Spatial, temporal and interindividual epigenetic variation of functionally important DNA methylation patterns

Eberhard Schneider¹, Galyna Pliushch¹, Nady El Hajj¹, Danuta Galetzka², Alexander Puhl³, Martin Schorsch⁴, Katrin Frauenknecht⁵, Thomas Riepert⁶, Achim Tresch⁷, Annette M. Müller⁸, Wiltrud Coerdt⁹, Ulrich Zechner² and Thomas Haaf^{1,*}

¹Institute of Human Genetics, Julius Maximilians University, Biozentrum, Am Hubland, 97074 Wuerzburg, ²Institute of Human Genetics, ³Department of Obstetrics and Gynecology, Johannes Gutenberg University, Langenbeckstrasse 1, 55131 Mainz, ⁴Fertility Center Wiesbaden, Mainzer Strasse 98-102, 65189 Wiesbaden, ⁵Department of Neuropathology, Johannes Gutenberg University, Langenbeckstrasse 1, 55131 Mainz, ⁶Institute of Legal Medicine, Johannes Gutenberg University, Am Pulverturm 3, 55131 Mainz, ⁷Gene Center, Department of Chemistry and Biochemistry, Ludwig Maximilians University, 81377 Munich, ⁸Department of Paediatric Pathology, University Clinic of Bonn, Sigmund-Freud Strasse 25, 53123 Bonn and ⁹Department of Paediatric Pathology, Johannes Gutenberg University, Langenbeckstrasse 1, 55131 Mainz, Germany

Received July 20, 2009; Revised and Accepted February 11, 2010

ABSTRACT

DNA methylation is an epigenetic modification that plays an important role in gene regulation. It can be influenced by stochastic events, environmental factors and developmental programs. However, little is known about the natural variation of genespecific methylation patterns. In this study, we performed quantitative methylation analyses of six differentially methylated imprinted genes (H19, MEG3, LIT1, NESP55, PEG3 and SNRPN), one hypermethylated pluripotency gene (OCT4) and one hypomethylated tumor suppressor gene (APC) in chorionic villus, fetal and adult cortex, and adult blood samples. Both average methylation level and range of methylation variation depended on the gene locus, tissue type and/or developmental stage. We found considerable variability of functionally important methylation patterns among unrelated healthy individuals and a trend toward more similar methylation levels in monozygotic twins than in dizygotic twins. Imprinted genes showed relatively little methylation changes associated with aging in individuals who are >25 years. The relative differences in methylation among neighboring CpGs in

the generally hypomethylated *APC* promoter may not only reflect stochastic fluctuations but also depend on the tissue type. Our results are consistent with the view that most methylation variation may arise after fertilization, leading to epigenetic mosaicism.

INTRODUCTION

It is becoming increasingly clear that DNA sequence variations and environmental factors alone cannot account for the phenotypic differences among individuals. Studies in genetically identical organisms, i.e. monozygotic twins, isogenic animal strains and cloned animals implicated non-DNA sequence and non-environment-based effects as a main source of phenotypic variation (1-4). After genes and environment, epigenetic factors, in particular DNA methylation, may form the molecular basis of this 'third component'. Much of the existing DNA methylation variation may result from stochastic events during epigenetic reprogramming in gametogenesis, early embryogenesis and later developmental processes. This spontaneous epimutation rate can be modified by genetic and environmental factors (5-7). It is estimated that epimutations occur 100 times more frequently than somatic DNA mutations (8,9).

© The Author(s) 2010. Published by Oxford University Press.

^{*}To whom correspondence should be addressed. Tel: +49 931 3188738; Fax: +49 931 3184069; Email: thomas.haaf@uni-wuerzburg.de

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/ by-nc/2.5), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Epigenetic information is not encoded by the DNA sequence itself but by reversible modifications of DNA and/or histones, that can be transmitted from cells to daughter cells and sometimes even from one generation to another (7,10). Methylation of the 5'-cytosine residues in CpG dinucleotides directly targets DNA. It is associated with posttranslational histone modifications that lead to a locally condensed inactive chromatin structure (11,12). One important component of epigenetic gene regulation are CpG islands (CGIs) that are 500-2000 bp long and associated with promoters in most mammalian genes. Most CpGs in promoters are protected from methylation in somatic tissues. Promoter methylation leads to stable gene silencing during development/differentiation and in disease processes (13-15). For example, pluripotency genes switch from a demethylated and transcriptionally active state in embryonic cells to a fully methylated repressed state in somatic cells (16,17). In contrast, tumor suppressor genes are demethylated and active in somatic cells; ectopic methylation begins early in tumorigenesis (18-20). Genomic imprinting is a parent-specific epigenetic modification in which allelespecific expression depends on male versus female germline transmission. The differentially methylated regions (DMRs) of imprinted genes are involved in the control of parental allele-specific gene activity (21,22).

Genomic DNA methylation patterns may exhibit considerable variation among human individuals, as well as intraindividual changes over time (23-26). Large-scale epigenome mapping revealed that CpG-rich regions exhibit substantially less interindividual variation of DNA methylation patterns than CpG-poor regions. Most CGIs showed relatively similar low methylation levels in all individuals, with a few methylated CpGs randomly distributed among the many unmethylated CpGs (27). This is consistent with the view that the density of methylated CpGs in a cis-regulatory region (rather than individual CpGs, i.e. in transcription factor binding sites) turns a gene 'on' or 'off' (15,28). It is wellknown that epigenetic gene silencing during tumor initiation and progression involves methylation of entire CGIs but not single critical CpGs (18-20). It is plausible to assume that epigenetic variation contributes significantly to both phenotypic variation among individuals and human disease (7,14,29). However, in order to associate unexpected methylation levels of developmentally important genes with disease processes, it is crucial to assess the normal range of methylation variation. Using quantitative methylation assays, we measured the methylation levels in cis-regulatory regions of imprinted, pluripotency and tumor suppressor genes in different tissues and/or developmental states.

MATERIALS AND METHODS

Tissue samples and DNA preparation

Blood samples were obtained with informed consent from 12 healthy monozygotic and 14 dizygotic twin pairs as well as from unrelated individuals. Chorionic villus samples (CVS) of 48 term placentas of healthy newborns were

obtained with informed consent of the mothers from the Department of Gynecology and Obstetrics, Mainz University. Sperm samples of 45 males with normal spermiograms (according to WHO guidelines) were obtained with informed consent of the donors from the Fertility Center Wiesbaden.

In addition, we analyzed native chorionic villi from 29 first-trimester spontaneous abortions and 12 cultured CVS. Ten fetal brain samples (frontal cortex) were obtained from spontaneous abortions and stillbirths that underwent pathomorphological examination at the Department of Paediatric Pathology, Mainz University. Following diagnostics, the excess samples were made anonymous. Placental and fetal tissues were taken only from pregnancies without detectable chromosome and fetal abnormalities. Tissues were dissected within 24 h after abortion/stillbirth and stored at -80° C until further analysis. Adult brain samples from 12 humans between 24 and 83 years of age were obtained from the Institute of Legal Medicine and the Department of Neuropathology, Mainz University. Causes of death reflected accidents and a wide spectrum of common diseases (e.g. cardiovascular disease). Medical history and histological examination revealed no evidence for hereditary brain disorders. After removal of meninges frontal cortex tissue (area A10) was dissected between 1-2 days postmortem and immediately frozen and stored at -80° C. Use of anonymized excess tissue materials (from pathomorphological diagnostics) for scientific analyses was approved by the local ethics committee [Aerztekammer Rheinland-Pfalz, Decisions No. 837.103.04 (4261) and 837.073.07 (5608)].

The DNeasy Blood and Tissue Kit (Qiagen, Hilden Germany) was used for genomic DNA isolation. Sperm samples were first purified with Pure Sperm 40/80 (Nidacon, Molndal, Sweden) and and then treated with 100 mM Tris–Cl, 10 mM EDTA, 500 mM NaCl, 1% SDS and 2% β -mercaptoethanol. After incubation for 2 h at 56°C with proteinase K, DNA was prepared, as described above.

Studied genes

The DMRs of two paternally methylated (H19 and *MEG3*) and four maternally methylated (*LIT1*, *NESP55*, *PEG3* and *SNRPN*) genes as well as the promoters of the pluripotency gene OCT4 and the tumor suppressor gene APC were analyzed by bisulfite pyrosequencing. Amplicons targeting methylation-dependent cis-regulatory regions of these well-studied genes were chosen from the literature (30–35). Although the gene-specific pyrosequencing assays are based on only a small number (2–7) of CpGs per gene, the analyzed sites are thought to be representative of a given DMR or promoter. Individual CpGs in larger CGIs cannot stably maintain methylation states that differ from those of the neighboring CpGs. Usually, the entire CGI is either methylated or demethylated (27,28). PCR and sequencing primers (Table 1) were designed with the Pyrosequencing Assay Design Software (Biotage). A single nucleotide polymorphism (SNP), rs2107425, with approximately

Gene	Primer	Sequence $(5'-3')$	Amplicon length (bp)	Chromosomal localization (bp)	Number of CpGs	Reference
H19	Outer forward Outer reverse Nested forward Nested reverse* Sequencing	TTTTTGGTAGGTATAGAGTT AAACCATAACACTAAAACCC TGTATAGTATA	231	Chromosome 11 1 977 647-1 977 878	4	El-Maarri et al. (30)
MEG3	Forward Reverse* Sequencing	GATITTTTTTATATATTGTGTTTG CTCATTTCTCTAAAAATAATTAACC GTGTTTGAATTTATTTTGTTTGG	220	Chromosome 14 100 361 907-100 362 129	3	Zechner et al. (35)
LIT1	Forward Reverse* Sequencing	AATTAGTAGGTGGGGGG CTAAAAAACTCCCTAAAAATC GGGGGTAGTYGGAG	122	Chromosome 11 2 677 751-2 677 873	2	Mackay et al. (31)
NESP55	Outer forward Outer reverse Nested forward Nested reverse* Sequencing	TTTTTATTTTATAGGGTGTATTT AAAATAAAATACTTAAACACCAC TTTTTGTAGAGTTAGAGGGTAGGT AAAAAAAAAA	343	Chromosome 20 56 848 506-56 848 849	3	El-Maarri et al. (30)
PEG3	Forward Reverse*	GGTGTAGAAGTTTGGGTAGTTG CTCACCTCACCTCAATACTAC TGTTATTTTGGGTTGGT	153	Chromosome 19 62 043 756-62 043 909	4	Mackay et al. (31)
SNRPN	Forward* Reverse Sequencing 1 Sequencing 2	AGGGAGTTGGGATTTTTGTATT CCCAAACTATCTCTTAAAAAAAAC ACACAACTAACCTTACCC CCAACCTACCT	237	Chromosome 15 22 751 105-22 751 342	$2 \times 3 = 6$	White et al. (32)
OCT4	Forward Reverse*	AAGTTTTTGTGGGGGGATTTGTAT CCACCCACTAACCTTAACCTCTA TGAGGTTTTGGAGGG	185	Chromosome 6 31 246 461-31 246 646	2	Deb-Rinker et al. (33)
APC	Forward* Reverse Sequencing	GGTTAGGGTTAGGTAGGTTGT ACTACACCAATACAACCACATATC CCACACCCAACCAA	193	Chromosome 5 112 101 274-112 101 467	7	Schatz et al. (34)

Table 1. Genes and primers for pyrosequencing

This table provides the primer sequences for bisulfite pyrosequencing of our studied genes, the length and chromosomal localization (according to ensembl 54, May 2009) of the amplified segments, and the number of CpG sites in the amplicons. Biotinylated primers are indicated by star symbols.

50% heterozygosity is located at the 4th base from 5'-end of the nested reverse primer for *H19*. In previously published methylation analyses (30) and also in our own experience, this A to G change did not cause a biased amplification of DNA methylation. The binding sites for the other studied genes do not contain known SNPs. Although we cannot entirely exclude that the one or other analyzed sample displays a novel allelic sequence variation, such rare variants, if any, cannot explain the observed methylation variation of imprinted genes.

Imprinting of the paternally expressed IGF2 gene and the maternally expressed non-protein-coding H19 gene on chromosome 11p15 is regulated by differentially methylated CCCTC binding factor sites in the 5'-region of H19 (36). Methylation disturbances in the H19 germline DMR lead to dysregulation of IGF2-H19 imprinting and intrauterine growth retardation (Silver–Russell syndrome) (37) and overgrowth (Beckwith-Wiedemann syndrome) (38), respectively. The germline DMR within intron 10 of the KCNQ1 gene regulates expression of the paternally expressed KCNQ10T1 (LIT1) transcript and the maternally expressed CDKN1C gene on chromosome 11p15. Epimutations in this *LIT1* DMR are also associated with Beckwith-Wiedemann syndrome (39). The methylation status of the germline DMR in the SNRPN promoter/ exon 1 is important for imprint establishment/maintenance of genes in the Prader-Willi/Angelman syndrome region (40). The PEG3 gene on chromosome 19q13.4 is paternally expressed in embryo and placenta and can induce apoptosis. It has a DMR in its promoter region (41). The maternally expressed non-coding *MEG3* RNA represents a growth suppressor. Hypomethylation of the *MEG3* promoter on chromosome 14q32.2 is associated with low-birth weight, muscular hypotonia and various dysmorphisms (35). The *GNAS* locus on chromosome 20q13.3 has a highly complex imprinting pattern. Its different products (Gsalpha, XLalphas and NESP55) are involved in early postnatal adaptions and neuroendocrine functions. *NESP55* is expressed from the maternal chromosome in restricted brain areas (42).

The transcription factor OCT4 (POU5F1) is the key gene for maintaining pluripotency in mammalian cells (43). It is highly expressed in human oocytes, down-regulated in early embryos, and then expressed *de novo* from the embryonic genome in blastocysts. Differentiation of the somatic tissues is associated with downregulation of OCT4 (44). Germline mutations in the *APC* tumor suppressor gene cause familial adenomatous polyposis, a hereditary cancer syndrome. Somatic inactivation of *APC* by promoter hypermethylation is frequently seen in sporadic colorectal cancers (18).

Bisulfite pyrosequencing

Bisulfite treatment of genomic DNA was performed using the EpiTect Bisulfite Kit (Qiagen, Hilden, Germany). The PCR reaction mixture consisted of $2.5 \,\mu$ l $10 \times$ PCR buffer, $2.5 \,\mu$ l $50 \,\text{mM}$ MgCl₂, $2.5 \,\mu$ l $10 \,\text{mM}$ dNTP mix, $1.0 \,\mu$ l (100 ng) of each forward and reverse primer, $0.5 \,\mu$ l ($2.5 \,\text{U}$) FastStart Taq DNA Polymerase (Roche Diagnostics, Mannheim, Germany), $14 \,\mu$ l PCR-grade water and 100 ng template DNA. PCR amplifications were carried out with an initial denaturation step at 94°C for 3 min, 35-45 cycles of 94°C for 30 s, primer-specific annealing temperature for 30 s, 72°C for 60 s, and a final extension step at 72°C for 10 min. Bisulfite pyrosequencing was performed on a PSQ96MA Pyrosequencing System (Biotage, Uppsala, Sweden) with the PyroGold SQA reagent kit (Biotage) (45). The Pyro Q-CpG software (Biotage) was used for data analysis. To demonstrate the reliability of our quantitative methylation assays, we performed duplicate tests for a subset of tissue samples (mainly samples representing outliers in the box plot analysis). The methylation difference between duplicate measurements was $2.7 \pm 1.6\%$ for H19 (12 samples tested), $1.4 \pm 0.9\%$ for MEG3 (13 samples), $2.5 \pm 1.5\%$ for *LIT1* (11 samples), $1.8 \pm 1.1\%$ for NESP55 (14 samples), $1.9 \pm 1.4\%$ for PEG3 (19 samples), $1.1 \pm 0.9\%$ for SNRPN (10 samples), $1.1 \pm 1.1\%$ for OCT4 (3 samples), and $1.1 \pm 0.6\%$ for APC (14 samples).

Data analysis

Quantitative methylation data were analyzed with SPSS version 17.0.1 (http://www.spss.com). Box plots were generated using the default parameters of SPSS. They display the location, dispersion and skewness of a set of data. The bottom and the top of the box indicate the 25th and 75th percentile, respectively. The T bars extend from the boxes to at most 1.5 times the height of the box. Outliers are samples that do not lie within these T bars, extreme outliers have values more extreme than three times the box length away from the median. Samples falling in the T bars were considered to be normally methylated, whereas outliers (and extreme outliers) may represent extreme methylation values. The χ^2 -test was used to compare the proportion of outliers in different tissues. A P < 5% was considered significant.

Statistical inference concerning the methylation data sets for the eight studied genes between different tissues was done with a multivariate analysis of variance (MANOVA), which is generalization of common analysis of variance. *T*-test was used to compare the methylation values of a particular gene between two groups.

Scatter plots were used to visualize the relationship between the DNA methylation levels (of a given gene) in unrelated blood samples and the age of the examined individuals. A linear regression was calculated to infer the dependence of methylation percentage on age. The goodness of fit measure (R^2) was used to quantify this relation. R^2 can range from +1 (perfect linear dependence) to 0 (no correlation). The *P*-value was calculated by ANOVA. It determines whether or not R^2 significantly differs from 0.

We plotted a receiver operating characteristics (ROC) curve for the six studied imprinted genes, comparing the pairwise methylation differences of monozygotic (MZ) and dizygotic (DZ) twins. Using the area under curve as a test statistic for discriminative power, the corresponding P-value is exactly that obtained by a Wilcoxon rank sum

test. In addition, a multivariate discrimination analysis (Fisher's linear discriminant analysis) including the methylation data sets of the six studied genes was performed. The non-parametric Wilcoxon test was then used to compare the pairwise methylation differences of two groups (MZ versus DZ).

PAM analysis (46) was used to assess whether the methylation pattern across seven neighboring CpG sites in the *APC* promoter is predictive of the tissue type. By pattern we mean the relative difference in the methylation values among the seven CpGs. The input data were the mean-centered methylation data of all samples.

RESULTS

Gene-, tissue- and developmental stage-dependent variation of DNA methylation

Here, we determined the methylation percentages of six DMRs (H19, MEG3, LIT1, NESP55, PEG3 and SNRPN) and two promoters (OCT4 and APC) in chorionic villi from 48 healthy newborns, in 10 fetal and 12 adult cortices, and 26 adult blood samples. As expected, imprinted genes were differentially methylated with median methylation values ranging from 37% to 53%, OCT4 was hypermethylated with median methylation values from 88% to 97%, and APC was hypomethylated with median methylation values from 4% to 7% (Table 2). The box plots in Figure 1 demonstrate that the median methylation of DMRs and promoters as well as the range of methylation variation are dependent on the gene, tissue type and developmental stage. For example, the median H19 DMR methylation ranged from 43% in adult blood to 52% in adult brain, whereas the median SNRPN methylation ranged from 36% in adult blood to 41% in CVS. Compared to the developing fetal brain, H19 and LIT1 were >5%upmethylated, whereas MEG3 was >5% downmethylated in adult brain.

When only looking at the six imprinted genes, which have been studied in all tissues, CVS displayed 5%, fetal brain 17%, adult brain 7% and adult blood samples 1% extreme DMR methylation values (Table 2). The number of outliers (including extreme outliers) in blood samples was significantly lower (χ^2 -test, P < 0.05) than in any other studied tissue. The difference between CVS and fetal brain was also significant. Three of 48 (6%) CVS, one of 10 (10%) fetal cortices, two of 13 (15%) adult cortices, but none (0%) of 25 adult blood samples displayed extreme DMR methylation values in two or more imprinted genes. In one fetal cortex 5 of the 6 studied DMRs were hypermethylated. Collectively, these data suggest that methylation imprints are relatively tightly regulated in blood, whereas CVS and brain exhibit considerably more methylation variation, consistent with a relaxation of imprinting (DMR methylation). With the notable exception of one CVS with extreme (40%) OCT4 hypomethylation, the pluripotency gene OCT4 was hypermethylated (83-98%) and the tumor suppressor gene APC was hypomethylated (2-11%) in all analyzed tissue samples.

Gene	Fetal cortex		Adult cortex		Newborn CVS		Adult blood	
	Median methylation	Number of outliers						
H19	47%	1/10	52%	0/13	44%	1/45	43%	0/22
MEG3	51%	0/9	45%	1/10	44%	1/47	46%	0/25
LIT1	46%	2/10	53%	0/13	48%	4/40	49%	0/20
NESP55	43%	1/10	39%	2/13	42%	0/33	43%	0/24
PEG3	40%	1/10	42%	1/13	45%	3/48	48%	1/25
SNRPN	40%	2/9	40%	1/13	41%	4/34	36%	0/25
OCT4	ND	ND	97%	1/13	88%	1/12	94%	1/12
APC	4%	0/10	7%	0/8	ND	ND	6%	1/26
Imprinted genes 7/58 (17%)		7/58 (17%)		5/75 (7%)		13/247 (5%)		1/141 (1%)
Samples with outliers $3/10(30\%)$		3/10 (30%)		3/13 (23%)		10/48 (21%)		3/25 (12%)
Samples with outliers in 1/10 (10%) multiple genes			2/13 (15%)		3/48 (6%)		0/25 (0%)	

Table 2. Median methylation and number of extreme methylation values in different tissues

This table presents the median methylation (of all analyzed CpGs) and the number of outliers (including extreme outliers) in the studied DMRs and promoters in chorionic villi (of healthy newborns), fetal and adult cortex, and adult blood. ND, not done.

In addition, we performed a MANOVA for each studied gene to test for differences between the methylation data sets of CVS, fetal brain, adult brain and adult blood. The null hypothesis that there are no differences between tissues was clearly rejected (Wilks' Lambda test, P < 0.001) for H19, MEG3, PEG3, SNRPN and APC. There was a trend (P = 0.05)for LIT1. Group-wise comparisons revealed significant (t-test, P < 0.05) H19 methylation differences between CVS and adult blood, between fetal cortex and adult blood, as well as between adult cortex and adult blood. MEG3 methylation differences were found between CVS and fetal cortex, between fetal and adult cortex, and between fetal cortex and adult blood; PEG3 methylation differences between CVS and adult cortex, between CVS and adult blood, and between adult cortex and blood; SNRPN methylation differences between CVS and adult blood and between adult cortex and blood; APC methylation differed between fetal and adult cortex and between fetal cortex and adult blood.

Methylation changes with age

A general decrease in DNA methylation with age has been reported in blood cells (25). When we plotted the measured methylation values of six imprinted genes in blood samples against age (between 21 and 72 years) of the donor and calculated the linear regression (Figure 2A), there was a significant decrease of methylation with age for H19 ($R^2 = 0.264$; P = 0.003 that there is no correlation between methylation and age), *MEG3* ($R^2 = 0.535$; P < 0.001), LIT1 ($R^2 = 0.241$; P = 0.006), NESP55 $(R^2 = 0.310; P = 0.001)$ and SNRPN $(R^2 = 0.133;$ P = 0.037). However, this age effect was completely lost when removing all data points between 21 and 25 years (Figure 2B). T-test revealed highly significant (P < 0.001) differences in H19, MEG3, LIT1, NESP55 and SNRPN methylation between younger individuals (up to 25 years) and those >25 years.

Methylation variation in monozygotic and dizygotic twins

Figure 3A displays the differences in DMR methylation in blood samples of 12 MZ and 14 DZ pairs. For all six studied imprinted genes, the (median) pairwise methylation difference between MZ was somewhat smaller than between DZ. Because of the small number of studied twin pairs, there was no significant difference between the MZ and DZ groups at the single gene level. However, when a ROC curve (Figure 3B) was calculated with the methylation data sets of all six studied genes (including outliers), there was a significant (Wilcoxon test, P = 0.038) difference between MZ and DZ. A multivariate discrimination analysis based on the pairwise methylation differences of the six studied genes correctly classified 86% of the MZ and 88% of the DZ. This supports the assumption that the methylation differences of imprinted genes between pairs are predictive of the mode of twinning. The age of the MZ ranged from 27 to 70 years (mean 43 years), whereas the age of the DZ ranged from 38 to 72 years (mean 55 years). Although we cannot exclude age as a confounding variable, the observation that DMR methylation is rather stable between 27 and 72 years (Figure 2B) argues against this possibility.

Methylation patterns of neighboring CpG sites

Our pyrosequencing assay for the tumor suppressor *APC* targets seven neighboring CpGs. With the notable exception of placenta (native CVS), the analyzed region of the *APC* promoter was hypomethylated in all tissues (Figure 4A). Native placental villi (n = 29) displayed a much higher CGI methylation (median 37%) than cultured CVS (n = 12, median 10%). The maternal *APC* allele has been reported to be methylated in the placenta (47), implying that cell culture leads to a loss of *APC* methylation (maternal imprinting). Adult cortex (n = 12), blood (n = 26), and sperm (n = 45) samples displayed median methylation values of 8%, 6% and 4%, respectively. By visual inspection of the fluctuation



Figure 1. Median methylation and range of methylation variation of two paternally methylated DMRs, *H19* and *MEG3*, and four maternally methylated DMRs, *LIT1*, *NESP55*, *SNRPN* and *PEG3*, of imprinted genes as well as of promoters of one pluripotency gene, *OCT4*, and one tumor suppressor gene, *APC*. The box plots of a given gene show the distribution of DMR or promoter methylation values in 48 CVS from healthy newborns, 10 fetal and 12 adult brain samples, and 26 adult blood samples from unrelated individuals. The median is represented by a horizontal line. The bottom of the box indicates the 25th percentile, the top the 75th percentile. Outliers are shown as open circles, extreme outliers as stars.

of methylation levels across the seven studied CpG sites (Figure 4A), CpGs 2, 3, 6 and 7 appeared to be most predictive of the tissue type. When we used Fisher's discriminant analysis including the methylation data sets of these four CpGs 80% of the adult blood, 82% of adult cortex, 58% of cultured CVS, 100% of native CVS and 76% of sperm samples were classified correctly from the other tissues. In addition, we performed a PAM analysis (Figure 4B) to assess whether the methylation pattern across the seven analyzed CpGs (relative to each other) is predictive of the tissue type. It turned out that native CVS can be distinguished properly from the rest of the tissues. Any further distinction based upon the shrunken centroid classifier was not possible, because in the remaining cases the classifier mostly decided for 'sperm', irrespective of the true tissue origin.

DISCUSSION

In contrast to the genome of an individual which is determined for all cells of the body at the time of conception, the epigenome is highly dynamic and differs from tissue to tissue. The most dramatic epigenetic changes occur during genome-wide reprogramming in gametogenesis and early embryogenesis (48,49). Genome-wide demethylation in the primordial germ cells erases essentially all methylation patterns to ensure an equivalent epigenetic state in germ cells of both sexes. Parent-specific methylation patterns according to the sex or the germline are then established during germ-cell differention (50). Genome-wide demethylation waves in the early embryo erase most germline methylation patterns, followed by *de novo* methylation and establishment of somatic methylation patterns (48,51). In addition to the carefully directed



Figure 2. Association between age and H19, MEG3, LIT1, NESP55, SNRPN and PEG3 methylation. (A) Regression lines and data points represent measurements in blood samples of unrelated healthy individuals between 21 and 72 years. R^2 is the linear correlation coefficient. (B) Regression lines when removing all samples from individuals up to 25 years.

epigenetic reprogramming processes during development, epigenetic modifications may occur randomly or in response to environmental influences and ageing (5–7,23–25,29). Epigenetic variation may not only contribute substantially to phenotypic differences among individuals but also play an important role in the etiology of common (complex) human disease (52–54). This is obvious in most cancers which display dramatically increased methylation of tumor suppressor genes (18–20).

Although there is no simple linear relationship between DNA methylation and gene activity and the regulatory mechanisms may differ from gene to gene, in general the level and/or pattern of methylation in *cis*-regulatory regions is thought to control chromatin conformation and transcriptional potential. Threshold models suggest that changes shifting CpG methylation above a critical density lead to gene silencing (27,55). Our quantitative study revealed considerable variation in DNA methylation in hypomethylated, differentially methylated or hypermethylated *cis*-regulatory regions that are directly or indirectly involved in the control of gene activity. Both the average methylation level and the range of methylation variation depend on gene locus, tissue

type and development stage, and age. It is noteworthy that the analyzed CVS and blood samples were from healthy newborns and adults; the brain samples were from fetuses and adults without detectable brain pathology. The lowest and the highest DMR methylation values in individual samples were 36% (in an adult blood) and 65% (in a fetal brain) for H19, 25% (in an adult brain) and 75% (in a fetal brain) for MEG3, 32% (in a CVS) and 67% (in a fetal brain) for LIT1, 24% and 68% (both in adult brain) for NESP55, 37% (in an adult brain) and 64% (in a fetal brain) for PEG3, and 21% (in a fetal brain) and 55% (in a CVS) for SNRPN. These functionally important methylation patterns exhibit a high degree of interindividual variability that by far exceeds DNA sequence variation. This enormous epigenetic variability cannot be explained by experimental noise, which is in the order of 1-2% discrepancy between the measured methylation levels of technical replicates.

Most epimutations underlying the variation in DNA methylation may occur after fertilization and probably represent stochastic errors in the establishment or maintenance of an epigenetic state. If an epimutation occurs



Figure 3. Differences in gene-specific methylation levels between pairs of monozygotic (MZ) and dizygotic twins (DZ). (A) The box plots display the pairwise difference in DMR methylation of H19, MEG3, LIT1, NESP55, SNRPN and PEG3 between MZ (12 pairs) and DZ (14 pairs). The median methylation difference is represented by a horizontal line. (B) ROC curve comparing the pairwise methylation differences of MZ and DZ. The area under the curve indicates the discriminative power of pairwise methylation differences in six studied genes between the MZ and DZ groups.

in the germline, all cells of the individual should be affected, leading to clear DMR hypomethylation (<10%) or hypermethylation (>90%) in our pyrosequencing assay. In contrast, if an epimutation occurs during embryogenesis or later in life, it affects only a subset of cells, resulting in somatic mosaicism. The wide range of methylation variation in *cis*-regulatory regions suggests the possibility that basically each cell in a given tissue may be unique in its epigenomic profile. Thus, epigenetic analyses are not only hampered by the fact that only few tissues such as blood are readily accessible, but also by the high degree of epigenetic mosaicism. Because bisulfite treatment heavily degrades DNA, methylation analyses of single cells are both technically challenging and time-consuming (56,57), and therefore not particularly suitable for studying cell-to-cell variability in larger cell populations. Currently available quantitative methylation assays such as bisulfite pyrosequencing provide only a rough estimate on the percentage of normally versus extremely (hypo- or hyper)methylated cells in a tissue. In addition, when interpreting the medical relevance of our data, i.e. whether or not the observed methylation differences influence gene expression, we have to take into account that the analyzed tissues consist of different cell types and heterogeneous cell populations. From single cell expression analyses it becomes increasingly clear that the average cell does not exist (58,59).

With the notable exception of native chorionic villi, which show differential methylation of the *APC* promoter (47), several hundred analyzed CpGs in cultured CVS, adult brain, blood and sperm all displayed <20% methylation. Average promoter methylation ranged from

4% in sperm to 10% in cultured CVS. Interestingly, the methylation patterns, that is the relative differences in methylation among seven analyzed neighboring CpGs also appeared to differ among tissues. For example, in adult blood with an average CGI methylation of 6% the median methylation decreased from 10% at CpG 6 to 4% at CpG 7, whereas in adult cortex with an average CGI methylation of 7% it increased from 7% at CpG 6 to 10% at CpG 7. This tissue-specific fluctuation between neighboring CpGs is much higher than the technical noise $(1.1 \pm 0.6\%)$ between replicates) of our APC pyrosequencing assay and, therefore may represent a true biological phenomenon, i.e. tissue- and developmental stage-dependent differences in nucleosome positioning (60) and/or binding of specific transcription factors (61). By discriminant analysis based on only a few CpG sites in the APC promoter, we can correctly predict the tissue type in most analyzed samples. More sophisticated algorithms including multiple genes might considerably improve discriminative power.

Collectively, our results highlight substantial variation of DNA methylation at *cis*-regulatory regions among apparently healthy individuals. Although the observed hypo- and hypermethylation values may represent normal variation, this does not necessarily imply that they are biologically irrelevant. In a conceptually related study on imprinted gene methylation in human abortions and stillbirths (fetal muscle samples), we found that *MEG3* hypermethylation was associated with downregulation of this gene and *H19* hypermethylation increased the likelihood for biallelic expression of the autocrine growth factor *IGF2* (62). In this context, it is noteworthy that altered DMR methylation patterns and



Figure 4. (A) Methylation patterns of seven neighboring CpG sites in the *APC* promoter region in adult blood (n = 26), adult cortex (n = 12), cultured (n = 12) and native (n = 29) CVS from first trimester abortions, and sperm (n = 45) samples. The box plots display the distribution of methylation values at each CpG site. Please note the different graduations of scale on the y-axis. (B) PAM analysis of methylation patterns across the seven CpGs. Each sample is represented by one column. PAM uses cross-validation to calculate posterior class probabilities for each sample. A classifier which is based upon this would vote for the class with the highest posterior in a given sample. Native CVS and sperm samples would be classified correctly to a large extent, whereas almost all other samples would be misclassified as sperm.

relaxation of imprinting have been described in normal populations. For example, about 10% of blood samples from healthy individuals showed biallelic *IGF2* expression (63). Loss of *IGF2* imprinting occurs in colorectal cancer and many other tumor types. The observation that *IGF2* was also dysregulated in normal mucosa and peripheral blood lymphocytes of these cancer patients (64) promotes the idea that relaxation of imprinting may predispose to tumorigenesis and, by extrapolation, other complex diseases. Most likely, the role of epigenetic changes for phenotypic variation and disease susceptibility is largely underestimated. Further research is needed to determine the functional implications of the enormous DNA methylation variation for the fine tuning of gene regulation.

ACKNOWLEDGEMENTS

We thank Daniela Weise for excellent technical assistance.

FUNDING

Funding for open access charge: German Research Foundation (HA 1374/8-1).

Conflict of interest statement. None declared.

REFERENCES

- Gärtner, K. and Baunack, E. (1981) Is the similarity of monozygotic twins due to genetic factors alone? *Nature*, 292, 646–764.
- Gärtner,K. (1990) A third component causing random variability beside environment and genotype. A reason for the limited success of a 30 year long effort to standardize laboratory animals?. *Lab Anim.*, 24, 71–77.
- 3. Yanagimachi, R. (2002) Cloning: experience from the mouse and other animals. *Mol. Cell. Endocrinol.*, **187**, 241–248.
- Wong,A.H., Gottesman,I.I. and Petronis,A. (2005) Phenotypic differences in genetically identical organisms: the epigenetic perspective. *Hum. Mol. Genet.*, 14, R11–18.
- Sutherland, J.E. and Costa, M. (2003) Epigenetics and the environment. Ann. NY Acad. Sci., 983, 151–160.

- Waterland, R.A. and Jirtle, R.L. (2003) Transposable elements: targets for early nutritional effects on epigenetic gene regulation. *Mol. Cell. Biol.*, 23, 5293–5300.
- 7. Horsthemke, B. (2006) Epimutation in human disease. Curr. Top. Microbiol. Immunol., 310, 45–59.
- Bennett-Baker, P.E., Wilkowski, J. and Burke, D.T. (2003) Age-associated activation of epigenetically repressed genes in the mouse. *Genetics*, 165, 2055–2062.
- 9. Goyal, R., Reinhardt, R. and Jeltsch, A. (2006) Accuracy of DNA methylation pattern preservation by the Dnmt1 methyltransferase. *Nucleic Acids Res.*, **34**, 1182–1188.
- Whitelaw, N.C. and Whitelaw, E. (2008) Transgenerational epigenetic inheritance in health and disease. *Curr. Opin. Genet. Dev.*, 18, 273–279.
- Jaenisch, R. and Bird, A. (2003) Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat. Genet.*, 33, 245–254.
- Vaissière, T., Sawan, C. and Herceg, Z. (2008) Epigenetic interplay between histone modifications and DNA methylation in gene silencing. *Mutat. Res.*, 659, 40–48.
- Rollins, R.A., Haghighi, F., Edwards, J.R., Das, R., Zhang, M.Q., Ju, J. and Bestor, T.H. (2006) Large-scale structure of genomic methylation patterns. *Genome Res.*, 16, 157–163.
- Feinberg, A.P. (2007) Phenotypic plasticity and the epigenetics of human disease. *Nature*, 447, 433–440.
- Weber, M., Hellmann, I., Stadler, M.B., Ramos, L., Pääbo, S., Rebhan, M. and Schübeler, D. (2007) Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat. Genet.*, 39, 457–466.
- Okita,K., Ichisaka,T. and Yamanaka,S. (2007) Generation of germline-competent induced pluripotent stem cells. *Nature*, 448, 313–317.
- 17. Reik,W. (2007) Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature*, **447**, 425–432.
- Esteller, M., Sparks, A., Toyota, M., Sanchez-Cespedes, M., Capella, G., Peinado, M.A., Gonzalez, S., Tarafa, G., Sidransky, D., Meltzer, S.J. *et al.* (2000) Analysis of adenomatous polyposis coli promoter hypermethylation in human cancer. *Cancer Res.*, 60, 4366–4371.
- Ting,A.H., McGarvey,K.M. and Baylin,S.B. (2006) The cancer epigenome - components and functional correlates. *Genes Dev.*, 20, 3215–3231.
- Esteller, M. (2008) Epigenetics in cancer. N. Engl. J. Med., 358, 1148–1159.
- 21. Bartolomei, M.S. and Tilghman, S.M. (1997) Genomic imprinting in mammals. *Annu. Rev. Genet.*, **31**, 493–525.
- 22. Ferguson-Smith,A.C. and Surani,M.A. (2001) Imprinting and the epigenetic asymmetry between parental genomes. *Science*, **293**, 1086–1089.
- 23. Siegmund,K.D., Connor,C.M., Campan,M., Long,T.I., Weisenberger,D.J., Biniszkiewicz,D., Jaenisch,R., Laird,P.W. and Akbarian,S. (2007) DNA methylation in the human cerebral cortex is dynamically regulated throughout the life span and involves differentiated neurons. *PLoS ONE*, 2, e895.
- 24. Bjornsson,H.T., Sigurdsson,M.I., Fallin,M.D., Irizarry,R.A., Aspelund,T., Cui,H., Yu,W., Rongione,M.A., Ekström,T.J., Harris,T.B. *et al.* (2008) Intra-individual change over time in DNA methylation with familial clustering. *J. Am. Med. Assoc.*, **299**, 2877–2883.
- 25. Bollati, V., Schwartz, J., Wright, R., Litonjua, A., Tarantini, L., Suh, H., Sparrow, D., Vokonas, P. and Baccarelli, A. (2009) Decline in genomic DNA methylation through aging in a cohort of elderly subjects. *Mech. Ageing Dev.*, **130**, 234–239.
- 26. Farcas, R., Schneider, E., Frauenknecht, K., Kondova, I., Bontrop, R., Bohl, J., Navarro, B., Metzler, M., Zischler, H., Zechner, U. *et al.* (2009) Differences in DNA methylation patterns and expression of the CCRK gene in human and non-human primate cortices. *Mol. Biol. Evol.*, **26**, 1379–1389.
- Bock,C., Walter,J., Paulsen,M. and Lengauer,T. (2008) Inter-individual variation of DNA methylation and its implications for large-scale epigenome mapping. *Nucleic Acids Res.*, 36, e55.

- Sontag,L.B., Lorincz,M.C. and Luebeck,E.G. (2006) Dynamics, stability and inheritance of somatic DNA methylation imprints. J. *Theor. Biol.*, 242, 890–899.
- Foley, D.L., Craig, J.M., Morley, R., Olsson, C.A., Dwyer, T., Smith, K. and Saffery, R. (2009) Prospects for epigenetic epidemiology. *Am. J. Epidemiol.*, 169, 389–400.
- 30. El-Maarri, O., Becker, T., Junen, J., Manzoor, S.S., Diaz-Lacava, A., Schwaab, R., Wienker, T. and Oldenburg, J. (2007) Gender specific differences in levels of DNA methylation at selected loci from human total blood: a tendency toward higher methylation levels in males. *Hum. Genet.*, **122**, 505–514.
- 31. Mackay, D.J., Boonen, S.E., Clayton-Smith, J., Goodship, J., Hahnemann, J.M., Kant, S.G., Njølstad, P.R., Robin, N.H., Robinson, D.O., Siebert, R. *et al.* (2006) A maternal hypomethylation syndrome presenting as transient neonatal diabetes mellitus. *Hum. Genet.*, **120**, 262–269.
- 32. White,H.E., Durston,V.J., Harvey,J.F. and Cross,N.C. (2006) Quantitative analysis of SNRPN gene methylation by pyrosequencing as a diagnostic test for Prader-Willi syndrome and Angelman syndrome. *Clin. Chem.*, **52**, 1005–1013.
- Deb-Rinker, P., Ly, D., Jezierski, A., Sikorska, M. and Walker, P.R. (2005) Sequential DNA methylation of the Nanog and Oct-4 upstream regions in human NT2 cells during neuronal differentiation. J. Biol. Chem., 280, 6257–6260.
- 34. Schatz, P., Distler, J., Berlin, K. and Schuster, M. (2006) Novel method for high throughput DNA methylation marker evaluation using PNA-probe library hybridization and MALDI-TOF detection. *Nucleic Acids Res.*, **34**, e59.
- 35. Zechner, U., Kohlschmidt, N., Rittner, G., Damatova, N., Beyer, V., Haaf, T. and Bartsch, O. (2009) Epimutation at human chromosome 14q32.2 in a boy with a upd(14)mat-like clinical phenotype. *Clin. Genet.*, **75**, 251–258.
- 36. Sasaki, H., Ishihara, K. and Kato, R. (2000) Mechanisms of Igf2/ H19 imprinting: DNA methylation, chromatin and long-distance gene regulation. J. Biochem., 127, 711–715.
- 37. Gicquel, C., Rossignol, S., Cabrol, S., Houang, M., Steunou, V., Barbu, V., Danton, F., Thibaud, N., Le Merrer, M., Burglen, L. *et al.* (2005) Epimutation of the telomeric imprinting center region on chromosome 11p15 in Silver-Russell syndrome. *Nat. Genet.*, **37**, 1003–1007.
- Cooper,W.N., Luharia,A., Evans,G.A., Raza,H., Haire,A.C., Grundy,R., Bowdin,S.C., Riccio,A., Sebastio,G., Bliek,J. *et al.* (2005) Molecular subtypes and phenotypic expression of Beckwith-Wiedemann syndrome. *Eur. J. Hum. Genet.*, **13**, 1025–1032.
- 39. Mitsuya,K., Meguro,M., Lee,M.P., Katoh,M., Schulz,T.C., Kugoh,H., Yoshida,M.A., Niikawa,N., Feinberg,A.P. and Oshimura,M. (1999) LIT1, an imprinted antisense RNA in the human KvLQT1 locus identified by screening for differentially expressed transcripts using monochromosomal hybrids. *Hum. Mol. Genet.*, 8, 1209–1217.
- Horsthemke, B. and Buiting, K. (2006) Imprinting defects on human chromosome 15. Cytogenet. Genome Res., 113, 292–299.
- 41. Huang, J.M. and Kim, J. (2009) DNA methylation analysis of the mammalian PEG3 imprinted domain. *Gene*, **442**, 18–25.
- 42. Plagge, A. and Kelsey, G. (2006) Imprinting at the GNAS locus. *Cytogenet. Genome Res.*, **113**, 178–187.
- 43. Surani, M.A., Hayashi, K. and Hajkova, P. (2007) Genetic and epigenetic regulators of pluripotency. *Cell*, **128**, 747–762.
- 44. Monk,M., Hitchins,M. and Hawes,S. (2008) Differential expression of the embryo/cancer gene ECSA(DPPA2), the cancer/ testis gene BORIS and the pluripotency structural gene OCT4, in human preimplantation development. *Mol. Hum. Reprod.*, 14, 347–355.
- 45. Tost,J., Dunker,J. and Gut,I.G. (2003) Analysis and quantification of multiple methylation variable positions in CpG islands by pyrosequencing. *Biotechniques*, 35, 152–156.
- 46. Tibshirani, R., Hastie, T., Narasimhan, B. and Chu, G. (2002) Diagnosis of multiple cancer types by shrunken centroids of gene expression. *Proc. Natl Acad. Sci. USA*, **99**, 6567–6572.
- Guilleret, I., Osterheld, M.C., Braunschweig, R., Gastineau, V., Taillens, S. and Benhattar, J. (2009) Imprinting of tumor-suppressor genes in human placenta. *Epigenetics*, 4, 62–68.

- 48. Reik,W., Dean,W. and Walter,J. (2001) Epigenetic reprogramming in mammalian development. *Science*, **293**, 1089–1093.
- 49. Haaf, T. (2006) Methylation dynamics in the early mammalian embryo: implications of genome reprogramming defects for development. *Curr. Top. Microbiol. Immunol.*, **310**, 13–22.
- Hajkova, P., Erhardt, S., Lane, N., Haaf, T., El-Maarri, O., Reik, W., Walter, J. and Surani, M.A. (2002) Epigenetic reprogramming in mouse primordial germ cells. *Mech. Dev.*, **117**, 15–23.
- Mayer, W., Niveleau, A., Walter, J., Fundele, R. and Haaf, T. (2000) Demethylation of the zygotic paternal genome. *Nature*, 403, 501–502.
- Abdolmaleky,H.M., Smith,C.L., Faraone,S.V., Shafa,R., Stone,W., Glatt,S.J. and Tsuang,M.T. (2004) Methylomics in psychiatry: modulation of gene-environment interactions may be through DNA methylation. *Am. J. Med. Genet. B*, **127**, 51–99.
- Hatchwell, E. and Greally, J.M. (2007) The potential role of epigenomic dysregulation in complex human disease. *Trends Genet.*, 23, 588–595.
- 54. Mill,J., Tang,T., Kaminsky,Z., Khare,T., Yazdanpanah,S., Bouchard,L., Jia,P., Assadzadeh,A., Flanagan,J., Schumacher,A. *et al.* (2008) Epigenomic profiling reveals DNA-methylation changes associated with major psychosis. *Am. J. Hum. Genet.*, **82**, 696–711.
- 55. Hsieh, C.L. (1994) Dependence of transcriptional repression on CpG methylation density. *Mol. Cell. Biol.*, **14**, 5487–5494.
- Kerjean, A., Vieillefond, A., Thiounn, N., Sibony, M., Jeanpierre, M. and Jouannet, P. (2001) Bisulfite genomic sequencing of microdissected cells. *Nucleic Acids Res.*, 29, e106.
- 57. Geuns, E., De Rycke, M., van Steirteghem, A. and Libaers, I. (2003) Methylation imprints of the imprint control region of the

SNRPN-gene in human gametes and preimplantation embryos. *Hum. Mol. Genet.*, **12**, 2873–2879.

- Levsky, J.M. and Singer, R.H. (2003) Gene expression and the myth of the average cell. *Trends Cell Biol.*, 3, 4–6.
- May, A., Kirchner, R., Müller, H., Hartmann, P., El Hajj, N., Tresch, A., Zechner, U., Mann, W. and Haaf, T. (2009) Multiplex RT-PCR expression analysis of developmentally important genes in individual mouse preimplantation embryos and blastomeres. *Biol. Reprod.*, 80, 194–202.
- Dodd, I.B., Micheelsen, M.A., Sneppen, K. and Thon, G. (2007) Theoretical analysis of epigenetic cell memory by nucleosome modification. *Cell*, **129**, 813–822.
- 61. Pant, V., Kurukuti, S., Pugacheva, E., Shamsuddin, S., Mariano, P., Renkawitz, R., Klenova, E., Lobanenkov, V. and Ohlsson, R. (2004) Mutation of a single CTCF target site within the H19 imprinting control region leads to loss of Igf2 imprinting and complex patterns of de novo methylation upon maternal inheritance. *Mol. Cell. Biol.*, 24, 3497–3504.
- Pliushch,G., Schneider,E., Weise,D., El Hajj,N., Tresch,A., Seidmann,L., Coerdt,W., Müller,A.M., Zechner,U. and Haaf,T. (2010) Extreme methylation values of imprinted genes in human abortions and stillbirths. *Am. J. Pathol.*, in press (PMID 20093482)
- Sakatani, T., Wei, M., Katoh, M., Okita, C., Wada, D., Mitsuya, K., Meguro, M., Ikeguchi, M., Ito, H., Tycko, B. et al. (2001) Epigenetic heterogeneity at imprinted loci in normal populations. Biochem. Biophys. Res. Commun., 283, 1124–1130.
- 64. Cui,H., Horon,I.L., Ohlsson,R., Hamilton,S.R. and Feinberg,A.P. (1998) Loss of imprinting in normal tissue of colorectal cancer patients with microsatellite instability. *Nat. Med.*, 4, 1276–1280.