SPANDIDOS PUBLICATIONS

Extensive demethylation of normally hypermethylated CpG islands occurs in human atherosclerotic arteries

SILVIA A. CASTILLO-DÍAZ[1], MARÍA E. GARAY-SEVILLA[2], MARTHA A. HERNÁNDEZ-GONZÁLEZ[3], MARTHA O. SOLÍS-MARTÍNEZ[2] and SILVIO ZAINA[2]

Departments of [1]Medicine and Nutrition and [2]Medical Sciences, Division of Health Sciences, Leon Campus, University of Guanajuato, 20 de Enero no. 929, 37000 Leon; [3]Research Unit UMAE 1, IMSS, Blvd. López Mateos S/N, 37000 Leon, Guanajuato, Mexico

Received May 25, 2010; Accepted July 21, 2010

DOI: 10.3892/ijmm_00000515

Abstract. Global DNA hypomethylation potentially leading to pro-atherogenic gene expression occurs in atherosclerotic lesions. However, limited information is available on the genomic location of hypomethylated sequences. We present a microarray-based survey of the methylation status of CpG islands (CGIs) in 45 human atherosclerotic arteries and 16 controls. Data from 10,367 CGIs revealed that a subset (151 or 1.4%) of these was hypermethylated in control arteries. The vast majority (142 or 94%) of this CGI subset was found to be unmethylated or partially methylated in atherosclerotic tissue, while only 17 of the normally unmethylated CGIs were hypermethylated in the diseased tissue. The most common functional classes among annotated genes adjacent to or containing differentially methylated CGIs, were transcription (23%) and signalling factors (16%). The former included *HOX* members, *PROX1*, *NOTCH1* and *FOXP1*, which are known to regulate key steps of atherogenesis. Expression analysis revealed differential expression of all CGI-associated genes analysed. Sequence analysis identified novel DNA motifs with regulatory potential, associated with differentially methylated CGIs. This study is the first large-scale analysis of DNA methylation in atherosclerosis. Our data suggest that aberrant DNA methylation in atherosclerosis affects the transcription of critical regulatory genes for the induction of a pro-atherogenic cellular phenotype.

Correspondence to: Dr Silvio Zaina, Department of Medical Sciences, Division of Health Sciences, Leon Campus, University of Guanajuato, 20 de Enero no. 929, 37000 Leon, Guanajuato, Mexico
E-mail: szaina@leon.ugto.mx

Abbreviations: AA, atherosclerotic artery; CA, control artery; CGI, CpG island; DM-CGI, CpG island that is differentially methylated between atherosclerotic and control arteries; HH-CGI, CpG island that is hypermethylated in both atherosclerotic and control arteries

Key words: DNA methylation, atherosclerosis, CpG island

Introduction

Atherosclerosis risk is in part determined by dietary and lifestyle-related factors, suggesting that diet- and environment-gene interactions are among the key phenomena in the aetiology and natural history of the disease. Epigenetics offers in principle a molecular model of such interactions, by envisioning that dietary and environmental risk factors modulate gene expression by laying down aberrant DNA methylation and histone post-translational modification patterns, which in turn convert permissive, relaxed chromatin into non-permissive, compact chromatin, or *vice versa*, depending on the specific gene involved (1). Accordingly, elegant animal models offer proof of principle for the idea that simple dietary interventions can have a significant impact on DNA methylation (2). As for atherosclerosis, potentially predisposing abnormal DNA methylation patterns have been detected in animal models prior to the appearance of vascular lesions (3). Furthermore, human and animal studies consistently indicate that global DNA hypomethylation accompanies advanced atherosclerosis (3,4). Importantly, 2 recent human studies have indicated that circulating levels of inflammation markers and exposure to environmental risk factors for atherosclerosis were associated with hypomethylation of highly repeated elements of leukocyte DNA (5,6). Identifying the sequences undergoing demethylation in atherosclerosis could provide important insights into the molecular mechanisms of atherogenesis and possibly of diet- or environment-gene interactions (7). We therefore conducted an array-based analysis of the DNA methylation status of CpG islands (CGIs) in control (CAs) and atherosclerotic arteries (AAs). CGIs are discrete sequences present both in intra- and intergenic regions, including the promoters of about half of the human genes (8). The observation that promoter CGI hypermethylation is usually associated with gene silencing in a number of physiological situations including cancer, suggests the pivotal role of CGI DNA methylation in transcription regulation (9). Our study is the first large-scale description of sequences undergoing aberrant DNA methylation in atherosclerosis.

Materials and methods

Tissue samples. All relevant procedures were approved by the local ethics committee. Atherosclerotic artery samples were obtained from ischaemic cardiopathy patients undergoing revascularization surgery. Although the atherosclerotic lesions were not histologically classified for severity, our samples consisted of advanced stage lesions, according to surgery recommendations. A fragment of the coronary artery at the lesion site, including the lesion and underlying media, was obtained during the application of the aorta-coronary bridge. The control samples were aortic fragments from patients undergoing aortic valve replacement surgery and with no coronary atherosclerotic lesions, as was indicated by cardiac catheterization and confirmed during surgery. Diabetes was the only exclusion criterion for both groups. Tissue samples were collected in RNAlater (Ambion) and transferred to -20˚C within 2 h from surgery.

Array analysis. Two pools of DNA were prepared, representing equal amounts of the control (n=16) or atherosclerotic (n=45) artery samples, respectively. The 2 groups did not differ in gender (2:1 male:female ratio) or age (22-79 years, p>0.05). Hypermethylated DNA fractions were prepared by digestion with the methylation-sensitive enzyme, *Hpa*II, and *Hpa*II-end containing (i.e. unmethylated) fragments were eliminated by exonuclease III treatment according to the method described by Shann *et al* (10), followed by amplification with the WGA system (Sigma). Reference DNA was prepared by mock digestions in which *Hpa*II was omitted. Primers were eliminated by passage through MinElute columns (Qiagen). Array analysis (labelling, hybridization and raw data generation) was performed at the UHN Microarray Centre, Toronto, Canada. The microarray used (HCGI15K; UHN; www.microarrays.ca) was a single-spotted array containing 14,923 CpG-island clones including the HCGI8.1K set of UHN and a new 6.8K set generated by sequence-specific primers targeting potential CpG-island clones that are currently not in any of the UHN collections. BLAT-annotated clones with overlapping sequences >20% of either clone were considered as duplicates and excluded. Clones with internal repeat sequences, with no or >1 BLAT hits and those mapped to the mitochondrial chromosome (M), were also excluded. Two series of arrays in triplicate, one for the control and the other for the atherosclerotic samples, were co-hybridized with Alexa647-labelled (red) *Hpa*II-digested DNA and Alexa555-labelled (green) reference DNA. The background-subtracted Alexa647 and Alexa555 intensities were normalized to the overall median Alexa555 intensity calculated on all the arrays, and the $\log_2$ of the Alexa647/Alexa555 ratio was calculated for each individual array as previously reported (10). Comparisons between the 2 array groups were conducted by ANOVA. Array data have been deposited into the GEO database (www.ncbi.nlm.nih.gov/ projects/geo/) with accession no. GSE15552.

Array data validation. Array data for 10 differentially methylated CGIs were validated by methylation-sensitive PCR (MS-PCR) in 10 more samples from each group with the same characteristics as the ones used for the array analysis. In brief, 10 ng DNA were digested with 1 U *McrBC* overnight and a 200-300 bp CGI portion was amplified by 35-cycle PCRs. Primers were designed with BiSearch (bisearch.enzim.hu) and their sequences are available upon request (11). Band intensity was quantified with Quantity One software (Bio-Rad) and normalized to the median intensity of the PCR products of the mock-treated DNA (input) made equal to the arbitrary value of 1. MS-PCRs for each gene were run and analysed on the same gel to reduce variability.

Gene expression analysis. RNA was extracted with the RNeasy system (Qiagen) and cDNA was prepared from 3 μg RNA pooled from the same 10 samples used for MS-PCR, using the Access RT-PCR system (Promega). For each gene, one-fiftieth of cDNA was amplified by using the LightCycler FastStart DNA Master^{PLUS} SYBR Green I (Roche) according to the manufacturer's instructions in a LightCycler 1.5 machine (Roche), with primers designed in qPrimerDepot (primerdepot.nci.nih.gov). Primer sequences are available upon request. Ct values were calculated using a fluorescence value equal to twice the baseline as the threshold. The expression was normalized by calculating the ratio between the Ct of a sample and the average Ct for *GAPDH* in the same group CAs or AAs.

Bioinformatics. For bioinformatics analysis, CGI annotations were as provided by the array supplier and were based on NCBI UniGene human build 217 and UCSC human build hg18. Genomic localization was determined with the NCBI Map Viewer resource (www.ncbi.nlm.nih.gov). CG content and CpG density (CpG count/CGI length ratio) were determined with the EMBOSS CG and CpG report tools (oz.ifc.unam.mx/EMBOSS). Gene functions were deduced according to the categories on the Gene Ontology database (www.geneontology.org). DNA motif search was conducted with MEME software (meme.sdsc.edu) using default parameters (12). DNA motifs were compared against the databases of known protein-binding DNA sequences, namely the TRANSFAC and JASPAR core databases, using TOMTOM (meme.nbcr.net/meme4_1/cgi-bin/tomtom.cgi).

Results

CGI methylation status in CAs and AAs. Our array-based survey yielded reliable data on the methylation status for 10,367 CGIs. We classified CGIs as unmethylated if the Alexa647/Alexa555 ratio was <0.3, or if $\log_2$ Alexa647/Alexa555 ($\log_2$r/g) was <-1.6, representing an excess hybridized reference DNA (i.e. not *Hpa*II digested) compared to filtered DNA. CGIs were considered to be hypermethylated if the Alexa647/Alexa555 ratio was >3, corresponding to $\log_2$r/g >1.6. Intermediate values were considered as partial hypo- or hypermethylation (Fig. 1A).

In accordance with the well established notion that the vast majority of CGIs are unmethylated (8), most of the (9,085 or 87.6%) CA CGIs fell within the unmethylated class and accordingly, the overall median CA CGI $\log_2$r/g was -2.4. In contrast, the hypermethylated CGIs were a minority (151

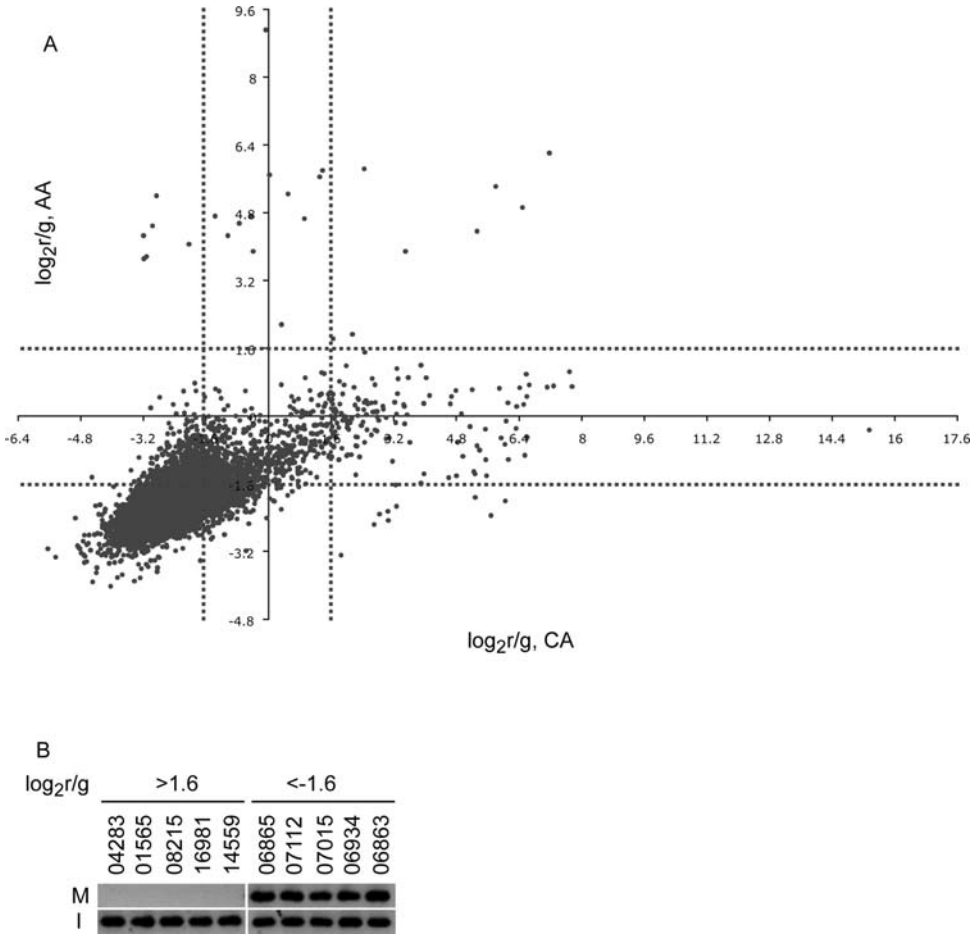

Figure 1. Global pattern of CGI methylation. (A) Plot of CGI $\log_2$r/g values in CAs and AAs. Each data point represents the methylation status of an individual CGI in CAs and AAs. Vertical and horizontal dashed lines represent the $\log_2$r/g values of -1.6 and 1.6 on each axis and therefore separate hypermethylated, unmethylated and intermediately methylated CGIs in either group. (B) Validation of array-deduced methylation status of 5 hypermethylated and 5 unmethylated CGIs (left and right panels, respectively) in CAs by *McrBC* digestion-based MS-PCR. CGIs are referred to using the last 5 digits of the UHN clone nomenclature, i.e. UHNhscpg00XXXXX. M and I, *McrBC*-digested and input (undigested) DNA, respectively.

or 1.4%) in this group. In order to corroborate the array results, the methylation status of 10 randomly chosen un- or hypermethylated CA CGIs was verified by MS-PCR in the DNA pool used for the hybridizations (Fig. 1B).

The comparison of the CGI methylation status between CAs and AAs revealed significant ($p<0.01$) changes. Of the 151 normally hypermethylated CGIs, 142 (or 94%) were found to be unmethylated or in an intermediate methylation status in AAs (Fig. 1A). Conversely, only 17 of the CGIs that were unmethylated or intermediately methylated in CAs, were hypermethylated in AAs. We validated the methylation status of 10 CGIs differentially methylated between AAs and CAs (DM-CGIs) adjacent to relevant genes by MS-PCR in 10 more age- and gender-paired samples not included in the array analysis (Fig. 2A). Validation examples of DM-CGI methylation status are shown in Fig. 2B. Further characterization was carried out of those differentially methylated CGIs that showed the most extreme changes in DNA methylation, i.e. were hypermethylated in one group ($\log_2$r/g >1.6) and incompletely methylated or unmethylated (i.e. showing an Alexa647/Alexa555 ratio <0.75, corresponding to $\log_2$r/g <-0.4) in the other group, with a p-value of <0.01. Excluding BLAT-unmatched clones, 45 CGIs hypermethylated in CAs and 10 in AAs met these criteria. These CGIs will be referred to as DM-CGIs.

Genomic location of DM-CGIs. A recent study of normal, non-vascular human tissues has shown that CGIs with tissue-specific methylation status are found in specific genomic locations, thus providing insights on the mechanisms and transcriptional impact of differential methylation (13). In order to assess whether comparable patterns occur in atherosclerosis, we analysed the genomic locations of DM-CGIs (n=45) and non-DM-CGIs that were hypermethylated in both CAs and AAs (HH-CGIs; n=8). Twenty-nine DM-CGIs were intragenic and 26 were intergenic, whereas the majority of HH-CGIs (6 out of 8) was intragenic. These CGIs differed in their distribution (Fig. 3). DM-CGIs unmethylated in AAs tended to be localized preferentially in promoters (i.e. 5' UTR, TSS and 1.0 kb upstream to the latter) and first introns. In contrast, DM-CGIs hypermethylated in AAs showed a broader distribution encompassing 5' and middle third gene portions ($p<0.03$ in comparison with the latter group). Intragenic HH-CGIs, however, were preferentially positioned at middle third gene portions ($p<0.02$ in comparison with either DM-CGI class distribution). All 3 CGI classes were

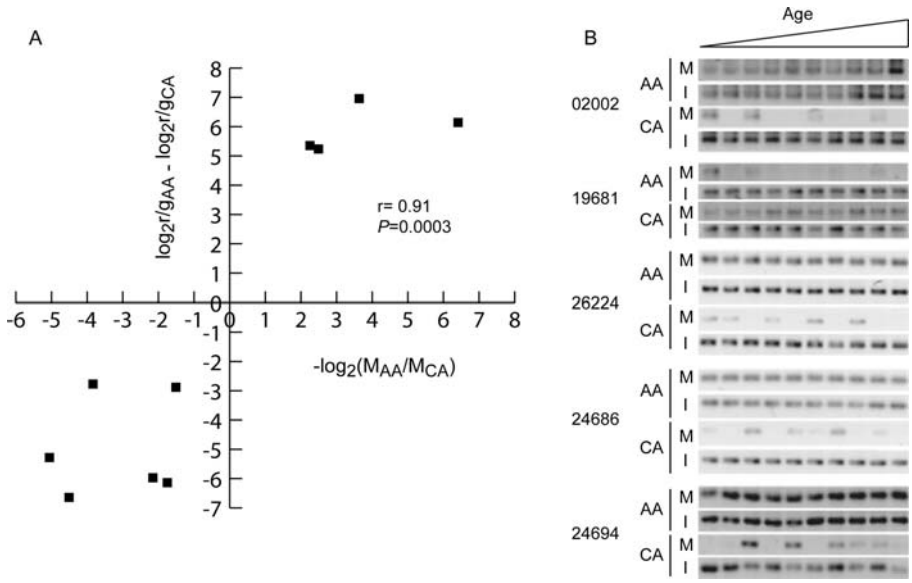

Figure 2. Validation of selected DM-CGI methylation status. (A) Correlation between MS-PCR data (X-axis) and array data (Y-axis) for 10 validated DM-CGIs. M_{CA} and M_{AA} are average normalized MS-PCR band intensities in CAs and AAs, respectively. The Alexa647/Alexa555 ratios in CAs and AAs are r/g_{CA} and r/g_{AA}, respectively. Negative and positive values indicate hypo- and hypermethylation in AAs relative to CAs, respectively. (B) MS-PCR analysis of 5 representative CGIs in 10 age- and gender-paired samples. Age difference between the paired samples was <5 years. Numbers on the left are the last 5 digits of the CGI clone notation of UHN, i.e. UHNhscpg00XXXXX. M and I, *McrBC*-digested and input (undigested) DNA, respectively. Absent or decreased bands indicate a relative hypermethylation.

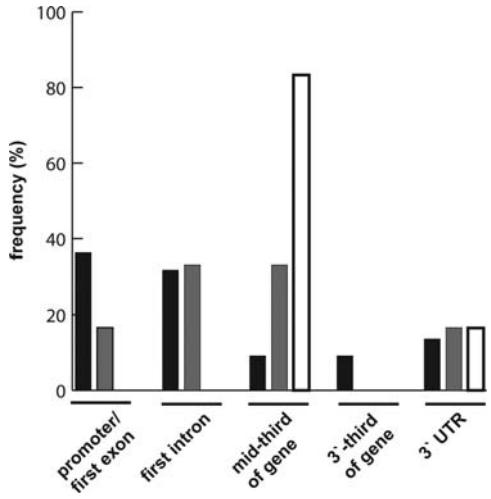

Figure 3. Distribution of intragenic DM-CGIs and HH-CGIs across different gene portions. Black, DM-CGIs unmethylated in AAs; grey, DM-CGIs hypermethylated in AAs; white, HH-CGIs.

similarly represented at 3' UTR regions (Fig. 3). Of the 20 DM-CGIs and HH-CGIs that were not within promoters or 3' UTR regions, the vast majority (17 or 85%) was located in introns rather than in protein coding sequences.

As for intergenic DM-CGIs, their distance from flanking genes varied (i.e. 1.2-120 kb). Of this CGI group, 14 (or 54%) were the only array-represented CGIs positioned between the respective up- or downstream genes, while 11 were the closest CGIs to either the respective down- or upstream gene. Furthermore, DM-CGIs that were unmethylated in AAs were in average ~100-fold more distant from downstream or upstream genes than DM-CGIs hypermethylated in AAs or intergenic HH-CGIs (350.5+678.1, 2.6+2.4 and 2.7+2.2 kb, respectively; p<0.001).

Identity and expression of DM-CGI-associated genes. The functional classification of annotated genes upstream or downstream to, or containing DM-CGIs, revealed several gene functions, with a bias towards transcription factors (TFs), representing the most numerous class, followed by signalling factors and genes involved in organelle homeostasis (Table I). A description of the DM-CGI genomic location and adjacent annotated genes is presented in Table II. Furthermore, we analysed the expression of TFs and atherosclerosis-related genes adjacent to, or containing the 10 DM-CGIs for which we validated the methylation status, by quantitative RT-PCR in the same 10 samples used for MS-PCR (Fig. 4, graphs on the right). A schematic representation of the regions containing validated DM-CGIs is shown in Fig. 4, left side. Differential DNA methylation coincided with changes in expression in all the cases analyzed (p<0.03 in comparison with normalized Ct). The methylation status of several DM-CGIs was inversely related with gene expression. Two intergenic DM-CGIs located in the *HOX* TF clusters were one notable example. One DM-CGI was located downstream to *HOXD10* between *HOXD4* and *HOXD8* in chromosome 2 and was hypermethylated in AAs, while the other was located in a region of chromosome 12 including *HOXC9*, *HOXC10* and *HOXC11*, and was unmethylated in AAs. The expression of *HOXC9*, *HOXC10* and *HOXD10* was inversely related to the methylation status of the associated DM-CGI (Fig. 4A and B). The same pattern was observed in a DM-CGI unmethylated in AAs occupying the *EGFL7* promoter and downstream to the TF gene, *NOTCH1* (Fig. 4C and D). Two other DM-CGIs lay ~1 Mb apart in a region of chromosome 1 in the vasohibin 2 (*VASH2)* promoter and

Functional classification of genes upstream or downstream to, or containing DM-CGIs, ordered by class size, with the percentage (%) of total gene count in each class.[a]

Gene function	DM-CGI methylation status in AAs		%
	Hypermethylated	Unmethylated	
Transcription	*HOXD4*, *HOXD8*, *ASCL1*, *PROX1*	*LHX1*, *DMRTA2*, *RUNX3*, *SMAD2*, *ZC3H13* *ZNF184*, *ZNF253*, *ZNF506*, *ZNF821*, *NOTCH1*, *CITED2*, *ZNHIT2*, *DIP2A*, *ZFP1*, *GSC*, *MORF4L2*, *HOXC11*, *HOXC10*, *FOXP1*	23.0
Signalling	*FKBP1A*, *RPS6KC1*	*CENTB5*, *GPR143 GPR20*, *GNB1L*, *ASZ1* *ESRRG*, *S100B*, *CAMK1D*, *TAOK1*, *MGC13024* *GAL*, *ASB3*, *EGFL7*	15.6
Organelles	*NSFL1C*	*GORASP1*, *WDR48*, *AP1G1*, *MUTED* *KIAA0174*, *IER3IP1*, *M6PRBP1*	8.3
Protein catabolism	*UBE2D4*	*FAU*, *ZNRF1*, *SERPINA13*, *HERC6*	5.2
Transport		*CLIC4*, *APXL*, *SLC45A4*, *KCNJ1*, *CFTR*	5.2
Apoptosis		*FAF1*, *SYVN1*, *MCPH1*, *P53AIP1*	4.2
Cell adhesion	*CDH4*, *DCC*	*ASAM*, *CADM1*, *NRXN1*	4.2
Chromatin	*MBD2*	*HIST1H2BL*, *TBL1X*, *SETD8*	4.2
Lipid metabolism		*ACBD5*, *PIGY*, *PIB5PA*, *AGPAT2*	4.2
Cytokine	*CTF1*	*LIF*, *OSM*	3.1
Cytoskeleton		*MYOM2*, *SMTN*, *CAPZB*	3.1
Mitochondrial ribosome	*MRPS24*	MRPL48, MRPS9	3.1
Non-lipid metabolism	*PAH*	*PUSL1*, *PRMT2*	3.1
Cell proliferation	*IGF1*	*APC*	2.1
Inflammation		*NAGA*, *TICAM1*	2.1
Redox	*FLJ39599*	*TXNRD*	2.1
Testis-specific		*HORMAD2*, *DNAJB13*	2.1
Angiogenesis	*VASH2*		1.0
Erythropoiesis	*FLVCR*		1.0
Hypoxia		*LDHD*	1.0
Micro-RNA		*DICER1*	1.0

[a]If multiple functions have been reported for a gene, the function relevant to atherogenesis is listed.

upstream to *PROX1*, another member of homeobox TFs, respectively (Fig. 4E and F). Hypermethylation of DM-CGIs in AAs coincided with transcriptional down-regulation of the latter 2 genes. Similarly, hypomethylation of a 3' UTR DM-CGI coincided with up-regulation of the *S100B* gene and promoter hypermethylation was associated with *MBD2* down-regulation in AAs (Fig. 4G and H). In contrast with the above patterns, hypomethylation of 2 intragenic DM-CGIs was associated with expression down-regulation in the case of smoothelin (*SMTN)* and TF *FOXP1* (Fig. 4I and J).

Furthermore, we confirmed the methylation status of 3 DM-CGIs previously reported to have a hypermethylation tendency in atherosclerosis (i.e. in the 5' region of *ESR2*, *FADS2* and *TFPI2)* (14-16). Log_2r/g values in CAs and AAs were -2.72, -1.40 for *ESR2*, and -3.14, -2.18 for *FADS2*, respectively, with a p-value of <0.05 in both cases. Corresponding values for *TFPI2* were -2.27 and -2.11 with no significant difference, and were in line with the observation that hypermethylation at that gene was present in a minority of lesions (16). Other promoters previously studied in atherosclerosis were not represented in the arrays used here.

Bioinformatic analysis of DM-CGI. We searched for sequence features in DM-CGIs and HH-CGIs that could yield further insights into the mechanisms of CGI differential methylation. The comparison of CGI size, CG content and CpG density (no. of CpG to CGI length ratio), a parameter shown to be related with CGI methylation status in somatic cells (17), did not reveal any significant difference. Furthermore, we screened for DNA motifs unique for any CGI class in 4 kb-sequences centered at the CGI midpoint. The presence of class-specific DNA motifs could signal the possible involvement of DNA-binding factors in the establishment of differential DNA methylation. The MEME software yielded a 29-bp motif (DMM) that was present in 55.5% of the DM-CGIs unmethylated in AAs (E=2.2e-118) but absent in the DM-CGIs hypermethylated in AAs or the HH-CGIs (Fig. 5A). The lack of DMM detection in HH-CGIs was not due to

Table II. DM-CGIs with genomic location, associated genes and average log_2r/g values in CAs and AAs.

CGI clone	Genomic location	Within	Upstream	Downstream	Average log_2r/g CA	Average log_2r/g AA
DM-CGIs hypermethylated in CAs, unmethylated in AAs						
UHNhscpg0000083	chr17:77921639-77922214		FLJ35767	FLJ22222	2.29070985	-0.935517059
UHNhscpg0000313	chr2:51112865-51113731		NRXN1	ASB3	2.226160184	-0.681858958
UHNhscpg0001565	chr3:39123016-39123538	GORASP1	WDR48	TTC21A	6.521206416	-0.903185685
UHNhscpg0002002	chr12:52663262-52664017		HOXC11	HOXC10	1.994891121	-0.886415227
UHNhscpg0003364	chr14:94304731-94304947	GSC	SERPINA13	DICER1	6.091264337	-0.998422156
UHNhscpg0005293	chr1:25055302-25055491		CLIC4	RUNX3	6.016275549	-1.983046692
UHNhscpg0005702	chr12:122434216-122434426		SBNO1	SETD8	1.768802278	-1.000565333
UHNhscpg0005813	chr13:45322805-45324127		NURIT	KIAA0853	2.97263872	-0.505776329
UHNhscpg0006769	chr1:1230036-1231291	CENTB5	CENTB5	PUSL1	3.237808178	-2.101301822
UHNhscpg0008215	chr3:71197019-71197553	FOXP1	EIF4E3	LOC100289131	5.359131933	-0.778913464
UHNhscpg0008474	chr11:128280580-128280993	C11orf45	KCNJ1	P53AIP1	1.69480582	-0.486156485
UHNhscpg0009592	chr18:43041077-43041115		IER3IP1	SMAD2	3.580301757	-0.595334021
UHNhscpg0013290	chr7:116906424-116907180		ASZ1	CFTR	2.195421389	-0.460685003
UHNhscpg0014151	chr2:104835710-104836266		CR936735	MRPS9	2.590023363	-0.42021637
UHNhscpg0014542	chr11:115092453-115092670		BC047021	MGC13125	2.078305966	-0.887925062
UHNhscpg0014599	chr10:27570656-27571247		ACBD5	LOC387646	5.520081361	-1.006617213
UHNhscpg0014796	chr6:140018073-140018373		CITED2	AK097143	2.416687107	-0.59569034
UHNhscpg0014912	chr11:122734054-122734895		ASAM	AK127457	5.242359626	-1.341504446
UHNhscpg0015048	chr19:4780161-4780685		TICAM1	M6PRBP1	3.047119619	-0.492737606
UHNhscpg0015225	chr1:215376726-215377012		ESRRG	GPATC2	2.199685837	-1.36364171
UHNhscpg0015321	chr17:24645319-24645686		NUFIP2	TAOK1	5.909880104	-0.493664785
UHNhscpg0015763	No BLAT hit				1.642190065	-0.684376507
UHNhscpg0016316	No BLAT hit				2.235284245	-0.470463224
UHNhscpg0017732	chrX:102829282-102829630		MORF4L2	TMEM31	1.634887846	-0.703610948
UHNhscpg0017800	chr1:19635271-19635541	CAPZB	AK096102	C1orf151	4.799718052	-0.589762553
UHNhscpg0018821	chr19:19793020-19793758	ZNF506	FLJ46230	ZNF253	4.571042478	-1.42677703
UHNhscpg0018976	chr6:27678451-27678629	FKSG56	ZNF184	HIST1H2BL	3.214585772	-1.601708402
UHNhscpg0019872	chr11:64645232-64646526	FAU	ZNHIT2	SYVN1	5.544692553	-1.381365105
UHNhscpg0020507	chr8:142314675-142315035		SLC45A4	GPR20	4.766596074	-1.382565412
UHNhscpg0020592	chrX:9693151-9694306	GPR143	TBL1X	APXL	3.858298486	-1.553981206
UHNhscpg0022803	chr11:73264949-73265794	CHCHD8	MRPL48	DNAJB13	5.651905748	-2.321222231
UHNhscpg0022884	chr11:68193727-68194620		SAPS3	GAL	1.961095511	-0.909580258
UHNhscpg0023790	chr16:73705572-73706745	LDHD	ZNRF1	ZFP1	5.211070651	-1.284382721
UHNhscpg0023931	chr17:32239295-32240467		MRM1	LHX1	2.66772773	-2.534658496
UHNhscpg0024686	chr21:46842417-46843516	S100B	DIP2A	PRMT2	3.940243041	-1.339843615
UHNhscpg0024689	chr22:18221859-18223031	GNB1L	FLJ21125	TXNRD2	1.632872597	-0.423449001
UHNhscpg0024694	chr22:29824296-29825316	SMTN	BC012026	PIB5PA	5.30028017	-0.671827614
UHNhscpg0024745	chr22:40799574-40800700	FAM109B	NAGA	LOC91689	5.252898787	-1.895368556
UHNhscpg0025462	chr6:8008992-8009814	MUTED	BC028580	MGC26597	6.572617804	-0.471029058
UHNhscpg0025899	chr8:2062660-2063586	MYOM2	KBTBD11	MCPH1	4.932429293	-0.926405025
UHNhscpg0026224	chr9:138679621-138680856	EGFL7	NOTCH1	AGPAT2	6.004100257	-0.637057408
UHNhscpg0027388	chr4:89597065-89598125	HERC5	HERC6	PIGY	1.846248553	-0.73415746
UHNhscpg0027650	chr10:12430778-12432679	CAMK1D	C10orf7	CCDC3	3.134990152	-1.040169498
UHNhscpg0027678	chr5:111782224-111784173		EPB41L4A	APC	3.015604131	-2.219980056
UHNhscpg0027684	chr22:28971571-28973279	LIF	HORMAD2	OSM	3.022969021	-2.441966395
UHNhscpg0028047	chr16:70451013-70451959	LOC55565	AP1G1	KIAA0174	2.200354593	-0.673368865
UHNhscpg0028140	chr1:50665901-50666652		AJ301580	FAF1	2.80695587	-2.283181688
DM-CGIs unmethylated in CAs, hypermethylated in AAs						
UHNhscpg0001071	chr1:212222309-212223400		RPS6KC1	PROX1	-1.068343579	4.284862484
UHNhscpg0003554	chr20:1320740-1321479	FKBP1A	AK000809	NSFL1C	-3.144878822	3.790954386
UHNhscpg0005354	chr7:43912050-43912251	URG4	MRPS24	UBE2D4	-0.425721857	3.918773175
UHNhscpg0011135	chr1:211190078-211190296	VASH2	FLVCR	VASH2	-0.479853218	4.75654165

Continued.

CGI clone	Genomic location	Within	Upstream	Downstream	Average log_2r/g CA	Average log_2r/g AA
UHNhscpg0014640	chr16:15435454-15435917	C16orf45	FLJ39599	LKAP	-0.780517895	4.578974225
UHNhscpg0015177	chr16:30840387-30840828	LOC283932	CTF1	MGC13024	-3.222082358	4.287964286
UHNhscpg0016809	chr18:49929963-49930163		DCC	MBD2	-1.392033622	4.748487633
UHNhscpg0019681	chr2:176723074-176723546		HOXD9	HOXD4	-3.214739959	3.743257903
UHNhscpg0024537	chr20:59552448-59553578	CDH4	LOC284757	CDH4	-2.064500246	4.083960366
UHNhscpg0028067	chr12:101834908-101835700	PAH	IGF1	ASCL1	-2.895150067	5.229689851

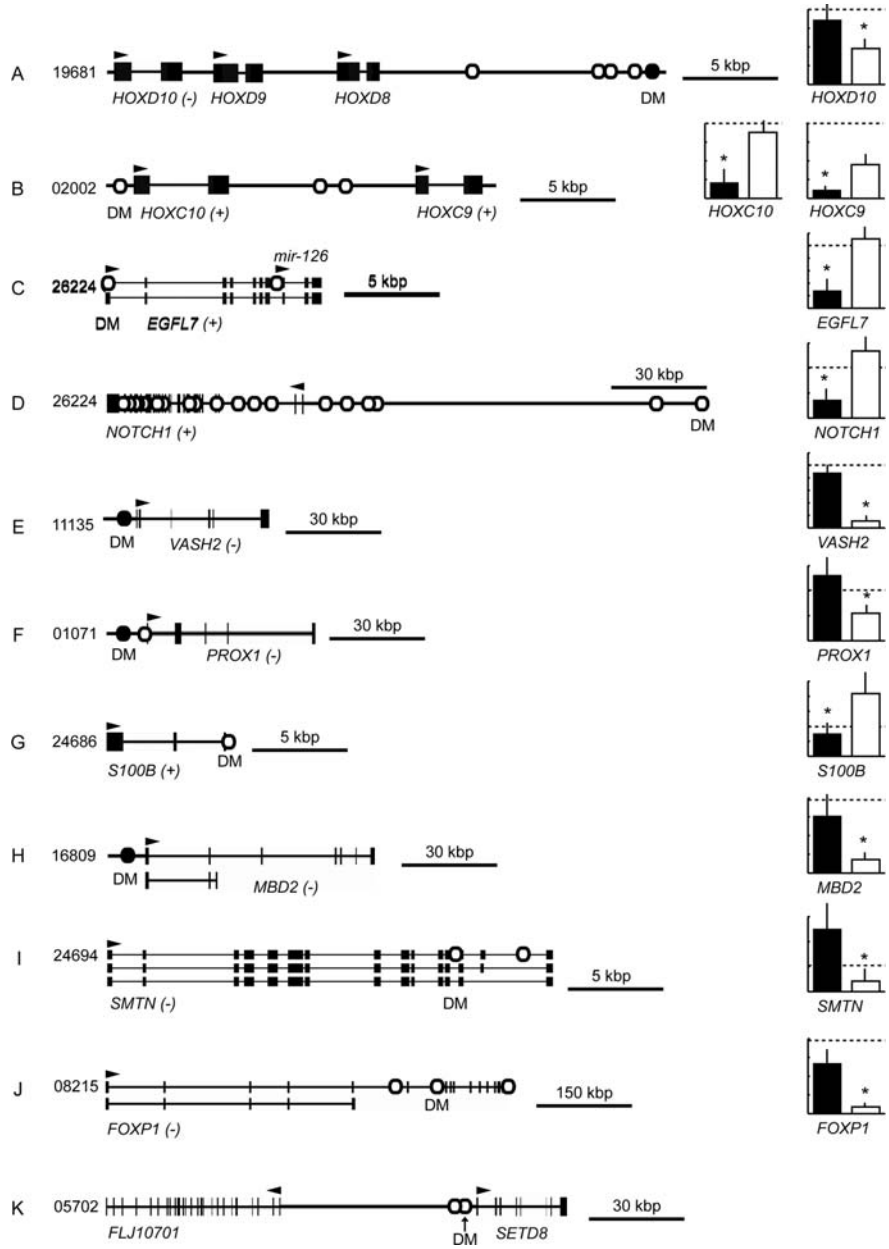

Figure 4. Schematic map of 10 validated DM-CGI-containing regions (left) and the expression of the relevant annotated genes (right). In the diagrams on the left, open and solid ovals are unmethylated and hypermethylated CGIs, respectively. CGIs that are DM-CGIs (indicated as DM) are marked by an arrow and their methylation status in AAs is shown. (+) or (-) beside the gene names indicates up- and down-regulation in AAs, respectively. In the case of unlabelled genes, expression analysis was not carried out. The arrowhead above the genes indicates the start site and the direction of transcription. CGIs are referred to using the notation of UHN as indicated in the legend of Fig. 1. Eleven sequences are shown as the position of CGI 26224 is shown both relative to *NOTCH1* and *EGFL7*. Right graphs, quantitative RT-PCR expression analysis of genes indicated below each graph. Mean and SD of normalized expression are shown. Black and white bars, CAs and AAs, respectively. For simplicity, numbers are not shown in the value axis, but the normalized expression value=1 is indicated as reference by the horizontal dashed lines. *$p<0.05$ in comparison between the CAs and AAs.

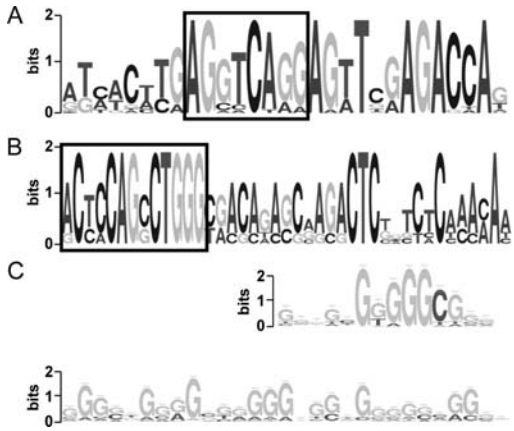

Figure 5. DNA motifs specific for either DM- or non-DM-CGIs. Motifs are represented as sequence logos produced by the MEME software. The height of a letter in a stack indicates the probability of the letter to be represented at that position. (A and B) DMM and non-DMM, respectively. Boxes contain motif portions that are conserved in the *M. musculus*. (C) Krox binding site-homologous motif. The sequence in the lower part of panel C is the motif query, the upper sequence indicates the Krox motif (accession no. M00982).

sample size, as repeated analyses of randomly chosen groups of n=8 DM-CGIs always identified the motif. Neither the DMM presence nor copy number correlated with DM-CGI size, but the DMM copy number was weakly correlated with CpG density (r=0.42, p<0.02). DMM was not preferentially associated with intra- or intergenic DM-CGIs, but the DMM-associated DM-CGIs were on average ~10-fold closer to downstream genes than non-DMM-associated DM-CGIs (56.8±50.0 and 644.2±511.1 kb, respectively, p<0.03). Furthermore, we detected a candidate non-DM-CGI motif (non-DMM) present in 6 out of the 8 HH-CGIs but not present in any of the DM-CGIs (E=3.3e-14; Fig. 5B). Both DMM and non-DMM were invariably located outside CGIs. The analysis of mouse CGIs with BLAST homology with DMM-associated DM-CGI revealed a conserved 5'-AGNNCAGG-3' sequence present in DMM, and a 5'-ACTCCAGNCTGGG-3' sequence present in non-DMM (Fig. 5A and B). Neither DMM nor non-DMM match any known binding site included in the TRANSFAC and JASPAR core databases. In addition, the MEME software yielded a 29-bp DNA motif present in 100% of the DM-CGIs hypermethylated in AAs and in 12 (or 26.6%) of the DM-CGIs unmethylated in AAs. This DNA motif shows homology with the Krox binding site according to the TRANSFAC database (p=8.6e-07; Fig. 5C). The 2 DM-CGI-specific DNA motifs were not mutually exclusive, as 4 DMM-associated DM-CGIs also contained a putative Krox binding site. We also detected CT- and A-rich motifs previously reported in normally hypermethylated CGIs but these were not significantly over-represented in any CGI class analysed here (18).

Discussion

By examining the DNA methylation status of ~10,000 CGIs in CAs and AAs, we hereby presented an initial large-scale analysis of the epigenetic changes associated with cardiovascular disease. The analysis of CA CGI methylation showed that a small portion of autosomal, not imprinted gene-associated CGIs is hypermethylated, thus supporting previous suggestions that this DNA methylation pattern could be a general feature of somatic tissues (18,19). The vast majority of these normally hypermethylated CGIs is found unmethylated in AAs, indicating that this net CGI demethylation contributes to the global DNA hypomethylation observed in atherosclerotic tissues in previous studies (20).

DM-CGIs were located within or relatively close to a number of genes known to be involved in atherogenesis. Approximately 40% of the DM-CGI-associated genes were regulatory (i.e. TFs and signalling factors), in accordance with the idea that epigenetic changes are upstream events in the establishment of an atherosclerosis-specific gene expression pattern. In general, CGI differential methylation was associated with up- or down-regulation of genes implicated in angiogenesis, vascular smooth muscle cells (VSMCs) phenotype modulation and inflammation. One remarkable example was the opposite effects on *HOXC* and *HOXD* members, known to promote vascular remodelling and inhibit adult angiogenesis, respectively (22). *HOXCs* located in physical proximity to a DM-CGI unmethylated in AAs were selectively expressed in AAs, whereas the precisely opposite pattern was observed in *HOXD10*. A recent study reported that aberrant homeobox CGI methylation is a frequent event in breast cancer, thus suggesting a partially overlapping epigenetic dysregulation of cell proliferation and migration in distinct diseases (23). Furthermore, we observed activation of *NOTCH1* and *EGFL7*, 2 inducers of angiogenesis that are physically associated with the same DM-CGI, and down-regulation of *VASH2*, an inhibitor of adult angiogenesis, in AAs (24-26). It has also been reported that demethylation of the *EGFL7* promoter activates *miR*-126, a pro-angiogenic micro-RNA transcribed from an *EGFL7* intron (27). Another landmark of atherogenesis is the transition of VSMC from a contractile, differentiated phenotype to a secretory, relatively immature one. This phenomenon is accompanied by a significant re-organization of organelles and secretory vesicles (28). Accordingly, we observed down-regulation of the differentiated VSMC marker, *SMTN,* in AAs, and organelle homeostasis and secretory genes were found associated with DM-CGIs (29). As for inflammation, we observed down-regulation of *FOXP1*, a forkhead family TF shown to inhibit macrophage activation, and up-regulation of the AGE receptor ligand, S100B (30,31). Relatively few genes involved in chromatin structure regulation were associated with DM-CGIs and their role in atherogenesis is poorly understood. One exception is *MBD2*, representing a methylcytosine-binding protein that can be induced in cultured VSMCs by stimulation with homo-cysteine, an amino acid implicated in atherogenesis (32).

The analysis of DM-CGI genomic location provided insights into the potential mechanisms of transcriptional control by CGI methylation. In the case of the 10 validated DM-CGIs, the methylation status of the intergenic or promoter DM-CGIs was inversely related with gene expression in accordance with the generally accepted role of DNA methylation in transcription regulation. In contrast, the opposite was observed between methylation and transcription for the intragenic DM-CGI-associated genes, *FOXP1* and *SMTN*. This observation suggests that intragenic DNA methylation

nplex impact on gene expression as hinted at by studies on plants (33). For all the validated intergenic DM-CGIs, differential CGI methylation was associated with changes in the expression of genes located as far as 120 kb and separated by up to 5 non-DM-CGIs from the relevant DM-CGI. This observation is reminiscent of the results of a previous study showing that regions of tissue-specific methylation are often distant from genes (34). Thus, our data suggest the testable hypothesis that at least some DM-CGIs could be important long-range epigenetic expression regulators.

Our data indicate that DM-CGIs and HH-CGIs show different intragenic distributions. DM-CGIs unmethylated in AAs and HH-CGIs are significantly more abundant in promoters and in the middle third introns of genes, respectively. These results differ from the ones of a recent large-scale bioinformatics-based study including a wide range of non-vascular normal human tissues (13). The authors showed that DM-CGIs are significantly more abundant in introns than in other gene compartments, while promoters preferentially include constitutively unmethylated CGIs. In the light of this data comparison, it is conceivable that during atherogenesis, the normal equilibrium between mechanisms that maintain selected (i.e. promoter-associated) CGIs in a constitutively unmethylated state, and other mechanisms establishing tissue-specific CGI hypermethylation is lost in favour of the former. Another difference between our studies is the lack of a substantial association of any DM-CGI or HH-CGI with protein coding regions in our samples. Further study is required to assess whether these discrepancies reflect distinct CGI methylation regulations in normal and diseased cells. Furthermore, our data do not show any significant preference for the inter- or intragenic location of DM-CGIs, whereas the study mentioned above indicates that DM-CGIs are mainly intragenic. Nevertheless, this likely reflects a bias against intergenic elements in the latter CGI collection (13).

DM-CGI nucleotide composition analysis provided limited hints on the mechanisms of differential CGI methylation between AAs and CAs. CpG density did not correlate with the DM- or HH-CGI status, indicating that CGI demethylation in CAs is mechanistically distinct from the preferential demethylation of CpG-dense CGIs in somatic cells (13,17). The presence of CGI class-specific DNA motifs suggests that differential CGI methylation could be regulated by DNA-binding factor(s). As for the yet uncharacterized DMM motif, our data on the distance between DM-CGIs and adjacent annotated genes suggest that this DNA motif could regulate the DNA methylation of CGIs involved in relatively short-range interactions with downstream genes. The functional significance of the over-representation of the DMM motif in high CpG density DM-CGIs is difficult to interpret, although it could be related to the preferential somatic demethylation of CpG-dense CGIs as mentioned above (13,17). The presence of putative Krox binding sites in a subset of DM-CGIs was expected, since early growth response TFs play a pivotal role in atherogenesis (35). Nevertheless, Krox binding site-associated CGIs showed inconsistent DNA methylation changes between CAs and AAs, making it difficult to assign a straightforward epigenetic function to this motif.

One limitation of the present study was the cell type heterogeneity of AAs compared to CAs, which raises the possibility that at least some of the observed DM-CGI methylation and gene expression patterns were not established *de novo* in atherosclerosis but were pre-existent in non-vascular cells recruited to the lesion, such as monocytes, macrophages, lymphocytes and others. Nonetheless, at least in the case of validated DM-CGI, the sharp between-group differences in DNA methylation and expression levels suggests that the observed changes involved the majority of cells in AAs, thus including VSMCs and endothelial cells. Another study is currently in progress to identify cell type-specific epigenetic changes in vascular lesions.

Acknowledgements

This study was funded by DINPO (University of Guanajuato), and the 'Program for the Improvement of the University Professorate' of the Mexican Ministry of Education (PROMEP-SEP) (S.Z.). We thank Gertrud Lund for critically reading the manuscript.

References

1. Gallou-Kabani C, Vigé A, Gross MS and Junien C: Nutri-epigenomics: lifelong remodelling of our epigenomes by nutritional and metabolic factors and beyond. Clin Chem Lab Med 45: 321-327, 2007.
2. Waterland RA and Jirtle RL: Transposable elements: targets for early nutritional effects on epigenetic gene regulation. Mol Cell Biol 23: 5293-5300, 2003.
3. Lund G, Andersson L, Lauria M, Lindholm M, Fraga MF, Villar-Garea A, Ballestar E, Esteller M and Zaina S: DNA methylation polymorphisms precede any histological sign of atherosclerosis in mice lacking apolipoprotein E. J Biol Chem 279: 29147-29154, 2004.
4. Hiltunen MO, Turunen MP, Hakkinen TP, Rutanen J, Hedman M, Makinen K, Turunen AM, Alto-Setala K and Yla-Herttuala S: DNA hypomethylation and methyltransferase expression in atherosclerotic lesions. Vasc Med 7: 5-11, 2002.
5. Baccarelli A, Tarantini L, Wright RO, Bollati V, Litonjua AA, Zanobetti A, Sparrow D, Vokonas P and Schwartz J: Repetitive element DNA methylation and circulating endothelial and inflammation markers in the VA normative aging study. Epigenetics 5: April 1, 2010 (Epub ahead of print).
6. Baccarelli A, Wright RO, Bollati V, Tarantini L, Litonjua AA, Suh HH, Zanobetti A, Sparrow D, Vokonas PS and Schwartz J: Rapid DNA methylation changes after exposure to traffic particles. Am J Respir Crit Care Med 179: 572-578, 2009.
7. Pogribny IP and Beland FA: DNA hypomethylation in the origin and pathogenesis of human diseases. Cell Mol Life Sci 66: 2249-2261, 2009.
8. Bird A, Taggart M, Frommer M, Miller OJ and Macleod D: A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA. Cell 40: 91-99, 1985.
9. Greger V, Passarge E, Höpping W, Messmer E and Horsthemke B: Epigenetic changes may contribute to the formation and spontaneous regression of retinoblastoma. Hum Genet 83: 155-158, 1989.
10. Shann YJ, Cheng C, Chiao CH, Chen DT, Li PH and Hsu MT. Genome-wide mapping and characterization of hypomethylated sites in human tissues and breast cancer cell lines. Genome Res 18: 791-801, 2008.
11. Tusnády GE, Simon I, Váradi A and Arányi T: BiSearch: Primer-design and search tool for PCR on bisulfite treated genomes. Nucleic Acids Res 33: e9, 2005.
12. Bailey TM and Elkan C: Fitting a mixture model by expectation maximization to discover motifs in biopolymers. In: Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology. AAAI Press, Menlo Park, CA, pp28-36, 1994.
13. Previti C, Harari O, Zwir I and del Val C: Profile analysis and prediction of tissue-specific CpG island methylation classes. BMC Bioinformatics 10: 116, 2009.

14. Kim J, Kim JY, Song KS, Lee YH, Seo JS, Jelinek J, Goldschmidt-Clermont PJ and Issa JP: Epigenetic changes in estrogen receptor beta gene in atherosclerotic cardiovascular tissues and in-vitro vascular senescence. Biochim Biophys Acta 1772: 72-80, 2007.
15. Devlin AM, Singh R, Wade RE, Innis SM, Bottiglieri T and Lentz SR: Hypermethylation of Fads2 and altered hepatic fatty acid and phospholipid metabolism in mice with hyperhomocysteinemia. J Biol Chem 282: 37082-37090, 2007.
16. Zawadzki C, Chatelain N, Delestre M, Susen S, Quesnel B, Juthier F, Jeanpierre E, Azzaoui R, Corseaux D, Breyne J, Torpier G, Staels B, Van Belle E and Jude B: Tissue factor pathway inhibitor-2 gene methylation is associated with low expression in carotid atherosclerotic plaques. Atherosclerosis 204: e4-e14, 2009.
17. Weber M, Hellmann I, Stadler MB, Ramos L, Pääbo S, Rebhan M and Schübeler D: Distribution, silencing and evolutionary impact of promoter DNA methylation in the human genome. Nat Genet 39: 457-466, 2007.
18. Shen L, Kondo Y, Guo Y, Zhang J, Zhang L, Ahmed S, Shu J, Chen X, Waterland RA and Issa J-PJ: Genome-wide profiling of DNA methylation reveals a class of normally methylated CpG island promoters. PLoS Genet 10: 2023-2036, 2007.
19. Ghazi H, Gonzales FA and Jones PA: Methylation of CpG-island-containing genes in human sperm, fetal and adult tissues. Gene 114: 203-210, 1992.
20. Turunen MP, Aavik E and Ylä-Herttuala S: Epigenetics and atherosclerosis. Biochim Biophys Acta 1790: 886-891, 2009.
21. Christensen BC, Houseman EA, Marsit CJ, Zheng S, Wrensch MR, Wiemels JL, Nelson HH, Karagas MR, Padbury JF, Bueno R, Sugarbaker DJ, Yeh RF, Wiencke JK and Kelsey KT: Aging and environmental exposures alter tissue-specific DNA methylation dependent upon CpG island context. PLoS Genet 5: e1000602, 2009.
22. Gorski DH and Walsh K: Control of vascular cell differentiation by homeobox transcription factors. Trends Cardiovasc Med 6: 213-220, 2003.
23. Tommasi S, Karm DL, Wu X, Yen Y and Pfeifer GP: Methylation of homeobox genes is a frequent and early epigenetic event in breast cancer. Breast Cancer Res 11: R14, 2009.
24. Takeshita K, Satoh M, Ii M, Silver M, Limbourg FP, Mukai Y, Rikitake Y, Radtke F, Gridley T, Losordo DW and Liao JK: Critical role of endothelial Notch1 signaling in postnatal angiogenesis. Circ Res 100: 70-78, 2007.
25. Campagnolo L, Leahy A, Chitnis S, Koschnick S, Fitch MJ, Fallon JT, Loskutoff D, Taubman MB and Stuhlmann H: EGFL7 is a chemoattractant for endothelial cells and is up-regulated in angiogenesis and arterial injury. Am J Pathol 167: 275-284, 2005.
26. Shibuya T, Watanabe K, Yamashita H, Shimizu K, Miyashita H, Abe M, Moriya T, Ohta H, Sonoda H, Shimosegawa T, Tabayashi K and Sato Y: Isolation and characterization of vasohibin-2 as a homologue of VEGF-inducible endothelium-derived angiogenesis inhibitor vasohibin. Arterioscler Thromb Vasc Biol 26: 1051-1057, 2006.
27. Fish JE, Santoro MM, Morton SU, Yu S, Yeh RF, Wythe JD, Ivey KN, Bruneau BG, Stainier DY and Srivastava D: miR-126 regulates angiogenic signaling and vascular integrity. Dev Cell 15: 272-284, 2008.
28. Thyberg J, Nilsson J, Palmberg L and Sjölund M: Adult human arterial smooth muscle cells in primary culture. Modulation from contractile to synthetic phenotype. Cell Tissue Res 239: 69-74, 1985.
29. Van der Loop FT, Gabbiani G, Kohnen G, Ramaekers FC and van Eys GJ: Differentiation of smooth muscle cells in human blood vessels as defined by smoothelin, a novel marker for the contractile phenotype. Arterioscler Thromb Vasc Biol 17: 665-671, 1997.
30. Shi C, Sakuma M, Mooroka T, Liscoe A, Gao H, Croce KJ, Sharma A, Kaplan D, Greaves DR, Wang Y and Simon DI: Down-regulation of the forkhead transcription factor Foxp1 is required for monocyte differentiation and macrophage function. Blood 112: 4699-4711, 2008.
31. Huttunen HJ, Kuja-Panula J, Sorci G, Agneletti AL, Donato R and Rauvala H: Coregulation of neurite outgrowth and cell survival by amphoterin and S100 proteins through receptor for advanced glycation end products (RAGE) activation. J Biol Chem 275: 40096-40105, 2000.
32. Yideng J, Jianzhong Z, Ying H, Juan S, Jinge Z, Shenglan W, Xiaoqun H and Shuren W: Homocysteine-mediated expression of SAHH, DNMTs, MBD2, and DNA hypomethylation potential pathogenic mechanism in VSMCs. DNA Cell Biol 26: 603-611, 2007.
33. Zhang X, Yazaki J, Sundaresan A, Cokus S, Chan SW, Chen H, Henderson IR, Shinn P, Pellegrini M, Jacobsen SE and Ecker JR: Genome-wide high-resolution mapping and functional analysis of DNA methylation in arabidopsis. Cell 126: 1189-1201, 2006.
34. 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 and Beck S: DNA methylation profiling of human chromosomes 6, 20 and 22. Nat Genet 38: 1378-1385, 2006.
35. Khachigian LM: Early growth response-1 in cardiovascular pathobiology. Circ Res 98: 186-191, 2006.