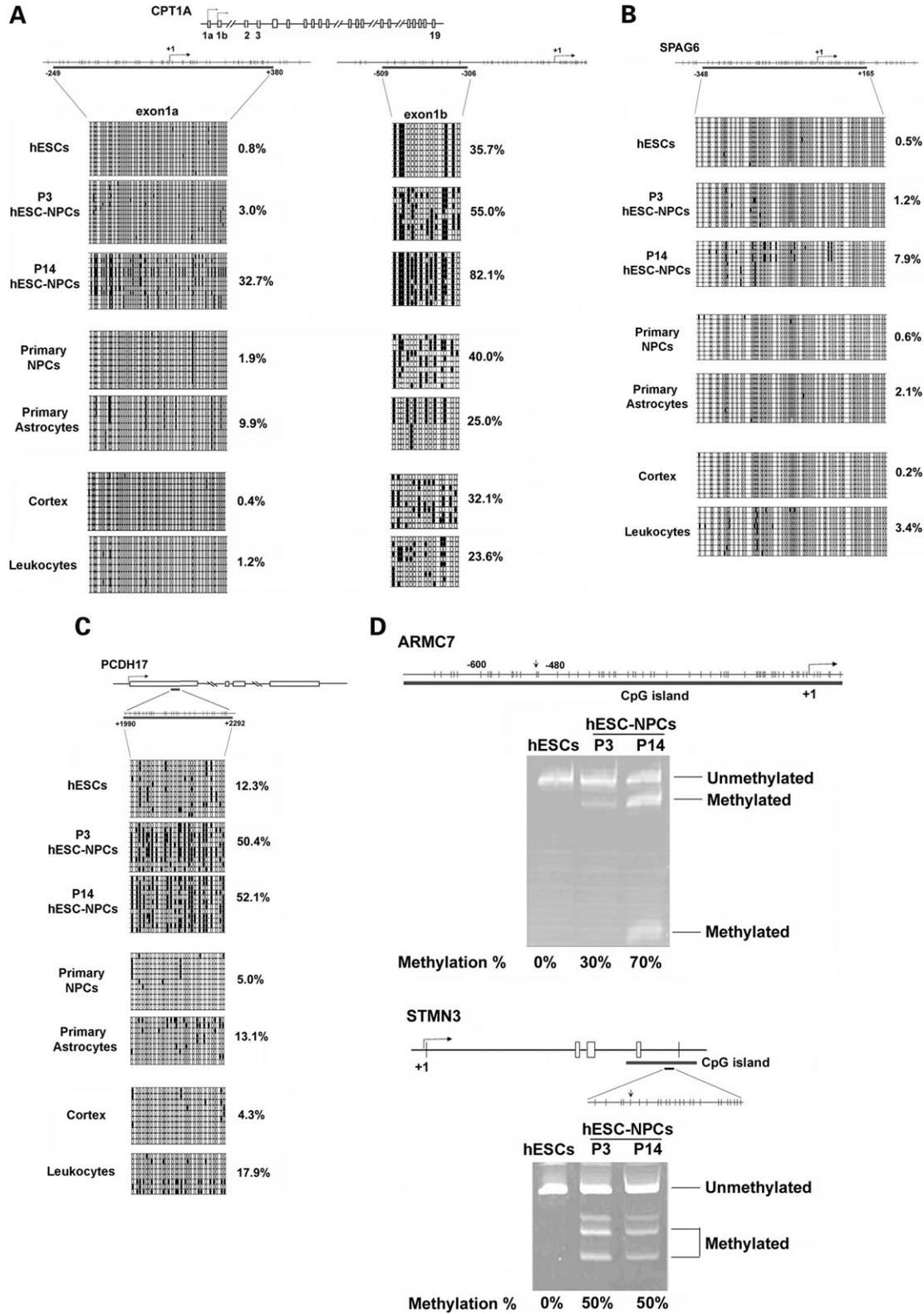
Clone ID was adopted from customary microarray. The ratio of hybridization signals was averaged from three to six datasets and expressed as mean  $\pm$  SD (hESC–NPCs probe intensity over hESCs probe intensity). Ref sequence, sequence annotation, and CpG island position in the gene structure were labeled according to UCSC Genome Browser on Human May 2004 Assembly. Clones in bold are those CpG islands that have been verified to exhibit increased methylation in P14 hESC–NPCs when compared with hESCs as shown in Figure 2 and online Supplementary Material Figure S2.

(17). Abnormal methylation in tumor suppressor genes or other cell-cycle control genes is one of the major causes underlying cancer formation (26,27). For example, a recent comprehensive genome-wide search for hypermethylated CpG islands in colon cancer cell line Caco-2 and prostate cancer cell line PC-3 as well as colon tumor samples has identified a total of 367 hypermethylated CpG islands (17). Among the 26 methylated CpG island promoters identified in our study (Table 1), there is no overlap with any of the 367 CpG island promoters reported for cancer cells (17). Our data suggests that the population of methylated CpG islands in hESC–NPCs is distinctively different from that found in cancer cells. We further confirmed the unmethylated status of five CpG islands in promoter regions of cancer-related genes (*VHL*, *THBS1*, *ERBIN*, *TMEFF2* and *P16*). As shown in Supplementary Material, Figure S3, CpG islands in these cancer-related genes (*VHL*, *THBS1*, *ERBIN*, *TMEFF2* and *P16*) do not exhibit any increase in DNA methylation in hESC–NPCs. Furthermore, we also found that undifferentiated hESCs maintain unmethylated CpG islands in either *CPT1A*, *SPAG6*, *PCDH17* genes, or these five cancer-related genes (*VHL*, *THBS1*, *ERBIN*, *TMEFF2* and *P16*) at moderate passage P54 and extremely late passage P170 (data not shown), suggesting that this undifferentiated hESC line (HSF6) is epigenetically stable at these gene loci in our culture conditions.

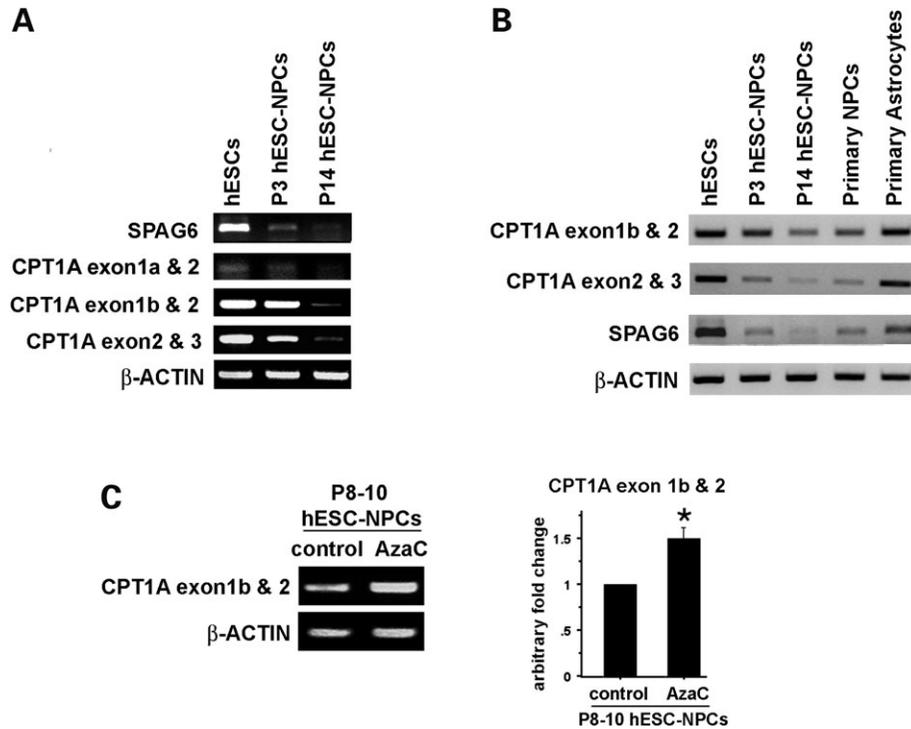
#### DNA hypermethylation and gene silencing

To understand the significance of hypermethylation of CpG islands on gene promoter activity during hESC differentiation, we performed semi-quantitative RT–PCR analysis of *CPT1A* and *SPAG6* transcripts. *CPT1A* is a mitochondrial protein involved in fatty acid metabolism, and mutation of this gene causes recurrent attacks of hypoketotic hypoglycemia (28,29). *SPAG6* gene is a component of the axoneme central apparatus in sperm cells and may also be involved in the motility of ependymal cilia in the nervous system (30,31). We found that *CPT1A* and *SPAG6* mRNA levels are high in hESCs, but significantly decreased or below detectable levels in P3 and P14 hESC–NPCs (Fig. 3A). This result is consistent with the notion that methylation of CpG islands in the promoter causes gene silencing. In contrast, when CpG methylation occurs in the coding exon of *PCDH17*, there is no correlation between the levels of CpG methylation and the level of *PCDH17* mRNA (data not shown).

Neural cells in the brain express both *CPT1A* and *CPT1C* isoforms (32). Indeed, we found that both *CPT1A* and *CPT1C* mRNA are detected in human fetal NPCs and astrocytes by RT–PCR (Fig. 3B and Supplementary Material, Fig. S4). In addition, the level of *CPT1C* isoform is significantly increased in hESC–NPCs, consistent with its abundant expression in differentiated neural lineage cells (32, Supplementary Material, Fig. S4). In contrast, the level of *CPT1A* in late passage P14 hESC–NPCs is significantly lower than in normal fetal NPCs and astrocytes (Fig. 3B), consistent with the possibility that DNA hypermethylation in *CPT1A* promoters could result in the imbalanced expression of *CPT1* isoforms in hESC–NPCs. To test whether the inhibition of *CPT1A* gene expression in late passage



**Figure 2.** Bisulfite sequencing analysis and COBRA assay of selective CpG island methylation in hESCs, hESC-NPCs, primary NPCs and astrocytes, and normal cortex and leukocytes. Methylation patterns in *CPT1A* exon1a, exon1b (A), *SPAG6* (B) gene promoters and *PCDH17* (C) exon CpG island. Each filled black dot represents one methylated CpG site and an open dot an unmethylated CpG. Schematic diagram of promoter and CpG sites in vertical bars are presented for each gene. Arrows indicate the transcription initiation site of human promoters examined. CpG sites in the regions analyzed by bisulfite genomic sequencing are underlined by black bars. (D) COBRA methylation assay of DNA methylation in *ARMC7* gene promoter and *STMN3* exon/intron region. The CpG site analyzed is indicated by arrow.



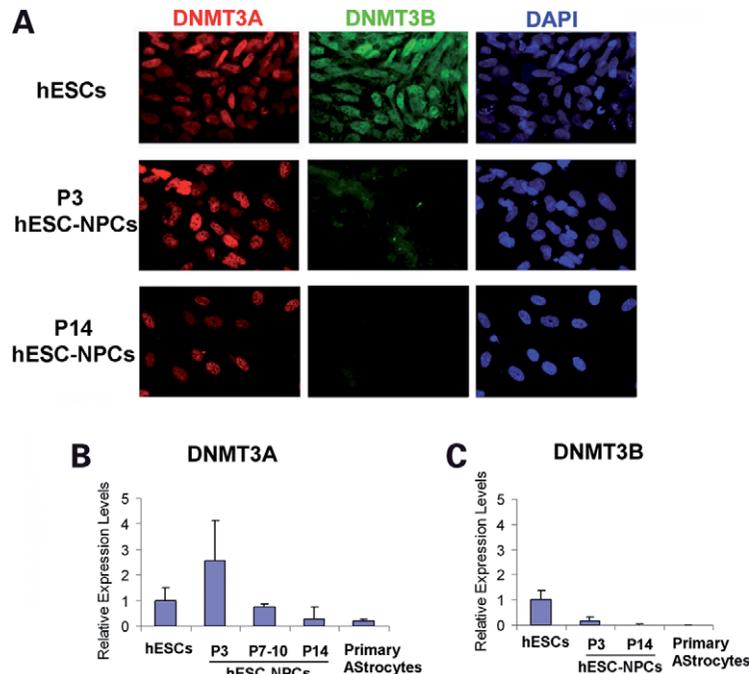
**Figure 3.** Semi-quantitative RT-PCR analysis of mRNA levels of *CPT1A* and *SPAG6* genes during hESC differentiation. (A) Down-regulation of mRNA transcripts from *CPT1A* exon1a and exon1b and *SPAG6* mRNAs upon hESC differentiation and prolonged passaging of hESC-NPCs. Exon-specific primers for *CPT1A* gene are labeled next to the gel picture. (B) Comparison of *CPT1A* and *SPAG6* mRNA levels among hESCs, hESC-NPCs, primary NPCs and astrocytes. (C) Treatment of AzaC (1  $\mu$ M) to late passages of hESC-NPCs (P8–10) for 6 days led to an increase in *CPT1A* gene expression. Ratio of densitometry reading of *CPT1A* transcripts over  $\beta$ -ACTIN is shown in bar graph to the right with *P*-value <0.01 (unpaired *t*-test).

hESC-NPCs can be reversed in culture, we treated late passage hESC-NPCs (P8–10) with the demethylating agent 5-azacytidine (AzaC) for 6 days in culture. AzaC treatment significantly increased mRNA levels of *CPT1A*, suggesting that DNA methylation-mediated gene silencing of *CPT1A* can be prevented in late passage hESC-NPCs by treating cultures with a DNA methyltransferase inhibitor such as AzaC (Fig. 3C).

#### Potential involvement of Dnmt3a with *de novo* methylation in CpG islands

We next examined which DNA methyltransferases carry out *de novo* DNA methylation activity in CpG islands during the conversion of hESCs into NPCs, and further passaging of hESC-NPCs. We first determined the expression of DNMT3A, DNMT3B and DNMT1 in hESCs and P3 and P14 hESC-NPCs by immunocytochemistry and RT-PCR analysis. *DNMT1* mRNA is expressed at comparable levels in hESCs, and P3 and P14 hESC-NPCs (data not shown), consistent with its maintenance role in DNA methylation. Both DNMT3A and DNMT3B are highly expressed in hESCs, raising the possibility that both enzymes may be involved in *de novo* methylation of CpG islands during the conversion of hESCs into early passage hESC-NPCs. However, immunocytochemistry shows that DNMT3B

protein is dramatically down-regulated in P3 and P14 hESC-NPCs (Fig. 4A), whereas DNMT3A enzyme is expressed at a similar level during all stages of cell differentiation. Quantification of *DNMT3A* and *DNMT3B* mRNAs levels by real-time RT-PCR shows a similar pattern for the stage-specific expression of *DNMT3A* and *DNMT3B* in hESCs, and P3 and P14 hESC-NPCs (Fig. 4B). In fact, *DNMT3A* mRNA level is transiently increased in P3 hESC-NPCs (Fig. 4B) and maintained at levels in P7–10 or P14 hESC-NPCs comparable to those in hESCs or in fetal astrocytes, respectively. The expression analysis of DNMTs suggests that DNMT3A could be the enzyme mediating the increase in DNA methylation in CpG islands in late passage hESC-NPCs. To determine if DNMT3A is directly associated with those CpG islands, we performed chromatin immunoprecipitation (ChIP) experiments with a rabbit antibody against DNMT3A in P3 and P14 hESC-NPCs. ChIP assays showed that gene promoters of *CPT1A* (both exon1a and exon1b promoters) and *SPAG6* are associated with DNMT3A (Fig. 5). In contrast, there is no DNMT3A association with the *ERBIN* gene promoter, which is unmethylated throughout the entire period of cell culture. Taken together, our data suggests that selective targeting of DNMT3A to specific gene promoters could be a leading mechanism underlying the hypermethylation of these CpG island promoters during the prolonged culture of NPCs.



**Figure 4.** Analysis of mRNA and protein levels of *DNMT3A* and *DNMT3B* genes during hESC differentiation. (A) Immunostaining of hESCs, and P3 and P14 hESC–NPCs with antibodies against DNMT3A (red) and DNMT3B (green). DAPI counter-staining shows nuclear morphology. (B) Real-time PCR assay for expression levels of *DNMT3A* and *DNMT3B*. Relative *DNMT3A* and *DNMT3B* levels were normalized with *GAPDH*.

#### Methylation status of repeat elements and non-CpG island promoters of *OCT4* and *GFAP* genes during hESC differentiation

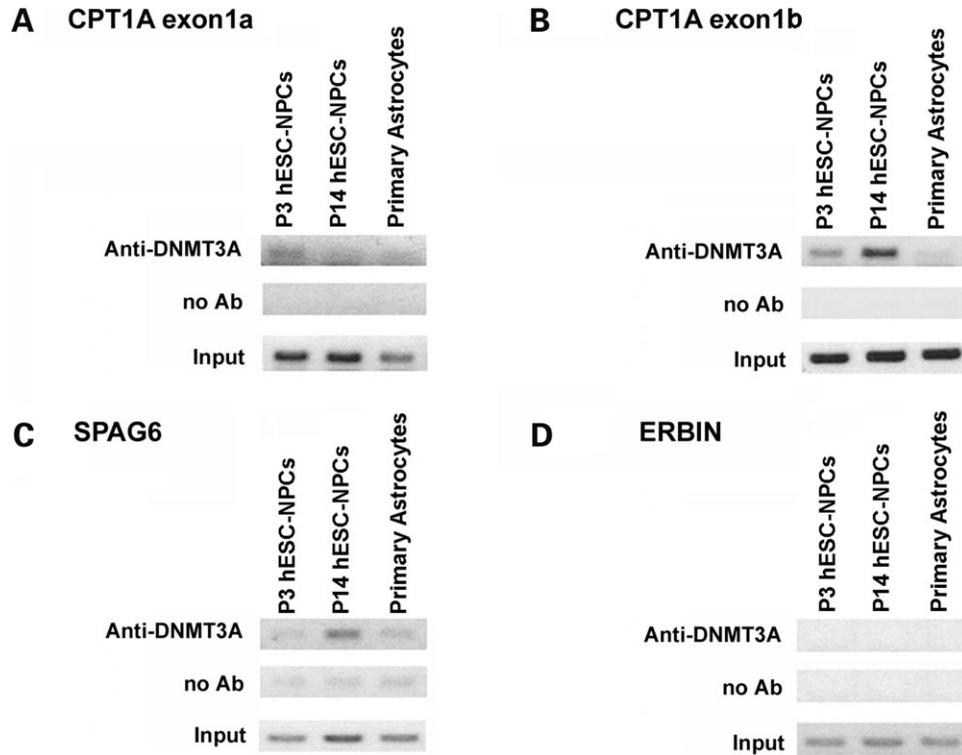
The bulk of DNA methylation in the mammalian genome occurs in various types of repeat elements including Alu repeats, retrotransposons, long- or short- interspersed repetitive elements (LINE and SINE) and pericentromeric repeats, which constitute ~40% of the human genome sequence (33,34). We surveyed different classes of repeat elements in the genome which were not covered in our initial DMH screen. Except for a slight transient increase in the MER52C class of repeat elements in P3 hESC–NPCs, we did not find any significant changes in DNA methylation patterns among hESCs, and P3 and P14 hESC–NPCs. Importantly, the methylation patterns in these repeat elements in hESCs and hESC–NPCs are similar to those in normal leukocytes (Fig. 6), suggesting that DNA methylation patterns in the bulk of repeat elements are stable during *in vitro* hESC differentiation.

It is also noted that *de novo* methylation can occur in non-CpG island gene promoters. Upon neural induction, the pluripotent marker gene *OCT4* is silenced in P3 hESC–NPCs. Indeed, our DNA methylation analyses show dramatic increases in DNA methylation, consistent with a previous report (35). Moreover, our DNMT3A ChIP assay demonstrates that the *OCT4* gene promoter is targeted by DNMT3A for *de novo* DNA methylation during *in vitro* hESC differentiation (Supplementary Material, Fig. S5). We also examined methylation status in a non-CpG island promoter region (–1640 to –1215 bp) within the human *GFAP* (*GFAP*) gene. In both hESCs and hESC–NPCs, this region

is heavily methylated (80–90%) and only a minor decrease in DNA methylation occurs in late passages of P14 hESC–NPCs when compared with hESCs and P3 hESC–NPCs (Supplementary Material, Fig. S6). In contrast, cortical NPCs exhibit a lower level of methylation (65.6%) *in vivo* and cultured primary NPCs only contain a minimal level of methylation (7.8%) in this region, suggesting that demethylation occurs rapidly in primary NPCs *in vitro* (Supplementary Material, Fig. S6). Our data suggest the DNA methylation in non-CpG island regions could also be differentially regulated between primary NPCs and hESC–NPCs during *in vivo* and *in vitro* differentiation.

#### DISCUSSION

Our present study provides the first glimpse of methylation changes in 4608 CpG islands during neural differentiation of hESCs. Using a DNA microarray-based methylation profiling technique, we demonstrate that ~1.4% of CpG islands undergo distinct *de novo* methylation during the conversion of hESCs into differentiated hESC–NPCs. One of the goals of this study is to determine if the current neural differentiation protocol will cause any abnormal methylation in cancer-related genes in hESC-derivatives. After comparing our list of hypermethylated CpG islands with the list of genes reported to become hypermethylated in different types of cancer cells (17), we did not find any overlap between these two groups of CpG island genes. Our results indicate that the event of *de novo* methylation during hESC differentiation is distinctively different from that in cancer cells. One caveat of the above



**Figure 5.** ChIP analysis of DNMT3A association with selective gene promoters in NPCs. ChIP primers were designed within the methylation assay region. *ERBIN* promoter, which displayed no methylation change during differentiation, was used as the negative control. DNMT3A was selectively associated with *CPT1A* exon1a and exon1b, and *SPAG6* promoters in P3 and P14 hESC-NPCs. Note that PCR conditions were slightly different for *CPT1A* exon1a promoter region (40 cycles) from *CPT1A* exon1b, *SPAG6* and *ERBIN* promoters (35 cycles), potentially due to the different PCR amplification efficiencies at different promoter regions.

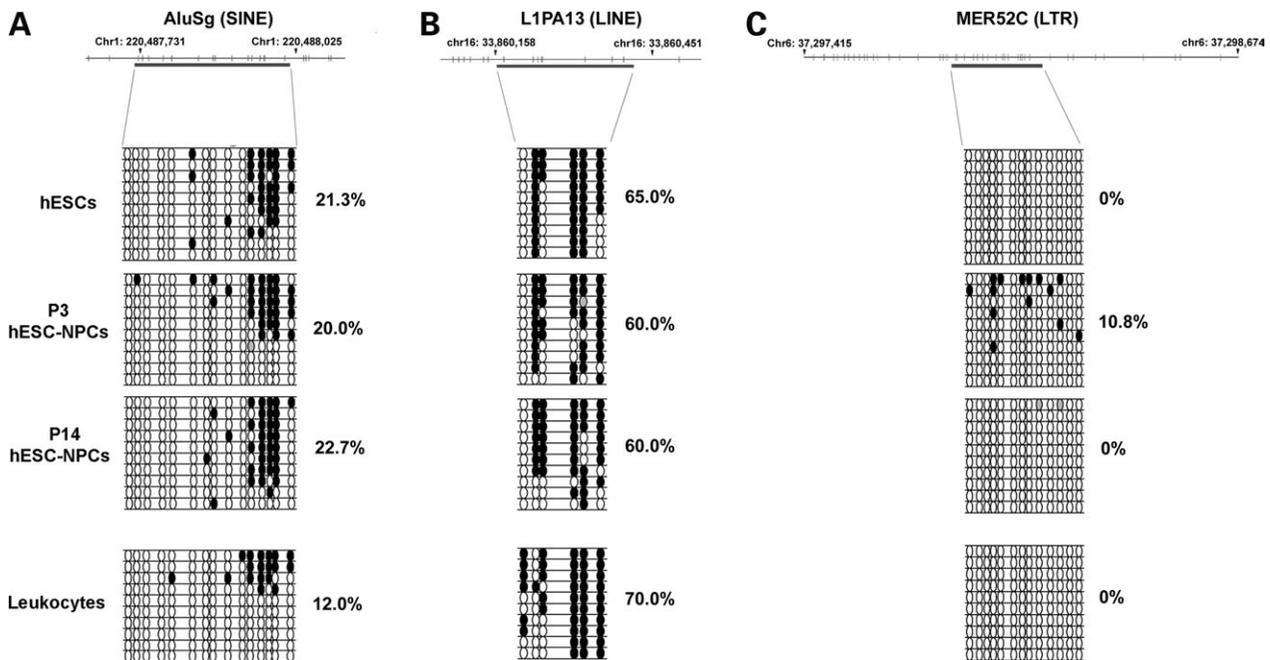
comparison is that hESC-NPCs are distinctively different by lineage from colon and prostate cancer cells. It may be more meaningful to compare our results with the profile of hypermethylated CpG islands in brain tumors that are derived from neural progenitors in future studies.

In addition, this study addresses whether in certain cases *de novo* methylation of CpG islands in hESC-NPCs are a result of a normal cell differentiation program that occurs *in vivo* or represents abnormal hypermethylation that takes place under our current differentiation/culture conditions. Bisulfite sequencing analysis in hESC-NPCs, primary fetal NPCs and astrocytes, and leukocytes indicates that levels of DNA methylation in *CPT1A* and *SPAG6* promoters are higher in hESC-NPCs than in primary cells, indicating that methylation levels in these two CpG islands are abnormally high. In fact, such an increase in promoter methylation can lead to transcriptional silencing (Fig. 3), which may affect cell physiology. Indeed, deficiency of *CPT1A* has been associated with hypoketotic hypoglycemia in humans (28,29), suggesting that reduced expression of *CPT1A* in hESC-NPCs can potentially render these cells less efficient in lipid metabolism. Our results indicate that monitoring methylation changes in hESC derivatives may be a necessary step with the current differentiation/culture conditions of hESCs. Preventing *de novo* DNA methylation at these specific gene loci may be required if expanded cultures of NPCs are to be used in therapeutic applications in the future.

It has been recently reported that although the DNA methylation pattern is relatively stable in imprinting loci during long-term culture of hESCs (41), an increase in methylation in the promoters of certain tumor suppressor genes is observed in several lines of hESCs (13). In our studies, we did not observe any increased methylation in five cancer-related genes in HSF6 hES cell line over a period of 3 years with 170 passages, suggesting that HSF6 cell line is epigenetically quite stable under current hESC culture conditions.

*De novo* DNA methyltransferases must be involved in CpG island methylation during hESC differentiation, as has been postulated previously for methylation of CpG islands in mouse ES cells (36). In our study, we found that expression of *de novo* methyltransferase DNMT3A is much higher than DNMT3B expression in hESC-NPCs, suggesting that DNMT3A could be the major enzyme for *de novo* CpG island methylation. This possibility was confirmed by our ChIP assays which demonstrated the association of DNMT3A with CpG island DNA sequences in hESC-NPCs. Therefore, DNMT3A could be a prime target for preventing DNA methylation in CpG islands through biochemical and pharmacological approaches (37,38).

Our current study is limited to profiling DNA methylation in a CpG island library containing 4608 fragments during neural differentiation of one cell line of HSF6 hESCs. It will be of interest to verify and extend our findings with additional



**Figure 6.** Bisulfite sequencing analysis of DNA methylation in repeat elements during hESC differentiation by (A) AluSg (SINE), (B) L1PA13 (LINE) and (C) MER52C (LTR). Each dot represents one CpG site with filled black dots as methylated and open dots as unmethylated. Chromosome locations of these repeat elements are shown in schematic diagrams. Analyzed regions are underlined by black bars, and CpG sites represented by vertical bars.

hESC cell lines and different differentiation paradigms such as lymphocyte and myocyte differentiation. Furthermore, dynamic DNA methylation changes could also occur in non-CpG island regions in cell lineage differentiation genes (7,8). Thus, with the advent of whole-genome tiling arrays (39) and new methods for genome-wide DNA methylation profiling such as DNA immunoprecipitation with 5' methylcytosine antibodies (17,40), it will be feasible to analyze the step-by-step changes in the DNA methylation pattern at all genomic regions during hESC differentiation. Because DNA methylation can directly or indirectly influence gene expression, it will be necessary to couple the genome-wide analysis of DNA methylation patterns with gene expression patterns. Comparing genome-wide methylation patterns with gene expression changes will allow us to determine if DNA methylation is causally linked to lineage-specific gene expression during development. This information will be extremely useful for developing novel strategies to control DNA methylation and gene expression during hESC differentiation for the ultimate goal of obtaining the most suitable hESC derivatives for cell therapy in the future.

## MATERIALS AND METHODS

### Cultures of hESCs and fetal astrocytes

Undifferentiated hESCs (HSF6 cell line with NIH code UC 06) were maintained on gamma-irradiated mouse feeder layer in DMEM high glucose supplement with 20% Knockout Serum Replacer (Invitrogen), 1 mM glutamine, 1% non-essential amino acids and 0.1 mM  $\beta$ -mercaptoethanol. hESCs were passaged every 3–4 days using Dispase and Collagenase IV

(Invitrogen) at a concentration of 1 mg/ml in KSR medium. Medium was changed everyday for optimal cell growth. Fetal astrocytes were purchased from AllCells LLC., and cultures were expanded as suggested by the vendor. Fetal brain tissue was obtained from the Surgical Pathology Laboratory at UCLA (Dr. Michael Teitell). Human primary NPCs were derived from primary cultures of fetal cortical tissues at the 11-week gestation stage using a protocol for culturing mouse NPCs (7). The use of hESCs and fetal brain tissues was approved by UCLA IRB.

### Directed neural differentiation of hESCs *in vitro*

hESCs were passaged onto gamma-irradiated mouse feeder cell layer using the standard culture conditions as described (10). To induce neural differentiation, hESCs between passages 50–90 were first replated for 1–2 h to get rid of most feeder cells, and then plated onto matrigel-coated plates with hESC medium and bFGF treatment for two consecutive passages. In the middle of the second passage, hESC culture medium was switched to DMEM/F12 serum-free medium/B27 supplement, which is suitable for the expansion of NPCs. Subsequently, partially differentiated hESCs were dissociated and plated onto poly-ornothine/fibronectin substrates with the serum-free medium and bFGF treatment (10 ng/ml) and continuously passaged to be converted into homogenous hESC-derived NPCs. In the first passage of hESC–NPCs (P1), we observed significant cell death in our cultures, suggesting a selection pressure for many non-NPC cell types. However, cell death was reduced to a minimum in the following two passages (P2 and P3) with enriched populations of hESC–NPCs. After NPCs became 80% confluent, they were sequentially passaged by scraping

them off the plate or with light trypsin–EDTA treatment with serum-free DMEM/F12–B27 medium. Sequentially passaged cells were fixed with 4% paraformaldehyde (PFA)/PBS for immunostaining or harvested for RNA and DNA purification. For AzaC treatment, late passaged hESC–NPCs (P8–10) were treated with AzaC at the concentration of 1  $\mu$ M for 6 days.

### Bisulfite genomic sequencing analysis and COBRA assay

Genomic DNA samples were obtained from hESCs and early (P3 hESC–NPCs) and late (P14 hESC–NPCs) passages of hESC–NPCs. DNA was digested with *Bgl*II overnight and treated with sodium bisulfite for 15 h as described previously (24). Converted DNA samples were cleaned using Wizard<sup>®</sup> DNA Clean Up Kit (Promega) and were amplified by PCR reactions. PCR products were either directly sequenced or cloned into the Topo Vector 4.0 (Invitrogen) for sequencing of allelic methylation patterns. The sequence and primer information are available upon request. The COBRA assay followed the previously published protocol (25). MethPrimer program (<http://www.urogene.org/methprimer/index1.html>) was used to design primers for bisulfite sequencing PCR or restriction PCR for *STMN3* (6F5) and *ARMC7* (33A11) gene promoter CpG islands. The primers for *STMN3* are 5'-TTTTAGGGTATTTTAAATATAATGAATTTA and 5'-ACAAAACAAACAACAAACAAAAAAC, which amplify the bisulfite-treated DNA into a 247 bp fragment with a HpyCH4IV site. Expected fragment sizes are 143 and 104 bp if the CpG site is methylated. The primers for *ARMC7* gene promoter CpG island are 5'-TTTTTTGTAAAGGGGTAAAGGATT and 5'-CAACAAAATCTAACAACACCCATATTA, which amplify the bisulfite-treated DNA into a 169 bp fragment restricted by *Taq*I into 132 and 37 bp fragments if the CpG site is methylated. The PCR reactions were performed with AmpliTaq Gold DNA polymerase from Applied Biosciences for 35 cycles with an annealing temperature of 55°C.

### RT–PCR assays

RNA samples were extracted from hESCs, hESC–NPCs and primary astrocytes using TRIzol<sup>®</sup> (Invitrogen), and converted to cDNA with Omniscript<sup>®</sup> RT Kit (Qiagen) or iScript cDNA Synthesis Kit (Bio-Rad) for real-time PCR. PCR primers were designed across the exons for the genes of interest. PCR was carried out to assay gene expression and  $\beta$ -ACTIN was used as a control. Real-time PCR was carried out with Bio-Rad iCycler real-time PCR machine using iQ SYBR Green Supermix (Bio-Rad). Primers were designed across the exons, and each primer pair was tested for PCR efficiency using a standard curve and melting curve of series diluted cDNA. Relative gene expression levels were calculated after they were normalized with expression levels of GAPDH.

### Immunohistochemistry

For SSEA4 (Chemicon) staining, hESCs on coverslips were transferred to 24-well plates and live-stained with primary antibody directly diluted in culture medium for 30 min at 37°C. The cells were washed three times with PBS and incubated with Cy2- or Cy3-labeled anti-mouse IgG (1:500) secondary antibody (Jackson ImmunoResearch, West Grove,

PA, USA) for 1 h at 37°C. Stained cells were then fixed with 4% PFA/PBS and counter-stained with DAPI. For OCT4, NESTIN, SOX2, TuJ1, DNMT3A and DNMT3B immunostaining, cultured cells were first fixed with 4% PFA/PBS for 20 min at room temperature and washed with PBS. Cells were then permeabilized with 0.4% Triton-X in TBST for 20 min and blocked in 10% milk and 1% normal goat serum for 1 h. Coverslips were incubated overnight at 4°C in primary antibodies diluted in 3% BSA in TBST [monoclonal mouse anti-OCT4 (1:20, Santa Cruz), polyclonal rabbit anti-NESTIN (1:200), polyclonal goat anti-SOX2 (1:200, Santa Cruz), monoclonal mouse anti-TuJ1 (1:2000, a gift from Dr. Frankfurter at University of Virginia), monoclonal mouse anti-Dnmt3a (1:400, Imgenex) and polyclonal rabbit anti-Dnmt3b (1:1000, a gift from Dr. En Li at Novartis Institute)]. After being washed three times with PBS, coverslips were incubated in fluorochrome-conjugated secondary antibodies for 1 h at room temperature with protection from light. DAPI or Hoechst dye #33342 was used to label cell nuclei.

### Chromatin immunoprecipitation

ChIP assays were performed according to Upstate protocol. Cells were crosslinked with 1% formaldehyde in PBS for 20 min at room temperature and scraped off the plate following three washes with cold PBS. After that, cells were lysed with SDS lysis buffer with proteinase inhibitors and 1 mM dithiothreitol (DTT). Cell lysates were sonicated to break down chromatin into 200–500 base pairs in length. Chromatin was 1:10 diluted, and precleared with protein A agarose beads (Upstate) for 1 h at 4°C. Approximately, 3% of the chromatin from each sample was saved as input. The rest of the chromatin was used for immunoprecipitation with either polyclonal rabbit anti-Dnmt3a (a gift from Dr. En Li at Novartis Institute) or no antibody as a control at 4°C overnight. Immunoprecipitated antibody–protein–DNA complexes were collected by protein A agarose beads at 4°C for 1 h and extensively washed. The complex was eluted and reverse crosslinked at 65°C for 6 h and proteinase K was added for further incubation at 45°C for 1 h. DNA samples were extracted using phenol:chloroform method. ChIP primers were designed within the corresponding methylation assay regions and PCR was performed for 35 (CPT1A exon1b, SPAG6, ERBIN) or 40 cycles (CPT1A exon1a) with the following cycling condition (94°C, 30 s; 56°C, 30 s; 72°C, 1 min).

### Microarray-based methylation profiling

A genomic library of human unmethylated CpG island fragments was constructed by size fraction of *Sma*I cut DNA samples pooled from healthy human leukocytes (Z. Wang, unpublished data). Individual clones were cultured in 96-well microplates and inserts from 4226 random clones (44 plates) were PCR amplified with M13 vector primers. The quality of this CpG island fragment library was first assessed by sequencing analysis of 58 randomly picked clones. The majority of these 58 clones from this *Sma*I genomic fragment library were from CpG islands: of the 63 *Sma*I fragments (five clones have double inserts), 45 (71%) were from gene promoter region CpG islands; eight were

from rRNA gene repeats; and the remaining were repetitive sequences (three inserts) and non-CGI sequences (seven inserts). Additional 384 CpG island fragments from 320 cancer-related genes were directly amplified from human genomic DNA. The total of 4608 PCR-amplified fragments were purified and spotted onto polylysine-coated microscope slides with an Affymetrix pin spotting machine.

For microarray hybridization, genomic DNA samples were treated with the methylation-sensitive restriction enzyme *SmaI* at 25°C overnight. A partially double-stranded linker with a blunt end was used to ligate the *SmaI*-digested fragments. For linker preparation, oligos (100 μM each) are mixed and incubated at 75°C for 2 min and cool to room temperature for 30–60 min. The oligo sequences for linker are:

- 5'-AGCACTCTCCAGCCTCTCACCGAC-3'
- 3'-AGAGTGGCTG-5'

Linkers were ligated to digested genomic DNA samples with T4 DNA ligase at 16°C overnight. The ligated genomic DNA were subjected to PCR amplification. PCR conditions are: 72°C, 5 min to fill the bottom strand of the linker, then 25 cycles of 95°C for 40 s, 72°C for 2 min using standard PCR reagents with the addition of betaine and DMSO. PCR products were run on agarose gel to check for amplification efficiency before labeling.

The PCR products were then purified by Qiagen PCR cleaning kit, labeled with either Cy3 or Cy5-dCTP, mixed and hybridized to the CpG island fragment microarray. The hybridized arrays were scanned with an Axon 4000B Microarray Scanner and the images were analyzed with GenPix Pro software. Three pairs of array results (hESCs versus P3 hESC-NPCs, hESCs versus P14 hESC-NPCs and P3 hESC-NPCs versus P14 hESC-NPCs) were normalized and the ratios of hybridization signals were averaged. For identifying CpG island clones exhibiting increased methylation, we initially set up a threshold of averaged ratio of hESC-NPCs probe intensity over hESCs probe intensity at <0.667 and hybridization signals must be present on at least three out of six slides (hESCs versus P3 NPCs and hESCs versus P14 NPCs). Additional candidate clones were included that are consistently below a ratio of 0.90 across all the arrays examined. Candidate clones were subsequently amplified and sequenced and their chromosomal locations were determined with UCSC BLAT genome browser (genome.ucsc.edu). All the sequences of CpG islands described in this article are available upon request.

## SUPPLEMENTARY MATERIAL

Supplementary Material is available at HMG Online.

## ACKNOWLEDGEMENTS

We thank Mike Teitell, Yi Sun, Karen Reue and Peyman Golshani for critically reading the article. We also thank Yan Hu for technical help in the initial phase of this study. Y.S. conducted most of the experiments and participated in paper writing. J.C. assisted in experiments and in paper review.

Z.W. assisted in microarray experiments and data analysis. G.F. conceived and designed experiments and wrote the article. This work was supported by NIH Grant RO1 NS44405 and a research contract grant from HealthBanks Biotech Co., at Taiwan to G.F. Funding to pay the Open Access publication charges for this article was provided by UCLA.

**Conflict of Interest statement.** The authors have declared that no competing interests exist.

## REFERENCES

1. Jaenisch, R. and Bird, A. (2000) Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat. Genet.*, **33**(suppl.), 245–254.
2. Grace-Goll, M. and Bestor, T.H. (2005) Eukaryotic cytosine methyltransferases. *Annu. Rev. Biochem.*, **74**, 481–514.
3. Li, E., Bestor, T.H. and Jaenisch, R. (1992) Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell*, **69**, 915–926.
4. Okano, M., Bell, D.W., Haber, D.A. and Li, E. (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for *de novo* methylation and mammalian development. *Cell*, **99**, 247–257.
5. Robertson, K.D. (2005) DNA methylation and human disease. *Nat. Rev. Genet.*, **6**, 597–610.
6. Xu, G.L., Bestor, T.H., Bourc'his, D., Hsieh, C.L., Tommerup, N., Bugge, M., Hulten, M., Qu, X., Russo, J.J. and Viegas-Pequignot, E. (1999) Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. *Nature*, **402**, 187–191.
7. Fan, G., Martinowich, K., Chin, M.H., He, F., Fouse, S.D., Hutnick, L., Hattori, D., Ge, W., Shen, Y., Wu, H. *et al.* (2005) DNA methylation controls the timing of astroglial development through regulation of JAK-STAT signaling. *Development*, **132**, 3345–3356.
8. Takizawa, T., Nakashima, K., Namihira, M., Ochiai, W., Uemura, A., Yanagisawa, M., Fujita, N., Nakao, M. and Taga, T. (2001) DNA methylation is a critical cell-intrinsic determinant of astrocyte differentiation in the fetal brain. *Dev. Cell*, **1**, 749–758.
9. Hochedlinger, K. and Jaenisch, R. (2003) Nuclear transplantation, embryonic stem cells, and the potential for cell therapy. *N. Engl. J. Med.*, **349**, 275–286.
10. Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S. and Jones, J.M. (1998) Embryonic stem cell lines derived from human blastocysts. *Science*, **282**, 1145–1147.
11. Keller, G. (2005) Embryonic stem cell differentiation: emergence of a new era in biology and medicine. *Genes Dev.*, **19**, 1129–1155.
12. Nistor, G.I., Totoiu, M.O., Haque, N., Carpenter, M.K. and Keirstead, H.S. (2005) Human embryonic stem cells differentiate into oligodendrocytes in high purity and myelinate after spinal cord transplantation. *Glia*, **49**, 385–396.
13. Maitra, A., Arking, D.E., Shivapurkar, N., Ikeda, M., Stastny, V., Kassaei, K., Sui, G., Cutler, D.J., Liu, Y., Brimble, S.N. *et al.* (2005) Genomic alterations in cultured human embryonic stem cells. *Nat. Genet.*, **37**, 1099–1103.
14. Bestor, T.H., Gundersen, G., Kolsto, A.B. and Prydz, H. (1992) CpG islands in mammalian gene promoters are inherently resistant to *de novo* methylation. *Genet. Anal. Tech. Appl.*, **9**, 48–53.
15. Bestor, T.H. (2000) The DNA methyltransferases of mammals. *Hum. Mol. Genet.*, **9**, 2395–2402.
16. Li, E. (2002) Chromatin modification and epigenetic reprogramming in mammalian development. *Nat. Rev. Genet.*, **3**, 662–673.
17. Keshet, I., Schlesinger, Y., Farkash, S., Rand, E., Hecht, M., Segal, E., Pikarski, E., Young, R.A., Niveleau, A., Cedar, H. and Simon, I. (2006) Evidence for an instructive mechanism of *de novo* methylation in cancer cells. *Nat. Genet.*, **38**, 149–153.
18. Ying, Q.L., Stavridis, M., Griffiths, D., Li, M. and Smith, A. (2003) Conversion of embryonic stem cells into neuroectodermal precursors in adherent monoculture. *Nat. Biotechnol.*, **21**, 183–186.
19. Conti, L., Pollard, S.M., Gorba, T., Reitano, E., Toselli, M., Biella, G., Sun, Y., Sanzone, S., Ying, Q.L., Cattaneo, E. and Smith, A. (2005) Niche-independent symmetrical self-renewal of a mammalian tissue stem cell. *PLoS Biol.*, **3**, e283.

20. Qian, X., Shen, Q., Goderie, S.K., He, W., Capela, A., Davis, A.A. and Temple, S. (2000) Timing of CNS cell generation: a programmed sequence of neuron and glial cell production from isolated murine cortical stem cells. *Neuron*, **28**, 69–80.
21. Huang, T.H., Perry, M.R. and Laux, D.E. (1999) Methylation profiling of CpG islands in human breast cancer cells. *Hum. Mol. Genet.*, **8**, 459–470.
22. Toyota, M., Ho, C., Ahuja, N., Jair, K.W., Li, Q., Ohe-Toyota, M., Baylin, S.B. and Issa, J.P. (1999) Identification of differentially methylated sequences in colorectal cancer by methylated CpG island amplification. *Cancer Res.*, **59**, 2307–2312.
23. Gardiner-Garden, M. and Frommer, M. (1987) CpG islands in vertebrate genomes. *J. Mol. Biol.*, **196**, 261–282.
24. Clark, S.J., Harrison, J., Paul, C.L. and Frommer, M. (1994) High sensitivity mapping of methylated cytosines. *Nucleic Acids Res.*, **22**, 2990–2997.
25. Xiong, Z. and Laird, P.W. (1997) COBRA: a sensitive and quantitative DNA methylation assay. *Nucleic Acids Res.*, **25**, 2532–2534.
26. Jones, P.A. and Baylin, S.B. (2002) The fundamental role of epigenetic events in cancer. *Nat. Rev. Genet.*, **3**, 415–428.
27. Reyes, T., Morrison, S.J., Clarke, M.F. and Weissman, I.L. (2001) Stem cells, cancer, and cancer stem cells. *Nature*, **414**, 105–111.
28. Gobin, S., Thuillier, L., Jögl, G., Faye, A., Tong, L., Chi, M., Bonnefont, J.P., Girard, J. and Prip-Buus, C. (2003) Functional and structural basis of carnitine palmitoyltransferase 1A deficiency. *J. Biol. Chem.*, **278**, 50428–50434.
29. Prip-Buus, C., Thuillier, L., Abadi, N., Prasad, C., Dilling, L., Klasing, J., Demaugre, F., Greenberg, C.R., Haworth, J.C., Droin, V. *et al.* (2001) Molecular and enzymatic characterization of a unique carnitine palmitoyltransferase 1A mutation in the Hutterite community. *Mol. Genet. Metab.*, **73**, 46–54.
30. Neilson, L.I., Schneider, P.A., Van Deerlin, P.G., Kiriakidou, M., Driscoll, D.A., Pellegrini, M.C., Millinder, S., Yamamoto, K.K., French, C.K. and Strauss, J.F., III. (1999) cDNA cloning and characterization of a human sperm antigen (SPAG6) with homology to the product of the *Chlamydomonas* PF16 locus. *Genomics*, **60**, 272–280.
31. Zhang, Z., Sapiro, R., Kapfhamer, D., Bucan, M., Bray, J., Chennathukuzhi, V., McNamara, P., Curtis, A., Zhang, M., Blanchette-Mackie, E.J. and Strauss, J.F., III. (2002) A sperm-associated WD repeat protein orthologous to *Chlamydomonas* PF20 associates with Spag6, the mammalian orthologue of *Chlamydomonas* PF16. *Mol. Cell. Biol.*, **22**, 7993–8004.
32. Price, N., van der Leij, F., Jackson, V., Corstorphine, C., Thomson, R., Sorensen, A. and Zammit, V. (2002) A novel brain-expressed protein related to carnitine palmitoyltransferase I. *Genomics*, **80**, 433–442.
33. Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W. *et al.* (2001) Initial sequencing and analysis of the human genome. *Nature*, **409**, 860–921.
34. Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A. *et al.* (2001) The sequence of the human genome. *Science*, **291**, 1304–1351.
35. Cowan, C.A., Atienza, J., Melton, D.A. and Eggan, K. (2005) Nuclear reprogramming of somatic cells after fusion with human embryonic stem cells. *Science*, **309**, 1369–1373.
36. Hattori, N., Abe, T., Suzuki, M., Matsuyama, T., Yoshida, S., Li, E. and Shiota, K. (2004) Preference of DNA methyltransferases for CpG islands in mouse embryonic stem cells. *Genome Res.*, **14**, 1733–1740.
37. Brueckner, B. and Lyko, F. (2004) DNA methyltransferase inhibitors: old and new drugs for an epigenetic cancer therapy. *Trends Pharmacol. Sci.*, **25**, 551–554.
38. Schneider-Stock, R., Diab-Assef, M., Rohrbeck, A., Foltzer-Jourdainne, C., Boltze, C., Hartig, R., Schonfeld, P., Roessner, A. and Gali-Muhtasib, H. (2005) 5-Aza-cytidine is a potent inhibitor of DNA methyltransferase 3a and induces apoptosis in HCT-116 colon cancer cells via Gadd45- and p53-dependent mechanisms. *J. Pharmacol. Exp. Ther.*, **312**, 525–536.
39. Kim, T.H., Barrera, L.O., Zheng, M., Qu, C., Singer, M.A., Richmond, T.A., Wu, Y., Green, R.D. and Ren, B. (2005) A high-resolution map of active promoters in the human genome. *Nature*, **436**, 876–880.
40. Schumacher, A., Kapranov, P., Kaminsky, Z., Flanagan, J., Assadzadeh, A., Yau, P., Virtanen, C., Winegarden, N., Cheng, J., Gingeras, T. and Petronis, A. (2006) Microarray-based DNA methylation profiling: technology and applications. *Nucleic Acids Res.*, **34**, 528–542.
41. Rugg-Gunn, P.J., Ferguson-Smith, A.C. and Pedersen, R.A. (2005) Epigenetic status of human embryonic stem cells. *Nat. Genet.*, **37**, 585–587.