Human Molecular Genetics, 2012, Vol. 21, No. 9 ***2086–2101***
doi:10.1093/hmg/dds026
Advance Access published on February 3, 2012

Periconceptional maternal micronutrient supplementation is associated with widespread gender related changes in the epigenome: a study of a unique resource in the Gambia

Batbayar Khulan[1,2,†,‡], Wendy N. Cooper[2,‡], Benjamin M. Skinner[1], Julien Bauer[1], Stephen Owens[3,4], Andrew M. Prentice[3,4], Gusztav Belteki[5,6], Miguel Constancia[2,7,8], David Dunger[6] and Nabeel A. Affara[1,*]

[1]Mammalian Molecular Genetics Group, Department of Pathology, University of Cambridge, Tennis Court Road, Cambridge CB2 1QP, UK, [2]Metabolic Research Laboratories, Department of Obstetrics and Gynaecology, University of Cambridge, Cambridge CB2 0SW, UK, [3]MRC International Nutrition Group, London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK, [4]MRC Keneba, The Gambia, [5]Neonatal Intensive Care Unit, Cambridge University Hospitals NHS FoundationTrust, Box 226 The Rosie Hospital, Robinson Way, Cambridge CB2 0QQ, UK, [6]Department of Paediatrics, University of Cambridge, Cambridge CB2 0QQ, UK, [7]National Institute for Health Research, Cambridge Biomedical Research Centre, Cambridge CB2 0QQ, UK and [8]Centre for Trophoblast Research, University of Cambridge, Cambridge CB2 3EG, UK

Received December 28, 2011; Revised and Accepted January 30, 2012

In addition to the genetic constitution inherited by an organism, the developmental trajectory and resulting mature phenotype are also determined by mechanisms acting during critical windows in early life that influence and establish stable patterns of gene expression. This is the crux of the developmental origins of health and disease hypothesis that suggests undernutrition during gestation and infancy predisposes to ill health in later life. The hypothesis that periconceptional maternal micronutrient supplementation might affect fetal genome-wide methylation within gene promoters was explored in cord blood samples from offspring of Gambian women enrolled into a unique randomized, double blind controlled trial. Significant changes in the epigenome in cord blood DNA samples were further explored in a subset of offspring at 9 months. Gender-specific changes related to periconceptional nutritional supplementation were identified in cord blood DNA samples, some of which showed persistent changes in infant blood DNA samples. Significant effects of periconceptional micronutrient supplementation were also observed in postnatal samples which were not evident in cord blood. In this Gambian population, the increased death rate of individuals born in nutritionally poor seasons has been related to infection and it is of interest that we identified differential methylation at genes associated with defence against infection and immune response. Although the sample size was relatively small, these pilot data suggest that periconceptional nutrition in humans is an important determinant of newborn whole genome methylation patterns but may also influence postnatal developmental patterns of gene promoter methylation linking early with disease risk.

*To whom correspondence should be addressed. Tel: +44 1223333700; Fax: +44 1223333346; Email: na@mole.bio.cam.ac.uk
†Present address: Endocrinolgy Unit, Center for Cardiovascular Science, Queen's Medical Research Institute, University of Edinburgh, 47 Little France Crescent, Edinburgh EH16 4TJ, UK.
‡Joint first authors with equal contribution.

INTRODUCTION

In addition to the genetic constitution inherited by an organism, the developmental trajectory and resulting mature phenotype are also determined by mechanisms acting during critical windows in early life that influence and establish stable patterns of gene expression. Such mechanisms involve epigenetic processes of which DNA methylation is the most stable and well-established mediator (1,2). The developmental plasticity achieved by this fusion of genetic and epigenetic mechanisms permits adjustment of gene activity in the developing fetus and may facilitate adaptation to external influences, and hence improved chances of survival. These epigenetic modifications to gene activity can persist and hence set the metabolic state of an organism essentially for life. Adverse fetal conditions may induce inappropriate levels of gene expression in the postnatal period and through to adult life that may foster conditions leading to pathology of various body systems (3,4). This is the crux of the developmental origins of health and disease hypothesis which suggests that the metabolic syndrome (coronary heart disease, type 2 diabetes, stroke, hypertension) is rooted in undernutrition during fetal life and infancy (1,5–8). Under this hypothesis, disease arises from a mismatch between a deprived intrauterine nutritional environment (for which epigenetic adaptations have been imposed in early development) and a comparatively rich postnatal environment (9–11).

Maternal nutrition during pregnancy has a strong influence on the intrauterine environment and manipulations in the rat (global maternal undernutrition, maternal low protein by isocaloric diets and uterine artery ligation) lead to offspring with intrauterine growth restriction (IUGR) associated with histological alterations to various organs, obesity, insulin resistance and endothelial dysfunction in later life (12–16). These phenotypic changes are accompanied by significant gene expression changes between offspring from normal and undernourished mothers (17). Some insight into the mechanistic basis of this pathology comes from studies on candidate genes involved in important metabolic pathways [peroxisomal proliferator-activated receptor (PPARa) and the glucocorticoid receptor] where changes in the methylation status of promoters of these genes and in their mRNA levels have been observed between dietary restricted and control offspring (18–20). Alterations to global histone and DNA methylation in rat brain have been observed following uterine artery ligation (14). These findings lend support to the proposition that epigenetic modifications of the fetal genome in response to the intrauterine environment establish altered patterns of gene expression that reprogram the metabolic state and are maintained into adulthood (17).

These experimental models are based on extreme global maternal nutritional distortion and are similar to the Dutch famine study where offspring conceived in this period showed decreased methylation at the *IGF2* DMR0 (21). There is, however, evidence from murine and other models that more subtle nutritional manipulations and supplementation can influence epigenetic regulation (22). Deficiencies in micronutrients such as folic acid, vitamin B6 and B12, zinc and selenium have been shown to result in alteration to both DNA and histone methylation (23).

In rural Gambia, the alternation of a single rainy season (June–October) with a long dry season has an important impact on nutrition patterns. A high incidence of deficiencies in several micronutrients has been demonstrated in pregnant women from the Kiang West region, particularly during the nutritionally depressed rainy season (24,25). This seasonal deficiency is associated with increased incidence of low birth weight babies, primarily the result of IUGR (26) and variations in infant and childhood morbidity and mortality (27) and young adult mortality (28,29).

Our study investigates a unique resource resulting from a double-blinded Pre- and Periconceptional Multiple Micronutrient Supplementation trial (ISRCTN 13687662) to assess the effect of supplementation on the offspring of women in rural Gambia. The study has compared the global methylation patterns of genomic DNA between groups of offspring whose mothers were taking micronutrient supplementation or placebo during the pre- and periconceptional period. The comparison has been made between cord blood from newborn babies from the two groups and also peripheral blood from infants (about 9 months old) in later life. The findings indicate changes in the methylation of CpG loci associated with micronutrient supplementation, age and supplementation and age alone and show that there are differences between males and females in all categories, suggesting that males and females potentially follow different developmental trajectories reflecting the physiological differences between the sexes both in the presence and absence of micronutrient supplementation.

RESULTS

Umbilical cord and peripheral blood DNA samples from offspring born to women who took part in the micronutrient supplementation trial were analysed using the HumanMethylation 27 Beadchips. The data from male and female samples (previously verified by polymerase chain reaction (PCR) analysis with *SRY* and *AMELXY*—amelogenenin on the X and Y—sex-specific primers to confirm sex status) were analysed separately to remove the confounding effects of X-linked loci that undergo X-inactivation and, hence, identify genuine sex-specific differences in methylation patterns (see Materials and Methods for details and quality control of the samples).

The data were filtered to select only those loci that showed $>10\%$ change in methylation status at false discovery rate (FDR) corrected P-values < 0.001 (the full raw data files have been uploaded to GEO at the NCBI website; accession number GSE34257). Several pairwise comparisons were performed to identify; first, the effects of micronutrient supplementation on methylation status of CpG loci in cord and infant samples; second, any methylation changes that occur between the cord and infant stages (postnatal changes) in both the placebo and micronutrient supplementation cohorts; third, to identify any differential methylation of CpG loci between the sexes (excluding X-linked genes) in cord and infant samples. Figure 1 summarizes the comparisons (and comparisons of comparisons) designed to identify different classes of methylation change together with the numbers of loci involved. Table 1 summarizes the numbers of loci that change in each of these pairwise comparisons indicating their chromosomal distribution (sex-linked or autosomal) and the relative direction of methylation change (delta beta

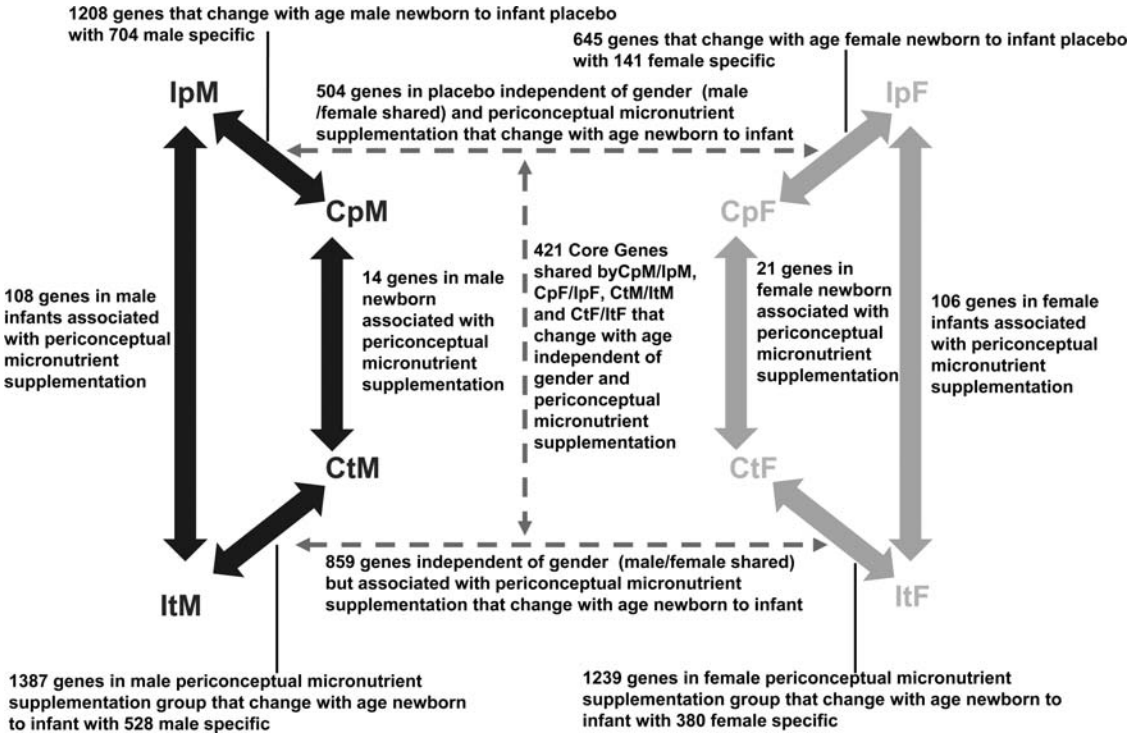

Figure 1. Schematic representation of pairwise comparisons between males and females of placebo and supplementation cohorts for newborn and infants. This figure presents a diagrammatic representation of the various pairwise comparisons made between the different newborn and infant cohorts. The numbers represent the number of loci found to be differentially methylated between states under comparison. I, infant; C, cord (newborn); p, placebo; t, supplemented; M, male; F, female. For example, IpM represents the infant placebo male cohort.

score—increase or decrease when delta delta betas were greater than $\pm 5\%$). The detailed lists of genes, CpG loci, relative measure of methylation change and DiffScores (FDR corrected *P*-value equivalents where 30 or -30 is equivalent to $P = 0.001$, 40 or -40 $P = 0.0001$ etc) are given in Supplementary Material, Dataset 1. There are several observations to emerge from this analysis.

Pairwise comparisons to identify methylation changes associated with micronutrient supplementation

(i) A small number of changes are observed in cord genomic DNA supplementation versus placebo comparisons: males showed 14 and females 21 differentially methylated loci, suggesting that periconceptional micronutrient supplementation programmed changes may only become evident at a later age (Table 1).

(ii) There is no overlap between genders in loci showing methylation changes in the cord comparisons : the majority of the changes are found among autosomal loci in both sexes for the cord placebo to supplementation comparisons with no overlap between the genders.

(iii) A greater number of changes are observed in infant genomic DNA supplementation versus placebo comparisons: males showed 108 and females 106 differentially methylated loci with only 5 (*XIRP1, MEOX1, GNA11, C1orf54, KRTAP21-1*) in agreement between male and female infants.

(iv) Persistence of methylation changes between cord and infants in supplementation versus placebo comparisons: despite the differences between sexes, within each sex a significant proportion of the micronutrient associated methylation changes (that meet the threshold criteria) observed in cord bloods are also observed in infants (males 7/14: *MEOX1, SPAG4L, SPATA22, NRN1L, C1orf54, PIP, PTPN20B*; females 8/21*: HSPC176, XIRP1, BPIL1, MGMT, CHIT1, C14orf152, KLRC2, DEFB123*).

Pairwise comparisons to identify postnatal methylation changes

(i) Postnatal changes in placebo comparisons: for both males and females, pairwise comparisons of the cord to infant blood DNA methylation within the placebo cohorts show that a large number of CpG loci change methylation postnatally, with more changes evident in males (1208) than females (645). Seventy eight percent (504/645) of the changes found in females are shared with males (Fig. 1). There are 704 male-specific and 141 female-specific changes postnatally.

(ii) Postnatal changes in supplementation comparisons: the same is true for comparison of cord to infant blood DNA methylation within the micronutrient supplementation cohorts for both sexes except that the disparity between males (1387) and females (1239) changes is reduced. Sixty-nine percent (859/1239) of the changes in females are shared with males, a slightly lower degree of overlap than found in the placebo cohorts. There are 528 male-specific and 380 female-specific changes postnatally.

(iii) Postnatal methylation changes independent of gender and micronutrient supplementation: a set of 421 'core'

Table 1. Summary of methylation changes in pairwise comparisons

Pairwise comparison	Increased methylation	Decreased methylation	Comment
CpM versus CtM			Effect of micronutrient supplementation in newborn males
Autosomal	6	7	
Sex linked	0	1	
CpF versus CtF			Effect of micronutrient supplementation in newborn females
Autosomal	7	14	
Sex linked	0	0	
IpM versus ItM			Effect of micronutrient supplementation in infant males
Autosomal	51	52	
Sex linked	2	3	
IpF versus ItF			Effect of micronutrient supplementation in infant females
Autosomal	29	76	
Sex linked	0	1	
CpM versus IpM			Comparison male newborn and infant placebo groups to identify postnatal changes
Autosomal	386	752	
Sex linked	12	58	
CpF versus IpF			Comparison female newborn and infant placebo groups to identify postnatal changes
Autosomal	352	278	
Sex linked	9	6	
CtM versus ItM			Comparison male newborn and infant micronutrient supplementation groups to identify postnatal changes associated with micronutrient supplementation
Autosomal	503	814	
Sex linked	13	57	
CtF versus ItF			Comparison female newborn and infant micronutrient supplementation groups to identify postnatal changes associated with micronutrient supplementation
Autosomal	499	716	
Sex linked	10	14	

I, infant; C, cord (newborn); p, placebo; t, supplemented; M, male; F, female. For example, IpM represents the male infant placebo cohort. The data show that more changes are demethylations than methylations in response to both age and supplementation and that the majority of the changes are autosomal. A significant methylation change was considered to be a delta beta (fractional change) $\geq \pm 0.1$ at a DiffScore $\geq \pm 30$ (FDR corrected *P*-value equivalent of 0.001).

methylation changes between cord blood and infant blood DNA (Fig. 1) is concordant across both male and female infants in both placebo and micronutrient supplementation cohorts. These are likely to be changes associated with early postnatal development rather than micronutrient supplementation; for example, changes in cellular composition of infant blood compared with cord blood (see Supplementary Material, Dataset 2 for details).

The 'core' set of changes predominantly involves increases in methylation, with ~65% of core loci showing an increase and 35% a decrease. In contrast, the sex-specific changes in both placebo and supplemented cohorts predominantly involve decreases in methylation.

There are a number of interesting points established by the above findings. First, periconceptional micronutrient supplementation is associated with differential methylation changes detectable at birth and in infants at 9 months. Secondly, there is marked difference between males and females in the loci undergoing methylation change. Thirdly, some 50% of the differentially methylated loci detected in cord blood DNA persist into infant blood DNA. Fourthly, in both the placebo and micronutrient supplementation cohorts, there are postnatal changes specific to males and females. Fifthly, a 'core' of postnatal changes are shared between males and females and are independent of micronutrient supplementation.

Analysis of non-core CpG loci changing postnatally

From the analyses summarized above, a group of core (associated with postnatal development rather than micronutrient supplementation) CpG loci have been identified. Removal of these core loci from the loci changing postnatally (newborn to infant) for male and female placebo and micronutrient supplementation cohorts has produced four groups of non-core CpG loci (CpM versus IpM—787 loci; CtM versus ItM—966 loci; CpF versus IpF—224 loci and CtF versus ItF—818 loci). Table 2 presents a comparison for males and females of the postnatal changes between the placebo and supplementation cohorts and has permitted fractionation of non-core loci into those that are shared, those specific to male or female placebo cohorts postnatally and those specifically associated with postnatal development and micronutrient supplementation in each sex. Where loci are shared between placebo and supplemented cohorts, delta betas (relative change in methylation levels) were compared and scored as equal if within 0.05. Using these criteria, the majority of shared changes are of similar direction and magnitude (see Supplementary Material, Datasets 3 and 4 for detailed listings of loci and associated genes). Table 3 compares these groups of loci between the sexes to permit more refined differentiation of categories of loci undergoing change in relation to age, sex and micronutrient supplementation in order to highlight sex differences (see Supplementary Material, Dataset 5 for detailed listings of loci and associated genes). Figure 2 depicts schematically the comparisons summarized in Tables 2 and 3. The main point to emerge from this is that the majority of postnatal changes for placebo and supplementation groups are gender specific with few shared loci, indicating different sex-specific trajectories.

Table 2. Different categories of non-core loci changing postnatally

Methylation change	CpM versus IpM/CtM versus ItM (non-core loci changing postnatally newborn to infant in placebo compared with non-core genes changing postnatally associated with micronutrient supplementation newborn to infant in males)				CpF versus IpF/CtF versus ItF (non-core loci changing postnatally newborn to infant in placebo compared with non-core genes changing postnatally associated with micronutrient supplementation newborn to infant in females)			
	Increase		Decrease		Increase		Decrease	
	X	A	X	A	X	A	X	A
Shared: treatment change = placebo	1	50	23	258	1	51	0	23
Shared: treatment change > placebo	1	15	5	37	0	9	0	6
Shared: treatment change < placebo	0	2	1	6	0	2	0	2
Placebo specific	3	79	27	279	1	50	4	75
Age and supplementation specific	4	196	26	341	2	197	12	513

See also Figure 4. Supplementation age change was considered equal to placebo age change when delta betas were within 0.05. I, infant; C, cord (newborn); p, placebo; t, supplemented; M, male; F, female. For example, IpM represents the infant placebo male cohort. The data show that more changes are occurring in males than females, that more changes are decreases than increases in methylation and are in autosomal loci. Additionally, where loci are shared between placebo and supplemented cohorts, the magnitude of the change is predominantly equal.

Table 3. Male and female specific and male and female shared non-core loci changing postnatally in placebo and supplemented groups

	M specific		F specific		Shared
	X	A	X	A	X
Placebo					
Increase	3	74	1	45	0
Decrease	26	260	3	56	1
Supplemented					
Increase	4	144	2	145	0
Decrease	25	259	11	431	1

See also Figure 4. Supplementation age change was considered equal to placebo age change when delta betas were within 0.05. I, infant; C, cord (newborn); p, placebo; t, supplemented; M, male; F, female. For example, IpM represents the infant placebo male cohort. The data show that more changes are occurring in males than females, that more changes are decreases than increases in methylation and are in autosomal loci. Additionally, where loci are shared between placebo and supplemented cohorts, the magnitude of the change is predominantly equal.

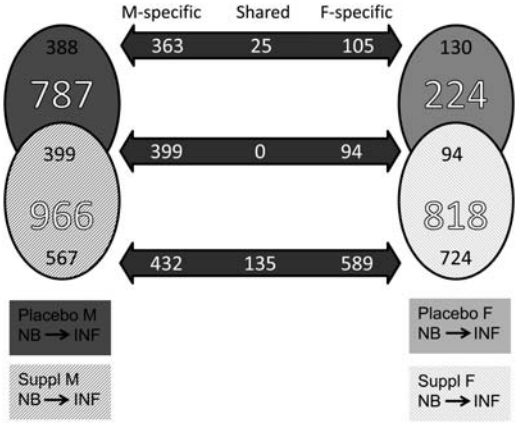

Figure 2. Schematic representation of the non-core loci comparisons detailed in Tables 2 and 3. In this figure, each oval represents the non-core loci changing with age (large white numbers) in a pairwise comparison. These have then been compared to identify placebo-specific, supplementation specific and shared loci for each gender (smaller black numbers; see also Table 2). The double-headed arrows show the overlaps and specificities between genders (smaller white numbers; see also Table 3). These data show that many of the loci are specific to both supplementation state and to gender.

It is interesting to note that in reciprocal comparisons (male placebo to female supplemented and female placebo to male supplemented postnatally), far greater numbers of loci are shared between the male placebo and female supplemented groups (418) than vice versa (72). This suggests that the male placebo group is more similar to both male and female supplemented groups than the female placebo group is to either male or female supplemented groups. It is possible that this reflects different postnatal nutritional environments for males and females where, for cultural reasons, male offspring may have a nutritionally privileged status.

Changes in loci associated with imprinted genes

A comprehensive list of 97 human and mouse imprinted genes (30) was compared against all loci and associated genes showing differential methylation in order to determine whether the methylation status of imprinted genes is influenced by nutrition or sex or changes during postnatal development. Table 4 summarizes the magnitude and direction of changes detected at specific CpG loci in 15 of these 97 genes, the variable associated with the change, the reported pattern of imprinting observed at these genes and the behaviour of individual CpG loci in males and females in the different comparisons represented. Several of the genes exhibit differential methylation at more than one CpG locus. This is illustrated by the complex *GNAS* locus implicated in a wide range of disorders where there are changes at three CpG loci in treated females and at a single different locus in treated males. Figure 3 summarizes the location of the altered CpG loci in relation to the differentially methylated regions (DMRs) associated with the *GNAS* gene (31). It is clear that most of the changes observed in imprinted genes occur postnatally and are not influenced by micronutrient supplementation.

Table 4. Imprinted genes showing significant differential methylation in response to supplementation and postnatal development

Gene	Expression	CpG locus	Comparison	Influenced by	CpFvIpF	CpMvIpM	CtFvItF	CtMvItM	IpFvItF	IpMvItM
ATP10A	M	cg14001035	CtMvItM; CpMvIpM	PD		−0.123		−0.156		
		cg08831522	CpMvIpM	PD		−0.123				
		cg22269180	CtMvItM	PD				−0.169		
ASB4	M in mouse	cg09375488	CpMvIpM	PD		−0.111				
		cg25692621	CpFvIpF; CtFvItF; CtMvItM; CpMvIpM	PD	−0.105	−0.151	−0.145	−0.160		
		cg26847490	CtMvItM	PD				−0.118		
CPA4	M in most tissues	cg19690404	CtFvItF; CtMvItM; CpMvIpM	PD		−0.120	−0.139	−0.210		
		cg01796223	CtMvItM	PD				−0.168		
DLK1	P in brain, M in placenta	cg09971646	CtFvItF; CtMvItM	PD			0.170	0.192		
GNAS	M and P	cg09437522	CtFvItF; IpFvItF	PD, treatment			0.115		0.132	
		cg24346429	CtFvItF	PD			−0.117			
		cg20582984	CtMvItM	PD				−0.181		
		cg01817393	CtFvItF	PD			−0.117			
GRB10	M and P in different tissues	cg03104936	CtFvItF; CtMvItM; CpMvIpM	PD		−0.126	−0.147	−0.172		
		cg09150232	CpMvIpM	PD		−0.109				
HYMAI	P	cg07018708	CpMvIpM	PD		−0.110				
KCNQ1	P	cg17229197	CpFvIpF	PD	−0.116					
		cg16465939	CtMvItM	PD				0.104		
		cg12949760	CpFvIpF; CtFvItF; CtMvItM; CpMvIpM	PD	0.106	0.135	0.214	0.151		
MEST	P	cg02490034	CpFvIpF; CtFvItF; CtMvItM; CpMvIpM	PD	0.154	0.164	0.185	0.207		
MKRN3	P	cg16131766	CtFvItF; CpMvIpM	PD		−0.103	−0.127			
		cg03969797	CpFvIpF; CtMvItM; IpMvItM	PD, treatment	−0.115			−0.195		−0.172
PEG10	P	cg21405195	CtFvItF; CtMvItM	PD			−0.110	−0.131		
SLC22A18	M	cg21019522	CtFvItF	PD			0.128			
		cg19497444	CtMvItM; CpMvIpM	PD		0.104		0.109		
		cg19906550	CpFvIpF; CtFvItF; CtMvItM; CpMvIpM	PD	0.145	0.116	0.171	0.151		
		cg03336167	CtFvItF; CtMvItM; CpMvIpM; IpMvItM	PD, treatment		0.149	0.129	0.284		0.183
SNRPN	P	cg24993443	CtFvItF; CtMvItM	PD			0.118	0.113		
WT1	M and biallelic in placenta	cg13641903	CtFvItF	PD			0.124			
ZIM2	P	cg22354595	CtMvItM	PD				−0.127		

This table shows the individual CpG loci within imprinted genes with significant differential methylation, the comparisons the loci were detected in and the direction and magnitude of the changes. The predominant influences are micronutrient supplementation and postnatal development. Of these, the greater influence is postnatal development, supporting the idea that a small number of methylation changes directly in response to supplementation influence later developmental trajectories. PD, postnatal development.

Functional annotation of genes associated with methylation changes

Genes associated with differentially methylated loci were subjected to functional annotation to assess (i) for enrichment of genes involved in particular biochemical functions and cellular processes, (ii) disease associations and involvement in inherited disease as indicated by OMIM, and (iii) patterns of up-regulated gene expression in different tissues. As some of the gene lists are extensive, the main tool used for such analysis was DAVID (see Materials and Methods) which returns considerable annotation data and individual gene information from a variety of databases together with an evaluation of gene enrichment linked to biochemical functions and cellular processes. The DAVID analysis is presented in Supplementary Material, Dataset 1 and shows the functional annotation available for all genes contained in the DAVID database and the gene enrichment data for those groups of genes showing a significant FDR corrected *P*-value. For the smaller gene lists to emerge from the cord and infant placebo to supplemented cohort comparisons and the male and female placebo postnatal changes (cord to infant), the associated gene annotation and disease associations were curated manually (using data from DAVID, GENE and OMIM databases) and are shown in Supplementary Material, Dataset 2. Given the smaller number of genes involved, it was not possible to obtain significant assessment of gene enrichment. In such cases, consideration of the annotations associated with each gene and disease involvement (included in the tables) provides an indication of the cellular functions and processes that may be impacted by differential methylation.

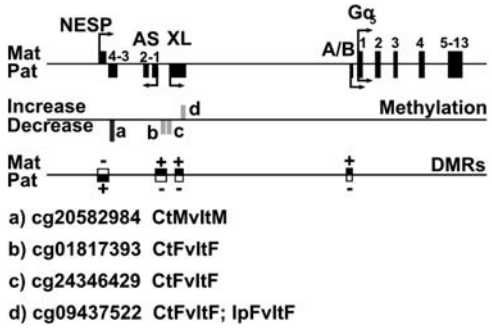

Figure 3. Methylation changes in the complex imprinted locus *GNAS*. Schematic diagram (modified from 30) of the complex imprinted locus *GNAS* showing the location of the differentially methylated loci in relation to the position and pattern of imprinting of the DMRs of the *GNAS* gene. The diagram details the CpG loci (a–d) undergoing change, the comparisons in which these changes were detected and the direction of the change of each locus in relation to sex. We cannot be certain that the CpG loci fall within the orbit of the closest DMR. However, assuming that they are influenced by the prevalent pattern of methylation for each parent, we would predict the following predominant effects for each CpG. (a) Demethylation of the paternal allele. (b) and (c) Demethylation of the maternal allele. (d) Increased methylation of the paternal allele.

From the various DAVID analyses, it is clear that there is enrichment in various lists for genes that encode signalling molecules, glycoprotein receptors and various types of membrane-linked proteins. The male supplementation-specific postnatal changes additionally show enrichment for genes involved in ion homeostasis. The DAVID analysis of core loci has shown enrichment for similar genes encoding signalling molecules, glycoproteins and membrane bound proteins but also features genes involved in immune response and defence against bacterial infection.

Table 5 represents a broad classification of the genes associated with loci undergoing methylation change included in Supplementary Material, Table S2 for the CpMvCtM, CpFvCtF, IpMvItM, IpFvItF, CpMvIpM non-core male-specific postnatal placebo changes, CpFvIpF non-core female-specific postnatal placebo changes and male/female shared postnatal placebo changes. Brief annotations are summarized in Supplementary Material, Table S2 and more detailed information and supporting references can be found in the Gene and OMIM databases. What is evident in these comparisons (particularly the postnatal changes observed in the male and female placebo cohorts) is the predominance of genes with an immune function and defence against infection.

Most promoter-associated CpG loci undergoing differential methylation are not found in CpG Islands

Most (72%) of the CpG loci on the Human Methylation 27 BeadChip are located in CpG islands. We have examined the distribution of differentially methylated CpG loci identified in the various comparisons reported in this study. We find that only 23 to 45% of the loci are located in CpG islands. Figure 4A summarizes the distribution data for various comparisons. This is significantly different to the distribution expected if these loci were randomly selected (Wilcoxon signed rank test $w = 36$, $n = 8$ and $P = 0.0078$), indicating that there is preferential methylation change at promoter CpG loci outside of CpG islands. Furthermore, as shown in Figure 4B, most of the genes associated with differentially methylated CpG loci lack CpG islands in contrast to the genes on the chip as a whole. Figure 4C presents a plot of the distribution of the differentially methylated CpG loci outside CpG islands with respect to the location of the transcription start site (TSS) of associated genes. Almost all CpG loci are within 1500 bp of the TSS, and the majority (63%) are within 500 bp of the TSS.

DISCUSSION

The findings from DNA methylation genome-wide analysis indicate that periconceptional maternal micronutrient supplementation leads to widespread changes to the epigenome primarily affecting autosomal loci, some of which may only be detected postnatally, with clear differences in patterns between the sexes. There are methylation changes associated with the placebo state that occur postnatally and these too show male- and female-specific patterns (see Materials and Methods for more detailed discussion of caveats and technical limitations).

Different genes respond directly to supplementation in males and females

This is the first study to interrogate individual CpG loci on a genome-wide basis at different time points following micronutrient supplementation provided over a defined window in early human development. Both in cord and infant blood placebo to supplementation comparisons, the male and female methylation changes involve almost non-overlapping sets of loci. Table 5 and Supplementary Material, Dataset 2 summarize the functions and disease associations of genes linked to differentially methylated loci determined in a number of the key comparisons. In Table 5, where clear gene functions and disease associations (indicated by mutations) have been defined, these have been classified under a series of broad categories. Genes involved in cancer and development have, in particular, been defined by the effects of mutation on phenotype. A small number of genes show differential methylation in more than one comparison. In most cases, this involves different CpG loci in relation to sex, supplementation or age, indicating very specific modification of sites within promoters.

For male newborns, there is a diverse spectrum of 14 genes that undergo methylation changes associated with micronutrient supplementation. Of these, seven (*MEOX1*, *SPAG4L*, *SPATA22*, *NRN1L*, *C1orf54*, *PIP*, *PTPN20B*) are also found in the infant supplemented group, with two genes being notable in relation to neuritogenesis (*NRN1L*) in the developing nervous system and skeletal development (*MEOX1*) where there could be an impact on growth. In contrast, the 21-gene methylation changes in female newborns highlight two clear groupings. First, those involved in immune (*SIGLEC 5, CD4, KLRC2*) and non-immune (*BPIL1, CHIT1, DEFB123*) defence against infection. Secondly, those involved in cardiovascular function and disease affecting cardiac morphogenesis

Table 5. Classification of gene associated with differentially methylated loci based on functional data and disease associations

Functional classification		CpMvCtM	CpFvCtF	IpMvItM	IpFvItF	CpMvIpM	CpFvIpF	Placebo shared
Haematological, immunological, inflammatory, non-immune defence		SIGLEC1, SP1R3	SIGLEC5, CD4, CHIT1, KLRC2, DEFB123, BPIL1	CD80, LRRFIP1, VAV1, SIGLEC5, SAA1, NLRP2, LPCAT2, CCR9, NLRP5	UBASH3A, REG3G, XCR1, TXK, SGLEC11, KIR3DX1, FGFBP2, BPIL1, HLA-DPA1, PF4, ICOS, CHIT1, IGK, KLRC2, GIMAP5, CREB3L3, DEFB123, IFNG	NLRP4, NCR2, NLRC5, PDCD1, STAP1, FDCSP, CCR6, CD209, CD4, CD86, CMKLR1, CREB3L3, CTSW, CYP4F3, DARC, FCRL2, FPR1, GBP7, GP9, HLA-DMA, IGJ, KLRC2, LAIR1, MAPK8, MAPK9, NLRP4, NCR2, NLRC5, OKL38, PDCD1, TFP1, TM7SF4, IGL@, IKBKAP	APOBEC4, BTNL2, CISH, FCER2, HLA-DQB2, IFI44, IL27RA, LILRB2, MMD, NLRP2, SLC11A1, SIGLEC1, TNFRSF13B, TRIM22, XCR1, ZBP1, TRG, IGK	LAT2, HTN3, GNLY
	Mutation			GP1BB	PRF1, HBD, GP5	MPO, OAS1, PRF1, CIITA, CYBB, EMP1, HPR, IL2RA, IL31RA, LYST, MASP1, RAG1, SIAE	CFI	CD19
Cancer and development		MEOX1, PIP	MGMT	SLC22A18, MEOX1, HOX5A, BRD7, PPP2R1B, RA12, TM4SF5, DEC1, ERV3, WNT3A	PTPRG, PUM2, S100A9,HOXB6, CHRNA3, MGMT	AMELX, GRB10, INHBC, MAPK8, MAPK9, MEOX1, MGMT, NOX3, OGFR, SHRM, SMARCA1, TMPRSS3, TNFSRF10A	ALX3, BCAS1, CASP8, CDC2L2, CHRNA5, DENND2D, MPZL2, FAP, HOXB6, HRASLS, LIMD1, PTPRG, RASSF6, S100A2, SMPD3, AURKA, TP53INP1	TM4SF1, RPA1, MGMT
	Mutation			SALL4, BMPR1A, NPR2, COL11A1, TMPRSS3, TNN12, NDUFS1	GNAS, CDC2L	ABCC12, APC, ARSE, ABS11, CTSK, DYM, MASP1, MKS1, MMP13, MYH3, NPR2, OFD1, PAX8, RB1, RNASEL, RUNX2, RUNX3, SALL4, SHOC2, WDR65, IGF1	BMP4, BCL2, GDF3, OFD1	
Cardiovascular			XIRP1, LPA	XRIP1, BIRC8, FKBP1B, GNA11, ELIMIN1	XRIP1, GNA11, SPP2	ACYP2, BIRC8, S1PR3, PDCD1, GLRX3, VCAM1		TM4SF1
	Mutation		FLVCR2, ELN	FLVCR2, TMEM43, KCNJ5	SLC12A3	ABCC9, KCNE1L, TCAP, TNNI3	KCNQ1	
Neurological		NRN1L		GRM7, PODXL2, NTSR1, ARNT2, MRGPRX3	APBA2, CAST, OR5P2, ATP13A4, NAV1, PION	AVPR1A, CCDC22, GRM5, MRGPRX3, NSF	OXT	IGDCC3
	Mutation			ATXN1, ATP1A2, PNPLA6	SH3TC2, MPZ	ASPA, IKBKAP, SH3TC2, TPH1		ATP1A2, SERPINI1
Diabetes and obesity				SLC6A14, MKRN3, AHSG, PROX1, PRG		BRS3, GCG, HYMAI, REG1A, REG1B, MKRN3	SCG5, MKRN3	
	Mutation			SPINK1	ADPN	GLIS3, HK2, IL2RA, SPINK1, SIAE, OAS1	AVP	

Continued

Table 5. *Continued*

Functional classification	CpMvCtM	CpFvCtF	IpMvItM	IpFvItF	CpMvIpM	CpFvIpF	Placebo shared
Metabolic and lipid metabolism		CYP2E1	HMGCS2		ADH1A, ADH1C, PSTK, NRK1, CPT1B, DEGS1, DGKA, FABP1, FABP5, ACF2, FXYD2, GLRX2, HCAR2, GSTA5, GSTP1, HAO1, LIPF, LYST, NLN, PPT2, RSC1A1, SCO1, SLC25A26, SLC25A5, SLC35D2, SLC6A14, SLCO1B1	ABCC2, PNPLA3, AVPR2, BHMT, MAN2C1, THRB	
	Mutation		ALAD, GNMT, HMBS, OTC	MCCC1, MOGAT2, NUDT14	ABCB11, AKR1D1, BCMO1, SLC25A20, SLC26A3, SLC4A1, XDH, STAR	AVP, MCCC1, OXCT2	KYNU
Myogenic				SMPX, MYOG	JSRP1, MBNL3, MYL9, PDE1A	MYL4	
	Mutation			TNNT1	MYH3, SMPX	ACTN3	
Ocular			NXNL1		MIP, RRH		
	Mutation	NYX	CRYAA	TACSTD2	RDH12, DCN, EGFL11		

The genes from each pairwise comparison have been grouped into functional categories from manual curation and their DAVID and OMIM analyses (see Supplementary Material, Datasets 1 and 2). They are split into those with and without known disease-causing mutations. The largest groups of genes are seen in the CpMvIpM comparison, and in the categories of haematological, immunological, inflammatory, non-immune defence and metabolic and lipid metabolism. The small numbers of genes in the placebo shared category highlights the different developmental trajectories of the sexes.

(*XIRP1*), brain vascularization (*FLVCR2*), vascular connective tissue (*ELN*) and promotion of thrombogenesis at atherosclerotic lesions (*LPA*). It is interesting to note that of the eight genes also found in infants (*HSPC176, XIRP1, BPIL1, MGMT, CHIT1, C14orf152, KLRC2, DEFB123*), four (*BPIL1, CHIT1, DEFB123, KLRC2*) are associated with fighting infection and one (*XIRP1*) with the development of the heart. It is striking that 50% of the supplementation-related changes seen in cord blood are also seen in infants, indicating early long-term reprogramming of epigenetic status associated with nutritional deficiency or supplementation. This is also demonstrated by the larger number of loci showing differential methylation in the analogous infant comparisons where very few genes (five genes: *XIRP1, MEOX1, GNA11, C1orf54, KRTAP21-1*) are found in common between males and females. In both male and female infants, the dominant groupings of genes are those associated with (i) immune response and defence against infection (more in females), (ii) cancer and development (more in males) and (iii) neurological function. In this latter group, it is of note that in females five of eight genes are associated with neurodegenerative disorders (*PION, CAST, APBA2* associated with Alzheimer disease and *MPZ* and *SH3TC2* with Charco-Marie Tooth disease) and in males two of eight (*ATXN1* associated with degeneration of the cerebellum, brain stem and spinal cord and *NTSR1* with dopamine-associated behavioural neurodegenerative and neuropsychiatric disorders). Interestingly, in infant males (in contrast to the newborn males), there is now a significant grouping of eight genes linked to cardiovascular function. One further grouping evident in males but less prominent in females is that of genes associated with obesity (*SLC6A14, MKRN3*—a paternally expressed gene mapping to the Prader–Willi syndrome (PWS) region on chromosome 15*, PROX1*) and diabetes (*SPINK1, AHSG, PGR*). Defects in the *SPINK1* gene are associated with hereditary pancreatitis and tropical calcific pancreatitis (TCP) commonly found in tropical countries. Fibrocalculous pancreatic diabetes is a form of diabetes secondary to TCP and is characterized by early onset diabetes mellitus and high incidence of pancreatic cancer.

A number of studies have examined epigenetic changes at individual loci affecting promoter and enhancer regions in response to nutrition at critical periods of gestation. Sandovici *et al*. (32) have shown that the enhancer associated with the *Hnf4a* promoter in pancreatic islets undergoes both DNA methylation and histone modification in response to nutritional manipulation in early rat development. Lillycrop *et al*. (13) have demonstrated methylation changes in the *PPARα* gene in the offspring of pregnant rats fed on a protein restricted diet, and Godfrey *et al*. (33) have produced data suggesting that methylation of the retinoic acid receptor alpha is linked to adiposity and maternal diet. Other studies have examined the impact of season of conception in the Gambia (and hence indirectly potential micronutrient deficiency) (34) and famine (22,35) during gestation and shown an association with alterations in methylation at multiple loci. In the case of the Dutch hunger famine (35), sex differences in differential methylation were also evident. The Waterland *et al*. (34) study is most comparable with the present investigation where a genome-wide selection for meta-stable epi-alleles (MEs—loci where an epigenetic state is established during

development that has phenotypic effects in later life) was used to produce a battery of probes to assess methylation states in individuals in relation to season of conception. Their findings show an increase in methylation at a number of loci that is associated with conception during the nutritionally poor rainy season. Very few of these loci are associated with genes (four genes; *PAX8, LTBP4, SEPT6* and *FREQ*) that are concordant with changes observed in the current study and we have observed both increases and decreases in methylation.

More methylation changes are seen between cord and infant states

Also striking is the large number of loci that show temporal changes in methylation when cord and infant states are compared, some of which will be related to postnatal development rather than the effects of sex or micronutrient supplementation. Candidates for a role in postnatal development have been labelled as 'Core Loci' in Figure 1, loci shared by males and females independent of supplementation. The non-core Loci can be divided within each sex into supplementation-specific, placebo-specific and shared loci, the latter being loci that change in methylation postnatally independently of supplementation (Fig. 2). Figure 2 also summarizes the cross gender comparison of these three categories and indicates that the majority of the supplementation and placebo-specific loci remain sex specific and there is no sharing between males and females of the placebo to supplement shared category. Thus, it is possible to identify four groups of methylation change postnatally; male and female placebo-specific changes, male and female supplementation-specific changes and male or female changes independent of supplementation that are potentially specific to male or female postnatal development.

This fractionation of the postnatal changes into distinct sets of loci suggests that males and females follow different developmental trajectories (in accordance with expected physiological differences) both in the presence of supplementation or when likely subject to micronutrient deficiency. The mechanisms of this differential response remain unknown but presumably involve some means of sensing the nutritional environmental cues. It is interesting to note that DAVID analysis (see Supplementary Material, Dataset 1) has shown significant enrichment in most lists for loci associated with genes that encode signalling molecules, glycoprotein receptors and various types of membrane-linked proteins that may be involved in sensing and transmitting environmental information. The enrichment in the 'Core Loci' group of genes involved in immune response and defence against infection is what may be predicted for loci important in early postnatal development and survival shared between males and females in placebo and supplemented groups and suggests that a programme of epigenetic modulation is an integral part of the developing immune response.

The male and female placebo-specific loci changing postnatally are intriguing (see Table 5 and Supplementary Material, Table S2). Clearly, supplementation alters the postnatal methylation patterns and therefore the patterns observed in the male and female placebo cohorts could be considered loci that characterize the pathological state arising from likely micronutrient deficiency. More loci undergo change in males than females but in both, the predominant groupings are again genes associated with haematopoietic and immune functions and cancer and development. Several genes are associated with susceptibility to bacterial and viral infection in both males (*CD209, FPR1, GBP7, OAS1, CYBB, IL2RA*—where mutations lead to an immunodeficiency with decreased numbers of peripheral T cells) and females (*APOBEC4, CISH, FCER2, SLC11A1, TNFRSF13B, TRIM22, ZBP1*), in some cases (*SLC11A1*—tuberculosis, leprosy and Buruli ulcer) particularly in tropical regions. Susceptibility to infection with *Plasmodium vivax* is influenced by the *DARC* gene (males) and to *Plasmodium falciparum* by the *HPR* (males) and *CISH* (females) genes. There are also a number of genes associated with inflammatory and autoimmune disorders (males—*SIAE, PDCD1*; females—*CFI, LILRB2, ZBP1, BTNL2, SLC11A1*). Finally, variants in three genes whose methylation is altered in males are associated with susceptibility to type I diabetes (*OAS1, IL2RA*—where reduced immune response may lead to predisposition ; *SIAE*—where mutations cause a spectrum of autoimmune disorders including type 1 diabetes).

Differences between males and females in the placebo state postnatally

Three other groupings are worthy of comment where there is marked difference between males and females.

(i) Genes involving metabolic functions and lipid metabolism are prominent in males where they have roles in fatty acid uptake, transport and metabolism (*FABP1, FABP5, CPT1B, ACSF2, HCAR2, HAO1, LIPF, SLC25A20*), hepatic function and detoxification (*ABCB11, AKR1D1, GSTA5, GSTP1, SCO1*), cholesterol metabolism (*STAR*), vitamin A synthesis (*BCMO1*) important for vision, embryonic development and cellular differentiation and glucose absorption in the intestine (*RSC1A1*).

(ii) A group of genes implicated in glucose metabolism, obesity and diabetes (*BRS3*, *GCG*, *HYMAI*, *REG1A*, *REG1B*, *GLIS3*, *HK2*, *IL2RA*, *MKRN3*—found in the imprinted PWS region on chromosome 15, is paternally expressed and believed to contribute to the PWS phenotype, *SPINK1*). *SPINK1, IL2RA, OAS1* and *SIAE* have already been discussed above; however, in relation to diabetes, these findings reveal further relevant genes involved in pancreatic islet cell differentiation (*GLIS3*, where mutations lead to neonatal diabetes and congenital hypothyroidism), islet cell regeneration and diabetogenesis (*REG1A, REG1B*), transient neonatal diabetes mellitus and intrauterine growth retardation, dehydration and failure to thrive due to a lack of normal insulin secretion (*HYMAI* also paternally expressed) and glucose metabolism (*GCG*—glucagon regulating glyconeogenesis and gluconeogenesis; *HK2*—mitochondrial skeletal muscle form where variants have been associate with insulin resistance in peripheral tissues and noninsulin-dependent diabetes mellitus).

(iii) Genes involved in cardiovascular function; either direct cardiac involvement leading to various cardiomyopathies (*ABCC9, KCNE1L*, *KCNQ1* in females—all potassium ion channels associated with atrial fibrillation and ventricular arrhythmias; *TCAP*—a cardiac sarcomere protein involved in

limb-girdle muscular dystrophy and sudden death; *TNNI3*—cardiac troponin involved in hypertrophic cardiomyopathy; *GLRX3*—pressure-induced overload cardiac hypertrophy; *PDCD1*—mouse model of cardiomyopathy) or relating to the vasculature and angiogenesis and endothelial cell function (*ACYP2*—genome-wide association with coronary artery disease; *BIRC8*—endothelial cell apoptosis; *S1PR3*—regulation of angiogenesis and endothelial cell function; *VCAM1*—endothelial cell adhesion during vascularization).

The comparison of male- and female-specific placebo postnatal changes highlights a number of points. First, it emphasizes that different developmental trajectories are followed by males and females under conditions of suboptimal maternal micronutrient status, such as prevail in the Gambia between the rainy and dry seasons. Secondly, the greater number of loci undergoing differential methylation in males under such conditions indicates that they are impacted more than females in early postnatal development. Thirdly, it is clear that changes in genes functioning in the immune system and defence against infection are a major category affected by micronutrient deficiency. This resonates with a recent study that has described differential methylation of immune-related genes in early postnatal, infant and early year development, suggesting a role for epigenetic modification in the development of the immune system (36). Our findings are also concordant with previous studies in rural Gambia correlating the effect of season of birth on immune function and mortality predominantly arising from infectious disease (29,37). These studies have shown that there is increased mortality from infectious disease and compromised immune function in young adults born in the nutritionally poor rainy season. This suggests that early postnatal life influences are critical to the developing immune system. Our data take this a step further, suggesting that the periconceptional period may also be important. Recent studies have shown that there are differential methylation patterns associated with different haematopoietic cell types (38–40). Cellular composition of infant blood may be influenced by different pathogen and allergen challenges, or by differential responses to those challenges. As this is a randomized trial and these children are raised in the same rural Gambian environment, we think it highly unlikely that there are significant and consistent differences in pathogen and allergen exposure between the placebo and supplementation cohorts. It is, however, possible that the infant cohorts show different cellular responses to their shared pathogen challenges dependent upon their periconceptional nutritional status. For this reason, we draw no inference as yet regarding the precise biological mechanism underpinning the observed difference in methylation between the infant placebo and supplemented cohorts.

Imprinted loci

A number of changes are found in loci associated with imprinted genes. These genes are involved in intrauterine growth (*GNAS, HYMAI, GRB10, PEG10*—in mice; *MEST*—in mice), abnormal behaviours *(MKRN3, SNRPN*—Angelman and PWS; *MEST*—maternal behaviour in mice), urogenital tract development (*WT1*) and various tumours (*WT1, SLC22A18*), obesity (*MKRN3, SNRPN*—potentially through modified eating behaviour; *DKL1*—adipocyte differentiation) and cardiac defects (*KCNQ1*). Most of the methylation changes occur postnatally in all loci and in three cases there is an effect of micronutrient supplementation (*GNAS, MKRN3* and *SLC22A18*). It is clear that males are impacted to a greater extent than females by alterations to the methylation status of imprinted genes (Table 4). It is also evident that different CpG loci associated with the same gene respond to different variables and in a sex-specific pattern. This may reflect subtle modulation of promoter activity. An example is the *GNAS* gene showing differential methylation patterns between supplemented males and females where there are maternally, paternally and biallelically expressed transcripts. This is a locus associated with many physiological processes where complex and different patterns of imprinting in males and females together with different mutations lead to a diverse spectrum of genetic disorders affecting several tissues especially skeletal development and endocrine function (pseudopseudohypoparathyroidism—short stature, obesity, round facies, subcutaneous ossifications, brachydactyly and other skeletal anomalies; pseudohypoparathyroidism type 1a and 1c—as above but with resistance to parathyroid hormone; pseudohypoparathyroidism type 1b—as above but isolated renal resistance to parathyroid hormone; McCune-Albright syndrome—polyostotic fibrous dysplasia of bone, severe developmental skeletal defects and defective endocrine function; progressive osseus heteroplasia—dermal ossification beginning in infancy, followed by increasing and extensive bone formation in deep muscle and fascia, and some pituitary tumours). Of particular note is that the deletion of the paternal locus is associated with severe pre- and postnatal growth retardation (41).

Distribution of CpG loci

From Figure 4, it can be seen that a minority of differentially methylated CpG loci are within CpG islands (23–45%) in contrast to the chip content as a whole (72%). The majority of the genes discussed above have changes (63 versus 27) in such lone CpG loci. This suggests that lone promoter CpG loci are preferentially responding to environmental factors and possibly indicate the location of important functional elements within the affected promoters. It should be borne in mind that we cannot predict the effect on gene expression of an increase or decrease in methylation at any given CpG locus. This will require *in vivo* correlation on the effects of methylation or demethylation on mRNA expression levels (ideally in the target tissue) or modelling of such changes in animal models where analogous methylation changes in homologous loci can be examined directly.

Conclusion

These data indicate that (i) micronutrient supplementation in early gestation is associated with methylation changes (generally demethylation) detectable in cord blood and early infancy DNA and hence may be important in programming gene activity in later life, (ii) there is little overlap between males and females in the loci undergoing methylation changes associated with supplementation, and (iii) postnatal changes involve

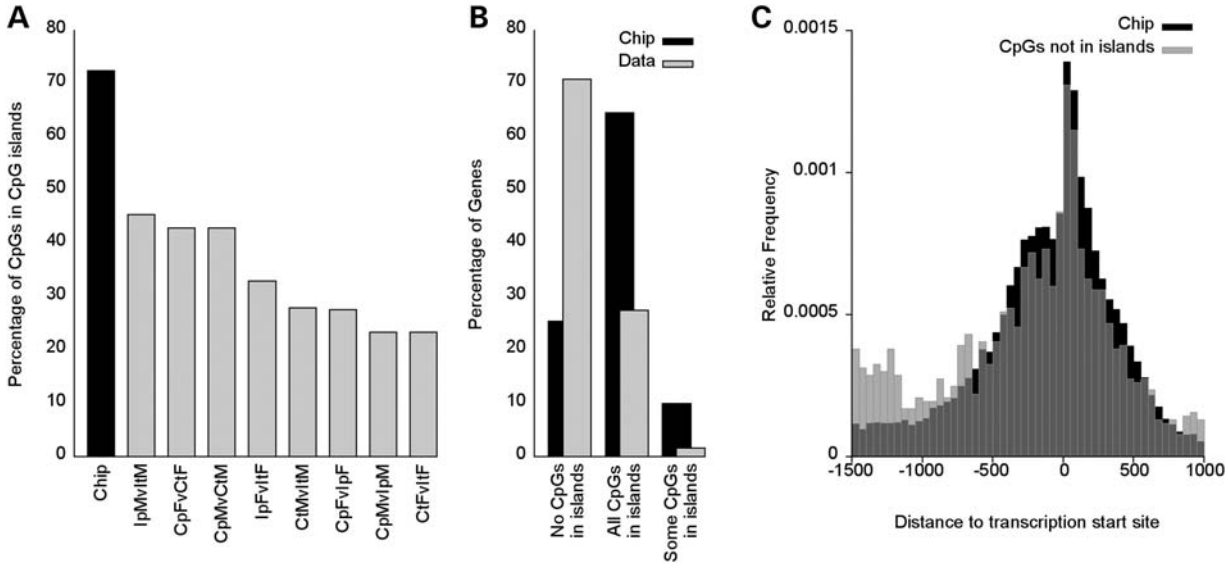

Figure 4. Proportion of CpG loci undergoing differential methylation within CpG islands. (**A**) 70% of CpG loci on the human methylation 27 BeadChip are located within CpG islands ('Chip'). The figure summarizes the distribution of CpG loci highlighted by our pairwise comparisons. In each case, a lower percentage than expected is found within CpG islands ($P = 0.0078$). (**B**) Most of the genes on the chip have all associated CpGs within CpG islands. The opposite is true of our data: most genes in our data do not have CpGs within CpG islands. (**C**) Most of the CpG loci on the chip (black) and in our data (light grey) are close to a gene TSS, with a slight upstream bias. Overlap between the two is shown in dark grey.

many more genes that highlight differences between the supplemented and placebo cohorts that potentially have a bearing on respective disease outcomes. Most of the observed methylation changes are relatively modest, suggesting that influences on phenotype probably represent the combined effect from many genes. The mechanisms by which these epigenetic modifications to the genome are achieved and the combined impact of small changes in many genes and pathways on phenotype remain to be elucidated.

MATERIALS AND METHODS

Trial, population and sample collection

Consenting non-pregnant women aged between 17 and 45 years from Kiang West, The Gambia were recruited to take part in a placebo-controlled randomized trial for Periconceptional Multiple Micronutrient Supplementation (ISRCTN 13687662). Maternal age, height, weight and upper arm circumference values in both groups were assessed with the Student *t*-test (two-tailed, unequal variance) to confirm that women were randomly assigned. Non-pregnant women took either UNIMMAP tablets [containing a balanced combination of 14 vitamins and minerals, specially formulated for use in pregnancy—Vitamin A (800 RE), Vitamin D (200 IU), Vitamin E (10 mg), Vitamin C (70 mg), Vitamin B1 (1.4 mg), Vitamin B2 (1.4 mg), Niacin (18 mg), Vitamin B6 (1.9 mg), Vitamin B12 (2.6 mg), Folic acid (400 μg), Iron (30 mg), Zinc (15 mg), Copper (2 mg), Selenium (65 g) and Iodine (150 μg)] or placebo, and once pregnancy was confirmed (averaging 9.5 week gestation), all women were switched to receive FeFol tablets (60 mg of elemental iron and 250 μg of folate). Women were excluded from this analysis if they were enrolled for <14 days before last menstrual period, or if they apparently consumed <70% of allocated tablets. The trial was approved by the MRC/Gambian Government Ethics Committee, and the current study was also approved by the MRC Laboratories Fajara Scientific Coordinating Committee (L2007.77, 21 September 2007) and the Gambian Government Ethics Committee (25 October 2007). Birth weight, length and head circumference of newborns and placental weight were measured. DNA was extracted from umbilical cord blood collected at birth from 23 offspring of treated mothers (10 females and 13 males) and 36 offspring of placebo mothers (18 females and 18 males). DNA was also extracted from circulating blood of some of the same newborn individuals at an infant age of ~9 months from 10 offspring of treated mothers (6 females and 4 males) and 15 offspring of placebo mothers (8 females and 7 males). It should be noted that the cellular composition of these blood samples will be different and differences in methylation that are reported must take this into consideration.

Microarray methylation analysis

The HumanMethylation27 BeadChip (Illumina, San Diego, CA, USA) containing 27 578 individual CpGs located in promoters was used to measure DNA methylation associated with some 14000 different genes. One microgram of DNA was treated with sodium bisulphite using the EZ DNA methylation kit (Zymo Research) as directed by the manufacturer's protocol. Whole genome amplification, labelling, hybridization and scanning (on the BeadArray scanner) were performed following the protocols given in the accompanying Illumina manual.

Bisulphite MassArray validation

For a selection of loci, microarray data were validated using the MassArray (Sequenom, San Diego, CA, USA) system in triplicate assays (three independent bisulphite conversions). Target regions containing the CpG of interest were amplified by PCR. Primers (listed in Supplementary material, Table S2) were designed using EpiDesigner (Sequenom), screened to avoid known single nucleotide polymorphisms (SNPs) and synthesized (Sigma) with tags added. The PCR tags are required for the MassCLEAVE chemistry (Sequenom) which was performed exactly according to the manufacturer's instructions. Samples were then spotted onto SpectroCHIPs (Sequenom) for analysis by mass spectrometry (MassARRAY Compact Analyzer, Bruker Sequenom). Complete conversion was verified using the R-package, MassArray (42) and data were output using the EpiTYPER software.

Sample quality control, validation of methylation changes and data analysis

Validation of the trial cohort samples was performed by multidimensional scaling [principal component analysis (PCA)] and unsupervised hierarchical clustering analysis on loci that returned a ratio from the methylation data above 0.1 for the standard deviation (SD)/mean (MEAN). This was done as a quality control measure to identify any rogue samples before detailed data analysis. These analyses are shown in Figure 5A–C. Figure 5A shows the separation of all samples (based on the data from 20446 CpG loci where the ratio SD/MEAN was >0.1) on the basis of two principal components; the first component separating males and females and the second broadly distinguishing cord and infant samples. There is one female infant sample that segregates with male samples and probably represents a mislabelling and three male cord samples that partially segregate with female samples and are likely to represent significant contamination of cord with maternal blood. Figure 5B shows the multidimensional scaling analysis of all female samples (based on 27542 CpG loci without filtering for loci where SD/MEAN was >0.1 in order to maximize discrimination between female samples that may have been mislabelled or contaminated with maternal blood) that effectively separates female cord and infant samples and has identified a further anomalous sample (571) identified as an infant but is clearly segregating with the cords. Figure 5C shows the result of unsupervised hierarchical clustering and has identified the same outliers (except 571). The boxed samples in each of these figures have been excluded leaving 56 newborns (21 in the micronutrient cohort: 11 males and 10 females; and 35 in placebo cohort: 17 males and 18 females) and 23 infants (9 in the micronutrient cohort: 4 males and 5 females; and 14 in the placebo cohort: 7 males and 7 females) for further data analysis.

Independent validation of the methylation changes detected by the Illumina platform was performed by comparing methylation levels at defined CpG loci observed on the arrays with those detected by Sequenom MassArray analysis (38) for 34 loci. The data were used to create the scatter plot shown in Figure 6. The R^2 value of 0.7174 ($P < 0.0001$) shows good correlation between the two methods of measuring methylation at a given CpG locus, indicating that the global data from the array platform renders a reliable assessment of methylation changes across the genome.

Microarray data from all samples that survived quality control assessment were then analysed using the Methylation Module in the Illumina Genome Studio software and the methylation ratios [beta value $=(C/(C+T+100)$ where C is the fluorescence associated with methylated sites and T the unmethylated sites] were calculated. This analysis uses the Illumina custom model that models variance as a polynomial function derived from experimental data as detailed in the Illumina manual (GenomeStudio Methylation Module v1.8 User Guide, page 51). A FDR correction to the data was applied by the software. Stringent thresholds were set to select significant changes in methylation represented by the difference in the beta values (delta beta) between the groups being analysed. Only those loci that showed a greater than $\pm 10\%$ delta beta at a P-value $= 0.001$ (equivalent to a DiffScore of >30 or <-30 on the Illumina significance threshold scale) were selected. Pairwise comparisons were used to detect major effects on methylation associated with a single variable; micronutrient supplementation, sex or postnatal development. This has its limitations as subtle interactions between such variables will not be detected. A number of further caveats should also be borne in mind. First, methylation changes have been studied in genomic DNA extracted from white blood cells in cord and infant blood and may not reflect all changes occurring in other tissues in addition to those in haematopoietic cells. Secondly, postnatal changes may arise due to differences in cellular composition of cord and infant blood. Thirdly, variation in methylation between individuals may be provoked by different pathogen challenges or genetic differences between individuals. Fourthly, the confounding effects of maternal contamination of cord blood. The placebo cohorts permit separation of methylation alterations associated with supplementation from those caused by other changes, such as cellular composition. PCA and hierarchical clustering have permitted removal of outliers that may have arisen through mislabelling or maternal contamination of cord blood. Fifthly, 907 of the loci on the HumanMethylation27 BeadChip overlap with SNPs (43). Such polymorphisms can confound the assessment of methylation change if the distribution of alleles is skewed between the placebo and micronutrient supplemented cohorts. As this study is based on homogeneous rural Gambian populations, asymmetric distribution of alleles will be insignificant and the differences between groups represent true methylation changes. Nonetheless, differentially methylated loci overlapping with SNPs are given in Supplementary Material, Table S2.

Annotation of gene lists

The lists of genes associated with differentially methylated loci for both male and female data sets were uploaded into DAVID (**D**atabase for **A**nnotation, **V**isualization and **I**ntegrated **D**iscovery at the NIH National Institute of Allergy and Infectious Disease—http://david.abcc.ncifcrf.gov/) to obtain functional annotation and to assess whether there was enrichment for particular categories of gene and associated

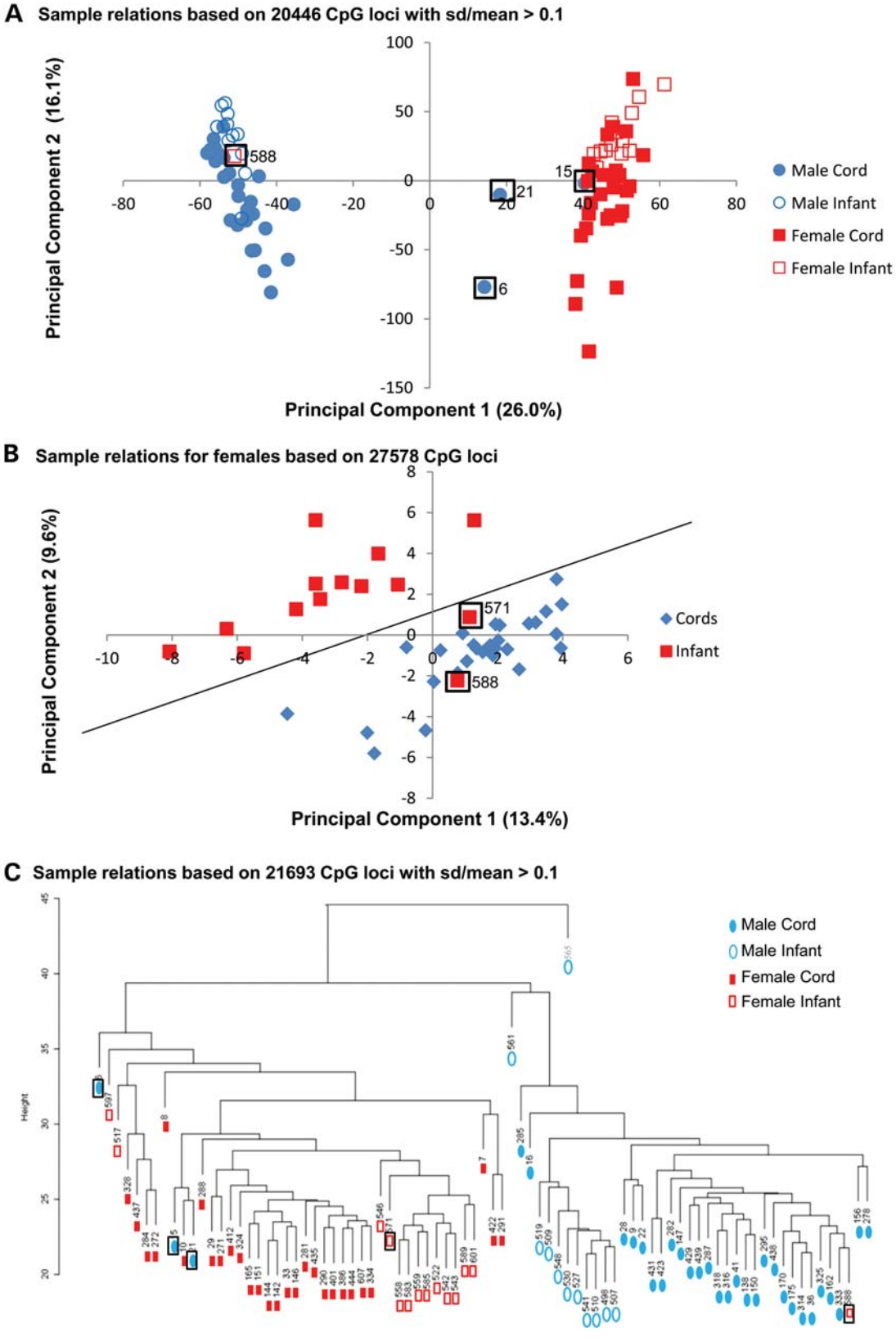

Figure 5. Multidimensional scaling (PCA) and hierarchical clustering analysis of sample set using methylation data. (**A**) Gender assignment of samples was confirmed by using loci where the ratio SD/MEAN was >0.1 to segregate samples by multidimensional scaling. Principal component 1 has separated the samples on the basis of sex and principal component 2 has permitted a broad distinction of cord and infant samples. Sample 588 is likely to represent a mislabelled male sample and samples 6, 15 and 21 male cord samples with significant maternal blood contamination. (**B**) Analysis of female samples using multidimensional scaling analysis. Separate analysis of female samples was used to ensure improved discrimination between samples to detect maternal possible maternal blood contamination or mislabelling of cord and infant cases. In this case, methylation data from all CpG loci was used in the analysis. The diagonal line marks the separation of cord and infant samples. Two infant samples clearly segregate with the cord samples; 588 (already identified as rogue from A) and 571. (**C**) Hierarchical clustering of all samples based on CpG loci SD/MEAN ratio >0.1.

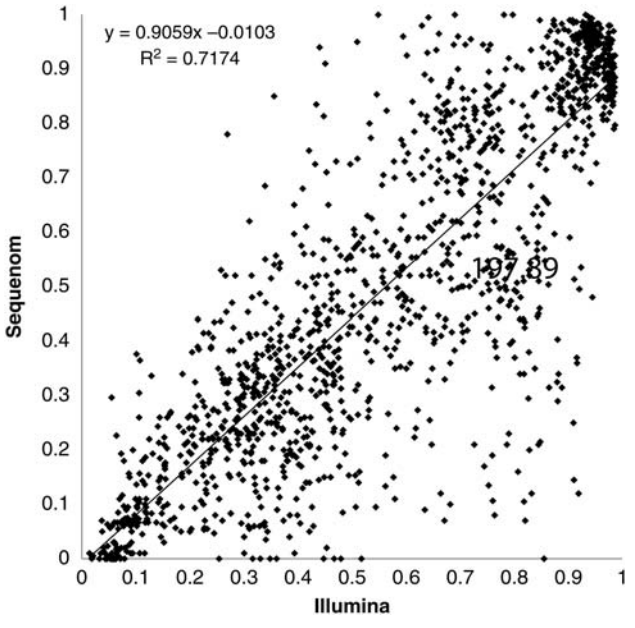

Figure 6. Scatter plot of methylation determined by illumina array and sequenom MassArray analysis. In order to validate the methylation data determined by the analysis on the Illumina HumanMethylation 27 Beadchip, 34 defined CpG loci were also assayed by Sequenom MassArray analysis on 78 DNA samples. The graph represents a scatter plot to measure the correlation between the two methods. The R^2 value of 0.7174 provides confidence in the genome-wide data extracted from the Illumina HumanMethylation 27 Beadchip.

functions. DAVID returns various classes of data indicating clusters of genes that may have related biological functions and an enrichment score for such clusters with *P*-values (corrected for multiple testing) for groups of genes in enrichment clusters. Where the *P*-values have been significant, the detailed functional clustering data of these genes are included in Supplementary Material, Dataset 1. Functional annotation files have also been included in Supplementary Material, Dataset 1. These files contain expression data, disease associations, disease classifications, genes included in OMIM (**O**nline **M**endelian **I**nheritance **I**n **M**an) associated with inherited disorders, Entrez gene entries, GeneRIF data, KEGG and various other pathway analyses for particular genes. Where no information on genes was available from DAVID, manual curation was performed using resources at NCBI that included the GENE and OMIM databases. The manually supplemented curation for various gene lists is assembled in Supplementary Material, Dataset 2.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

ACKNOWLEDGEMENTS

We would like to acknowledge Dr Emily Clemente, Christopher Reitter and Kerry Harvey of Cambridge Genomic Services (Department of Pathology) for running the methylation arrays. We thank the mothers and babies who participated in the trial.

Conflict of Interest statement. None declared.

FUNDING

This work was supported by grant number BB/F001355/1 from the BBSRC.

REFERENCES

1. Gluckman, P.D. and Hanson, M.A. (2004) Living with the past: evolution, development, and patterns of disease. *Science*, **305**, 1733–1736.
2. Bateson, P. (2001) Fetal experience and good adult design. *Int. J. Epidemiol.*, **30**, 928–934.
3. Gluckman, P.D., Hanson, M.A., Spencer, H.G. and Bateson, P. (2005) Environmental influences during development and their later consequences for health and disease: implications for the interpretation of empirical studies. *Proc. Biol. Sci.*, **272**, 671–677.
4. Bateson, P., Barker, D., Clutton-Brock, T., Deb, D., D'Udine, B., Foley, R.A., Gluckman, P., Godfrey, K., Kirkwood, T., Lahr, M.M. *et al.* (2004) Developmental plasticity and human health. *Nature*, **430**, 419–421.
5. McMillen, I.C. and Robinson, J.S. (2005) Developmental origins of the metabolic syndrome: prediction, plasticity, and programming. *Physiol. Rev.*, **85**, 571–633.
6. Hales, C.N. and Barker, D.J. (2001) The thrifty phenotype hypothesis. *Br. Med. Bull.*, **60**, 5–20.
7. Barker, D.J. and Osmond, C. (1986) Infant mortality, childhood nutrition, and ischaemic heart disease in England and Wales. *Lancet*, **1**, 1077–1081.
8. Barker, D.J., Winter, P.D., Osmond, C., Margetts, B. and Simmonds, S.J. (1989) Weight in infancy and death from ischaemic heart disease. *Lancet*, **2**, 577–580.
9. Gluckman, P.D., Hanson, M.A., Morton, S.M.B. and Pinal, C.S. (2005) Life-long echoes—a critical analysis of the developmental origins of adult disease model. *Biol. Neonate*, **87**, 127–139.
10. Barker, D.J. (2001) A new model for the origins of chronic disease. *Med. Health Care Philos.*, **4**, 31–35.
11. Gluckman, P.D., Hanson, M.A. and Pinal, C. (2005) The developmental origins of adult disease. *Matern. Child Nutr.*, **1**, 130–141.
12. Vickers, M.H., Breier, B.H., Cutfield, W.S., Hofman, P.L. and Gluckman, P.D. (2000) Fetal origins of hyperphagia, obesity, and hypertension and postnatal amplification by hypercaloric nutrition. *Am. J. Physiol. Endocrinol. Metab.*, **279**, E83–E87.
13. Lillycrop, K.A., Phillips, E.S., Jackson, A.A., Hanson, M.A. and Burdge, G.C. (2005) Dietary protein restriction of pregnant rats induces and folic acid supplementation prevents epigenetic modification of hepatic gene expression in the offspring. *J. Nutr.*, **135**, 1382–1386.
14. Ke, X., Lei, Q., James, S.J., Kelleher, S.L., Melnyk, S., Jernigan, S., Yu, X., Wang, L., Callaway, C.W., Gill, G. *et al.* (2006) Uteroplacental insufficiency affects epigenetic determinants of chromatin structure in brains of neonatal and juvenile IUGR rats. *Physiol. Genomics*, **25**, 16–28.
15. Desai, M., Crowther, N.J., Lucas, A. and Hales, C.N. (1996) Organ-selective growth in the offspring of protein-restricted mothers. *Br. J. Nutr.*, **76**, 591–603.
16. Bieswal, F., Ahn, M.-T., Reusens, B., Holvoet, P., Raes, M., Rees, W.D. and Remacle, C. (2006) The importance of catch-up growth after early malnutrition for the programming of obesity in male rat. *Obesity (Silver Spring)*, **14**, 1330–1343.
17. Morris, T.J., Vickers, M., Gluckman, P., Gilmour, S. and Affara, N. (2009) Transcriptional profiling of rats subjected to gestational undernourishment: implications for the developmental variations in metabolic traits. *PLoS ONE*, **4**, e7271.
18. Bertram, C., Trowern, A.R., Copin, N., Jackson, A.A. and Whorwood, C.B. (2001) The maternal diet during pregnancy programs altered expression of the glucocorticoid receptor and type 2 11beta-hydroxysteroid dehydrogenase: potential molecular mechanisms underlying the programming of hypertension in utero. *Endocrinology*, **142**, 2841–2853.

19. Drake, A.J., Tang, J.I. and Nyirenda, M.J. (2007) Mechanisms underlying the role of glucocorticoids in the early life programming of adult disease. *Clin. Sci.*, **113**, 219–232.
20. Woodall, S.M., Breier, B.H., Johnston, B.M. and Gluckman, P.D. (1996) A model of intrauterine growth retardation caused by chronic maternal undernutrition in the rat: effects on the somatotrophic axis and postnatal growth. *J. Endocrinol.*, **150**, 231–242.
21. Heijmans, B.T., Tobi, E.W., Stein, A.D., Putter, H., Blauw, G.J., Susser, E.S., Slagboom, P.E. and Lumey, L.H. (2008) Persistent epigenetic differences associated with prenatal exposure to famine in humans. *Proc. Natl Acad. Sci. USA*, **105**, 17046–17049.
22. Waterland, R.A., Lin, J.-R., Smith, C.A. and Jirtle, R.L. (2006) Post-weaning diet affects genomic imprinting at the insulin-like growth factor 2 (Igf2) locus. *Hum. Mol. Genet.*, **15**, 705–716.
23. Ulrey, C.L., Liu, L., Andrews, L.G. and Tollefsbol, T.O. (2005) The impact of metabolism on DNA methylation. *Hum. Mol. Genet.*, **14**(Spec no. 1), R139–R147.
24. Bates, C.J., Villard, L., Prentice, A.M., Paul, A.A. and Whitehead, R.G. (1984) Seasonal variations in plasma retinol and carotenoid levels in rural Gambian women. *Trans. R. Soc. Trop. Med. Hyg.*, **78**, 814–817.
25. Bates, C.J., Prentice, A.M. and Paul, A.A. (1994) Seasonal variations in vitamins A, C, riboflavin and folate intakes and status of pregnant and lactating women in a rural Gambian community: some possible implications. *Eur. J. Clin. Nutr.*, **48**, 660–668.
26. Rayco-Solon, P., Fulford, A.J. and Prentice, A.M. (2005) Differential effects of seasonality on preterm birth and intrauterine growth restriction in rural Africans. *Am. J. Clin. Nutr.*, **81**, 134–139.
27. Brewster, D.R. and Greenwood, B.M. (1993) Seasonal variation of paediatric diseases in The Gambia, west Africa. *Ann. Trop. Paediatr.*, **13**, 133–146.
28. Moore, S.E., Cole, T.J., Poskitt, E.M., Sonko, B.J., Whitehead, R.G., McGregor, I.A. and Prentice, A.M. (1997) Season of birth predicts mortality in rural Gambia. *Nature*, **388**, 434.
29. Moore, S.E., Cole, T.J., Collinson, A.C., Poskitt, E.M., McGregor, I.A. and Prentice, A.M. (1999) Prenatal or early postnatal events predict infectious deaths in young adulthood in rural Africa. *Int. J. Epidemiol.*, **28**, 1088–1095.
30. McEwen, K.R. and Ferguson-Smith, A.C. (2010) Distinguishing epigenetic marks of developmental and imprinting regulation. *Epigenetics Chromatin*, **3**, 2.
31. Maupetit-Méhouas, S., Mariot, V., Reynès, C., Bertrand, G., Feillet, F., Carel, J.-C., Simon, D., Bihan, H., Gajdos, V., Devouge, E. *et al.* (2011) Quantification of the methylation at the GNAS locus identifies subtypes of sporadic pseudohypoparathyroidism type Ib. *J. Med. Genet.*, **48**, 55–63.
32. Sandovici, I., Smith, N.H., Nitert, M.D., Ackers-Johnson, M., Uribe-Lewis, S., Ito, Y., Jones, R.H., Marquez, V.E., Cairns, W., Tadayyon, M. *et al.* (2011) Maternal diet and aging alter the epigenetic control of a promoter-enhancer interaction at the Hnf4a gene in rat pancreatic islets. *Proc. Natl Acad. Sci. USA*, **108**, 5449–5454.
33. Godfrey, K.M., Sheppard, A., Gluckman, P.D., Lillycrop, K.A., Burdge, G.C., McLean, C., Rodford, J., Slater-Jefferies, J.L., Garratt, E., Crozier, S.R. *et al.* (2011) Epigenetic gene promoter methylation at birth is associated with child's later adiposity. *Diabetes*, **60**, 1528–1534.
34. Waterland, R.A., Kellermayer, R., Laritsky, E., Rayco-Solon, P., Harris, R.A., Travisano, M., Zhang, W., Torskaya, M.S., Zhang, J., Shen, L. *et al.* (2010) Season of conception in rural gambia affects DNA methylation at putative human metastable epialleles. *PLoS Genet.*, **6**, e1001252.
35. Tobi, E.W., Lumey, L.H., Talens, R.P., Kremer, D., Putter, H., Stein, A.D., Slagboom, P.E. and Heijmans, B.T. (2009) DNA methylation differences after exposure to prenatal famine are common and timing- and sex-specific. *Hum. Mol. Genet.*, **18**, 4046–4053.
36. Martino, D.J., Tulic, M.K., Gordon, L., Hodder, M., Richman, T., Metcalfe, J., Prescott, S.L. and Saffery, R. (2011) Evidence for age-related and individual-specific changes in DNA methylation profile of mononuclear cells during early immune development in humans. *Epigenetics*, **6**, 1085–1094.
37. Moore, S.E., Collinson, A.C., N'Gom, P.T., Aspinall, R. and Prentice, A.M. (2006) Early immunological development and mortality from infectious disease in later life. *Proc. Nutr. Soc.*, **65**, 311–318.
38. Bocker, M.T., Hellwig, I., Breiling, A., Eckstein, V., Ho, A.D. and Lyko, F. (2011) Genome-wide promoter DNA methylation dynamics of human hematopoietic progenitor cells during differentiation and aging. *Blood*, **117**, e182–e189.
39. Calvanese, V., Fernández, A.F., Urdinguio, R.G., Suárez-Alvarez, B., Mangas, C., Pérez-García, V., Bueno, C., Montes, R., Ramos-Mejía, V., Martínez-Camblor, P. *et al.* (2012) A promoter DNA demethylation landscape of human hematopoietic differentiation. *Nucleic Acids Res.*, **40**, 116–131.
40. Rakyan, V.K., Down, T.A., Maslau, S., Andrew, T., Yang, T.-P., Beyan, H., Whittaker, P., McCann, O.T., Finer, S., Valdes, A.M. *et al.* (2010) Human aging-associated DNA hypermethylation occurs preferentially at bivalent chromatin domains. *Genome Res.*, **20**, 434–439.
41. Geneviève, D., Sanlaville, D., Faivre, L., Kottler, M.-L., Jambou, M., Gosset, P., Boustani-Samara, D., Pinto, G., Ozilou, C., Abeguilé, G. *et al.* (2005) Paternal deletion of the GNAS imprinted locus (including Gnasxl) in two girls presenting with severe pre- and post-natal growth retardation and intractable feeding difficulties. *Eur. J. Hum. Genet.*, **13**, 1033–1039.
42. Thompson, R.F., Suzuki, M., Lau, K.W. and Greally, J.M. (2009) A pipeline for the quantitative analysis of CG dinucleotide methylation using mass spectrometry. *Bioinformatics*, **25**, 2164–2170.
43. Chen, Y.A., Choufani, S., Ferreira, J.C., Grafodatskaya, D., Butcher, D.T. and Weksberg, R. (2011) Sequence overlap between autosomal and sex-linked probes on the Illumina HumanMethylation27 microarray. *Genomics*, **97**, 214–222.