Human Molecular Genetics, 2007, Vol. 16, No. 5 **547–554**
doi:10.1093/hmg/ddm010
Advance Access published on March 5, 2007

Heritable rather than age-related environmental and stochastic factors dominate variation in DNA methylation of the human *IGF2/H19* locus

Bastiaan T. Heijmans[1,*], Dennis Kremer[1], Elmar W. Tobi[1], Dorret I. Boomsma[2] and P. Eline Slagboom[1]

[1]Molecular Epidemiology Section, Leiden University Medical Centre, Leiden 2333 ZC, The Netherlands and [2]Biological Psychology, Vrije Universiteit, Amsterdam 1081 BT, The Netherlands

Received December 7, 2006; Revised and Accepted January 26, 2007

Epigenetic variation may significantly contribute to the risk of common disease. Currently, little is known about the extent and causes of epigenetic variation. Here, we investigated the contribution of heritable influences and the combined effect of environmental and stochastic factors to variation in DNA methylation of the *IGF2/H19* locus. Moreover, we tested whether this locus was subject to age-related degeneration of epigenetic patterns as was previously suggested for global methylation. We measured methylation of the *H19* and *IGF2* differentially methylated regions (DMRs) in 196 adolescent and 176 middle-aged twins using a recently developed mass spectrometry-based method. We observed substantial variation in DNA methylation across individuals, underscoring that DNA methylation is a quantitative trait. Analysis of data in monozygotic and dizygotic twins revealed that a significant part of this variation could be attributed to heritable factors. The heritability of methylation of individual CpG sites varied between 20 and 74% for the *H19* DMR and was even higher, between 57 and 97%, for the *IGF2* DMR. Remarkably, the combined influence of environmental and stochastic factors on DNA methylation was not greater in middle-age than in adolescence, suggesting a limited role for age-related degeneration of methylation patterns at this locus. Single nucleotide polymorphisms in the *IGF2/H19* locus were significantly associated with DNA methylation of the *IGF2* DMR ($P = 0.004$). A preliminary analysis suggested an association between *H19* DMR methylation and body size ($P < 0.05$). Our study shows that variation in DNA methylation of the *IGF2/H19* locus is mainly determined by heritable factors and single nucleotide polymorphisms (SNPs) *in cis*, rather than the cumulative effect of environmental and stochastic factors occurring with age.

INTRODUCTION

Epigenetic mechanisms are intimately involved in phenomena that rely on a stable control of gene expression such as cell differentiation, X chromosome inactivation and genetic imprinting (1). In mammals, the molecular basis of epigenetic mechanisms includes a range of histone modifications, including acetylation, and the methylation of DNA, particularly methylation of cytosines of CpG dinucleotide sequences. A genomic region with deacetylated histones and methylated DNA is not accessible for the transcriptional machinery and gene expression is silenced (2). Although generally stable, the epigenetic state of a genomic region may undergo changes that result from transient environmental circumstances or from stochastic events related to inaccuracies of the machinery maintaining epigenetic patterns (3). Such changes may eventually have significant phenotypic consequences as the epigenetic state is heritable during mitosis and thus an increasing number of cells may display altered gene expression. Recent proof of this concept came from experiments with Agouti A^{vy} mice (4). Mice exposed to prenatal methyl-donor deficiency showed a reduced methylation of the *agouti* gene across tissues which was associated with increased expression of the gene product, which results in a yellow coat color. This epigenetically induced phenotype was stable over the lifetime.

*To whom correspondence should be addressed at: Molecular Epidemiology Section, Leiden University Medical Centre, Postal Zone S-05-P, PO Box 9600, 2300 RC Leiden, The Netherlands. Tel: +31 715269785; Fax: +31 715268280; Email: b.t.heijmans@lumc.nl

Human studies commonly investigate the role of epigenetic mechanisms in cancer where an epigenetic event in a single cell contributes to tumor formation (2). Studies on more widespread epigenetic changes that may affect tissue function and contribute to other diseases are, however, scarce and generally focus on the influence of diet and age. Intervention studies indicated an effect of diet on average genomic (i.e. global) DNA methylation. Several weeks of moderate depletion of the methyl-donor folate resulted in significant global hypomethylation of genomic DNA from lymphocytes in postmenopausal women (5,6). Global hypomethylation could afterwards be corrected by a diet containing sufficient folate (5). More recently, monozygotic (MZ) twins were comprehensively investigated for epigenetic differences. With increasing age, MZ twins were found to display significantly more epigenetic differences particularly at the level of global DNA methylation, methylation of repetitive sequences and global histone acetylation (7). The phenotypic consequences of such global epigenetic changes, however, are as yet unclear.

These data may indicate the widespread occurrence of diet and age-dependent degeneration of epigenetic patterns, which, in turn, might contribute to the phenotypic changes that arise during aging. An alternative interpretation, however, is that the epigenetic changes uncovered using global assessments may be enriched for well-tolerated alterations with limited phenotypic consequences. One approach to differentiate between these interpretations is the investigation of specific loci that are under epigenetic control and produce phenotypic consequences when this control is lost. The maternally imprinted insulin-like growth factor II (*IGF2*) gene on chromosome 11p15.5 is one of the best-characterized epigenetically regulated loci. It involves both the *IGF2* region itself, of which the maternal allele is methylated, and the *H19* region located ~150 kb p-ter, of which the paternal allele is methylated. Moreover, the loss of epigenetic control, specifically demethylation of the *IGF2* differentially methylated region (DMR), leads to IGF2 overexpression and is associated with a ~5-fold increased risk of colorectal neoplasia (8,9). *IGF2* imprinting defects also underlie the Beckwith–Wiedeman syndrome (10). Furthermore, *IGF2* is a major fetal growth factor and is implicated in body composition (11,12) and atherosclerosis (13).

The aim of the current study is 3-fold: first to evaluate the inter-individual variation in DNA methylation of the *H19* (14) and *IGF2* (8) DMRs in the general population, second to assess the influence of heritable versus environmental and stochastic factors on this variation and, third to examine whether the effect of these factors depends on age. To this end, DNA methylation was assessed in monozygotic (MZ) twins, who are identical at the DNA sequence level, and in dizygotic (DZ) twins, who on average share half of their genome identical-by-descent. In total, 372 twins were studied, of whom 196 were adolescent (mean age 17 years) and 176 middle-aged (mean age 45 years). DNA methylation was measured using a recently developed method by employing mass spectrometry that allows quantifying the methylation of individual CpG sites (15). In addition to the measurement of DNA methylation, SNPs were genotyped in the *IGF2*/*H19* region (*in cis*) and the *MTHFR* Ala222Val variant (*in trans*), which is associated with disturbed methyl-donor metabolism and global DNA hypomethylation (16,17), to test their association with methylation of the *H19* and *IGF2* DMRs. Finally, the association of methylation of the *H19* and *IGF2* DMRs with body size was considered in a preliminary analysis.

RESULTS

Characteristics of twins and *IGF2/H19* DMRs

DNA methylation of the *IGF2*/*H19* locus was analyzed in 196 adolescent and 176 middle-aged twins (in total, 372 individuals). The adolescent twins had a mean age of 16.7 years (SD, 2.0) and consisted of 54 MZ and 44 DZ twin pairs. The middle-aged twins had a mean age of 44.8 years (SD, 6.8) and consisted of 48 MZ and 40 DZ twin pairs.

DNA methylation was assessed at two locations: at the 5′ region of the *H19* transcript overlapping a CpG island, which is part of the *H19* DMR, and in exon 3 of the *IGF2AS* transcript, which is part of the *IGF2* DMR (Fig. 1). In the *H19* DMR, the methylation of 12 CpG sites was measured: six individually and the other six in pairs because they were directly adjacent and could not be resolved with the technology used. In the *IGF2* DMR, six CpG sites were assessed: four individually and two as a pair. Figure 2 shows the mean methylation of the various CpG sites separately for adolescents and middle-aged twins. The mean methylation of CpGs in the *H19* DMR varied between 23 and 30%. The mean methylation of the *IGF2* DMR was higher and varied around 50% (43–62%). For both regions, the average methylation was similar in adolescent and middle-aged individuals and independent of sex and age.

Inter-individual variation in DNA methylation

Considerable inter-individual variation in methylation was observed for each CpG site (Fig. 3). Methylation for the majority of sites was normally distributed, others were bimodally distributed, which could be explained by the presence of an SNP at the CpG site itself. For example, CpG site located at 382 bp in the *H19* DMR was non-methylated in a subset of individuals (Fig. 3).

Correlation in DNA methylation

DNA methylation of different CpG sites was correlated (Fig. 4). Particularly within the *H19* region, methylation was highly correlated (up to 0.92), although the two q-ter CpG sites showed lower correlations with the sites p-ter (0.42–0.73). Within the *IGF2* DMR, the bimodally distributed CpG site at 296 bp (Fig. 3) was uncorrelated with other sites. Of interest, there was a lower correlation up to 0.25 ($P = 10^{-5}$) between sites in the *H19* and the *IGF2* DMRs, which are located ~134 kb apart.

Principal components analysis (PCA) was used to reduce the dimensionality in the data by obtaining component scores for methylation. PCA extracts components from a correlated set of phenotypes and is not affected by factors merely influencing a single CpG site, such as an SNP abolishing a CpG site. Summarizing the correlated data into a smaller number of components also reduces the testing burden. Methylation of the *H19* DMR could be described by a single

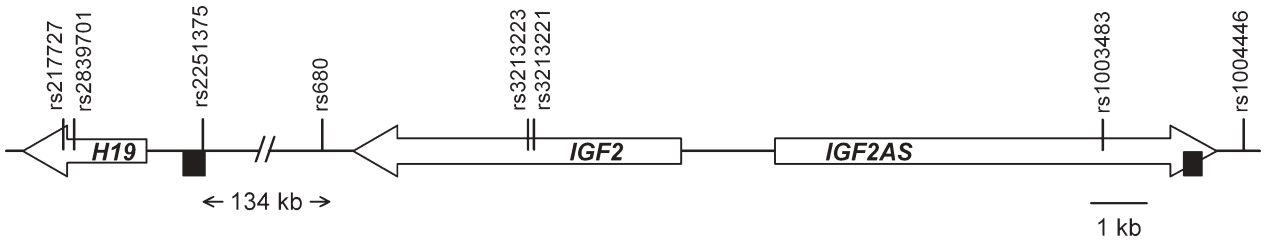

Figure 1. Overview of the *H19* and *IGF2*/*IGF2AS* genomic regions with, in black, the *H19* and *IGF2* DMRs assayed for methylation and the SNPs measured.

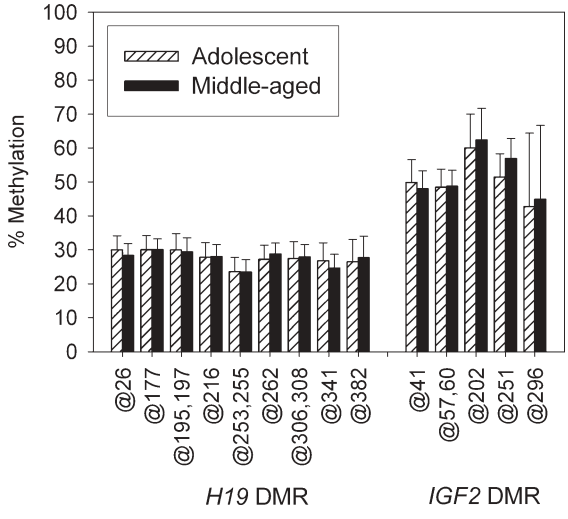

Figure 2. Mean DNA methylation for individual CpG sites in the *H19* and *IGF2* DMRs as observed in 196 adolescent (mean age 16.7) and 176 middle-aged twins (mean age 44.8).

main principal component describing 73% of the total variation brought about by the 12 CpG sites measured in this DMR. For the *IGF2* DMR (six CpG sites), 53% of the variation was captured with a single principal component.

Heritability of DNA methylation

The component scores obtained using PCA were used to estimate the genetic and environmental influences on DNA methylation in a twin design. The heritability of *H19* methylation was significant and virtually identical in the adolescent (0.33; 95% CI, 0.11–0.52) and middle-aged twins (0.35; 95% CI, 0.04–0.59), thus suggesting that the genetic control on DNA methylation at this locus was not compromised in the age range tested. There was no evidence for the influence of common environmental (e.g. familial) factors ($P = 0.87$).

The heritability of the component score describing the *IGF2* DMR methylation was higher (and thus the environmental influence smaller): 0.80 (95% CI, 0.68–0.87) in adolescent and 0.75 (95% CI, 0.59–0.84) in middle-aged twins. Again, the heritability was similar in both groups ($P = 0.82$), indicating that the influence of environmental factors did not increase with age. As for the *H19* DMR, an influence of common environment was absent ($P = 0.91$).

Figure 5 shows the heritability of all CpG sites individually for both adolescent and middle-aged twins separately as well as combined. As for the component scores, the heritabilities were substantially higher for the *IGF2* DMR than for the *H19* DMR and the heritabilities in adolescent and middle-aged twins were not statistically different with one exception (at 382 bp). Particularly, high heritabilities were observed for the three non-normally distributed CpG sites (*H19* at 382 kb *IGF2* at 202 and 296 bp). This high heritability is compatible with the effect of an SNP allele abolishing the CpG site and thus DNA methylation. When the analysis was repeated for these sites after excluding individuals with a low methylation, the heritability decreased toward that of other CpG sites in the two DMRs.

SNPs and DNA methylation

Subsequently, we investigated whether the heritability could be attributed to genetic variation at the *IGF2*/*H19* locus itself and the Ala222Val variant of the *MTHFR* gene (rs1801133), which is known to influence global DNA methylation (16,17). We genotyped three SNPs in the *H19* region and five SNPs in the *IGF2*/*IGF2AS* region (Fig. 1). The three *H19* SNPs together accounted for 1.4% of the total variance in the component score describing methylation of the *H19* region, which was not significant ($P = 0.40$, df = 3). Nevertheless, SNP rs2839701 was associated with four individual *H19* CpG sites (Fig. 6). The association of rs217727 with the non-normally distributed CpG site *H19* at 382 bp ($P = 0.0002$) was lost when individuals not methylated at this site were removed ($P = 0.37$). This is presumably explained by linkage disequilibrium (LD) between the SNP genotyped and an unknown, rare SNP abolishing the CpG site at 382 bp.

The associations of *IGF2*/*IGF2AS* SNPs with methylation of the *IGF2* DMR were more striking. The five SNPs combined explained 6.5% of the variance in the component score ($P = 0.004$, df = 5). A similar percentage of the variance of CpGs at 41 bp and 57,60 bp could be attributed to the SNPs (Table 1). Individual SNPs showing the strongest association with methylation of these two sites were rs680 ($P = 5 \times 10^{-4}$ and 10^{-3}, respectively) and rs1003483 ($P = 5 \times 10^{-4}$ and 4×10^{-4}, respectively; Fig. 6). The two SNPs were in modest LD ($D' = 0.92$, $r^2 = 0.32$) and statistical tests could not exclude an independent association of the SNPs.

As expected, the strongest association was observed for CpG site at 296 whose bimodal distribution was explained by rs1004446 ($P = 7 \times 10^{-13}$), which is in perfect LD ($r^2 = 1$) with rs3741209 that abolishes the CpG site.

In contrast to SNPs in cis, the *MTHFR* SNP was not associated with either *H19* or *IGF2* DMR methylation (Fig. 4).

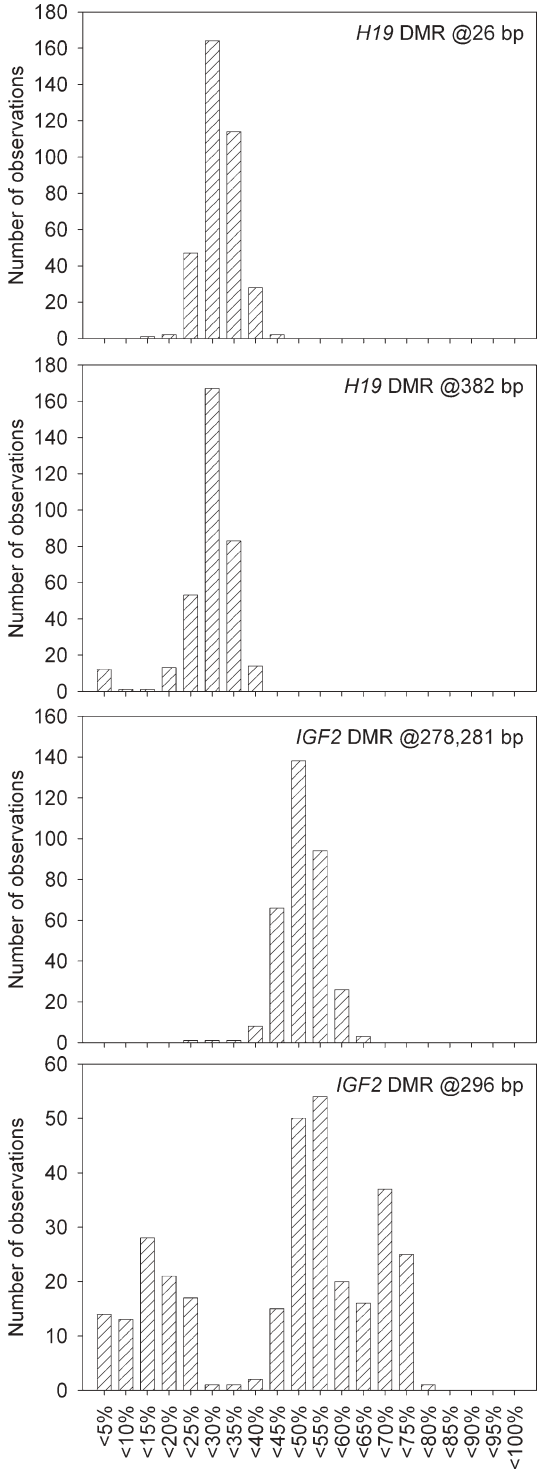

Figure 3. Distribution methylation of four representative CpG sites. SNPs underlie the non-normal distributions observed. *IGF2* DMR at 296 bp: rs3741209 at 297 bp; *IGF2* DMR at 202 bp: rs4929963 at 202 bp (data not shown); *H19* at 382 bp unknown SNP.

Methylation and body size

Measures of body size were available for the middle-aged twins. Higher methylation of the *H19* DMR as estimated using the

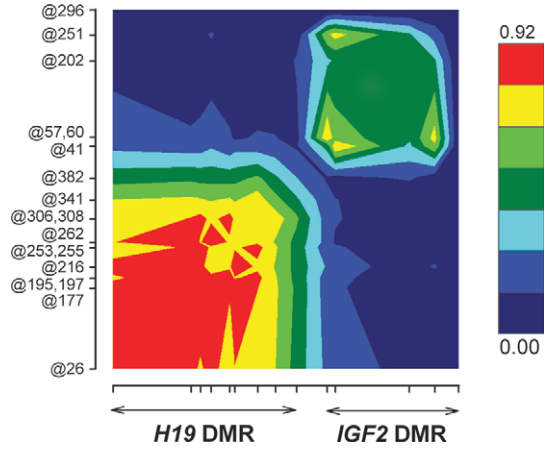

Figure 4. Graphical representation of the strength of the correlation between CpG sites.

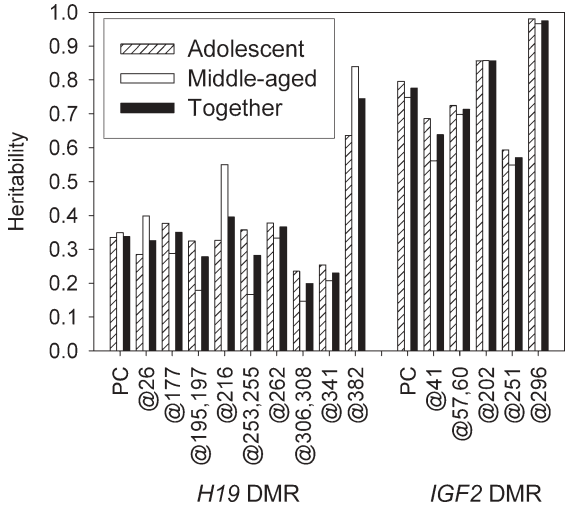

Figure 5. Heritability estimates for CpG sites and component scores based on PCA separately for adolescent and middle-aged twins as well as combined. For one CpG site (*H19* at 382 bp), the heritabilities between the two groups were significantly different ($P = 0.01$) and for another CpG site (*IGF2* DMR at 202 bp), there was evidence for a significant contribution of common environment ($P = 0.02$). PC denotes the methylation score describing the methylation per DMR obtained using PCA.

component score was associated with a smaller waist circumference ($P = 0.046$) and a lower waist–hip ratio (WHR) ($P = 0.023$), but not with body mass index (BMI). Methylation of the *IGF2* DMR was not associated with body size.

DISCUSSION

We observed substantial variation in DNA methylation of the *IGF2*/*H19* locus in both adolescent and middle-aged twins underscoring that DNA methylation is a quantitative trait (18). The variation could be attributed to heritable factors, particularly in case of the *IGF2* DMR for which the heritability was 75–80%. With 35%, the heritable influences on variation in *H19* methylation were smaller. One of the most striking

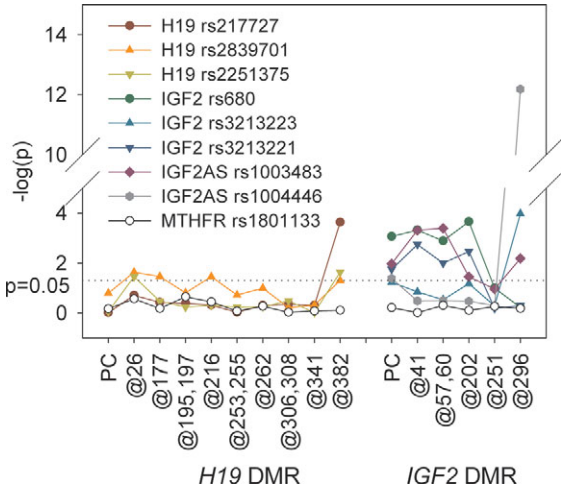

Figure 6. Association of SNPs in the *H19* and *IGF2*/*IGF2AS* genomic regions and the *MTHFR* gene with DNA methylation of *H19* and *IGF2* DMRs.

Table 1. Total variance of DNA methylation of the *H19* and *IGF2* DMRs explained by SNPs in cis

	Variance explained (%)	*P*-value
H19 DMR methylation and three SNPs in cis		
PC	1.4	0.398
At 26 bp	2.3	0.121
At 177 bp	1.9	0.179
At 195,197 bp	1.3	0.428
At 216 bp	2.0	0.180
At 253,255 bp	2.7	0.115
At 262 bp	1.4	0.366
At 306,308 bp	1.5	0.314
At 341 bp	0.9	0.532
At 382 bp	6.6	0.002
IGF2 DMR methylation and five SNPs in cis		
PC	6.5	0.004
At 41 bp	7.9	0.001
At 57,60 bp	7.1	0.004
At 202 bp	4.9	0.037
At 251 bp	2.6	0.244
At 296 bp	22.0	10^{-10}

Three SNPs were genotyped in the *H19* DMR (test with df = 3) and five in the *IGF2*/*IGF2AS* region (df = 5; see Fig. 1).

findings of the current study was that the combined influence of environmental and stochastic factors did not increase from adolescence to middle age at the expense of the heritability. This suggests that age-related degeneration of methylation patterns plays a minor role at this locus. A moderate but highly significant part of the heritability of DNA methylation was accounted for by SNPs in the *IGF2*/*H19* locus itself. In contrast, the *MTHFR* Ala222Val variant, which is known to be associated with global DNA hypomethylation (16,17), was not associated.

It has previously been speculated that variation in DNA methylation and other epigenetic characteristics may be partly heritable (3,19). This is the first study to show a substantial heritability for epigenetic variation at the *IGF2*/*H19* locus using a large series of twins. Previous data already hinted at the involvement of genetic factors in *IGF2*/*H19* methylation. Indications for familial aggregation were observed for *H19* methylation assayed using methylation-sensitive restriction endonucleases (20) and loss-of-imprinting of the *IGF2* DMR was not associated with a series of specific environmental factors including diet (9). We observed the highest heritability for the *IGF2* DMR whose methylation status was previously shown to be directly correlated with *IGF2* expression, particularly CpG sites at 41 and 57,60 bp (8).

In the current study, we reported on the methylation of an imprinted locus. Future studies are needed to establish whether our findings extend to histone modifications and non-imprinted loci. In addition, our study was limited to individuals with a maximum age of 62 years. It will be valuable to examine the occurrence of age-related degeneration of epigenetic patterns in very old age.

Methylation of the *IGF2*/*H19* locus was associated with SNPs *in cis*. For three CpG sites, the association was simply explained by a rare allele of an SNP abolishing the CpG site and thus its potential to become methylated. The other SNPs found to be associated, however, did not affect a CpG site in the DMRs and are of particular interest. Their further investigation may be a promising starting point for elucidating more sequence motifs guiding DNA methylation (21,22). The association of specific SNPs with DNA methylation suggests that its heritability has a genetic and not an epigenetic origin. This would be in line with the thought that DNA methylation of imprinted loci is erased during gametogenesis after which parent-of-origin specific patterns are established (23).

Our study indicated that the influence of environmental factors on *IGF2*/*H19* methylation was not greater in middle age than in adolescence. This implies that age-related degeneration is of limited relevance to this locus, even though its epigenetic state was hypothesized to be especially susceptible to environmental dysregulation (24). Fraga *et al*. (7) observed a greater discordance in global epigenetic measures among MZ twin pairs of increasing age. This greater discordance was observed in MZ twins from the age of 28 years and onwards (range 28–70 years), which is comparable with the age of the middle-aged twins we studied (range 34–62 years). Our study suggests that some loci may be relatively resistant to age-related changes in methylation. Fraga *et al*. (7) found age-related differences particularly in global DNA methylation, global histone modifications and repetitive sequences. Previous studies indicated that such global changes can be readily induced as well as reversed by diet (5,6) and their phenotypic consequences are currently unclear. It might be hypothesized that epigenetic patterns with greater stability during aging are more likely to be associated with phenotypic consequences if changed. The putative phenotypic consequences of quantitative variation in methylation of the *IGF2*/*H19* locus as reported here have yet to be characterized. Imprinting defects of the *IGF2*/*H19* locus, however, are associated with severe epigenetic disease (25) and an increased risk of colorectal adenoma (8,9).

A limitation of both our and the Fraga *et al*. study is that they relied on genomic DNA extracted from whole blood so that heterogeneity in cell populations may have contributed to the outcomes (26). Our study likely is less sensitive to such

heterogeneity because the epigenetic state of imprinted loci is less dependent on cell differentiation and, importantly, a previous study showed that when demethylation of the *IGF2* DMR was observed in peripheral blood lymphocytes of an individual, this was also found in colon tissue (8), which has a distinct embryologic origin (endoderm and mesoderm, respectively).

In pigs, IGF2 promotes muscle growth at the expense of fat deposition (12). In a preliminary analysis, we observed an association of higher *H19* DMR methylation with a smaller waist circumference and a lower WHR in middle-aged twins. Although the notion that common epigenetic variation might influence a phenotype in the general population is intriguing, the result should be interpreted with care. The sample size was relatively small and an explanation for the finding is speculative. It is currently unknown whether modestly increased *H19* DMR methylation upregulates *IGF2* expression. Also, it has not been established whether such quantitative differences measured in lymphocytes mark a soma-wide phenomenon as was suggested for loss-of-imprinting of the *IGF2* DMR (8).

We show the presence of substantial epigenetic variation at the *IGF2*/*H19* locus in the general population. Once established, DNA methylation of this locus is mainly genetically determined and appears to be relatively stable during aging. Our data highlight the potential of studying genetic and epigenetic variation in concert as a means to identify novel genomic determinants of common disease.

MATERIALS AND METHODS

Subjects

The adolescent and adult twins studied are part of The Netherlands Twin Register. Detailed characteristics of these samples have been described previously (27,28). The adolescent twins (age range 13–22 years) were recruited between 1985 and 1988 and all lived in the same household with their parents; the middle-aged twins (age range 34–62 years) were recruited between 1992 and 1995 and generally lived apart with their own families. Zygosity of the twins was initially determined by standardized questionnaires, typing of blood group markers and was confirmed with short tandem repeat markers measured as part of a genome-wide linkage scan (29). In the middle-aged twins, measures of body size were assessed during a visit to the research centre. Measurements included weight and height for calculation of BMI and waist and hip circumference for calculation of WHR. Informed consent was obtained from all participants and the study was approved by the Medical Ethics Committee of the Vrije Universiteit Amsterdam.

DNA methylation measurement

DNA methylation measurements were performed on genomic DNA extracted from whole blood. Genomic DNA was treated with sodium bisulfite using the EZ methylation kit (Zymo-Research). Sodium bisulfite converts unmethylated cytosines to uracils, whereas it does not affect methylated cytosines. One microgram of genomic DNA was bisulfite-treated. Methylation of individual CpG sites was assessed using the MALDI-TOF mass spectrometry-based method (Epityper, Sequenom) (15). Briefly, 10 ng of bisulfite-treated DNA is PCR-amplified with primers containing a T7-promoter. After PCR, the reverse strand of the product is transcribed using T7-polymerase and the resulting RNA is cleaved with RNase A. RNase A cleaves 3′ of every uracil thus producing different sized fragments, part of which will contain one or more original CpG sequences. Note that the reverse strand of the bisulfite-treated DNA is transcribed so that original CpG sequences are either present as CpA if unmethylated or CpG if methylated. As fragments originating from a methylated CpG sequence contain a G instead of an A-base, they will have a 16 Da higher molecular weight. This mass difference can be readily detected using a MALDI-TOF mass spectrometer (Autoflex, Bruker Daltonics). The spectra produced by the mass spectrometer were analyzed using the software QMA 1.1 (Sequenom). This software uses the genomic sequence of the amplicon to predict the fragments that occur after RNase A fragmentation and matches the expected and observed peaks taking into account that fragments are 16 Da heavier for every methylated CpG sequence they contain. The percentage methylation of individual CpG sites is calculated using the area under the peak of the signal from the unmethylated and methylated fragments. As DNA was extracted from whole blood, the percentage methylation reflects the average methylation in a population of cells.

The quantitative nature of the assay was shown previously (15). To achieve optimal quantification within the limitations of the technology, all samples were measured in triplicate. If the triplicate measurements had a standard deviation equal to or greater than 0.10, all data for the sample involved were discarded (removing 3.5% of measurements). In addition, DNA samples for which methylation could be established for $<60\%$ of the CpG sites were removed prior to analyses (removing 4.0% of measurements). Finally, CpG sites with a success rate lower than 80% after applying the previous criteria were not analyzed (discussed subsequently).

DNA methylation was assessed in two regions (Fig. 1): in the 5′ region of the *H19* transcript overlapping a CpG island, which is part of the *H19* DMR (14), and in exon 3 of the *IGF2AS* transcript, which is part of the *IGF2* DMR (8). The *H19* amplicon was amplified using the primers (tags in lower case) aggaagagag-GGGTTTGGGAGAGTTTGTGAGGT (forward) and cagtaatacgactcactatagggagaaggct-ATACCTACTACTCCCTAC CTACCAAC (reverse) and encompassed 413 bp (NCBI build 34, chr11: 1 975 948–1 976 360). All primers were designed on the reverse complement strand and anneal to sequences devoid of CpG dinucleotides. CpG sites are numbered according to the position of its cytosine in the amplicon counting from p-ter onwards. The *H19* DMR amplicon harbored 25 CpG sites, 12 of which were successfully measured (three fragments contained two adjacent CpG sites). Six CpG sites on four fragments could not be measured independently because the fragments had the same molecular weight and were overlapping in the spectrum; six CpG sites could not be assessed because of a success rate $<80\%$; and one site could not be measured because of an adjacent SNP (rs12292822) compromising the correct prediction of fragments by the software. The *IGF2* DMR amplicon was amplified using the primers aggaagagag-TGGATAGGA-GATTGAGGAGAAA (forward) and cagtaatacgactcactata-gggagaaggct-AAACCCCAACAAAAACCACT (reverse) and encompassed 338 bp (NCBI build 34, chr11: 2 126 035–

2 126 372; note that base 2 126 154 is a T instead of a C in >99% of individuals of European descent, thus abolishing a CpG site). The amplicon harbored seven CpG sites, six of which could be measured successfully (one fragment contained two CpG sites). CpG sites at 41 bp and 57,60 bp coincided with those studied by Cui *et al*. (8) and were highly correlated with *IGF2* expression. Measurement of the remaining CpG sites was compromised by an adjacent SNP (rs17883338). The mean success rate for the 18 successfully assayed CpG sites (distributed over 14 fragments) was 92.5%.

Genotyping

Using the HapMap phase II data, SNPs were selected in the *H19* and *IGF2*/*IGF2AS* regions using a pairwise tagging strategy ($r^2 > 0.8$; Fig. 1). Two SNPs commonly genotyped in other studies were also selected [rs217727 and rs680 (11,30)]. In addition, the Ala222Val (=677C/T= rs1801133) variant of the *MTHFR* gene was genotyped. Genotyping of the nine SNPs was done using the mass spectrometry-based hME assay (Sequenom) in two multiplex reactions. Average genotyping success rate was 99.1%; for 13 of 378 samples with DNA methylation data, there was insufficient DNA for genotyping, resulting in an effective average success rate of 95.7%. The observed minor allele frequencies were 0.18 (rs217727), 0.47 (rs2839701), 0.27 (rs2251375), 0.31 (rs680), 0.20 (rs3213223), 0.41 (rs3213221), 0.46 (rs1003483), 0.38 (rs1004446) and 0.36 (rs1801133).

Statistical analysis

Statistical analyses were carried out with four goals (1): to reduce the high dimensional data set of correlated CpG methylation data into a smaller number of independent factors using PCA (2), to estimate the heritability of variation in methylation of the individual sites as well as the component scores obtained using PCA (3), to assess the association between SNPs and methylation and (4) to test for association between DNA methylation and body size. Prior to PCA, correlations between methylation of different CpG sites were visualized using GOLD (31). PCA was performed on correlated methylation data of individual CpG sites separately for the *H19* and *IGF2* DMRs using SPSS 11. PCA indicated one factor with an eigenvalue exceeding 1 for either DMR. In the case of the *H19* DMR, this single factor explained 73% of the variance in methylation, whereas it explained 53% of the variance in *IGF2* DMR methylation. To assess the heritability, estimates of additive genetic V(A), common environmental V(C) and unique environmental and stochastic V(E) variance in DNA methylation were obtained using variance components analysis implemented in Mx 1.61 (http://www.vcu.edu/mx/). The analysis makes use of the differential genetic resemblance of monozygotic and dizygotic twins to estimate the effects of genes and environment on phenotypic variation. Differences in variance components and in heritability between adolescent and middle-aged twins were assessed by testing whether the model with the same absolute estimates for V(A), V(C) and V(E) in adolescent and middle-aged twins had a significantly worse fit to the data than a model with separate estimates for the two age groups. The association of SNPs with DNA methylation was tested as additive genotypic effects within the same variance components setting, thereby taking into account the dependency in the observations from twins. Pairwise linkage disequilibria between SNPs were estimated using HaploView (32). Testing for association between DNA methylation and body size was also done in the variance components setting with body size as outcome measure, the component scores for the two DMRs as explaining variables and age and sex as confounders. Prior to statistical analysis, methylation data were adjusted for bisulfite batch to exclude potential confounding effects.

ACKNOWLEDGEMENTS

We wish to thank Gabor Abbas for excellent technical assistance and Drs Matthias Ehrich and Dirk van den Boom from Sequenom, San Diego CA, USA for making available software for DNA methylation measurements. Financial support from The Netherlands Heart Foundation (grants 86.083, 88.042 and 90.313), The Netherlands Organization for Scientific Research NWO (grant 911-03-016) and the Centre for Medical Systems Biology (CMSB), a centre of excellence approved by the Netherlands Genomics Initiative/Netherlands Organization for Scientific Research (NWO), is gratefully acknowledged.

Conflict of Interest statement. No conflicts of interests declared.

REFERENCES

1. Jaenisch, R. and Bird, A. (2003) Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat. Genet.*, **33** (suppl.), 245–254.
2. Jones, P.A. and Baylin, S.B. (2002) The fundamental role of epigenetic events in cancer. *Nat. Rev. Genet.*, **3**, 415–428.
3. Petronis, A. (2006) Epigenetics and twins: three variations on the theme. *Trends Genet.*, **22**, 347–350.
4. 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.
5. Jacob, R.A., Gretz, D.M., Taylor, P.C., James, S.J., Pogribny, I.P., Miller, B.J., Henning, S.M. and Swendseid, M.E. (1998) Moderate folate depletion increases plasma homocysteine and decreases lymphocyte DNA methylation in postmenopausal women. *J. Nutr.*, **128**, 1204–1212.
6. Rampersaud, G.C., Kauwell, G.P., Hutson, A.D., Cerda, J.J. and Bailey, L.B. (2000) Genomic DNA methylation decreases in response to moderate folate depletion in elderly women. *Am. J. Clin. Nutr.*, **72**, 998–1003.
7. Fraga, M.F., Ballestar, E., Paz, M.F., Ropero, S., Setien, F., Ballestar, M.L., Heine-Suner, D., Cigudosa, J.C., Urioste, M., Benitez, J. *et al*. (2005) Epigenetic differences arise during the lifetime of monozygotic twins. *Proc. Natl Acad. Sci. USA*, **102**, 10604–10609.
8. Cui, H., Cruz-Correa, M., Giardiello, F.M., Hutcheon, D.F., Kafonek, D.R., Brandenburg, S., Wu, Y., He, X., Powe, N.R. and Feinberg, A.P. (2003) Loss of IGF2 imprinting: a potential marker of colorectal cancer risk. *Science*, **299**, 1753–1755.
9. Cruz-Correa, M., Cui, H., Giardiello, F.M., Powe, N.R., Hylind, L., Robinson, A., Hutcheon, D.F., Kafonek, D.R., Brandenburg, S., Wu, Y. *et al*. (2004) Loss of imprinting of insulin growth factor II gene: a potential heritable biomarker for colon neoplasia predisposition. *Gastroenterology*, **126**, 964–970.
10. Robertson, K.D. (2005) DNA methylation and human disease. *Nat. Rev. Genet.*, **6**, 597–610.
11. O'Dell, S.D., Miller, G.J., Cooper, J.A., Hindmarsh, P.C., Pringle, P.J., Ford, H., Humphries, S.E. and Day, I.N. (1997) Apal polymorphism in

insulin-like growth factor II (IGF2) gene and weight in middle-aged males. *Int. J. Obes. Relat. Metab. Disord.*, **21**, 822–825.
12. Van Laere, A.S., Nguyen, M., Braunschweig, M., Nezer, C., Collette, C., Moreau, L., Archibald, A.L., Haley, C.S., Buys, N., Tally, M. *et al*. (2003) A regulatory mutation in IGF2 causes a major QTL effect on muscle growth in the pig. *Nature*, **425**, 832–836.
13. Zaina, S., Pettersson, L., Ahren, B., Branen, L., Hassan, A.B., Lindholm, M., Mattsson, R., Thyberg, J. and Nilsson, J. (2002) Insulin-like growth factor II plays a central role in atherosclerosis in a mouse model. *J. Biol. Chem.*, **277**, 4505–4511.
14. Vu, T.H., Li, T., Nguyen, D., Nguyen, B.T., Yao, X.M., Hu, J.F. and Hoffman, A.R. (2000) Symmetric and asymmetric DNA methylation in the human IGF2-H19 imprinted region. *Genomics*, **64**, 132–143.
15. Ehrich, M., Nelson, M.R., Stanssens, P., Zabeau, M., Liloglou, T., Xinarianos, G., Cantor, C.R., Field, J.K. and van den, B.D. (2005) Quantitative high-throughput analysis of DNA methylation patterns by base-specific cleavage and mass spectrometry. *Proc. Natl Acad. Sci. USA*, **102**, 15785–15790.
16. Friso, S., Choi, S.W., Girelli, D., Mason, J.B., Dolnikowski, G.G., Bagley, P.J., Olivieri, O., Jacques, P.F., Rosenberg, I.H., Corrocher, R. *et al*. (2002) A common mutation in the 5,10-methylenetetrahydrofolate reductase gene affects genomic DNA methylation through an interaction with folate status. *Proc. Natl Acad. Sci. USA*, **99**, 5606–5611.
17. Castro, R., Rivera, I., Ravasco, P., Camilo, M.E., Jakobs, C., Blom, H.J. and de Almeida, I.T. (2004) 5,10-methylenetetrahydrofolate reductase (MTHFR) 677C–>T and 1298A–>C mutations are associated with DNA hypomethylation. *J. Med. Genet.*, **41**, 454–458.
18. Sandovici, I., Kassovska-Bratinova, S., Loredo-Osti, J.C., Leppert, M., Suarez, A., Stewart, R., Bautista, F.D., Schiraldi, M. and Sapienza, C. (2005) Interindividual variability and parent of origin DNA methylation differences at specific human Alu elements. *Hum. Mol. Genet.*, **14**, 2135–2143.
19. Bjornsson, H.T., Fallin, M.D. and Feinberg, A.P. (2004) An integrated epigenetic and genetic approach to common human disease. *Trends Genet.*, **20**, 350–358.
20. Sandovici, I., Leppert, M., Hawk, P.R., Suarez, A., Linares, Y. and Sapienza, C. (2003) Familial aggregation of abnormal methylation of parental alleles at the IGF2/H19 and IGF2R differentially methylated regions. *Hum. Mol. Genet.*, **12**, 1569–1578.
21. Bock, C., Paulsen, M., Tierling, S., Mikeska, T., Lengauer, T. and Walter, J. (2006) CpG island methylation in human lymphocytes is highly correlated with DNA sequence, repeats, and predicted DNA structure. *PLoS Genet.*, **2**, e26.
22. Feltus, F.A., Lee, E.K., Costello, J.F., Plass, C. and Vertino, P.M. (2006) DNA motifs associated with aberrant CpG island methylation. *Genomics*, **87**, 572–579.
23. Morgan, H.D., Santos, F., Green, K., Dean, W. and Reik, W. (2005) Epigenetic reprogramming in mammals. *Hum. Mol. Genet.*, **14** (Spec no. 1), R47–R58.
24. Waterland, R.A. and Jirtle, R.L. (2004) Early nutrition, epigenetic changes at transposons and imprinted genes, and enhanced susceptibility to adult chronic diseases. *Nutrition*, **20**, 63–68.
25. Jiang, Y.H., Bressler, J. and Beaudet, A.L. (2004) Epigenetics and human disease. *Annu. Rev. Genomics Hum. Genet.*, **5**, 479–510.
26. Martin, G.M. (2005) Epigenetic drift in aging identical twins. *Proc. Natl Acad. Sci. USA*, **102**, 10413–10414.
27. Boomsma, D.I., Kaptein, A., Kempen, H.J., Gevers Leuven, J.A. and Princen, H.M. (1993) Lipoprotein(a): relation to other risk factors and genetic heritability. Results from a Dutch parent-twin study. *Atherosclerosis*, **99**, 23–33.
28. Snieder, H., van Doornen, L.J. and Boomsma, D.I. (1997) The age dependency of gene expression for plasma lipids, lipoproteins, and apolipoproteins. *Am. J. Hum. Genet.*, **60**, 638–650.
29. Heijmans, B.T., Beekman, M., Putter, H., Lakenberg, N., van der Wijk, H.J., Whitfield, J.B., Posthuma, D., Pedersen, N.L., Martin, N.G., Boomsma, D.I. *et al*. (2005) Meta-analysis of four new genome scans for lipid parameters and analysis of positional candidates in positive linkage regions. *Eur. J. Hum. Genet.*, **13**, 1143–1153.
30. Petry, C.J., Ong, K.K., Barratt, B.J., Wingate, D., Cordell, H.J., Ring, S.M., Pembrey, M.E., Reik, W., Todd, J.A. and Dunger, D.B. (2005) Common polymorphism in H19 associated with birthweight and cord blood IGF-II levels in humans. *BMC Genet.*, **6**, 22.
31. Abecasis, G.R. and Cookson, W.O. (2000) GOLD—graphical overview of linkage disequilibrium. *Bioinformatics*, **16**, 182–183.
32. Barrett, J.C., Fry, B., Maller, J. and Daly, M.J. (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*, **21**, 263–265.