Hum Genet (2007) 122:505–514
DOI 10.1007/s00439-007-0430-3

ORIGINAL INVESTIGATION

Gender specific differences in levels of DNA methylation at selected loci from human total blood: a tendency toward higher methylation levels in males

Osman El-Maarri · Tim Becker · Judith Junen · Syed Saadi Manzoor · Amalia Diaz-Lacava · Rainer Schwaab · Thomas Wienker · Johannes Oldenburg

Received: 4 June 2007 / Accepted: 30 August 2007 / Published online: 13 September 2007

Abstract Abnormal patterns of DNA methylation are observed in many diseases such as tumors and imprinting disorders. Little is known about inter-individual and gender specific variations. Here, we report on accurate and sensitive quantitative measurements of methylation in DNA from total blood in 96 healthy human males and 96 healthy human females. Global methylation was estimated by studying two repetitive DNA elements, namely *Line-1* and *Alu* repeats, while single loci were investigated for three differentially methylated regions (DMRs) at *PEG3*, *NESP55* and *H19* imprinted genes and two additional loci at Xq28 (*F8* gene) and at 19q13.4 (locus between *PEG3* and ubiquitin specific protease 29). We observed inter-individual correlations in the degree of methylation between *Alu* and *Line-1* repeats. Moreover, all studied CpGs showed slightly higher methylation in males ($P < 0.0003$–0.0381), with the exception of DMRs at imprinted genes ($P = 0.0342$–0.9616) which were almost equally methylated in both sexes with only a small tendency towards higher methylation in males. This observed difference could be due to the process of X chromosome inactivation or merely to the presence of an additional X chromosome in female cells or could be a result of downstream effects of sex determination.

Electronic supplementary material The online version of this article (doi:10.1007/s00439-007-0430-3) contains supplementary material, which is available to authorized users.

O. El-Maarri (✉) · J. Junen · S. S. Manzoor · R. Schwaab · J. Oldenburg
Institute of Experimental Hematology and Transfusion Medicine, University of Bonn, Sigmund-Freud-Str. 25, 53127 Bonn, Germany
e-mail: osman.elmaarri@ukb.uni-bonn.de

T. Becker · A. Diaz-Lacava · T. Wienker
Institute of Medical Biometry, Informatics and Epidemiology (IMBIE), University of Bonn, Sigmund-Freud-Str. 25, 53127 Bonn, Germany

Introduction

DNA methylation constitutes an essential epigenetic mark that influences a wide variety of biological processes including gene expression, chromosomal stability, imprinting and cellular differentiation. Abnormal methylation patterns are associated with many human diseases such as tumors and imprinting disorders. In addition to diseases, inter-individual variations are also reported in healthy individuals at *Alu* repeats from whole blood DNA (Sandovici et al. 2005). On the germ cell level, Flanagan et al. (2006) reported on intra- and inter-individual epigenetic variation in human sperm cells within promoter CpG islands, pericentromeric satellites repeats and even age related variations at several loci. Moreover, an international project aiming at deciphering the methylation patterns in the human genome, the "Human Epigenome Project", provided evidence of considerable inter-individual variations and of tissue specific methylations (Rakyan et al. 2004).

In the context of the above mentioned variations in healthy individuals, the effect of gender has also been widely discussed. However, conflicting data exist on sex-dependent variations in levels of DNA methylation (Fuke et al. 2004; Sandovici et al. 2005; Sarter et al. 2005; Eckhardt et al. 2006). The available data could be divided into two groups depending on the type of analysis: the global methylation content and the locus specific one.

The global content of 5-methyl cytosine could be accurately measured by methods like total digestion of genomic

DNA to nucleotide or nucleoside levels and further separation by HPLC or by capillary electrophoresis (Kuo et al. 1980; Ramsahoye 2002; Stach et al. 2003; Fraga et al. 2005); data derived from such global methods show a tendency towards higher methylation levels in males (Fuke et al. 2004; Shimabukuro et al. 2006). These results are surprising when considering that in female cells, the inactivation of one X chromosome is accompanied by methylation of CpGs islands on the inactive X chromosome (Norris et al. 1991). However, this controversy could be explained by recent studies that showed hypomethylation on the inactive X chromosome, especially at gene poor-regions, hypermethylation of CpG islands at gene rich regions while the active X chromosome is hypermethylated in the body of genes; but still *Line-1* were found equally methylated on both active and inactive X-chromosome in female somatic cells (Hansen 2003; Weber et al. 2005; Wilson et al. 2006; Hellman and Chess 2007).

While a small number of studies addressed the accurate quantitative measurements of methylation at site specific loci and its correlation with gender, not all results were consistent. Sandovici et al. (2005) found no sex or age effect on the methylation of specific *Alu* repeats. Accordingly, Eckhardt et al. (2006), studying 2,524 different loci on chromosome 6, 20 and 22 in 12 different tissues and from 43 different samples, could not find any statistical differences between male and female samples. However, the heterogeneity of the type of tissues and the small number of samples make it difficult to draw conclusive results on the sex or age difference, especially when taking into consideration the possibilities of inter-individual differences that could mask small trends of sex or age dependent variations. On the other hand, Sarter et al. (2005) observed statistically significant differences between male and female lymphocytes. Using a quantitative, sensitive assay (Eads et al. 2000) they studied promoter regions of four different autosomal loci (*MTHFR*, *CALCA*, *MGMT* and *ESR1*) from a relatively large number of individuals (134 males and 157 females). The first three loci showed higher methylation for males, while *ESR1* methylation levels were not gender dependent as it was nearly unmethylated in both gender.

Furthermore, some indirect evidence also points to a trend of lower levels of methylation in female cells. Zvetkova et al. (2005) showed that ES cell lines derived from female mice are hypomethylated in comparison to ES cell lines from male mice. In addition, imprinted genes in primordial germ cells (PGC) were shown to be higher methylated in XY cells than in XX cells (Durcova-Hills et al. 2004, 2006). In these studies, the cells tested were in embryogenic stages where both X chromosomes in XX cells are still active. The authors proposed the presence of a locus on the X chromosome that is responsible for this lower methylation activity.

Therefore, to resolve this controversial issue, in this study we addressed the question whether gender specific differences exist in human somatic cells at specific loci. We analyzed homogenous samples derived from uncultured blood cells originated from matched age individuals (about 24-year old; 96 males and 96 females). The regions studied include two repetitive elements *Alu* and *Line-1* that cover the whole genome, three imprinted loci *PEG3*, *NESP55* and *H19* and two single loci gene: *F8* at Xq28 and a CpG rich region between paternally expressed gene 3 and ubiquitin specific protease 29 at second exon of a theoretical 4 exons gene at 19q13.4 (Djuric et al. 2006). The two single loci at Xq28 and 19q13.4 were chosen based on our previous knowledge that these two regions are not fully methylated, are mosaic in their patterns of methylation and show some inter-individual variations. In addition, we studied global methylation levels at CCGG sequences by the LUMA method (Karimi et al. 2006). As a result of this study, we report the novel observations that repetitive elements, namely *Line-1* and *Alu* repeats, and the two single loci showed a slight but clear tendency and statistically significant bias towards higher methylation in males while a bias towards higher methylation in males at imprinted loci was observed, but this was not statistically significant. Moreover, clear inter-individual associations in the level of methylation between *Alu* and *Line-1* repeats were observed.

Materials and methods

DNA from blood donors

Blood samples collected from 96 healthy women and 96 healthy men were obtained from blood donors at the Institute of Experimental Hematology and Transfusion Medicine, Bonn, Germany. DNA was extracted by a salting out method according to standard procedures. The individuals were matched for their age with averages of males and females of 24.80 ± 3.37 and 24.30 ± 3.54, respectively (Supplementary Table 1). The Ethics Committee of the Medical Faculty at the University of Bonn approved the use of the above-mentioned material for research purposes (Ethics Committee approval Nr. 106/05).

Global genomic methylation contents by the LUMA method

Luminescence methylation assay (LUMA) was essentially performed as previously described by Karimi et al. (2006). Briefly, 500 ng of DNA to be measured was cleaved with *Hpa*II + *Eco*R1 or *Msp*I + *Eco*R1 (5 U from each enzyme) in two separate tubes. The total volume of reaction was

20 µl. The tubes were incubated at 37°C for exactly 4 h. Next, 20 µl annealing buffer (20 mM Tris–acetate, 2 mM Mg–acetate pH 7.6) was added and samples were measured in a PSQ96™MA pyrosequencing machine (Biotage AB, Uppsala, Sweden). *Eco*R1 leaves 5′-TTAA sticky ends on the DNA that can be extended by As and Ts, while *Hpa*II and *Msp*I leaves 5′-GC that can be extended by Cs and Gs. Therefore, the machine was programmed to first add As, then Gs and Cs. The first would generate peaks corresponding to the *Eco*RI digestion which can be used as internal control for the quantity of DNA and the efficiency of cleavages, while the second would give peaks corresponding to all CCGG sites (for *Msp*I) or for the unmethylated portion of CCGG (for *Hpa*II). Next, the percentage of methylated CCGG sites was calculated according to the formula: Meth% = [1 − (peak *Hpa*II/peak *Eco*R1)/(peak *Msp*I/peak *Eco*R1)] × 100.

Bisulfite modification and PCRs

Bisulfite treatment of the DNA was performed essentially as previously described (El-Maarri et al. 2004). Primers used in this study are listed in Supplementary Table 2.

Site-specific methylation by SIRPH analysis

The site-specific methylation measurements at selected CpG sites were done by the SIRPH method as previously described by El-Maarri et al. (2002), El-Maarri (2004). SNuPE primers used in this study are given in Supplementary Table 2.

Pyrosequencing

The pyrosequencing analysis was performed according to Tost et al. (2003). The PCRs were generated using the same primers used in the bisulfite PCR (Supplementary Table 2) with a biotin labeled reverse primer that is needed to generate single stranded DNA for analysis. The reactions were performed using the Pyro Gold reagents from Biotage (Uppsala, Sweden) on a PyroMark ID system. The primers used for pyrosequencing are the same as SN1 (for *H19*, *PEG3* and *NESP55*) used by the SIRPH primer (Supplementary Table 2).

Cloning of individual PCR products

Individual PCR products were purified to remove unreacted primers and excess dNTPs by column purification using the QIAquick PCR purification kit (catalogue number: 28106; QIAGEN, Hilden, Germany). The purified PCR product was then ligated to a pGEM-T cloning vector (pGEM-T Vector System 1, catalogue number A3600, Promega, Mannheim, Germany). Individual positive clones were isolated by PCR using M13 primers and sequenced using standard protocols.

Statistical analysis

Methylation values were determined for the first group of 96 males and 96 females in a two 96-well plate format, and for the second group of 48 males and 48 females in one 96-well plate. In order to exclude bias due to experimental variation, for each methylation variable y, the measurement values y_i of y were corrected for trend within each gender group as follows: the samples were assigned natural numbers according to the order in which their methylation values were determined. These ordinals were treated as a variable x. For each methylation variable y, we computed the linear regression of y on x, separately for males and females. The measurements y_i of the variable y were adjusted according to $y_{i,\text{new}} = y_{\text{avg}} + \text{diff}_{yi,y(xi)}$. Here, y_{avg} is the average of the measurements y_i of variable y and $\text{diff}_{yi,x(yi)}$ is the difference between the observed value y_i to the value $y(x_i)$ that is expected according to the trend regression line $y(x)$. This procedure guarantees that there is no linear trend within the adjusted measurements of each subgroup. Finally, the complete sample was reconstructed from the corrected measurements of the male and female subgroup.

From the adjusted data sample, Pearson correlation coefficients $r_{y,z}$ were computed for all pairs (y,z) of methylation variables and for all pairs (y, sex) using the function *proc corr* from the SAS software package (SAS for Windows, version 9.1.; SAS Institute Inc., Cary, NC, USA; http://www.sas.com). This procedure was also used to compute P values for the null hypotheses H_0: $r_{,yz} = 0$ and H_0: $r_{y,\text{sex}} = 0$. In this way, the significance of the correlation coefficients was assessed. We applied a Bonferroni correction of 128, which is the number of considered pairs. Thus, uncorrected P values below 0.05/128 = 0.00039 were considered to be significant at size 0.05 after correction for multiple testing.

Statistical significance of sequencing data was assessed with the non-parametric Wilcoxon–Mann–Whitney test. The analysis was based on SAS software package using the function "proc npar1way". For each of the five sequencing regions (*NESP55*, *PEG3*, *H19*, *F8*, CpG island at 19q13.4), the number sites with positive methylation measurements per clone was computed. We tested the hypothesis of no difference in degree of methylation between sex groups. In case of imprinted regions *NESP55*, *PEG3*, and *H19*, clones were divided in two classes, strong methylation (more than 90% of positive methylated sites) and low methylation (rest, which shows less than 10% of positive methylated sites).

Results

In this study, we investigated locus specific methylation at selected CpGs at repetitive *Alu* and *Line-1* repeats, at three DMRs associated with imprinted loci *PEG3* (maternal alleles methylated), *NESP55* and *H19* (paternal alleles methylated) and at two single loci, in *F8* and at a locus on chromosome 19q13.4 between paternally expressed gene 3 and ubiquitin specific protease 29 at second exon of a theoretical 4 exons gene. Methylation levels at selected CpG sites were quantitatively evaluated by treating the DNA with bisulfite and subsequent quantification of the PCR for methylation by either the SNuPE combined with ion pair reverse phase HPLC (SIRPH) assay or by the pyrosequencing technique. Both methods are highly reproducible and able to accurately detect differences in methylation (El-Maarri et al. 2002; El-Maarri 2004; Tost et al. 2003).

Repetitive elements

The investigated repetitive elements were *Line-1* and *Alu* repeats. We did not amplify specific repeats at specific genomic loci but used degenerate primers that will amplify different subclasses of repeats dispersed over the whole genome; this was proven by sequencing of individual clones of cloned PCR products, no two clones were the same, this indicate the variability of the content of the PCR product that cover amplified region over all genome (Supplementary Table 3). Although such an assay is not specific for a given genomic locus, it provides a global estimate of methylation at the repeats under investigation. Using the SIRPH protocol we measured methylation at two CpG sites, SN2 and SN4, in the *Alu* consensus sequence, while four different CpG sites, SN1, SN3, SN8 and SN9, were analyzed from *Line-1* repeat sequence (Supplementary Table 1). All sites from these groups of healthy total blood DNA showed higher methylation levels for male samples than for female samples. The lowest methylation difference was 0.30% for *Alu* SN4 ($P = 0.0003$), while the highest difference was 5.80% for *Line-1* SN9 ($P < 0.0003$) (Table 1). The differences in methylation were statistically significant at all sites.

Imprinted differentially methylated regions

Two CpGs were investigated at *NESP55* and one CpG for each *H19* and *NESP55* (Supplementary Table 1). All four loci showed some small bias toward higher methylation levels in males, however, none of them reached statistical significance (*P* values ranges from 0.0342 for *PEG3* SN1 to 0.9616 for *NESP55* SN4) (Table 1). At the *H19* locus we observed some limited variations in the methylation levels that fluctuated between 35 and 48%.

Single loci

Two single loci were investigated in this study: a non-clustered group of CpG sites located in the exon 14 of the *F8* gene at Xq28 and a CpGs cluster located between *PEG3* and ubiquitin specific protease 29 at second exon of a theoretical 4 exons gene at 19q13.4. We had shown earlier that

Table 1 Mean methylation values of 96 healthy males and 96 healthy females at *NESP55*, *PEG3*, *H19*, *F8*, two CpGs at 19q13.4, *Alu* repeats and *Line-1* repeats

Region	Primer	CpG #	Males		Females		Mean diff	*P* values	Holm adjusted *P*
			Mean	SD	Mean	SD			
NESP55	SN1	3	35.26	2.38	34.84	2.96	0.42	0.283492	0.850000
	SN4	29	45.02	4.15	44.99	4.80	0.03	0.961617	0.962000
PEG3	SN1	12	30.28	3.22	29.35	2.79	0.93	0.034288	0.171000
H19	SN1	4	42.19	2.90	41.82	2.88	0.37	0.380858	0.762000
F8	SN2	4	89.00	3.80	85.05	4.75	3.95	<0.000001	<0.000001
	SN1	7	90.94	2.92	88.26	4.68	2.68	0.000006	0.000045
	SN3	8	84.96	6.43	68.77	8.46	16.19	<0.000001	<0.000001
19q3.4	SN2	8	68.42	7.77	62.97	7.25	5.45	0.000001	0.000013
	SN3	11	62.16	10.18	59.71	4.88	2.45	0.038188	0.153000
ALU	SN2	16	37.29	1.18	35.43	1.19	1.86	<0.000001	<0.000001
	SN4	10	32.50	0.48	32.20	0.63	0.30	0.000303	0.002120
L1	SN9	1	58.82	2.06	53.02	2.00	5.80	<0.000001	<0.000001
	SN1	13	58.11	1.60	56.50	1.14	1.61	<0.000001	<0.000001
	SN8	15	73.69	3.63	69.19	4.62	4.50	<0.000001	<0.000001
	SN3	24	61.76	1.60	60.04	1.32	1.72	<0.000001	<0.000001
LUMA	–	–	0.71	0.08	0.73	0.07	−0.03	0.023650	0.142000

These values are measured by the SIRPH protocol. The values obtained by LUMA are shown in the last row. The CpG number in the amplified region is given in the third column. Statistical significance between the male/female mean differences are measured by *t*-test and *P* values are given in the last column

the later locus is hypermethylated (Djuric et al. 2006). The first locus is not known to be subject for methylation upon X chromosome inactivation as it is already methylated in about 85% in both males and females. The locus at 19q13.4 is lying close to known imprinted region containing the *PEG3/ZIM2* imprinted genes and since the extent of this imprinted domain in human is not known yet, this locus may be regarded as being present in a potentially extensive imprinted domain. Three CpGs were investigated at the former, while two were investigated at the latter locus (Supplementary Table 1). All five CpGs showed a bias towards higher methylation levels in males with differences of methylation ranging from 2.45% for SN3 at 19q13.4 ($P = 0.0381$) to 16.19% for SN3 at *F8* ($P < 0.0003$) (Table 1). Four of the five sites were found to be statistically significant for the differences between male and female samples, while SN3 at 19q13.4 was not significant ($P = 0.0381$).

Confirmation of differences in a subgroup of 48 males and 48 females

In order to assess whether the above mentioned observed gender specific differences are due to a random inter-individual variation, we repeated the bisulfite treatment and the methylation measurements by pyrosequencing in a subgroup of 48 males and 48 females (from the same healthy blood donors) for *F8* and *Line-1* repeats as well as *NESP55, PEG3*, and *H19* imprinted regions (Supplementary Table 4) (Tost et al. 2003). Here, most loci again showed higher methylation levels in males (Table 2). The significance of the differences at *Line-1* was higher in the first cohort of the 96 males and 96 females than in the subgroup of 48 males and 48 females. This was specially noticeable at the second measured position and even more at the third; the reason for this could be due to the highly variable sequences amplified by our PCR assay and the limitation in the pyrosequencing to analyze only the input sequence that we specify, in this case corresponding to the consensus sequence. The CpGs sites that are preceded directly by stretches of Ts were excluded from the analysis due to inaccuracy of correct estimation of methylation due to merging the values corresponding to the stretches of Ts with the T corresponding to unmethylated portion of DNA.

Methylation levels at CpG sites that were studied by both pyrosequencing and SIRPH methods did not show concordance in the mean levels of methylations (Tables 1, 2). Although both methods are based on single nucleotide primer extension assays however, the experimental conditions in both assays are not the same and this explain the observed differences in the values. For example the pyrosequencing requires the isolation of single stranded biotin labeled PCR product, while the SIRPH use directly the bisulfite PCR product, the reaction temperature in the pyrosequencing is 28°C and involve no cycling while the SIRPH use 70 cycles of annealing at 40°C and extension at 52°C.

Sequencing of individual clones

To visualize the nature of these gender differences, we cloned and sequenced large numbers of clones from most studied regions (*NESP55*, *PEG3*, *H19*, *F8*, *Line-1*, and CpG Island at 19q13.4) from an average of eight females and eight males. The sequencing clones show largely mosaic patterns at the non-imprinted regions, while imprinted regions show a clear division between largely methylated and largely unmethylated clones with sporadic methylated CpGs at the non-methylated clones and sporadic unmethylated CpGs at the methylated clones (Fig. 1).

Table 2 Mean methylation values of 48 healthy males and 48 healthy females at *NESP55*, *PEG3*, *H19*, *F8*, *Line-1* repeats

Region	CpG #	Males		Females		Mean diff	*P* values	Holm adjusted *P*
		Mean	SD	Mean	SD			
NESP55	3	44.61	4.25	44.12	4.54	0.48	0.595043	1.000000
	5	58.24	2.32	57.64	2.22	0.60	0.204764	1.000000
	6	56.33	5.29	55.69	5.19	0.65	0.551644	1.000000
PEG3	12	44.90	5.49	44.57	5.36	0.32	0.781735	1.000000
	14	47.76	4.91	47.17	5.14	0.59	0.585872	1.000000
H19	4	49.16	2.65	48.63	3.07	0.53	0.385526	1.000000
	6	50.37	2.3	50.31	2.31	0.05	0.911652	1.000000
	8	52.29	2.56	52.21	2.93	0.09	0.883743	1.000000
F8	7	64.65	3.12	61.21	4.13	3.45	0.000096	0.001050
	8	68.25	2.99	63.96	2.86	4.29	0.000000	0.000000
L1_SNP9	15	37.20	3.28	32.88	3.99	4.32	0.000083	0.000992
	17	53.54	5.11	48.43	5.43	5.11	0.001160	0.011600
	18	23.29	3.32	22.08	2.29	1.22	0.166757	1.000000

These samples are the first 48 groups of samples reported also in Table 1, however, the data were measured from completely different PCR products and by pyrosequencing techniques. Statistical significance between the male/female mean differences are measured by *t*-test and *P* values are given in the last column

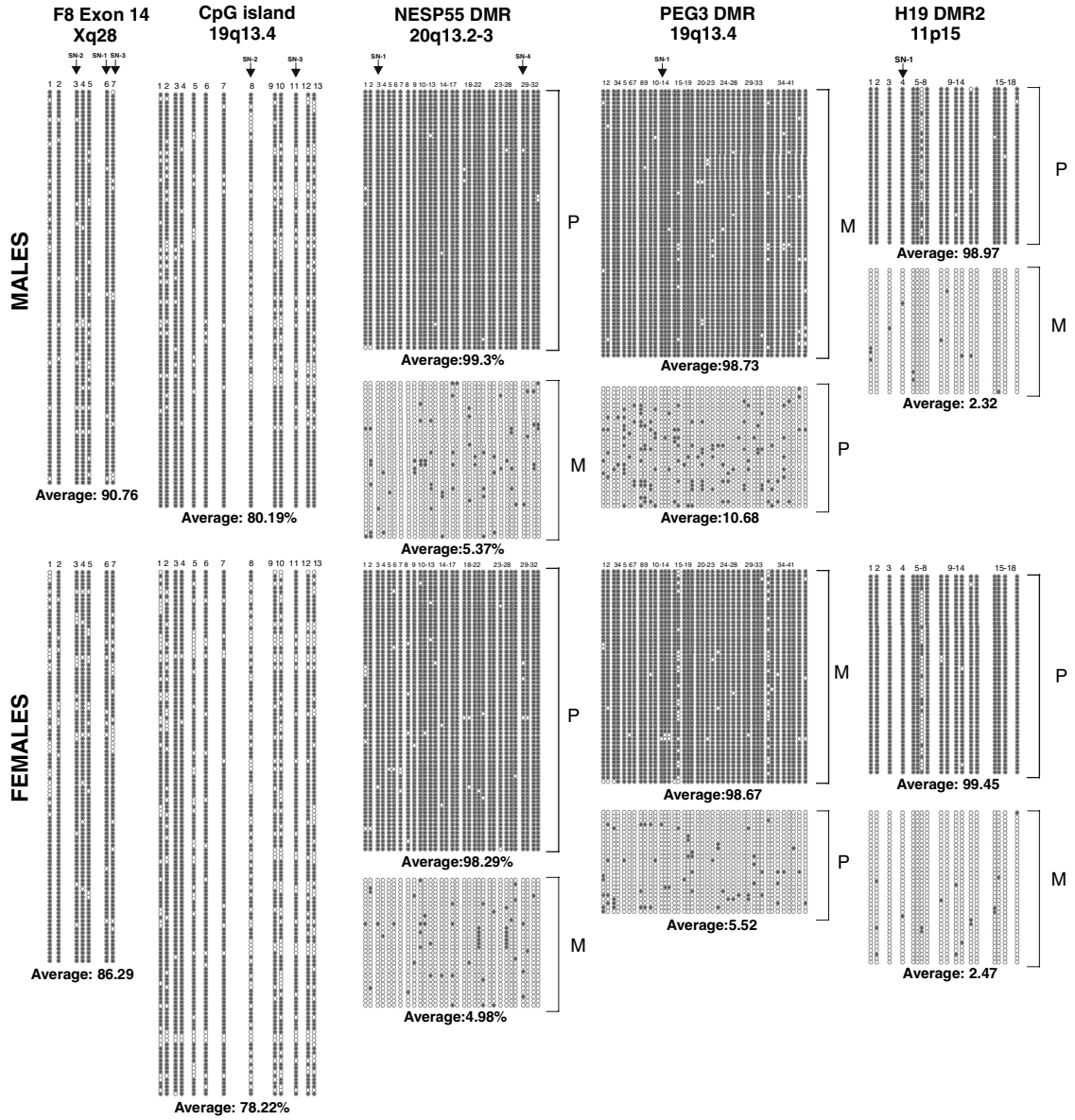

Fig. 1 Summary of the bisulfite sequencing results for males and females in two single loci (at Xq28 and 19q13.4) and three imprinted loci (*NESP55*, *PEG3*, *H19*). PCR products from eight different individuals for every gender were pooled together, cloned into a TA cloning vector and large numbers of individual clones were sequenced. Individual CpGs are represented as *filled* and *open circles* for methylated and unmethylated, respectively. Average methylation of all CpGs in all sequenced clones in a given region is provided under each group of clones, positions with known polymorphisms like CpGs 16 and 34 in *PEG3* and CpG 7 in *H19* were excluded from the calculated average. For imprinted regions, the methylated and unmethylated clones are grouped separately assuming that one parental alleles are largely methylated while the others are largely unmethylated (*P* paternal alleles; *M* maternal alleles). The positions of the CpGs studied by SIRPH reaction are indicated by *vertical arrows*, the SIRPH primer used and the number of CpGs analyzed is corresponding to Table 1

No sign of uncompleted conversion was observed as more than 99% of Cs at non CpG was converted to Ts. The average methylation difference showed a bias towards higher methylation levels in males at most loci; however, not all loci reach statistical significance. At *F8* we observed a male female difference of 4.47% ($P < 0.05$) and at the CpG island at 19q13.4 a mean differences of 1.97% ($P = 0.7540$). Regarding the unmethylated clones of imprinted differentially methylated regions (DMRs) of *NESP55* and *PEG3*, we observed higher unspecific methylation averages in males with mean differences of 0.39% ($P = 0.8818$) and 5.16% ($P < 0.05$), respectively (Fig. 1). Methylated clones were also, on average, higher methylated in males with mean differences of 1.01% ($P < 0.05$) and 1.54% ($P = 0.2651$) for *NESP55* and *PEG3*, respectively (Fig. 1). At *H19*, both the largely methylated group of clones ($P = 0.2532$) and the largely unmethylated group of clones ($P = 0.9943$) showed slightly lower methylation

in males. PCR product from *Line-1* repeat region was also cloned and individual clones were sequenced. As *Line-1* repeats showed great variety in the genome due to mutagenesis of CpG sites, no two clones were identical. This reflects the great variety of the amplified target that is not covering a specific locus but rather the whole genome. To calculate the average methylation, we considered sites that are either CG or TG at a given CpG site, from a total of 82 male and 106 female clones. The average methylations were 78.74 and 76.56% for males and females, respectively (Supplementary Table 3) clearly indicating a 2.18% higher methylation in males. Thus the sequencing results goes in line with the observed trends that we see in our quantitative results and were not unexpected.

Total methylation measurements by LUMA method

The luminescent methylation assay, or LUMA, can estimate the methylation levels in a given DNA sample at CCGG sites. LUMA analysis did reveal a slightly excessive, but not statistically significant higher methylation in females (Supplementary Table 1) (Table 1). However, it is important to mention that this assay, although global, is still limited by the CCGG sequence representing only a small fraction of all CpG sites. Hence, it does not represent the whole genome.

Inter-individual correlation between *Alu* and *Line-1* repeats

As detailed methylation values at different loci were generated during this study, we investigated the correlation with methylation levels or trends between these loci (Table 3). A significant association was observed between two repetitive elements: *Line-1* and *Alu* (*Alu* SN2 and *Line-1* SN1, SN3 and SN8) (Fig. 2; Table 3). The correlations were not restricted to one sex. Strong correlation was observed between different CpGs in the same repeat, i.e. between *Alu* SN2 and *Alu* SN4 and between the 4 CpGs at *Line-1* repeats, while less significance was observed between *Alu* SN4 and different CpG sites at *Line-1*. No significant correlation was shown with the LUMA methylation measurements, which is partially due to the fact that CCGG sites tested by the LUMA assay are not represented at the CpGs studied in the repetitive elements by the SIRPH reaction.

Discussion

About 0.8–1% of the human DNA consists of 5-methylcytosine. This level can vary slightly from one tissue to another with placenta having the lowest level, while brain tissue harbors the highest levels (Gama-Sosa et al. 1983). It is also widely accepted that a considerable portion of tumors are accompanied by a decrease in the total content of 5-methyl cytosine (Ehrlich 2006). However, it remains uncertain whether methylation levels, differ in the sexes as different studies report different findings (Fuke et al. 2004; Sandovici et al. 2005; Sarter et al. 2005; Eckhardt et al. 2006). The variability of the conclusions in different studies could be partially due to the differences in methods used, tissue samples and DNA loci studied in each report. For example, a small number of studied individuals could mask the sex dependent differences due to small inter-individual as well as assay variability. The advantage of our study is the homogeneity of the samples used in term of matching age group, the large number of samples and the "uncul-

Table 3 Summary table showing the significant correlation coefficient (Pearson correlation coefficient) of methylation at *Line-1* and *Alu* different CpGs

	Alu SN2		Alu SN4		L1 SN1		L1 SN3		L1 SN9		L1 SN8	
	F	M	F	M	F	M	F	M	F	M	F	M
Alu SN4	**0.52485** **<0.0001**	**0.5675** **<0.0001**										
L1 SN1	**0.59614** **<0.0001**	**0.51937** **<0.0001**	0.35539 0.0005	0.22735 0.0284								
L1 SN3	**0.63965** **<0.0001**	**0.47274** **<0.0001**	0.35856 0.0005	0.21712 0.0366	**0.6842** **<0.0001**	**0.8948** **<0.0001**						
L1 SN9	**0.59756** **<0.0001**	0.35338 0.005	0.36033 0.0004	0.21293 0.0416	**0.5077** **<0.0001**	**0.8297** **<0.0001**	**0.7794** **<0.0001**	**0.8115** **<0.0001**				
L1 SN8	0.1523 0.1615	0.10421 0.3256	-0.00324 0.9763	0.12194 0.2496	0.3094 0.004	**0.5144** **<0.0001**	**0.5165** **<0.0001**	**0.5928** **<0.0001**	**0.419** **<0.0001**	**0.6949** **<0.0001**		
LUMA	−0.14303 0.1691	−0.1194 0.2543	−0.20976 0.0413	−0.09339 0.3733	−0.06096 0.5638	−0.13876 0.1871	−0.14899 0.1587	−0.07806 0.4596	−0.10648 0.3097	−0.12753 0.2283	−0.03951 0.7163	−0.18707 0.0775

The methylation values at different CpG sites at *Alu* and CpG sites at *Line-1* were measured from two bisulfite PCR products. Data are derived from the 96 males and 96 females with methylation values measured by SIRPH protocol. *P* values are given below each correlation coefficient. The first column for a given region is for females (indicated by F), followed by males (indicated by M). Highly significant correlations are in bold

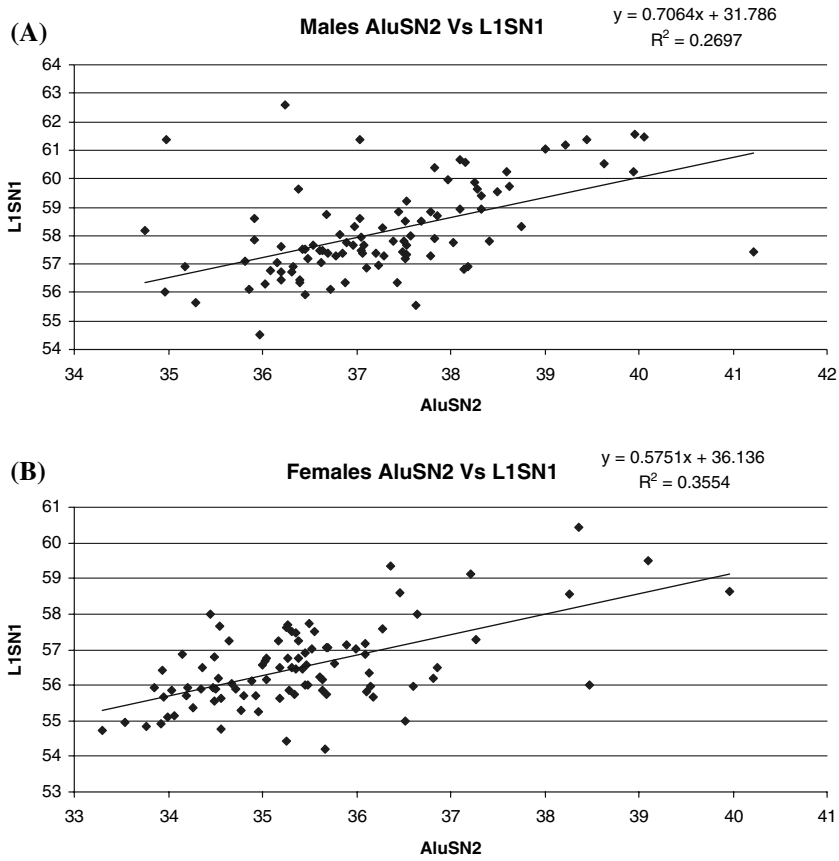

Fig. 2 Methylation values at one CpG site in *Alu* repeats (SN2) plotted against methylation values at one CpG site in *Line-1* repeats (SN1), **a** males and **b** females. 96 males and 96 females are represented; each *dot* represents the values for one healthy individual. A *trend line* is also shown together with regression coefficient. The same data are also represented in Table 2 column 1 and 2 and row 2

tured" source of DNA. Our results show slight but clear differences between both groups of male and female samples. At most loci, the average methylation in males was higher, however the degree of differences was not the same at all loci. This possibly reflects a sequence-dependent variation as a pronounced flanking sequence preference of Dnmt3a and Dnmt3b was previously detected (Handa and Jeltsch 2005).

Possible candidates responsible for the observed sex specific differences are the sex chromosomes. The presence of two X chromosomes in females could be acting in two directions: (1) a factor/gene on the X chromosome that is inducing a hypo-methylation phenotype; however, during X chromosome inactivation, this factor is reduced but not to the same level as in males due to leaky expression; or (2) The process of X chromosome inactivation by itself requires "considerable epigenetic resources" to inactivate one complete X chromosome resulting in less resources to properly methylate autosomal loci. This was suggested by Blewitt et al. (2005). Yet, no direct evidence for such a scenario is available.

Recent reports showed that CpGs in the body of genes (in contrast to CpG islands at promoter region) are more methylated on the active X chromosome (Hellman and Chess 2005; Weber et al. 2005). This would partially explain the higher male methylation we observed at the studied *F8* CpGs in exon 14. However, since the applied micro array approaches that were used in the above studies are qualitative rather than quantitative estimation of methylation levels, a direct comparison to our data is not possible. Yet, the higher male methylation at repetitive elements (*Line-1* and *Alu*) could also be influenced by the differential distribution of these repeats on the XY chromosomes (Abrusan and Krambeck 2006), but *Line-1* sequences were found to be equally methylated on both active and inactive X chromosomes in female somatic cells (Hansen 2003), therefore the difference distribution of these interspersed sequences on the sex chromosomes may not explain the observed methylation differences. What also supports this view is the lack of hypomethylated *Line-1* clones from the females sequenced clones (Supplementary Table 3).

But why DMRs of imprinted genes are not significantly affected? We can hypothesize that any factor influencing the gender specific difference in methylation levels (if any) acts during development or during the cell cycle at a time when methylation or the epigenetic marks at imprinted regions are already established. Another possibility is that the position of the imprinted genes in some specific chromosomal domains allows little variations in methylation. Such position effect was previously observed for inter-individual differences at *Alu* repeats (Sandovici et al. 2005).

In our study, we also observed a strong correlation in the level of methylation between CpG sites at *Line-1* and *Alu* repeats. Such observations have been reported in tumor samples (Weisenberger et al. 2005). Still ours is the first report on such small variations in a large number of healthy samples. These novel observations are highly suggestive for genetic and/or environmental influences that could include dietary effects. The level of methylation was shown to be affected by the levels of methyl donor supplementation in the diet (e.g. folic acid and vitamin B12). This was clearly shown during early embryogenesis at the *agouti* and at the *axin fused* loci in mice (Wolff et al. 1998; Waterland et al. 2006). Moreover, exposure to environmental pollutants, e.g., benzene, was also shown to induce slight but significant hypomethylation at *LINE-1* and *Alu*I repeats in humans (Bollati et al. 2007). The DNA samples we used were derived from young and healthy individuals who adhered to an adequate diet and whom we do not expect to have deficiencies in methyl-donor group uptake. However, neither this nor any other environmental factor could be totally excluded. A most probable scenario is that environmental as well as genetic factors play a role in modulating methylation levels in humans.

In summary, we are reporting on two novel observations. The first one establishes a gender specific differences in methylation levels with males being slightly but significantly higher methylated at the studied CpG sites with the exception of imprinted genes. The second observation indicates a strong inter-individual correlation in the level of methylation at *Alu* and *Line-1* repeats. The reasons behind both of these phenomena are not fully explored. Further studies are required to address possible factors that could be responsible.

Acknowledgments We would like to thank the blood donors for participating in this study and Doris Schmidt for her technical support with the pyrosequencing analysis.

Conflict of interest The authors declare no competing of interests.

References

Abrusan G, Krambeck HJ (2006) The distribution of L1 and Alu retroelements in relation to GC content on human sex chromosomes is consistent with the ectopic recombination model. J Mol Evol 63:484–492

Blewitt ME, Vickaryous NK, Hemley SJ, Ashe A, Bruxner TJ, Preis JI, Arkell R, Whitelaw E (2005) An *N*-ethyl-*N*-nitrosourea screen for genes involved in variegation in the mouse. Proc Natl Acad Sci USA 102:7629–7634

Bollati V, Baccarelli A, Hou L, Bonzini M, Fustinoni S, Cavallo D, Byun HM, Jiang J, Marinelli B, Pesatori AC, Bertazzi PA, Yang AS (2007) Changes in DNA methylation patterns in subjects exposed to low-dose benzene. Cancer Res 67:876–880

Djuric U, El-Maarri O, Lamb B, Kuick R, Seoud M, Coullin P, Oldenburg J, Hanash S, Slim R (2006) Familial molar tissues due to mutations in the inflammatory gene, NALP7, have normal post-zygotic DNA methylation. Hum Genet 120:390–395

Durcova-Hills G, Burgoyne P, McLaren A (2004) Analysis of sex differences in EGC imprinting. Dev Biol 268:105–110

Durcova-Hills G, Hajkova P, Sullivan S, Barton S, Surani MA, McLaren A (2006) Influence of sex chromosome constitution on the genomic imprinting of germ cells. Proc Natl Acad Sci USA 103:11184–11188

Eads CA, Danenberg KD, Kawakami K, Saltz LB, Blake C, Shibata D, Danenberg PV, Laird PW (2000) MethyLight: a high-throughput assay to measure DNA methylation. Nucleic Acids Res 28:e32

Eckhardt F, Lewin J, Cortese R, Rakyan VK, Attwood J, Burger M, Burton J, Cox TV, Davies R, Down TA, Haefliger C, Horton R, Howe K, Jackson DK, Kunde J, Koenig C, Liddle J, Niblett D, Otto T, Pettett R, Seemann S, Thompson C, West T, Rogers J, Olek A, Berlin K, Beck S (2006) DNA methylation profiling of human chromosomes 6, 20 and 22. Nat Genet 38:1378–1385

El-Maarri O, Herbiniaux U, Walter J, Oldenburg J (2002) A rapid, quantitative, non-radioactive bisulfite-SNuPE- IP RP HPLC assay for methylation analysis at specific CpG sites. Nucleic Acids Res 30:e25

El-Maarri O (2004) SIRPH analysis: SNuPE with IP-RP-HPLC for quantitative measurements of DNA methylation at specific CpG sites. Methods Mol Biol 287:195–205

El-Maarri O, Kueper M, Oldenburg J, Walter J (2004) Quantitative DNA-methylation analysis by the bisulfite conversion method. In: Weissensteiner T, Griffin HG, Griffin A (eds) PCR technology: current innovations, 2nd edn. CRC press, London, pp 175–185

Ehrlich M (2006) Cancer-linked DNA hypomethylation and its relationship to hypermethylation. Curr Top Microbiol Immunol 310:251–274

Flanagan JM, Popendikyte V, Pozdniakovaite N, Sobolev M, Assadzadeh A, Schumacher A, Zangeneh M, Lau L, Virtanen C, Wang SC, Petronis A (2006) Intra- and interindividual epigenetic variation in human germ cells. Am J Hum Genet 79:67–84

Fraga MF, Ballestar E, Paz MF, Ropero S, Setien F, Ballestar ML, Heine-Suner D, Cigudosa JC, Urioste M, Benitez J, Boix-Chornet M, Sanchez-Aguilera A, Ling C, Carlsson E, Poulsen P, Vaag A, Stephan Z, Spector TD, Wu YZ, Plass C, Esteller M (2005) Epigenetic differences arise during the lifetime of monozygotic twins. Proc Natl Acad Sci USA 102:10604–10609

Fuke C, Shimabukuro M, Petronis A, Sugimoto J, Oda T, Miura K, Miyazaki T, Ogura C, Okazaki Y, Jinno Y (2004) Age related changes in 5-methylcytosine content in human peripheral leukocytes and placentas: an HPLC-based study. Ann Hum Genet 68:196–204

Gama-Sosa MA, Wang RY, Kuo KC, Gehrke CW, Ehrlich M (1983) The 5-methylcytosine content of highly repeated sequences in human DNA. Nucleic Acids Res 11:3087–3095

Handa V, Jeltsch A (2005) Profound flanking sequence preference of Dnmt3a and Dnmt3b mammalian DNA methyltransferases shape the human epigenome. J Mol Biol 348:1103–1112

Hansen RS (2003) X inactivation-specific methylation of LINE-1 elements by DNMT3B: implications for the Lyon repeat hypothesis. Hum Mol Genet 12:2559–2567

Hellman A, Chess A (2007) Gene body-specific methylation on the active X chromosome. Science 315:1141–1143

Karimi M, Johansson S, Stach D, Corcoran M, Grander D, Schalling M, Bakalkin G, Lyko F, Larsson C, Ekstrom TJ (2006) LUMA (LUminometric methylation assay)–a high throughput method to the analysis of genomic DNA methylation. Exp Cell Res 312:1989–1995

Kuo KC, McCune RA, Gehrke CW, Midgett R, Ehrlich M (1980) Quantitative reversed-phase high performance liquid chromato-

graphic determination of major and modified deoxyribonucleosides in DNA. Nucleic Acids Res 8:4763–4776

Norris DP, Brockdorff N, Rastan S (1991) Methylation status of CpG-rich islands on active and inactive mouse X chromosomes. Mamm Genome 1:78–83

Ramsahoye BH (2002) Measurement of genome wide DNA methylation by reversed-phase high-performance liquid chromatography. Methods 27:156–161

Rakyan VK, Hildmann T, Novik KL, Lewin J, Tost J, Cox AV, Andrews TD, Howe KL, Otto T, Olek A, Fischer J, Gut IG, Berlin K, Beck S (2004) DNA methylation profiling of the human major histocompatibility complex: a pilot study for the human epigenome project. PLoS Biol 2:e405

Sandovici I, Kassovska-Bratinova S, Loredo-Osti JC, Leppert M, Suarez A, Stewart R, Bautista FD, Schiraldi M, Sapienza C (2005) Interindividual variability and parent of origin DNA methylation differences at specific human Alu elements. Hum Mol Genet 14:2135–2143

Sarter B, Long TI, Tsong WH, Koh WP, Yu MC, Laird PW (2005) Sex differential in methylation patterns of selected genes in Singapore Chinese. Hum Genet 117:402–403

Shimabukuro M, Sasaki T, Imamura A, Tsujita T, Fuke C, Umekage T, Tochigi M, Hiramatsu K, Miyazaki T, Oda T, Sugimoto J, Jinno Y, Okazaki Y (2006) Global hypomethylation of peripheral leukocyte DNA in male patients with schizophrenia: a potential link between epigenetics and schizophrenia J Psychiatr Res 41(12):1042–1046

Stach D, Schmitz OJ, Stilgenbauer S, Benner A, Dohner H, Wiessler M, Lyko F (2003) Capillary electrophoretic analysis of genomic DNA methylation levels. Nucleic Acids Res 31:e2

Tost J, Dunker J, Gut IG (2003) Analysis and quantification of multiple methylation variable positions in CpG islands by pyrosequencing. Biotechniques 35:152–156

Waterland RA, Dolinoy DC, Lin JR, Smith CA, Shi X, Tahiliani KG (2006) Maternal methyl supplements increase offspring DNA methylation at axin fused. Genesis 44:401–406

Weber M, Davies JJ, Wittig D, Oakeley EJ, Haase M, Lam WL, Schübeler D (2005) Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. Nat Genet 37:853–862

Weisenberger DJ, Campan M, Long TI, Kim M, Woods C, Fiala E, Ehrlich M, Laird PW (2005) Analysis of repetitive element DNA methylation by MethyLight. Nucleic Acids Res 33:6823–6836

Wilson IM, Davies JJ, Weber M, Brown CJ, Alvarez CE, MacAulay C, Schubeler D, Lam WL (2006) Epigenomics: mapping the methylome. Cell Cycle 5:155–158

Wolff GL, Kodell RL, Moore SR, Cooney CA (1998) Maternal epigenetics and methyl supplements affect agouti gene expression in Avy/a mice. FASEB J 12:949–957

Zvetkova I, Apedaile A, Ramsahoye B, Mermoud JE, Crompton LA, John R, Feil R, Brockdorff N (2005) Global hypomethylation of the genome in XX embryonic stem cells. Nat Genet 37:1274–1279