

(51) International Patent Classification:
C12Q 1/68 (2006.01)(21) International Application Number:
PCT/GB2014/053308(22) International Filing Date:
6 November 2014 (06.11.2014)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
1319576.3 6 November 2013 (06.11.2013) GB

(71) Applicants: UNIVERSITY OF SOUTHAMPTON [GB/GB]; Highfield, Southampton SO17 1BJ (GB). AGENCY FOR SCIENCE, TECHNOLOGY AND RESEARCH [GB/SG]; 1 Fusionopolis Way, #20-10 Connexis, 138632, Singapore (SG).

(72) Inventors: GODFREY, Keith Malcolm; c/o University of Southampton, Highfield, Southampton SO17 1BJ (GB). LILLYCROP, Karen Ann; c/o University of Southampton, Highfield, Southampton SO17 1BJ (GB). HOLBROOK, Joanna Dawn; c/o Agency for Science, Technology and Research, 1 Fusionopolis Way, #20-10 Connexis, Singapore 138632 (SG). HANSON, Mark Adrian; c/o University of Southampton, Highfield, Southampton SO17

1BJ (GB). COOPER, Cyrus; c/o University of Southampton, Highfield, Southampton SO17 1BJ (GB).

(74) Agent: LORD, Hilton David; Marks & Clerk LLP, 90 Long Acre, London, WC2E 9RA (GB).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: PHENOTYPE PREDICTION BY DETERMINING THE METHYLATION STATUS OF GENES

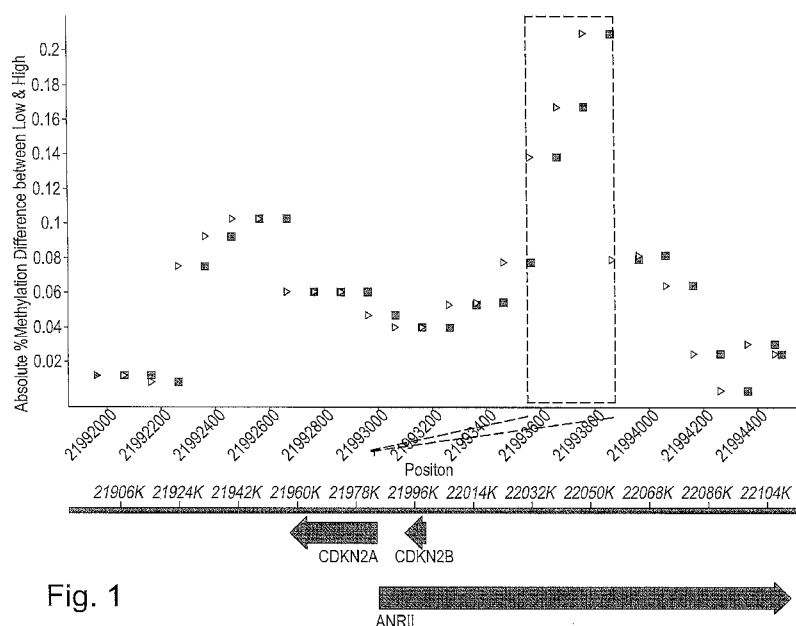


Fig. 1

(57) Abstract: Phenotypic characteristics selected from obesity and/or adiposity, bone health, and cardiovascular disease, may be predicted by determining the methylation status of one or more CpG dinucleotides of the, SLC6A4, CDKN2A, FEN1 and BACH1 genes.

**Published:**

- with international search report (Art. 21(3))
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))
- with sequence listing part of description (Rule 5.2(a))

PHENOTYPE PREDICTION BY DETERMINING THE METHYLATION STATUS OF GENES

Field of the Invention

The present invention relates to the use of epigenetic markers as a means for predicting the future development of one or more phenotypic characteristics in an individual. In particular, the present invention relates to a method of predicting the future development of one or more of obesity and adiposity, bone health and cardiovascular disease in an individual.

Background of the Invention

Fixed genomic variations, such as single nucleotide polymorphisms and copy number variations, explain only a fraction of the risk of developing obesity and/or metabolic disease in humans (1-4). There is now substantial evidence from both human and animal studies that the early life environment of an individual, through an epigenetic mechanism, can affect susceptibility to adiposity and metabolic disease in later life. For instance, in humans, an adverse intrauterine environment has been associated with the development of a range of chronic diseases in later life, including obesity, type-2 diabetes, cardiovascular disease, asthma, osteoarthritis and some forms of cancer (5). These findings have also been replicated in a variety of animal models where impaired nutrition, such as a low protein diet, caloric restriction, or even high fat diet during pregnancy, has been shown to induce long term changes in gene expression and metabolism in the offspring, and which persist throughout the life course (6,7).

The mechanism by which the early life environment can induce persistent phenotypic changes has been suggested to involve the epigenetic regulation of genes. Epigenetics is defined as processes that induce heritable changes in gene expression without a change in the DNA sequence. The major epigenetic mechanisms are DNA methylation, histone modification and non-coding RNAs. These processes control genomic imprinting and X chromosomal inactivation, and have also been implicated in the process of developmental plasticity, the mechanism by which an organism adjusts its developmental programme in response to environmental cues in early life, with long term effects on gene expression and phenotype.

Of particular relevance is the methylation of genomic regions called CpG dinucleotides. CpG dinucleotides, or CpG sites, are regions of DNA where a cytosine nucleotide occurs next to a guanine nucleotide. Cytosines in the CpG dinucleotide can be methylated to form 5-

methylcytosine. This increase in epigenetic methylation can cause genes influenced by the CpG dinucleotide to become less actively expressed or even silenced, or “turned off”. In contrast, a decrease in epigenetic methylation, i.e. hypomethylation, has been shown to be associated with over-expression of certain genes. Some genomic regions have a higher concentration of CpG dinucleotides than others. These regions are known as CpG Islands and are often located within the promoter region of a gene.

Feeding rats a modest reduction in protein during pregnancy induced the hypomethylation of the promoter regions of the nuclear receptors GR and PPAR α in the liver of the offspring, which was accompanied by increased GR and PPAR α expression and alterations in the metabolic processes that they control (6). In humans, differences in the methylation of a number of imprinted and non-imprinted genes have been found in the peripheral blood of individuals whose mothers were exposed to famine during pregnancy (8), suggesting that, as in the animal studies, DNA methylation changes induced by environmental cues in early life are largely maintained throughout the life-course.

Although DNA methylation patterns are often tissue-specific, recent studies have indicated that epigenetic traits in tissues such as cord, blood, and buccal tissues, may provide useful proxy markers of future disease risk in more disease-relevant cell types. For instance Godfrey *et al.*, have shown that methylation of a CpG within the promoter of the RXRA gene in umbilical cord at birth is associated with childhood adiposity at 6 and 9 years in two independent cohorts (9).

The CDKN2A gene locus encodes two potent inhibitors of cell growth, p16^{INK4a}, which is an inhibitor of cyclin-dependent kinase 4, and p14^{ARF} (alternative reading frame relative to p16). Both inhibitors are involved in cell cycle regulation. p14^{ARF} induces cell cycle arrest, by interacting with MDM2 and promoting its degradation, thereby preventing MDM2 from carrying out its usual function of destabilising p53. The resultant increase in p53 levels triggers p53-mediated cell cycle arrest at both G1 and G2/M phases. CDKN2A is located in a region adjacent to the CDKN2B gene, encoding a cyclin-dependent kinase inhibitor known as p15^{INK4b}. The encoded protein functions as a cell growth regulator. Both p16^{INK4a} and p15^{INK4b} are well established tumour suppressors that play key roles in cell proliferation and differentiation by binding specifically to CDK4 and CDK6, blocking cell proliferation. This shared region is frequently mutated in a wide variety of tumours.

The CDKN2A gene locus also shares a divergent promoter region with CDKN2B-AS, also known as ANRIL (antisense non-coding RNA in the INK4 locus) (10). ANRIL is a 3.8k

transcript which is thought to be involved in transcriptional repression. ANRIL negatively regulates both p16^{INK4a} and p15^{INK4b}. Single nucleotide polymorphisms within this region are associated with many types of cancers. The genes within the CDKN2A locus have also been shown to be important for a number of other cellular processes including cellular differentiation, ageing and in driving cellular senescence through irreversible cell cycle arrest (11-22).

The SLC6A4 (solute carrier family 6) gene encodes an integral membrane protein that functions as a serotonin transporter, transporting serotonin from the synaptic spaces into presynaptic neurons. This gene has previously been implicated in conditions such as Alzheimer's disease and depression.

The BACH1 gene encodes a transcription factor that belongs to the cap'n'collar type of basic region leucine zipper factor family. The encoded protein is thought to play a role in transcription activation and repression.

The FEN1 gene is located between the fatty acid desaturase 1 and 2 (FADS1 and FADS2) genes on chromosome 11 and encodes the enzyme Flap endonuclease 1. This enzyme is involved in DNA synthesis and repair.

With the escalating prevalence of lifestyle related disorders worldwide, such as obesity, particularly in early childhood and cardiovascular disease, the identification of an improved marker for determining whether or not an individual will develop such undesirable characteristics would be extremely valuable in order to assess individuals at risk from non-communicable disease, and would provide evidence for the role of early life environment within the normal range rather than extremes of environment in the development of obesity.

The current invention helps to solve this problem by providing a method for predicting development of an undesirable phenotypic characteristic in an individual, in particular, one or more of obesity, bone health and/or cardiovascular disease, and may be of assistance in identifying external factors relevant to the incidence of a particular characteristic in an individual, or in a population in general. In this way, information may be obtained which could be used to reduce the incidence of the condition and any associated metabolic disorders. Once the susceptibility for a particular characteristic has been determined in an individual, the lifestyle of that individual may be managed accordingly, in order to reduce the risk of the actual occurrence of the characteristic.

Summary of the Invention

Thus, in a first aspect, the invention provides a method of predicting the development of one or more phenotypic characteristics in an individual, wherein the phenotypic characteristic is selected from one or more of:

- (i) obesity and/or adiposity;
- (ii) bone mineral content and/or bone health, and
- (iii) cardiovascular disease,

wherein the method comprises determining the methylation status of one or more CpG dinucleotides of a gene in a sample, wherein said gene is selected from the group comprising SLC6A4, CDKN2A, FEN1 and BACH1 or a combination thereof, and wherein the methylation status of the one or more CpG dinucleotides provides an indication of the development of said phenotypic characteristic, or characteristics, in said individual.

In the alternative, or in addition, said method comprises detecting ANRIL, 14^{ARF} or p16^{INK4a} transcript, or a combination thereof, in a sample, wherein the presence/absence of ANRIL transcript is indicative of the development of said phenotypic characteristic.

Preferably, the level of gene expression of ANRIL is detected.

Preferably, the level of gene expression of 14^{ARF} and/or p16^{INK4a} is detected.

Advantageously, the methylation status is the level, or amount, of methylation, or the presence or absence of methylation. It may be determined by pyrosequencing, for example, as described hereinbelow, or by any suitable means known in the art. For example, the degree of DNA methylation may be also detected by, but not limited to, bisulfite sequencing, single molecular real time sequencing, methylation-sensitive PCR amplification, sequenom, bisulfite conversion followed by primer extension, or a combination thereof.

Preferably, the one or more phenotypic characteristic is one or more of obesity, adiposity, total body fat mass, percentage body fat mass and/or body fat distribution, and the one or more CpG dinucleotide is selected from the group comprising CpG5, CpG1 and CpG2 of the SLC6A4 gene, CpG1, CpG2, CpG3, CpG4, CpG5, CpG6, CpG7, CpG8, and CpG9 of the CDKN2A gene, CpG1 and 3 of the FEN1 gene and CpG1, CpG4, CpG5, CpG6 and CpG8 of the BACH1 gene. The genomic coordinates of these sites is provided hereinbelow, and was determined through the UCSC Genome Bioinformatics Group home page, at <http://genome.ucsc.edu>.

More preferably, the one or more phenotypic characteristic is one or more of obesity, adiposity, total body fat mass, percentage body fat mass and/or body fat distribution, and wherein the one or more CpG dinucleotide is selected from the group comprising CpG5, CpG1 and CpG2 of the SLC6A4 gene, preferably, CpG5 of the SLC6A4 gene.

Also preferred, the one or more phenotypic characteristic is one or more of obesity, adiposity, total body fat mass, percentage body fat mass and/or body fat distribution and the one or more CpG dinucleotide is selected from the group comprising CpG1, CpG2, CpG3, CpG4, CpG5, CpG6, CpG7, CpG8, and CpG9 of the CDKN2A gene locus.

Also preferred, the one or more phenotypic characteristic is one or more of obesity, adiposity, total body fat mass, percentage body fat mass and/or body fat distribution, and, wherein the one or more CpG dinucleotide is selected from the group comprising CpG1 and 3 of the FEN1 gene.

Alternatively, the one or more phenotypic characteristic is one or more of obesity, adiposity, total body fat mass, percentage body fat mass and/or body fat distribution, and, wherein the one or more CpG dinucleotide is CpG1, CpG4, CpG5, CpG6 and CpG8 of the BACH1 gene.

Preferably, wherein the one or more condition is bone mineral content /or bone health and the one or more CpG dinucleotides are selected from the group consisting of CpG4, CpG5, CpG6, CpG7, CpG8, and CpG9 of the CDKN2A gene, CpG1 of the SLC6A4 gene, and CpG1, CpG3, CpG4, CpG5, CpG7, and CpG8 of the BACH1 gene.

Preferably, wherein the one or more condition is cardiovascular disease and the one or more CpG dinucleotides are selected from the group consisting of CpG5 of the CDKN2A gene, CpG1 and CpG5 of the SLC6A4 gene, and CpG6 of the BACH1 gene.

Preferably, the method comprises determining the degree of methylation in genomic DNA from a tissue sample obtained from a patient.

Advantageously, the tissue sample is umbilical cord or another perinatal tissue sample including cord blood, placenta, hair, saliva, and buccal smear.

Alternatively, the tissue sample may be adipose tissue, for example.

Preferably, the method of the present invention comprises comparing methylation status with the methylation status of an individual with said phenotypic characteristic at a selected older age.

Preferably, the method of the present invention comprises comparing the methylation status with the methylation status of an individual without said phenotypic characteristic at a selected older age.

A further aspect of the present invention provides the use of the methylation status of one or more SLC6A4 CpG dinucleotides for predicting the development one or more phenotypic characteristics in an individual, wherein the phenotypic characteristic is selected from the following:

- (i) one or more of obesity, adiposity, total body fat mass, percentage body fat mass and/or body fat distribution;
- (ii) bone mineral content and/or bone health;
- (iii) cardiovascular disease.

Preferably, the one or more phenotypic characteristic is one or more of obesity, adiposity, total body fat mass, percentage body fat mass and/or body fat distribution, and wherein the one or more CpG dinucleotide is selected from the group comprising CpG5, CpG1 and CpG2 of the SLC6A4 gene, preferably, CpG5 of the SLC6A4 gene.

An alternative aspect of the present invention provides the use of the methylation status of one or more CDK2NA CpG dinucleotides for predicting the development one or more phenotypic characteristics in an individual, wherein the phenotypic characteristic is selected from the following:

- (i) one or more of obesity, adiposity, total body fat mass, percentage body fat mass and/or body fat distribution;
- (ii) bone mineral content and/or bone health;
- (iii) cardiovascular disease.

According to a further aspect, the present invention provides the use of the methylation status of one or more FEN1 CpG dinucleotides for predicting the development one or more phenotypic characteristics in an individual, wherein the phenotypic characteristic is selected from one or more of obesity, adiposity, total body fat mass, percentage body fat mass and/or body fat distribution.

According to a further aspect, the present invention provides the use of the methylation status of one or more BACH1 CpG dinucleotides for predicting the development one or more phenotypic characteristics in an individual, wherein the phenotypic characteristic is selected from the following:

- (i) one or more of obesity, adiposity, total body fat mass, percentage body fat mass and/or body fat distribution;
- (ii) bone mineral content and/or bone health;
- (iii) cardiovascular disease.

According to a further aspect, the present invention provides the detection of ANRIL transcript for predicting the development one or more phenotypic characteristics in an individual, wherein the phenotypic characteristic is selected from the following:

- (i) one or more of obesity, adiposity, total body fat mass, percentage body fat mass and/or body fat distribution;
- (ii) bone mineral content and/or bone health;
- (iii) cardiovascular disease.

Advantageously, said method comprises detecting ANRIL transcript in a sample, wherein the presence/absence of ANRIL transcript is indicative of the development of said phenotypic characteristic.

Preferably, said method comprises detecting ANRIL transcript in a sample, wherein the level of expression of ANRIL is indicative of the development of said phenotypic characteristic.

Preferably, the one or more phenotypic characteristic is adiposity.

According to a further aspect, the present invention provides the detection of p14^{ARF} transcript for predicting the development one or more phenotypic characteristics in an individual, wherein the phenotypic characteristic is selected from the following:

- (i) one or more of obesity, adiposity, total body fat mass, percentage body fat mass and/or body fat distribution;
- (ii) bone mineral content and/or bone health;
- (iii) cardiovascular disease

Advantageously, said method comprises detecting p14^{ARF} transcript in a sample, wherein the presence/absence of p14^{ARF} transcript is indicative of the development of said phenotypic characteristic or characteristics, in said individual.

Preferably, said method comprises detecting p14^{ARF} transcript in a sample, wherein the level of expression of p14^{ARF} is indicative of the development of said phenotypic characteristic or characteristics, in said individual.

Preferably, the one or more phenotypic characteristic is adiposity.

According to a further aspect, the present invention provides the detection of p16^{INK4a} transcript for predicting the development one or more phenotypic characteristics in an individual, wherein the phenotypic characteristic is selected from the following:

- (i) one or more of obesity, adiposity, total body fat mass, percentage body fat mass and/or body fat distribution;
- (ii) bone mineral content and/or bone health;
- (iii) cardiovascular disease.

Advantageously, said method comprises detecting p16^{INK4a} transcript in a sample, wherein the presence/absence of p16^{INK4a} transcript is indicative of the development of said phenotypic characteristic, or characteristics, in said individual.

Preferably, said method comprises detecting p16^{INK4a} transcript in a sample, wherein the level of expression of p16^{INK4a} is indicative of the development of said phenotypic characteristic, or characteristics, in said individual.

Preferably, the one or more phenotypic characteristic is adiposity.

Preferably, the method comprises comparing the level/presence/absence, with the level/presence/absence in an individual with said phenotypic characteristic at a selected older age. Preferably, the method comprises comparing the level/presence/absence with the level/presence/absence in an individual without said phenotypic characteristic at a selected older age.

In a further aspect, the invention provides a method of predicting the development of one or more phenotypic characteristics in an individual, wherein the phenotypic characteristic is selected from one or more of:

- (i) obesity and/or adiposity;

wherein the method comprises determining the expression of one or more genes in a sample, wherein said gene is selected from the group comprising ANRIL, p14^{ARF} and p16^{INK4a} or a combination thereof, and wherein the expression, or the level of expression, of the one or more genes provides an indication of the development of said phenotypic characteristic, or characteristics, in said individual.

Preferably, the method comprises comparing expression with the expression in an individual with said phenotypic characteristic at a selected older age. Preferably, the method comprises comparing the expression with the expression in an individual without said phenotypic characteristic at a selected older age.

Preferably, the method comprises determining the expression from a tissue sample obtained from a patient.

Advantageously, the tissue sample is umbilical cord or another perinatal tissue sample including cord blood, placenta, hair, saliva, and buccal smear.

Alternatively, the tissue sample may be adipose tissue, for example.

A further aspect of the invention provides a method of managing one or more of obesity, adiposity, total body fat mass, percentage body fat mass and/or body fat distribution, bone mineral content and/or cardiovascular disease in an individual comprising the following:

- (i) determining the methylation status of one or more CpG dinucleotides of a gene in a sample, wherein said gene is selected from the group comprising SLC6A4, CDKN2A, FEN1 and BACH1 or a combination thereof, according to the method of the present invention
- (ii) managing the diet, behaviour or physical activity of the individual, wherein the methylation status indicates the development of said characteristic in the individual.

The CpG sites of the present invention are listed in Table 1, below. The sites are provided as coordinates in relation to the hg19 build of the human genome sequence. The relationship to selected phenotypic characteristics of each methylated CpG site is described in the right three columns, with 'inverse' indicating that the methylated site is associated with a negative trend, such as increased likelihood of obesity.

Table 1

CpG ID	Human Genome Coordinates	Adiposity, obesity	Bone health	Cardiovascular disease
CDKN2A CpG1	Hg19 chr9: 21993721-	Inverse		
CDKN2A CpG2	Hg19 chr9: 21993697-	Inverse		
CDKN2A CpG3	Hg19 chr9: 21993694-	Inverse		
CDKN2A CpG4	Hg19 chr9: 21993654-	Inverse	Inverse	
CDKN2A CpG5	Hg19 chr9: 21993645-	Inverse	Inverse	Positive
CDKN2A CpG6	Hg19 chr9: 21993638-	Inverse	Inverse	
CDKN2A CpG7	Hg19 chr9: 21993629-	Inverse	Inverse	
CDKN2A CpG8	Hg19 chr9: 21993603-	Inverse	Inverse	
CDKN2A CpG9	Hg19 chr9: 21993583-	Inverse	Inverse	
SLC6A4 CpG1	Hg19 chr17: 28561601-	Positive	Positive	Positive
SLC6A4 CpG2	Hg19 chr17: 28561578-	Positive		
SLC6A4 CpG3	Hg19 chr17: 28561505-			
SLC6A4 CpG4	Hg19 chr17: 28561490-			
SLC6A4 CpG5	Hg19 chr17: 28561468-	Inverse		Positive
BACH1 CpG1	Hg19 chr21: 30668941	Inverse	Positive	
BACH1 CpG2	Hg19 chr21: 30668995			
BACH1 CpG3	Hg19 chr21: 30669026		Positive	
BACH1 CpG4	Hg19 chr21: 30669196	Inverse	Inverse	
BACH1 CpG5	Hg 19 chr21: 30669217	Inverse	Inverse	
BACH1 CpG6	Hg19 chr21: 30669219	Inverse		Inverse
BACH1 CpG7	Hg19 chr21: 30669245		Positive	
BACH1 CpG8	Hg19 chr21: 30669259	Positive	Positive	
FEN1 CpG1	Hg19 chr11: 61561170	Positive		
FEN1 CpG2	Hg19 chr11: 61561183			
FEN1 CpG3	Hg19 chr11: 61561220	Positive		
FEN1 CpG4	Hg19 chr11: 61561291			

It will be appreciated that it is generally more practical to assay for the presence of methylation on a selected gene, or selected sites of a gene, rather than on all of the genes relevant to the present invention. Accordingly, it is preferred to assay for methylation of CpG sites in one of CDKN2A, SLC6A4, BACH1, and FEN1. It will also be appreciated that assaying for methylation of each, or a selection of, the genes will provide a more comprehensive prognosis of the possible development of undesirable phenotypic

characteristics. In such cases, it will generally be preferred to analyse or assay the CpG sites of each gene separately, and to subsequently consider the results in the round.

From Table 1, it can be seen that the CpG sites of the invention are not all associated each with the three adverse indication columns, with only CpG5 being associated, either positively or negatively, with all three. Accordingly, the methylation state of each site provides an incremental indication of the likelihood of the development of a phenotypic condition, so that if all of the sites labelled 'inverse' under obesity are found to be methylated, then the chances of the individual having weight problems later in life are substantially higher than if none is methylated, if no action is taken to compensate the possibility.

Where neither 'positive' nor 'inverse' is indicated, then the methylation state of the CpG site in question has no statistical correlation with the column in question.

Brief Description of the Figures

Figure 1 illustrates differential methylation at the CDKN2A gene;

Figure 2 illustrates differential methylation at the SLC6A4 gene;

Figure 3 illustrates the correlation between perinatal DNA methylation for CDKN2A and SLC6A4 and DXA %fat age 6 years;

Figure 4 displays the results of luciferase and EMSA assays which examine the role of the CDKN2A region of interest in gene expression;

Figure 5 illustrates differential methylation at the BACH1 gene;

Figure 6 illustrates differential methylation at the FEN1 gene;

Figure 7 displays a graph showing the correlation between perinatal DNA methylation for BACH1 and FEN1 and DXA %fat age 6 years;

Figure 8 illustrates the ponderal index at age 18 months in the GUSTO cohort according to umbilical cord methylation of CDKN2A CpG1 methylation;

Figure 9 A - D display graphs showing p14ARF expression in (A) SW-873 (B) Saos-2 (C) A549 and (D) JAR cell lines;

Figure 9E displays the results of electrophoretic mobility shift assays showing binding to the CpG sites within the CDKN2A DMR.

Figure 10 illustrates the association between SLC6A4 CpG5 methylation total and percentage fatmass at 4 and 6 years in SWS cohort. Methylation has been divided into 4 equal groups according to rank; means and standard error of the mean are plotted for each group.

Figure 11 illustrates the association between SLC6A4 CpG5 methylation and tricep skinfold thickness from birth to 6 years in the SWS cohort. Methylation has been divided into 4 equal groups according to rank; means and standard error of the mean are plotted for each group.

Figure 12A illustrates the association between SLC6A4 CpG5 methylation and total fatmass at 4 years in the SWS cohort. Methylation has been divided into 4 equal groups according to rank; means and standard error of the mean are plotted for each group.

Figure 12B illustrates the association between SLC6A4 CpG5 methylation and child's tricep skinfold thickness at birth in the SWS cohort. Methylation has been divided into 4 equal groups according to rank; means and standard error of the mean are plotted for each group.

Figure 13A compares SLC6A4 CpG 5 methylation in adipose tissue from obese individuals and with lean individuals from the BIOCLAIMS cohort. Methylation from both lean and obese individuals is shown for CpG 5 (n=22). Whiskers show min to max. * p 0.01-0.05, ** p $\leq$ 0.01.

Figure 13B illustrates SLC6A4 transcript expression in adipose tissue from obese individuals and from lean individuals from an adult cohort. Whiskers show min to max. (n=63) * p 0.01-0.05, ** p $\leq$ 0.01.

Figure 14 illustrates the top two pathways enriched amongst the DMRs; (A) DNA replication, recombination and repair, and (B) cell death and survival, gene expression, and embryonic survival. The DMRs validated by pyrosequencing are indicated by the arrows.

Figure 15A illustrates the concordance of methylation values with percent body fat for CDKN2A. The X-axis shows percent body fat and y-axis shows % methylation as estimated by BATMAN. Sample data points are shaded by percent body fat group groups (dark = low, lowest percent body fat; medium = medium, medium percent body fat; light = high, highest percent body fat).

Figure 15B illustrates the correlation between DNA methylation for CDKN2A CpG1 ($\beta=-0.004$; $p=0.003$), CpG3 ($\beta=-0.005$; $p=0.007$) and CpG6 ($\beta=-0.004$; $p=0.01$), and DXA %fat age 6 years.

Figure 15C illustrates the correlation between DNA methylation for CDKN2A CpG1 ($\beta=-0.09$; $p=0.007$), CpG3 ($\beta=-0.07$; $p=0.1$), CpG4 ($\beta=-0.09$; $p=0.007$) and CpG6 ($\beta=-0.06$; $p=0.01$), and DXA %fat age 4 years.

Figure 15E illustrates the association between CDKN2A CpG1 ($p=0.015$), CpG2 ($p=0.047$) and CpG3 ($p=0.035$) methylation at birth and Subscapular skinfold thickness at age 2 years in the GUSTO cohort (dark shading = female; light shading = male).

Figures 16A, B and D illustrate the concordance of methylation values with percent body fat for SLC6A4 (A), FEN1 (B) and BACH1 (D). The X-axis shows percent body fat and y-axis shows % methylation as estimated by BATMAN. Sample data points are shaded by percent body fat group groups (dark = low, lowest percent body fat; medium = medium, medium percent body fat; light = high, highest percent body fat).

Figure 16C illustrates the validation of the association observed between the DMR methylation of BACH1 at birth and % fat mass at 6 years of age ($\beta=-0.009$; $p=0.038$). Methylation has been divided into four equal groups according to rank; means and standard errors are plotted for each group.

Detailed Description

The current invention is based upon the surprising finding that the degree of methylation of particular CpG sites within the CDK2NA gene at birth correlate with the propensity of an individual to develop one or more of obesity, low bone mineral content and poor bone health and cardiovascular disease in later childhood. We have also discovered associations of the degree of methylation of the CpG loci within SLC6A4 and BACH1 with the propensity of an individual to develop one or more of obesity, low bone mineral content and poor bone health

and cardiovascular disease, and of the degree of methylation of the CpG loci within FEN1 with the propensity of an individual to develop later obesity, thereby demonstrating that the early life environment is an important determinant of adiposity, bone health and the risk of cardiovascular disease, and may influence future health through the altered epigenetic regulation of gene expression.

In its broadest aspect, the present invention provides a method of predicting one or more of the following conditions:

- (i) obesity or adiposity;
- (ii) bone health; and
- (iii) cardiovascular disease.

The method comprises determining the degree of methylation of one or more CpG dinucleotides of a selected gene in a sample, wherein the degree of methylation correlates with propensity for said condition. In a preferred embodiment, the condition is obesity, total body fat mass, percentage body fat mass and/or body fat distribution. In another preferred embodiment, the condition is bone mineral content and/or poor bone health. In a further preferred embodiment, the condition is cardiovascular disease.

A method also comprises detecting the transcript, or determining the expression, of a selected gene in a sample, wherein the presence or absence of the transcript, or the level of expression, correlates with propensity for said condition. In a preferred embodiment, the condition is obesity, adiposity total body fat mass, percentage body fat mass and/or body fat distribution. In another preferred embodiment, the condition is bone mineral content and/or poor bone health. In a further preferred embodiment, the condition is cardiovascular disease. In a still preferred embodiment the gene is ANRIL, p14^{ARF} and/or p16^{INK4A}.

It will be appreciated that such a method of prediction is preferably applied to a neonate or young child, typically under the age of 12 months, and preferably which does not already exhibit one of, or the, phenotypic factor(s) being prognosticated. Samples are typically collected in the perinatal period, but the test may be done at a later age. If the test is performed later, the sample will typically be preserved, such as by freezing.

Where the sample used for DNA analysis is umbilical cord, or another perinatal tissue sample, it is preferred that the method further comprises a step of comparing the degree of methylation with the degree of methylation of variance of the condition in an individual at a selected older age.

The current method may be used to manage propensity for the occurrence of undesirable phenotypic characteristics in humans. For example, such use of epigenetic markers enables the targeting of corrective strategies where propensity for obesity is identified.

The tissue to be used for determination is desirably a tissue readily available in early life, and may be obtained from a neonate, infant, child or young adult. Desirably, the tissue is a perinatal tissue sample containing genomic DNA of the individual, taken at, or soon after, birth, typically within 12 months of parturition. Preferably, the tissue may be umbilical cord, adipose tissue, blood (including umbilical cord blood), placenta, chorionic villus biopsy, amniotic fluid, hair, or buccal smears. Preferably, such tissue will be from a newborn or taken from an infant within a few weeks of birth, more preferably within a few days of birth (less than a week), although samples taken later, e.g. within 6 months, or longer, may prove useful.

The method of the present invention may be used to provide a targeted intervention to modify the effect of the condition on health is possible.

Obesity can be associated with type 2 diabetes, metabolic syndrome, and other conditions for which obesity is recognised to be a risk factor.

Bone mineral content prediction is of interest in identifying individuals who may be susceptible to impaired linear growth, and disease states linked with low bone mineral content such as osteoporosis in later life.

Once the propensity for exhibiting phenotypes has been determined in an individual, management of that individual's lifestyle (diet, behaviour, exercise, etc.) can be undertaken to reduce the risk of the actual occurrence of the undesirable characteristic such as obesity, low mineral content, or cardiovascular disease.

The methods of the present invention may be used to identify external factors that may contribute to the incidence of the characteristics in the individual or population in general. In this way, information may be obtained useful in directing initiatives aimed at reducing the incidence of the condition and thereby the associated disorders. The utility of this invention is that it allows selection of appropriate ways of managing an individual for optimal health and wellbeing.

The method of the present invention is important, as obesity is also associated with metabolic disorders in later life, e.g. diabetes.

The accompanying Examples are for illustration only, and do not restrict the scope of the present invention.

EXAMPLES

EXAMPLE 1

Materials and Methods

Southampton Women's Survey: Participants

The Southampton Women's Survey (SWS) is a prospective cohort study of a large group of non-pregnant women aged 20 to 34 years living in the city of Southampton, UK. Comprehensive details of the study have been published previously (23). Women were recruited through General Practices across the city between April 1998 and December 2002. Each woman was invited to take part by letter, followed by a telephone call when an interview date was arranged; 12,583 women agreed to take part, 75 % of all women contacted. Trained research nurses visited the women at home and collected information about their health, diet and lifestyles, as well as taking anthropometric measurements. Women who subsequently became pregnant were followed up at 11, 19 and 34 weeks gestation and their offspring were studied in infancy and childhood. Amongst women who became pregnant, smoking status in pregnancy was ascertained at the 11 and 34 week interviews. A total of 1981 women became pregnant and delivered a live-born singleton infant before the end of 2003. Six infants died in the neonatal period and two had major congenital growth abnormalities, which left 1973 mother-offspring pairs. Table 2 outlines participant characteristics.

Table 2

Characteristic	group/number	% or Median (InterQ range)
Qualification level	none	0.9%
	CSE	9.9%
	O levels	27.0%
	A levels	26.4%
	HND	9.1%
	Degree	27.6%

Social class	Professional	6.4%
	Management/technical	42.2%
	Skilled non-manual	33.2%
	Skilled manual	6.9%
	Partly skilled	10.4%
	Unskilled	0.9%
Current smoker	No	77.0%
	Yes	23.0%
Pregnancy	First pregnancy	47.7%
	Pregnant before	52.3%
Sex of children	Male	52.8%
	Female	47.2%
Birth order	1 st	47.7%
	2 nd	38.6%
	3 rd or higher	13.6%
Woman's age at birth of child	352	30.7 (28.1 to 33.4)
Early pregnancy women's BMI	279	24.8 (22.7 to 28.2)
Birthweight (kg)	349	3.4 (3.2 to 3.8)
Gestation (wks)	352	40.0 (39.0 to 40.9)
Infant total fat (g) at birth	147	497.9 (364.3 to 644.3)
Offspring total fat (kg) at 4 years	283	4.0 (3.4 to 4.8)
Offspring total fat (kg) at 6 years	264	4.5 (3.6 to 6.0)
Infant percentage fat at birth	147	14.3 (11.6 to 17.3)
Offspring percentage fat at 4 years	283	27.8 (24.3 to 32.5)
Offspring percentage fat at 6 years	263	24.2 (19.9 to 29.7)
Height (cm) at 4 years	350	103.9 (101.2 to 106.9)
Height (cm) at 6 years	282	120 (116.7 to 124.3)

Body composition measurements in the SWS offspring

A sub-set of children from the SWS study with dietary data in infancy, who were born before 2003, were invited to take part in follow-up at age 4 and 6 years to assess their body composition. Adiposity measurements were made by dual-energy X-ray absorptiometry (Hologic Discovery, paediatric scan mode, Hologic Inc., Bedford, MA)(24).

Instruments were calibrated daily; coefficients of variation were <1 %. Follow-up of the children and umbilical cord sample collection/analysis was carried out under Institutional Review Board approval (Southampton and SW Hampshire Research Ethics Committee) with written informed consent. Clinical investigations were conducted according to the principles expressed in the Declaration of Helsinki. Umbilical cord is only collected in the immediate perinatal period; after collection it is typically frozen for later processing.

Growing Up in Singapore Towards Healthy Outcome (GUSTO): Participants

Umbilical cord specimens were collected from babies born at the KK Women's and Children's Hospital (KKH) and the National University Hospital (NUH), in Singapore. These hospitals are part of the GUSTO birth cohort study (25). Written parental consent to participate in the study was given and hard copies are stored by the GUSTO data team. Ethical approval for the study and the consent forms and contents was granted, by the ethics boards of both KKH and NUH, which are centralized Institute Review Board (CIRB) and Domain Specific Review Board (DSRB), respectively. We screened 3751 families and 2034 met eligibility criteria. Of the 1247 women (response rate 61.3 %) recruited, 1162 conceived naturally, while 85 conceived through *in vitro* fertilisation (IVF). At baseline, 55.9 % were Chinese, 26.1 % Malay and 18 % Indian. Mean maternal age at recruitment was 30.6 years (range: 18 to 46). Gestational age was defined from a dating ultrasound (10–12 weeks) followed by an additional scan at 18–22 weeks. The average birth weight in the GUSTO cohort was 3081 g, which is comparable to the average across a larger Singaporean sample of 3183 g for a term infant (unpublished data). Umbilical cord tissue samples were collected at the time of delivery, flushed with saline to remove foetal blood and flash-frozen in liquid nitrogen within 30 min of collection.

At age 18 months weight was recorded to the nearest gram using a calibrated scale (SECA-334, SECA Corp. Hamburg) and recumbent crown-heel length measured to the nearest 0.1 cm using a SECA-210 Mobile Measuring Mat. For reliability, measurements were taken

in duplicate. Ponderal index [weight/(length)³] was derived as a measure of adiposity. Table 3 outlines the characteristics of the children that participated in this study.

Table 3 GUSTO subject characteristics

Mother	Mean (SD) or %	n
Household monthly income (SGD)		196
0-1999	18.4	
2000-5999	61.7	
≥6000	19.9	
Maternal highest education %		203
Vocational/Technical training	9.9	
Primary	5.4	
Secondary	31.5	
GCE A level	27.6	
University	25.6	
Age at birth, years	30.5 (5.6)	209
Smoker	3.4	208
Maternal BMI	26.4 (4.6)	206
Infant		
Sex (%)		209
Male	50.2	
Female	49.8	
Birth Order %		208
1 st	44.2	
2 nd	30.2	
3 rd or higher	25.5	
Birth weight, kg	3.1 (0.4)	209
Gestational age, weeks	38.7 (1.3)	208
18 month ponderal index	20.5 (1.9)	177

The analysis in this table is based on those seen at age 18 months who had an umbilical cord sample available

Whole genome methylation analysis

Genomic DNA from umbilical cord samples was obtained from twenty-one children from the SWS cohort, chosen to represent a range of % fat values at 6 years (as measured by DXA

scan) between the 5th and the 95th percentiles. The DNA from each subject was sheared by sonication to produce fragments of between 200 and 500bp. A proportion of this DNA (input DNA) was purified and the remaining DNA incubated with a His-tagged MBD2b (methyl binding domain of MeCP2) protein to form a complex which was then captured on nickel coated magnetic beads (MethylCollector kit, Active Motif). Beads were washed to remove unmethylated DNA fragments before the DNA was eluted using an elution buffer designed to simultaneously elute methylated DNA while degrading the MBD2b protein (26). To ensure the purity of the methylated DNA fraction, real-time PCR was carried out using primers specific to promoter regions of β -actin, which is unmethylated and the H19 Imprinting control region (H19ICR) which is known to be methylated.

Methylated and input samples were then hybridised to the Agilent Human Promoter Whole-Genome ChIP-on-chip array (G4489A)(27). The Agilent Human Promoter Whole Genome Chip on chip array (G4489A) contains probes (60mers) spaced every 100-300bp across the promoter regions of the best-defined human Refseq gene promoters (21,000) from -8kb to +2kb downstream of the TSS. Microarray hybridisation was carried out by Oxford Gene Technology (OGT, Oxford UK) in accordance with the company's quality control procedures using standard protocols for labelling, hybridisation and washing. Microarray slides were scanned at 5 μ M resolution using the extended dynamic range (High 100 %, Low 10 %). The slides were then feature-extracted using Agilent Feature Extraction Software (v.9.5.3.1). All arrays were normalised per spot and per chip using an intensity dependent normalisation (Lowess normalisation) using Genespring.

Data analysis

The log₂ of Cy3/Cy5 values were obtained for each probe after background subtraction. The log₂ ratios were then passed through Bayesian Tool for Methylation Analysis (BATMAN)(28). BATMAN is an algorithm that generates Bayesian distributions of % methylation values for every 100 nucleotide region covered on the array. Log₂ ratios of the overlapping probes and CpG densities in the probe of 100nt of flanking genomic sequence are inputted to the likely beta value distributions. The mode of the distribution for each 100 nt region returned by BATMAN was used for further analysis. Sample quality control was performed by examining the frequency distribution of the BATMAN output as well as the raw log₂ ratios, as suggested by (26). Sample quality control was performed by examining the frequency distribution of the BATMAN output as well as the raw log₂ ratios, as suggested by (29)

Identification of differentially methylated regions (DMRs)

Statistical tests to identify Differentially Methylated Regions

Percentage methylation values at each 100 nt region were subjected to robust regression analysis percent fat to correct for heteroscedasticity (30). Fisher Exact tests were performed to identify larger chromosomal regions (differentially methylated regions, DMRs) that tiled sequentially on the array. These regions were tested for significant enrichment of differential methylation amongst the 100nt regions, within their span. This principal is similar to that of (29) but is especially designed for the custom array design.

Test for robustness to failed probe signals

It was observed that a portion of the probes included in the raw data had a log2 value approaching 0. These probes were assumed to be failed probes and were removed from the data set and the modified dataset was subjected to the BATMAN algorithm. Any DMRs that were not robust between the two data sets were removed from consideration. Finally, to enrich for true positives and to identify biological pathways implicated in developmental programming of obesity, pathway enrichment analysis and *de novo* network discovery were performed using the Ingenuity network analysis suite.

Pyrosequencing

The levels of methylation of individual CpG dinucleotides in DMRs identified by BATMAN analysis were measured by sodium bisulfite pyrosequencing. Genomic DNA was prepared and bisulphite conversion was carried out using the EZ DNA methylation kit (ZymoResearch, Irvine, California, USA). The pyrosequencing reaction was carried out using primers listed in Table 4. Modified DNA was amplified using Hotstart Plus DNA polymerase (QIAGEN). PCR products were immobilised on streptavidin–sepharose beads (GE Healthcare UK Ltd., Amersham, Buckinghamshire, UK), washed, denatured and released into annealing buffer containing the sequencing primers. Pyrosequencing was carried out using the SQA kit on a PSQ 96MA machine (Biotage, Uppsala, Sweden) and the % methylation was calculated using the Pyro Q CpG software (Biotage). Within assay precision was between 0.8 and 1.8 % and detection limits were 2 - 5 % methylation. It will be appreciated that any suitable primer may be used to target CpG sites, provided that it binds the sequence with sufficient avidity under the experimental protocol used.

Table 4: Primers

<u>Pyrosequencing Primer</u>	<u>Sequence</u>	<u>SEQ ID NO.</u>
CDKN2A 1-3 F	AGTAGGAAAGGTGTATTTTAAGTATATTT	1
CDKN2A 1-3 S	AGAATTATTGTTAATTATTTAAGTT	2
CDKN2A 4-9 F	TGGGGAGAATTATTGTTAATTATTTAAGTT	3
CDKN2A 4-9 S	TAGGAGAGTGGAGGA	4
CDKN2A 8-9 S	GTAGGTAGAGATTTTTTGAAATGT	5
CDKN2A 1-3 R Bio	TATCTCACCAATCCTCCACTCTCCTAAA	6
CDKN2A 4-9 R Bio	AAAAACCCATTTCCCTATTAECTACA	7
APOL5 F	AGTAATGGAATTGGTAGTTTTTAGAT	8
APOL5 P	TGTTTTATTGTGTTTTTTTATTTAG	9
BACH1 R1	ATCCCACCCCCAAAAATACTA	10
BACH1 P1	ACCTAAAATTTAAACTTCATCC	11
BACH1 F2	GGGATGAAGTTTAAATTTTAGGTGTTGTG	12
BACH1 P2	TGTATGAGGTTTTTTTTTATATTGA	13
BACH F3	TTTTTTGGGGGTGGGATTTAGGAA	14
BACH P3.1	TTTTTTTTTTTTTTGAGATAGAGT	15
BACH P3.2	GATTTTAGTTTATTGTAA	16
BACH P3.3	GATTTTTTTGTGTTAGT	17
BACH P3.4	TGTTTGGTTAATTTTTTTGTATTTTA	18
BACH P3.5	ATTATGTTGGGGTTTAGT	19
APOL5 R Bio	CCCAAAATACAACAATAAAATACTATATTC	20
BACH1 R2 Bio	ATTCCTAAATCCCACCCCCAAAA	21
BACH1 F1 Bio	GTGGTTTTTGGTTTAAAGTAAGAGAAT	22
BACH R3 Bio	CATACCTATAATCCCAACACTTTCA	23
SLC6A4 F	GGGATTTGTGTTTAATGAAATAGAGAATTG	24
SLC6A4 R	ACCCTAAAACCTCAATAACTTCACAC	25
SLC6A4 S	TGTTTAATGAAATAGAGAATTGT	26
FEN1 F	GAGAAAATAAATTTTAGTGTTTGGTAATG	27
FEN1 S	GGGATGGATTTGGAAG	28
FEN1 R Bio	TATTCTAATCCCAACCTAACCTACA	29

<u>Realtime primer</u>	<u>Sequence</u>	<u>SEQ ID NO.</u>
ANRIL unspliced F	CAGTGGCTTCCTGTTTCATGC	30

ANRIL unspliced R	GGGCTTGACGTCTGATCTGT	31
ANRIL F1	TGAAAAACACACATCAAAGGAG	32
ANRIL R1	GATTCCACCACACCTAACAG	33
P14ARF F1	CCCTCGTGCTGATGCTACTGAG	34
P14ARF R1	GTGAGAGTGGCGGGGTCGGC	35
P16INK4a F1	CGAATAGTTACGGTCGGAGGC	36
P16INK4a R1	TCGGGTGAGAGTGGCGGGGTC	37

<u>Primer</u>	<u>sequence</u>	<u>SEQ ID NO.</u>
ANRIL F1	TTGGGTACCACCTCTAACTCACAAAGAAAGC	38
ANRIL R1	GCCAAGCTTTGGGAATGACTAAGACACAC	39
ANRIL F2	GAGAAGCTTCCCAGGATATTTCGGGACTCA	40
ANRIL R2	TTACCATGGTGTGTCAGGTGACGGATGTAGC	41

Cloning of promoter regions

PCR was used to amplify the region containing the ANRIL promoter, and the region identified by the BATMAN analysis. Two sets of PCR primers were used, the first set amplified the ANRIL promoter (-926 to +20 relative to ANRIL TSS). A HindIII restriction site was added to the forward primer, and an NcoI restriction site to the reverse primer for cloning into pGL Basic (Promega, UK) to create pGL -951. The second PCR primer pair amplified a region of interest (ROI), immediately adjacent to the already cloned ANRIL promoter region (-1281 to -925). A KpnI restriction site was added to the forward primer, and a HindIII restriction site to the reverse primer for cloning into pGL -951. The completed plasmid, pGL ANRIL, contained the full genomic sequence from -1281 to +20 of the ANRIL promoter with a HindIII site inserted after the ROI to allow subsequent patch methylation. All PCR amplification was carried out using Hot Star High Fidelity DNA polymerase (QIAGEN). Primers are listed in Table 4 above. The base pair sequence of the cloned region was confirmed by sequencing (GATC, Germany).

Patch Methylation

A 351bp fragment, containing the nine CpG dinucleotides identified in the BATMAN analysis was excised from pGL ANRIL by restriction digest with KpnI and HindIII (Thermo Scientific). The fragment was treated with 8 units of SssI methylase (NEB UK) for 4 hours at 37°C. An additional 4 units of SssI, along with further SAM, was then added and the reaction incubated at 37°C overnight, followed by heat inactivation. Complete methylation of the fragment was confirmed by restriction digest using the methylation sensitive restriction

enzymes Acil and HpaII. A mock reaction was carried out in parallel, identical except for addition of water in place of SssI.

The methylated fragment was then re-ligated into the pGL ANRIL vector using the rapid ligation kit (NEB UK), following the manufacturer's recommendations. The completed plasmid was purified by phenol/chloroform extraction followed by ethanol precipitation. 1µl of purified plasmid DNA was restricted with NcoI and run on a TAE agarose gel for quantification.

Cell culture and transfection

SW872 cells were cultured in DMEM 4.5g Glucose +10% Fetal Bovine Serum + 1% Penicillin/Streptomycin. Cells were cultured in 96-well plates for 24hrs prior to transfection. 100ng of prepared plasmid DNA was transfected per well for both the methylated and mock reactions, with six replicates per transfection. The pGL CMV Renilla plasmid (Promega UK) was co-transfected as a control. Transfections were carried out using FuGENE 6 (Promega UK) following manufacturers' guidelines. Transfected cells were cultured for 48hrs prior to addition of lysis buffer. Luciferase assays were carried out using the Dual-Luciferase® Reporter Assay System (Promega UK), on a VarioSkan Flash Luminometer (ThermoScientific).

Electrophoretic mobility shift assays

Nuclear extracts were prepared from human neuroblastoma cell line IMR-32. Cell pellets were re-suspended in 800µl of lysis buffer (10mM Hepes, pH 7.9, 10mM KCl, 0.1mM EDTA, 1mM DTT, 0.5mM PMSF, 1mM leupeptin) and incubated on ice for 10 min. Then 50µl of NP-40 was added, the cells vortexed for 10 seconds and spun at 11,000g for 30 seconds. The supernatant was discarded and the pellet containing nuclei re-suspended in 50µl of nuclear extraction buffer (20mM HEPES, pH 7.9, 0.4M NaCl, 1mM EDTA, 1mM DTT, 1mM PMSF). This was then incubated on ice for 30 min and nuclear debris pelleted at 4°C for 5 minutes at 11,000g. The protein concentration of nuclear extracts was determined with the BCA protein assay kit (Pierce) using the manufacturer's instructions. The nuclear extracts were stored at -80°C until use.

For EMSA, double-stranded DNA oligonucleotides (biomers.net) labelled with biotin at the 5' termini of both strands were used. Sense strand sequences were as follows: Single-stranded oligonucleotide probes were annealed by heating equimolar amounts of complementary strands to 95°C for 5 minutes and slowly cooling the reaction mixture to room temperature.

Electrophoretic mobility shift assays were performed using LightShift Chemiluminescent EMSA Kit (Thermo Scientific). All binding reactions were carried out in the presence of 2pmol of the non-specific DNA Poly (dI*dC), x fmol of biotin-labelled probe, 1x binding buffer, 2.5 % glycerol, 0.05 % NP-40, 5mM MgCl₂ (20µl final volume) and incubated on ice for 20 minutes. For reactions carried out in the presence of nuclear extract (5µg) total protein was utilised. Competition was performed in the presence of 500-fold excess of unlabelled oligonucleotide (methylated or non-methylated) and incubating on ice for 30minutes prior to addition of labelled probe. Complexes were resolved on pre-run 4 % non-denaturing polyacrylamide gel in 0.5X Tris-borate-EDTA for 45 minutes at 100V followed by semi dry transfer (200mA 1hr) to a positive nylon membrane and UV cross-linking. After incubation in blocking buffer for 15 min at room temperature the membrane was incubated with streptavidin-HRP conjugate for 15 minutes. The membrane was then washed and visualised with a chemiluminescent substrate (Thermo Scientific).

Statistical analysis

Statistical analysis was carried out using Stata (Statacorp) versions 11.2 and 12.1. Histograms of all continuous variables were plotted to check for normality. The distributions of some measures of adiposity were observed to be skewed and were therefore transformed using a log e transformation. DXA adiposity measurements were calculated without including the child's head in order to minimise the effect of head movement on the measurement. Regression models were built using child's adiposity measurement (at either 4 or 6 years) as the outcome and CpG methylation as the predictor. Models were adjusted for child's sex and age as appropriate for each time point. Adiposity was measured as %fat and total fat in grams, converted to Z-scores to facilitate interpretation of the effect size, and methylation measurements as %s; the regression coefficient can therefore be interpreted as the standard deviation change in %fat (or total fat) for each % change in methylation. Results are presented as regression coefficients (β) for each 10% change in methylation with their associated P-values.

Results

Identification of Differentially Methylated Regions at birth, associated with later adiposity

To identify perinatal epigenetic biomarkers associated with later body composition outcomes, we carried out a genome wide DNA methylation analysis using a methyl binding domain (MBD) array in umbilical cord, comparing DNA methylation levels at birth with differences in percentage fat mass in the children aged 6 years. Array data was analysed

using the validated algorithm BATMAN (28), designed to handle MBD data. Genome wide DNA methylation data from high and low fat mass groups were compared at 100nt resolution, looking for regions of high differential methylation between these groups. This comparison identified 93 differentially methylated regions (DMRs) associated with percentage fat mass in the children at 6 years of age (Table 5).

Table 5

List of DMRs

GeneSymbol	Fisher_Exact_p	Chr	Start	End	Alt Start	Alt End
FEN1	1.00E-06	11	61561031	61561530		
CCND1	1.00E-06	11	69456477	69456776		
GAPDH	1.00E-06	12	6644416	6644715		
NGB	1.00E-06	14	77736478	77736777		
GUSB	1.00E-06	7	65445733	65445832		
CCDC147	6.00E-06	10	106114677	106115476		
ZNF17	6.00E-06	19	52075315	52075614		
GATA6	2.40E-05	18	19747012	19747411		
MPV17	2.90E-05	19	18301439	18301838		
AKAP8	3.20E-05	19	15492244	15492743		
TNFRSF19	4.20E-05	13	24154000	24154399		
ABCA4	4.70E-05	1	94587826	94588425		
TMEM116	5.50E-05	12	112450339	112450838	112446439	112446638
SETD4	5.60E-05	21	37432113	37432512		
SEZ6	6.50E-05	17	27337616	27337915		
ZNF56	7.60E-05	19	37178039	37178338		
APP	9.30E-05	21	27541903	27542102		
CREBL2	0.000121	12	12763984	12764683		
ZSWIM1	0.000146	20	44509911	44510510		
SP1	0.000228	12	53774930	53775429		
FBXO15	0.000281	18	71812958	71813357		
OR4C11	0.000282	11	55371373	55371672		
C17orf58	0.000297	17	65987793	65988192		
SAA1	0.00037	11	18279872	18280171		
SIAH1	0.000379	16	48418280	48418579		
COPZ1	0.000479	12	54733063	54733521		
LIN52	0.000655	14	74553677	74554676		
C1orf116	0.000813	1	207192230	207192629		
SLC6A4	0.000894	17	23561381	28561680		
ACTR6	0.000942	12	100594927	100595126		
MED16	0.000948	19	893402	893901		
C22orf39	0.000948	22	19435894	19436093		
FYN	0.001036	6	112040878	112041277		
MLEC	0.001079	12	121125691	121125990		

ID1	0.001079	20	30190922	30191421	30190622	30190921
AMZ2	0.00129	17	66246576	66246875		
FGF14	0.001357	13	102571875	102572274		
BACH1	0.001549	21	30668863	30669262		
EXD2	0.00176	14	69674108	69674407		
CBLN4	0.001803	20	54578308	54578807		
MEG8	0.001841	14	101379064	101379663		
NDUFS8	0.002252	11	67799076	67799375		
PCBP2	0.002439	12	53846447	53846946		
SNORD17	0.002526	20	17948153	17948452	17945153	17945552
ZNF664	0.00283	12	124458896	124459195		
OR11H6	0.003012	14	20691887	20692386		
OR4K5	0.003253	14	20384663	20385162		
SLC39A6	0.003518	18	33707646	33708045		
BRSK2	0.003586	11	1412220	1412419		
PRRC1	0.003859	5	126858993	126859692		
C14orf142	0.003905	14	93674364	93677363		
APOL5	0.003905	22	36112644	36113243	36108844	36109443*
WWC2	0.004009	4	184023608	184024107		
DDX10	0.004046	11	108534698	108534997		
DPF3	0.004774	14	73365990	73366261		
GNG4	0.004802	1	235812299	235812598		
LAMA5	0.005051	20	60941631	60941930		
DSCR9	0.005272	21	38593139	38593438		
RNF43	0.005397	17	56433328	56434227		
DGKZ	0.005776	11	46349163	46349562		
TRIM13	0.006775	13	50567722	50568621		
FAM174B	0.006871	15	93198365	93198664		
TOP2A	0.007924	17	38573149	38573648		
NGFR	0.007924	17	47575125	47575281		
C10orf107	0.009278	10	63417842	63418241		
PCNA	0.010237	20	5099301	5099900		
PALB2	0.012353	16	23649185	23649584		
DHX34	0.012353	19	47851748	47852047		
CD1E	0.013548	1	158324116	158326284		
HSF5	0.014521	17	56565374	56565573		
INTS7	0.014585	1	212205435	212205634		
IFNK	0.014585	9	27518797	27519196		
LOC100130987	0.014747	11	67155334	67155633		
ZNF420	0.014747	19	37567943	37568642		
VN1R4	0.016052	19	53775191	53775690		
OR5D18	0.0181	11	55587297	55587596		
TMEM109	0.018377	11	60682108	60682707		
CDKN2A	0.018945	9	21993565	21993864		
HLA-C	0.019681	6	31240679	31240978		

LTBR	0.020236	12	6492667	6493066		
TMEM61	0.022593	1	55441680	55441979		
MTNR1B	0.02273	11	92704083	92704582		
CD40	0.025506	20	44742246	44742845		
CLIC6	0.029338	21	36042842	36043341		
MRPL16	0.032981	11	59577672	59578071	59575872	59576471
FAR1	0.03451	11	13689600	13689799		
PSMD11	0.036848	17	30773381	30773979		
GNAS	0.036848	20	57467617	57467916		
CADM3	0.038332	1	159139108	159139407		
TMEM49	0.04337	17	57916371	57916970		
FGF14	0.045241	13	103052282	103052581		
KRTAP	0.045241	21	31875320	31875619		
OAS1	0.04992	12	113344320	113344754		

Pathway analysis

Pathway enrichment analyses of the 93 DMRs was performed using the Ingenuity software programme to search for known interactions between the genes and to identify pathways enriched amongst the DMRs. The top pathways enriched amongst the DMRs were DNA replication and repair, cell death and survival and cell to cell signalling. Within the DNA replication and repair category were DMRs within Cyclin dependent Kinase 2A (CDKN2A), cyclin D (CDND1), flap structure-specific endonuclease 1 (FEN1), BTB and CNC homology1 gene (BACH1), E3 ubiquitin-protein ligase (SIAH1), proliferating cell nuclear antigen (PCNA), A-kinase anchor protein 8 (AKAP8) (Figure 14B) and SLC6A4 (Figure 14).

Validation of Differentially Methylated Regions by pyrosequencing

To validate the association between the methylation of selected DMRs at birth with percentage fat mass in the children age 6 years, pyrosequencing-based bisulfite PCR analysis was carried out across the DMRs of CDKN2A, FEN1, BACH1 and SLC6A4, on umbilical cord DNA from 247 SWS subjects for which 6 year DXA measurements of adiposity were available. DMRs were selected from the top two pathways enriched amongst the DMRs, namely DNA repair and replication, and gene expression and cell survival, based upon the level of differential methylation between the high and low fat mass groups, and CpG dinucleotide clustering.

The region of CDKN2A identified as a DMR lies within intron 1 of p14^{ARF} and 922 bp upstream of the TSS of the long non-coding RNA ANRIL, which is transcribed in the opposite direction from p14^{ARF} (Figure 1). Figure 1 shows differential methylation at the

CDKN2A loci. Percentage adiposity values for SWS subjects were divided into three groups from low (group 1) to high (group 3) fat mass. The percentage methylation difference between group 1 and group 3 was then compared for each 100 nucleotide region to identify regions of high differential methylation.

An inverse association was observed on the MBD array between the methylation of the CDKN2A DMR at birth and %fat mass age 6 years (Figure 1 and Figure 15A). To validate this association, primers were designed across the 9 CpGs in the identified DMR and methylation analyzed by sodium bisulfite pyrosequencing on an extended sample set of subjects from the SWS cohort (n=247). This revealed that lower % methylation of CDKN2A CpGs 1, 2, 3, 4 and 6 was associated with greater %fat mass at age 6 years (all $p < 0.05$) (Table 6, Figure 3A and Figure 15B). As DXA measurements for the children at birth and 4 years of age were also available, the association between umbilical cord CDKN2A methylation and body composition outcomes of the children at these earlier time points was also examined. The methylation of CDKN2A CpGs 1, 2, 4 and 7 were significantly associated with % fat mass at age 4 years (all $p < 0.05$) (Table 6). There were also associations between the methylation of CDKN2A CpGs 1-9 with total fat mass ages 4 and 6 years (all $p < 0.05$) (Table 6, Figure 3A and Figures 15C & 15D). No associations were observed between CDKN2A methylation and %fat mass or total fat mass at birth (data not shown).

Table 6

CDKN2A	4 yr DXA: Percentage fat			4 yr DXA: total fat			6 yr DXA: Percentage fat			6 yr DXA: total fat		
CpG	n	b	p-value	n	b	p-value	n	b	p-value	n	b	p-value
1	221	-0.088	0.007**	221	-0.004	0.007**	208	-0.004	0.003**	209	-0.005	0.017*
2	215	-0.09	0.007**	215	-0.005	0.003**	204	-0.004	0.003**	205	-0.005	0.011*
3	192	-0.074	0.111	192	-0.005	0.024*	184	-0.005	0.007**	184	-0.007	0.014*
4	247	-0.092	0.007**	247	-0.005	0.001**	231	-0.004	0.005**	232	-0.007	0.001**
5	246	-0.054	0.101	246	-0.004	0.018*	230	-0.002	0.101	231	-0.005	0.022*
6	246	-0.06	0.06	246	-0.004	0.005**	230	-0.004	0.012*	231	-0.007	0.002**
7	244	-0.076	0.04*	244	-0.004	0.015*	228	-0.003	0.042	229	-0.006	0.016*
8	228	-0.057	0.096	228	-0.004	0.017*	212	-0.002	0.236	213	-0.005	0.048*
9	218	-0.074	0.089	218	-0.005	0.011*	203	-0.003	0.122	204	-0.007	0.022*

Pyrosequencing analysis of the DMRs associated with SLC6A4, FEN1 and BACH1 on the larger subset of SWS children also confirmed both the associations and direction of the associations between the methylation of the DMRs and %fat mass of the children age 6 years (Figures 2, 3B, 6, 7B, 5 and 7A, and Figure 16). Of the 5 CpGs measured within the DMR of SLC6A4 (Figure 2), higher methylation of SLC6A4 CpGs 1 and 2 was associated with higher %fat mass at age 6 years ($p < 0.004$) (Table 7, Figure 3B). Figure 2 shows differential methylation at the SLC6A4 loci. Percentage adiposity values for SWS subjects were divided into three groups from low (group 1) to high (group 3) fat mass. The percentage methylation difference between group 1 and group 3 was then compared for each 100 nucleotide region to identify regions of high differential methylation. Figure 3 shows the correlation between DNA methylation for CDKN2A and SLC6A4 and DXA %fat at age 6 years in the SWS cohort. DXA fat measurements plotted against percentage methylation (fourths) illustrating (A) the negative correlation of increasing DXA %fat with increased methylation levels at CDKN2A CpG4. (B) the positive correlation of increasing DXA %fat with increased methylation levels at SLC6A4 CpG1.

A positive association between SLC6A4 CpG 2 methylation was also observed with total fat mass ($p = 0.013$) at 6 years of age (Table 7). There were no associations between SLC6A4 methylation and %fat or total fat mass at birth or 4 years. For BACH1, the methylation of CpGs within the DMR was measured (Figure 5); an inverse association was observed between BACH1 CpG1 methylation and % fat mass at age 6 ($p = 0.038$) and 4 years ($p = 0.029$) (Table 8, Figure 7A); there were trends for inverse correlation between BACH1 CpG4, CpG5 and CpG6 methylation and child's %fat and/or fat mass, and for a positive correlation between BACH1 CpG8 methylation and child's %fat and fat mass (Table 8). For FEN1 the methylation of 4 CpGs within the DMR was measured (Figure 6); of the four CpGs examined within the DMR, CpG1 showed association with % fat mass at 6 years ($p = 0.02$) and CpG 3 methylation was associated with total fat mass age 4 years ($p = 0.045$) (Table 9, Figure 7B). Figure 5 shows differential methylation at the BACH1 gene. Percentage adiposity values for SWS subjects were divided into 3 groups from low (group 1) to high (group 3) fat mass. The percentage methylation difference between group 1 and group 3 was then compared for each 100 nucleotide region across the BACH1 locus. Figure 6 shows differential methylation at the FEN1 gene. Percentage adiposity values for SWS subjects were divided into 3 groups from low (group 1) to high (group 3) fat mass. The percentage methylation difference between group 1 and group 3 was then compared for each 100 nucleotide region across the FEN1 locus. Figure 7 shows the correlation between DXA Fat and DNA methylation for BACH1 and FEN1 in the SWS cohort. DXA %fat measurements plotted against percentage methylation (fourths) illustrating (A) the negative correlation of increasing methylation levels at BACH1 CpG1 with increased DXA %fat at age 6 years. (B) The positive correlation of increasing methylation levels at FEN1 CpG1 fat with increased DXA %fat at age 6 years.

Table 7

SLC6A4	4 yr DXA: Percentage fat			4 yr DXA: total fat			6 yr DXA: Percentage fat			6 yr DXA: total fat		
	n	b	p-value	n	b	p-value	n	b	p-value	n	b	p-value
CpG	204	0.064	0.631	204	0.003	0.618	192	0.016	0.004**	192	0.015	0.061
1	193	0.062	0.617	193	0.001	0.851	185	0.015	0.003**	185	0.018	0.013*
2	181	0.023	0.792	181	-0.001	0.879	176	0	0.953	176	0	0.966
3	179	0.014	0.882	179	0.001	0.803	175	0	0.968	175	0.001	0.867
4	178	-0.091	0.211	178	-0.006	0.123	175	-0.005	0.116	175	-0.009	0.073

Table 8

BACH1	4 yr DXA: Percentage fat			4 yr DXA: total fat			6 yr DXA: Percentage fat			6 yr DXA: total fat		
	n	b	p-value	n	b	p-value	n	b	p-value	n	b	p-value
CpG	267	-0.251	0.029*	267	-0.006	0.289	248	-0.009	0.038*	249	-0.006	0.346
1	274	-0.052	0.756	274	0.003	0.69	252	0.004	0.551	253	0.004	0.726
2	222	-0.169	0.192	222	-0.013	0.052	205	-0.007	0.23	206	-0.011	0.191
4	217	-0.107	0.138	217	-0.007	0.065	201	0	0.982	202	0	0.938
5	213	-0.215	0.074	213	-0.009	0.124	197	-0.008	0.209	198	-0.011	0.227
6	199	-0.032	0.853	199	-0.002	0.826	181	-0.004	0.625	182	0.002	0.87
7	195	0.239	0.114	195	0.012	0.122	178	0.012	0.115	179	0.021	0.055

Table 9

FEN1	4 yr DXA: Percentage fat			4 yr DXA: total fat			6 yr DXA: Percentage fat			6 yr DXA: total fat		
	n	b	p-value	n	b	p-value	n	b	p-value	n	b	p-value
CpG	271	0.222	0.095	271	0.006	0.32	254	0.013	0.02*	255	0.013	0.112
1	271	0.184	0.122	271	0.002	0.701	254	0.003	0.575	255	-0.002	0.805
2	271	0.234	0.045*	271	0.006	0.268	254	0.01	0.056	255	0.01	0.191
3	246	0.062	0.317	246	0.001	0.818	229	0.005	0.062	230	0.005	0.186

Association between CDKN2A methylation and ponderal index in the GUSTO cohort at 18 months and 24 months

Because of the strong associations found between methylation of specific CpGs within the DMR of CDKN2A and later fat mass, we sought to further replicate associations between umbilical cord CDKN2A CpG methylation and adiposity in a second independent cohort where body composition had been measured at age 18 months and 24 months. Lower methylation of CpG1, 3 and 9 was associated with greater infant adiposity assessed by ponderal index at age 18 months (CpG1; $p=0.015$, controlling for infant gender; Figure 8 and CpG9, ponderal index at day 1 $p=0.015$, ponderal index at 18 months $p=0.014$). Figure 8 shows the ponderal index at age 18 months in the GUSTO cohort according to umbilical cord methylation of CDKN2A CpG1 methylation.

Methylation of CDKN2A Differentially Methylated Region of Interest (DMROI) is correlated with ANRIL expression in umbilical cord

The CDKN2A DMR lies within the first intron of p14ARF and less than 1kb upstream of the transcriptional start site of ANRIL, transcribed in the opposite direction from the other genes in this locus. To determine if methylation of the DMR was associated with the expression of either p14ARF or ANRIL, real-time PCR was used to quantify their expression within umbilical cord samples from the GUSTO cohort. Real-time assays were designed that would detect the most commonly expressed splice variants of ANRIL as well as the unspliced form of the ANRIL transcript. Comparison of transcript levels with methylation at the CDKN2A DMR showed a significant inverse association between the methylation of CDKN2A CpGs 1,2, 4-8 with the spliced form of ANRIL ($p < 0.05$) and between CpGs 4-9 with the unspliced variant ($p < 0.02$, Table 10). There was also a positive association between CDKN2A CpG9 methylation and p14ARF expression levels ($p=0.049$) within the umbilical cord.

Table 10

		Δ ct					
		Univariate			Adjusted for Gender		
Transcript	CpG methylation correlated with:	Slope	p	FDR	Slope	p	FDR
ANRIL	CpG1	-1.11	0.016*	0.11	-1.04	0.021*	0.12
ANRIL	CpG2	-0.85	0.05*	0.19	-0.79	0.064*	0.21
ANRIL	CpG4	-1.80	0.023*	0.11	-1.74	0.028*	0.12
ANRIL	CpG5	-1.63	0.024*	0.11	-1.56	0.028*	0.12
ANRIL	CpG6	-1.78	0.025*	0.11	-1.72	0.029*	0.12
ANRIL	CpG7	-1.71	0.019*	0.11	-1.65	0.022*	0.12
ANRIL	CpG8	-2.28	0.001**	0.03	-2.22	0.002**	0.03
ANRIL unspliced	CpG4	-2.83	0.003**	0.04	-2.91	0.002**	0.03
ANRIL unspliced	CpG5	-1.93	0.025*	0.11	-2.00	0.019*	0.12
ANRIL unspliced	CpG6	-3.02	0.001**	0.03	-3.09	0.001**	0.03
ANRIL unspliced	CpG8	-2.07	0.016*	0.11	-2.14	0.011*	0.12
ANRIL unspliced	CpG9	-1.41	0.029*	0.12	-1.53	0.017*	0.12
p14ARF	CpG9	0.77	0.061	0.21	0.80	0.049*	0.18

Regression analysis of methylation level against qPCR Δ ct for all amplicons. Only those tests with nominal p value ≤ 0.061 are shown.

The CDKN2A ROI regulates ANRIL expression

Having shown associations between the methylation of specific CpGs sites within the CDKN2A gene locus at birth and later measurements of adiposity in infancy and childhood, we next investigated whether methylation of these CpG loci may have functional consequences for gene expression. The region of CDKN2A, identified through the MBD array, whose methylation status was associated with later adiposity in both the SWS and GUSTO cohorts, lies within intron 1 of p14ARF (273bp downstream of the 5' splice site) and 922bp upstream from the TSS of the ANRIL (Figure 1). Differential methylation of the DMR therefore has the potential to influence ANRIL and/or p14ARF expression. To determine whether the DMR identified within the CDKN2A locus is important for the expression of ANRIL, the promoter region of ANRIL (-951 to +20) plus or minus the DMR (-1281 to -950) was fused to the luciferase reporter gene in the vector pGL3basic and

transfected into the SW-872 liposarcoma cell line. Deletion of the DMR led to a three-fold increase in ANRIL promoter activity ($p=0.01$) (Figure 4A). Figure 4 is a luciferase assay examining the role of the CDKN2A region in gene expression and EMSA examining protein complex binding within the CDKN2A DMR. (A) Comparison of luciferase activity for the pGL ANRIL plasmid containing the ANRIL promoter region including the ROI in its natural genomic context upstream of the ANRIL TSS, and the pGL 951 plasmid containing just the ANRIL promoter. (B) EMSA with methylated and unmethylated competitors for CpG 4-7 examining binding in liposarcoma cell extract.

The CpG sites within the CDKN2A DMR regulates ANRIL and p14^{ARF} promoter activity

To determine whether the CpG sites associated with later adiposity may play a role in regulating the expression of ANRIL and/or p14^{ARF}, the promoter region of ANRIL (-1281bp to +20bp relative to TSS) and the 5' portion of the p14^{ARF} gene (-500bp to +1125bp relative to TSS) were fused to the reporter gene luciferase in the vector pGL3basic, and CpG sites 1, 2, 3, 4 and 8, mutated (CpG>TpG). Mutation of CpGs 1-4, or CpG 8, led to a decrease in ANRIL promoter activity (all $p\leq 0.008$) in the liposarcoma cell line SW-873. In contrast only mutation of CpG 3 led to a decrease in p14^{ARF} promoter activity in SW-873 cells, while mutation of the other CpG sites had no effect on expression. Mutation of these five CpG sites also led to a decrease in ANRIL promoter activity in SAOS-2 cells (Figure 9B) and in the CpGs examined in A459 and JAR cells. However in A459, SAOS-2 and JAR mutagenesis of the sites CpG sites led to an increase in p14^{ARF} expression (Figure 9A-D).

Methylation of the CpG 4 blocks DNA binding

Having shown that CpG loci associated with later adiposity are important in regulating the level of ANRIL and p14^{ARF} promoter activity, we investigated whether transcription factors may bind to these CpG sequences, and whether DNA methylation affected transcription factor binding. Nuclear extracts from SW-873 liposarcoma cells showed a single complex bound to oligonucleotides containing either CpGs 1 or CpGs 8-9, while two complexes were seen bound to the oligonucleotides containing CpGs 2-3 and 4-7. The higher molecular weight complexes bound to CpG 2-3 and 4-7 seen in SW-873 cells, were not detected in SAOS-2 cells (Figure 9E), suggesting that binding in this region occurs in a tissue specific manner.

To determine whether methylation of the CpG loci within the DMR may affect the binding of these complexes to the DNA, the unmethylated labelled probe containing CpGs 4-7 was incubated with nuclear extracts from the liposarcoma cells with a 50, 100 and 500 -fold excess of either: the unmethylated competitor sequence, or a competitor sequence containing a methylated cytosine at position 4 (CpG4). While binding to the sequence was substantially reduced in the presence of 50

fold excess of unmethylated specific competitor, binding was only reduced in the presence of a 500 fold excess of the methylated competitor (Figure 4B).

Conclusion

We identified over 93 DMRs in umbilical cord samples taken at birth that are associated with adiposity in later life at 6 years of age. The DMRs identified in these studies were found in both the 5' promoter regions of genes as well as within gene bodies. 41 of the 93 DMRs identified from the genome wide analysis were associated with genes which play key roles in DNA replication and repair. In animal models investigating the developmental programming of obesity, persistent alterations to the methylation and expression of cell cycle control genes have been reported in the offspring. Such changes may allow an organism to adjust its developmental programme by regulating cell number, turnover and/or differentiation in response to environmental cues in early life during the process of developmental plasticity.

In particular, we have shown that methylation of CDKN2A, SLC6A4, FEN1 and/or BACH1 at birth is associated with later adiposity at 6 years of age.

Increased methylation of CpGs within the CDKN2A gene locus measured at birth was strongly correlated with lower adiposity in childhood. For CDKN2A, there was also an association seen with total fat mass at 6 years and % fat mass and total fat mass at 4 years. An association between the increased methylation of CpGs within the DMR of CDKN2A and lower ponderal index in children at 18 months was seen in the GUSTO cohort, a second independent mother offspring cohort. A negative association between the methylation of CpG1 to 3 and subscapular skinfold thickness at 24 months was also seen in the GUSTO cohort. Interestingly no associations were seen between the methylation of this region and measures of adiposity at birth in either cohort, suggesting that methylation of these CpG loci may be linked with the gain of fat postnatally, or act as a marker for this increasing adiposity. These results demonstrate that the early life environment is an important determinant of adiposity and may influence future health through the altered epigenetic regulation of gene expression.

A methylation array was used to identify a DMR within intron1 of p14ARF and upstream of ANRIL. This region was shown to be important for the expression of ANRIL, and that its deletion leads to an increase in ANRIL promoter activity, suggesting that the DMR acts to inhibit ANRIL expression.

There was also a strong inverse correlation between the methylation of these CpGs and the expression of ANRIL, supporting the experiments which show that the DMR is important for ANRIL

expression and showing that methylation of CpGs within the DMR act to inhibit ANRIL expression. DMR methylation was also positively associated with p14ARF expression, demonstrating that methylation within the DMR has differential effects on these two genes and, in the case of p14ARF, enhances expression. The differential effects of methylation on ANRIL and p14ARF expression shown by these experiments may reflect the different location of the DMR with respect to the TSS of ANRIL and p14ARF, and show that lower ANRIL expression is associated with lower adiposity.

In liposarcoma cells, mutagenesis of CpG sites 1, 2, 3, 4 or 8 all led to a decrease in ANRIL promoter activity, however only mutation of CpG 3 led to a decrease in p14ARF expression, while mutation of the remaining CpG sites had no effect on p14ARF expression. CpG mutagenesis was also seen to decrease ANRIL promoter activity but increase p14ARF expression in A459, SAOS-2 and JAR cells. The differential effects of CpG mutagenesis within ANRIL and p14^{ARF} may reflect the different location of the CpG sites with respect to the TSS of ANRIL and p14^{ARF}. Interestingly, an inverse correlation was seen between the methylation of CpGs within the CDKN2A DMR and the expression of ANRIL, and a positive correlation between DMR methylation and p14^{ARF} expression in the GUSTO cohort, consistent with the DMR having opposing effects on ANRIL and p14^{ARF}. The CDKN2A DMR may also mediate tissue specific effects, as the effects of CpG mutagenesis was tissue dependent and there were also tissue specific differences in protein binding to this region. Binding, at least in liposarcoma cells, was also affected by methylation. Methylation of CpG4, one of the CpGs most strongly associated with later adiposity, reduced the binding of a specific protein complex to this sequence, suggesting that CpG methylation within the DMR may alter the regulatory capacity of this sequence by disrupting transcription factor binding to this sequence.

The MBD array also identified a number of DMRs associated with later adiposity, including SLC6A4, BACH1 and FEN1, the association of which with later adiposity was also confirmed by pyrosequencing. Interestingly, these are genes which have previously been linked to obesity or adipocyte function. SLC6A4 is a serotonin transporter, and serotonin plays an important role in the control of energy balance and body weight. Depletion of central serotonin results in hyperphagia and obesity, whereas agents that increase serotonin activity in the CNS inhibit food intake and promote weight loss. A promoter polymorphism in SLC6A4 has been associated with eating disorders, obesity and type 2 diabetes mellitus. Methylation of the SLC6A4 promoter is also associated with obesity in monozygotic (MZ) twins. The region identified in this screen was downstream of the differentially methylated region identified in the study that examined MZ twins. The role of the identified region is, at present, unknown and whether methylation affects the binding of a regulatory protein, or it is a marker of gene transcription is yet to be elucidated.

BACH1 encodes a transcription factor that belongs to the cap'n'collar type of basic region leucine zipper factor family. BACH1 plays a key role in adipocyte differentiation; it also regulates heme oxygenase which plays a critical role in inflammation, insulin signalling diabetes and obesity. The FEN1 gene is located between the fatty acid desaturase 1 and 2 (FADS1 and FADS2) genes, on chromosome 11. A recent genome-wide association study for plasma polyunsaturated fatty acids (PUFAs) showed strong evidence for association with this region of chromosome 11, most significant association between the SNP rs174537 (FEN1 10154G>T) near FADS1 and arachidonic acid (AA, 20:4 ω 6). AA is synthesised primarily in the liver and then mobilised to inflammatory cells *via* blood lipoproteins. AA is the precursor of prostaglandins, leukotrienes, and related compounds, all of which have important roles in inflammation. Thus, the FEN1 10154G>T SNP may represent a candidate gene of importance in obesity and cardiovascular disease.

Genome-wide association studies have suggested that genetic changes account for only a small proportion of the variations in fat mass. These findings suggest that the early life environment, through epigenetic mechanisms, may also make an important contribution to variation in adiposity. Without being bound by theory, it is believed that a 10 percent increase in perinatal CDKN2A methylation is associated with a decrease in %fat mass of around 0.20 standard deviations (95% CI -0.33, -0.06) at 4 years of age and 0.18 SD (95% CI -0.31, -0.05) age 6 years, controlling for sex.

The data indicate that epigenetic measures at birth may have prognostic value and suggest that it is possible to detect epigenetic marks in readily available tissues such as cord at birth and use these marks as predictors of later phenotype in more disease relevant tissues.

The association between CDKN2A methylation at birth and later adiposity indicates that such epigenetic measurements provide useful markers of body composition in later life. This is important, as obesity is casually associated with insulin resistance in children and suggests that such epigenetic markers may be of long term clinical significance. CpG loci within the CDKN2A DMR are important for the regulation of the long non-coding RNA ANRIL and p14^{ARF}, indicating that epigenetic processes, as well as genetic alterations, are important determinants of future disease risk within the CDKN2A locus.

EXAMPLE 2CDKN2A and associations with bone mineral content and bone density measurements

The same populations used in Example 1 were also used in Examples 2 and 3.

- Relating the umbilical cord CDKN2A measurements to DXA measurements of bone health at ages 4 and 6 years, there were strong associations between methylation of CDKN2A CpGs 4-9 and bone area and BMC at 4 and 6 years of age.

At ages 4 and 6 years, there were numerous associations between the methylation of CpGs4-9 and bone area and BMC, strongest for CpG9, which was also associated with BMD.

Table 11: Association of CDKN2A CpGs within the identified DMROI with total bone area, total BMC, total BMD, total size-corrected, total area adjusted BMC at ages 4 years and 6 years.

	4 yr DXA: Total bone area (cm sq), without heads, adjusted for sex			4 yr DXA: Total BMC (g), without heads, adjusted for sex			4 yr DXA: Total BMD (g/cm sq), without heads, adjusted for sex		
	n	b	p-value	n	b	p-value	n	b	p-value
CDKN2A CpG1	263	-0.163	0.589	263	-0.145	0.609	263	0	0.798
CDKN2A CpG2	256	-0.272	0.381	256	-0.311	0.286	256	0	0.358
CDKN2A CpG3	231	-0.245	0.558	231	-0.444	0.259	231	0	0.188
CDKN2A CpG4	294	-0.654	0.027	294	-0.636	0.024	294	0	0.1
CDKN2A CpG5	292	-0.658	0.024	292	-0.621	0.025	292	0	0.096
CDKN2A CpG6	292	-0.381	0.177	292	-0.526	0.05	292	0	0.052
CDKN2A CpG7	290	-0.633	0.056	290	-0.61	0.052	290	0	0.157
CDKN2A CpG8	273	-0.621	0.047	273	-0.633	0.032	273	0	0.11
CDKN2A CpG9	263	-1.099	0.005	263	-1.037	0.005	263	-0.001	0.043

	4 yr DXA: Total Size-corrected BMC (g), without heads, adjusted for sex			4 yr DXA: Total Area adjusted BMC (g), without heads, adjusted for sex		
	n	b	p-value	n	b	p-value
CDKN2A CpG1	261	0.11	0.302	263	-0.016	0.906
CDKN2A CpG2	254	0.038	0.733	256	-0.096	0.487
CDKN2A CpG3	230	0.011	0.938	231	-0.251	0.177
CDKN2A CpG4	292	-0.028	0.803	294	-0.12	0.406
CDKN2A CpG5	290	-0.01	0.928	292	-0.102	0.469
CDKN2A CpG6	290	-0.035	0.743	292	-0.226	0.098
CDKN2A CpG7	288	-0.03	0.811	290	-0.111	0.487
CDKN2A CpG8	271	0.011	0.928	273	-0.143	0.348
CDKN2A CpG9	261	-0.047	0.754	263	-0.171	0.363

	6 yr DXA: Total area (cm sq), without heads, adjusted for sex and age			6 yr DXA: Total BMC (g), without heads, adjusted for sex and age			6 yr DXA: Total BMD (g/cm sq), without heads, adjusted for sex and age		
	n	b	p- value	n	b	p- value	n	b	p- value
CDKN2A CpG1	216	0.026	0.954	216	-0.088	0.86	216	0	0.7
CDKN2A CpG2	212	-0.332	0.475	212	-0.518	0.321	212	0	0.26
CDKN2A CpG3	191	0.135	0.832	191	-0.3	0.675	191	0	0.324
CDKN2A CpG4	241	-0.567	0.214	241	-0.909	0.078	241	-0.001	0.067
CDKN2A CpG5	240	-0.833	0.068	240	-1.195	0.02	240	-0.001	0.022
CDKN2A CpG6	240	-0.545	0.216	240	-0.908	0.068	240	-0.001	0.047
CDKN2A CpG7	238	-0.653	0.194	238	-0.986	0.081	238	-0.001	0.08
CDKN2A CpG8	221	-0.703	0.147	221	-0.983	0.07	221	-0.001	0.084
CDKN2A CpG9	212	-1.494	0.015	212	-1.818	0.008	212	-0.001	0.023

	6 yr DXA: Total Size-corrected BMC (g), without heads, adjusted for sex and age			6 yr DXA: Total Area adjusted BMC (g), without heads, adjusted for sex and age		
	n	b	p-value	n	b	p-value
CDKN2A CpG1	213	0.058	0.741	216	-0.117	0.583
CDKN2A CpG2	209	0.018	0.921	212	-0.177	0.425
CDKN2A CpG3	188	-0.075	0.763	191	-0.441	0.149
CDKN2A CpG4	238	0.013	0.944	241	-0.323	0.173
CDKN2A CpG5	237	-0.073	0.699	240	-0.338	0.153
CDKN2A CpG6	237	0.019	0.919	240	-0.346	0.129
CDKN2A CpG7	235	0.008	0.969	238	-0.313	0.229
CDKN2A CpG8	218	0.05	0.8	221	-0.26	0.292
CDKN2A CpG9	209	-0.063	0.804	212	-0.276	0.377

CDKN2A and associations with lean mass

- Relating the umbilical cord CDKN2A measurements to DXA measurements of lean mass at ages 4 and 6 years, there were associations for CDKN2A CpGs 1-9 with %lean mass at ages 4 and 6 years – these are likely to primarily reflect associations of these CpGs with greater fat mass.

At ages 4 and 6 years, there were no associations between cord CDKN2A methylation and total lean mass, apart from a borderline association for CpG6 & 6 year lean mass; there were however numerous positive associations between methylation of CpGs 1-9 and % lean mass at ages 4 & 6.

Table 12: Association of CDKN2A CpGs within the identified DMROI with total lean mass and % lean mass at 4 years and 6 years.

	4 yr DXA: Total lean (g), without heads, adjusted for sex			4 yr DXA: % lean, without heads, adjusted for sex		
	n	B	p-value	n	b	p-value
CDKN2A CpG1	221	3.125	0.723	221	0.086	0.008
CDKN2A CpG2	215	-2.051	0.82	215	0.088	0.008
CDKN2A CpG3	192	-10.066	0.409	192	0.071	0.12
CDKN2A CpG4	247	-8.265	0.363	247	0.091	0.007
CDKN2A CpG5	246	-11.345	0.195	246	0.054	0.097
CDKN2A CpG6	246	-12.245	0.148	246	0.059	0.062
CDKN2A CpG7	244	-5.788	0.555	244	0.076	0.037
CDKN2A CpG8	228	-12.051	0.181	228	0.056	0.096
CDKN2A CpG9	218	-15.929	0.162	218	0.073	0.085

	6 yr DXA: Total lean (g), without heads, adjusted for sex & age			6 yr DXA: % lean, without heads, adjusted for sex & age		
	n	B	p-value	n	b	p-value
CDKN2A CpG1	209	5.076	0.7	208	0.108	0.002
CDKN2A CpG2	205	1.189	0.931	204	0.111	0.003
CDKN2A CpG3	184	-3.8	0.841	184	0.143	0.005
CDKN2A CpG4	232	-18.667	0.166	231	0.107	0.005
CDKN2A CpG5	231	-25.372	0.06	230	0.069	0.074
CDKN2A CpG6	231	-25.55	0.049	230	0.094	0.011
CDKN2A CpG7	229	-19.59	0.187	228	0.088	0.037
CDKN2A CpG8	213	-27.257	0.06	212	0.055	0.19
CDKN2A CpG9	204	-33.49	0.066	203	0.083	0.109

BACH1 and associations with bone mineral content and density

Umbilical cord BACH1 CpG1 methylation was associated with bone area and BMC at age 6 years, with similar associations at age 4 years. CpG3 methylation was associated with bone area at age 4 years, but not age 6 years. A small number of other associations were of borderline significance.

Table 13: Association of BACH1 CpGs within the identified DMROI with total bone area, total BMC, total BMD, total size-corrected, total area adjusted BMC at ages 4 years and 6 years.

	4 yr DXA: Total bone area (cm sq), without heads, adjusted for sex			4 yr DXA: Total BMC (g), without heads, adjusted for sex			4 yr DXA: Total BMD (g/cm sq), without heads, adjusted for sex		
	n	b	p-value	n	b	p-value	n	b	p-value
BACH1 CpG1	315	1.556	0.097	315	1.817	0.041	315	0.001	0.058
BACH1 CpG2	324	0.825	0.553	324	0.715	0.585	324	0	0.758
BACH1 CpG3	323	1.879	0.042	323	1.409	0.107	323	0.001	0.346
BACH1 CpG4	261	-1.364	0.251	261	-1.881	0.09	261	-0.002	0.064
BACH1 CpG5	255	-1.098	0.093	255	-0.76	0.213	255	0	0.501
BACH1 CpG6	251	-0.8	0.477	251	-0.754	0.476	251	-0.001	0.508
BACH1 CpG7	233	-0.177	0.908	233	0.222	0.88	233	0	0.733
BACH1 CpG8	229	-0.265	0.839	229	-0.505	0.689	229	0	0.686
	4 yr DXA: Total Size-corrected BMC (g), without heads, adjusted for sex			4 yr DXA: Total Area adjusted BMC (g), without heads, adjusted for sex					
	n	b	p-value	n	b	p-value			

BACH1 CpG1	313	0.356	0.309	315	0.59	0.187
BACH1 CpG2	322	-0.289	0.574	324	0.065	0.922
BACH1 CpG3	321	-0.239	0.493	323	-0.072	0.872
BACH1 CpG4	259	-0.714	0.098	261	-0.805	0.137
BACH1 CpG5	253	0.037	0.874	255	0.105	0.724
BACH1 CpG6	249	-0.146	0.722	251	-0.123	0.812
BACH1 CpG7	231	0.359	0.537	233	0.361	0.629
BACH1 CpG8	227	-0.934	0.062	229	-0.296	0.646

	6 yr DXA: Total area (cm sq), without heads, adjusted for sex and age			6 yr DXA: Total BMC (g), without heads, adjusted for sex and age			6 yr DXA: Total BMD (g/cm sq), without heads, adjusted for sex and age		
	n	b	p- value	n	b	p- value	n	b	p- value
BACH1 CpG1	258	3.624	0.01	258	3.564	0.025	258	0.002	0.111
BACH1 CpG2	262	0.894	0.692	262	0.04	0.988	262	-0.001	0.671
BACH1 CpG3	261	0.774	0.594	261	1.145	0.486	261	0.001	0.426
BACH1 CpG4	215	-1.484	0.418	215	-2.648	0.19	215	-0.002	0.148
BACH1 CpG5	211	-0.431	0.672	211	-0.115	0.92	211	0	0.888
BACH1 CpG6	207	-1.707	0.343	207	-2.388	0.235	207	-0.001	0.261
BACH1 CpG7	191	1.332	0.598	191	3.267	0.252	191	0.003	0.135
BACH1 CpG8	188	3.997	0.074	188	2.947	0.245	188	0.001	0.758

	6 yr DXA: Total Size-corrected BMC (g), without heads, adjusted for sex and age			6 yr DXA: Total Area adjusted BMC (g), without heads, adjusted for sex and age			
	n	b	p-value	n	b	p-value	
BACH1 CpG1	255	0.15	0.795	258	-0.168	0.814	
BACH1 CpG2	260	-0.586	0.521	262	-0.895	0.431	
BACH1 CpG3	259	-0.514	0.381	261	0.354	0.629	
BACH1 CpG4	212	-0.581	0.433	215	-1.105	0.223	
BACH1 CpG5	208	0.256	0.531	211	0.329	0.518	
BACH1 CpG6	204	-0.416	0.578	207	-0.618	0.495	
BACH1 CpG7	190	1.742	0.096	191	1.888	0.147	
BACH1 CpG8	187	-1.036	0.267	188	-1.181	0.308	

BACH1 and associations with lean mass

Umbilical cord BACH1 CpG1 methylation was associated with total lean mass and % lean mass at both ages 4 and 6 years.

Table 14: Association of BACH1 CpGs within the identified DMROI with total lean mass and % lean mass at 4 years and 6 years.

	4 yr DXA: Total lean (g), without heads, adjusted for sex			4 yr DXA: Percentage lean, without heads, adjusted for sex		
	n	b	p-value	n	b	p-value
BACH1 CpG1	267	68.112	0.025	267	0.245	0.03
BACH1 CpG2	274	43.779	0.317	274	0.053	0.744
BACH1 CpG3	274	10.585	0.713	274	-0.1	0.35
BACH1 CpG4	222	-40.348	0.251	222	0.174	0.171
BACH1 CpG5	217	-16.007	0.408	217	0.108	0.128
BACH1 CpG6	213	14.37	0.657	213	0.219	0.064
BACH1 CpG7	199	0.268	0.995	199	0.031	0.854
BACH1 CpG8	195	-6.566	0.87	195	-0.225	0.131
	6 yr DXA: Total lean (g), without heads, adjusted for sex and age			6 yr DXA: Percentage lean, without heads, adjusted for sex and age		
	n	B	p-value	n	b	p-value
BACH1 CpG1	249	93.605	0.026	248	0.254	0.03

BACH1 CpG2	253	-13.758	0.84	252	-0.105	0.582
BACH1 CpG3	252	54.497	0.217	251	-0.004	0.971
BACH1 CpG4	206	-42.821	0.431	205	0.163	0.274
BACH1 CpG5	202	5.463	0.856	201	-0.013	0.877
BACH1 CpG6	198	-9.762	0.862	197	0.186	0.229
BACH1 CpG7	182	71.265	0.37	181	0.1	0.666
BACH1 CpG8	179	78.958	0.251	178	-0.288	0.156

EXAMPLE 3

Introduction

A recognised cause of increasing systolic blood pressure with age is aortic stiffness, which is a significant predictor of future cardiovascular risk and all-cause mortality. Pulse wave velocity (PWV) is an indirect measure of vascular stiffness, with greater PWV also a recognised marker of cardiovascular risk. Such risk is recognised to be set in part by early developmental factors such as unbalanced prenatal nutrition although the use of PWV to assess such risk has been little utilised.

Rationale

Pulse wave velocity (PWV) is an indirect measure of vascular stiffness and higher PWV is an established cardiovascular risk marker.

Objective

To examine the association between umbilical cord methylation status and vascular structure in the child at age 9 years using magnetic resonance imaging (MRI) measurements of aortic stiffness.

Method

Magnetic resonance imaging (MRI) measurement of aortic PWV was performed on 234 children aged 9 years who were participants in the Southampton Women's Survey (SWS), a prospective study of developmental influences on later health and disease. Ethical approval was obtained from the local research ethics committee. The participants and parent/guardian

provided informed assent and consent respectively. MRI compatibility of the child and parent/guardian was determined as per local rules.

Aortic stiffness was assessed in the descending aorta. Flow measurements were made in the plane perpendicular to the long axis of the aorta on sagittal and coronal images, both in the proximal descending aorta at the level of the pulmonary trunk, and in the distal descending aorta above the bifurcation. A phase contrast flow mapping sequence was employed with free breathing and retrospective ECG gating. A velocity encoding gradient was applied in the through plane direction with a VENC of 150-200 cm/s. Images were acquired at 30 phases throughout the cardiac cycle. Velocity flow curves were generated using open source imaging software (Osirix) and PWV calculated in m/s using Matlab software (The Mathworks, Inc., Natick, MA) using the transit time method from $\Delta d / \Delta t$ where Δd is the distance between the two flow acquisition sites in the central aorta, and Δt is the transit time between the arrival of the systolic wave front at the two sites.

Right Brachial Blood Pressure was measured and recorded immediately following the flow sequence acquisitions with a MR compatible patient monitor (in vivo) using a paediatric cuff.

Results

PWV values were within previously reported childhood ranges. Umbilical cord CDKN2A CpG5 methylation and SLC6A4 CpG1 and CpG5 methylation were positively associated with MRI aortic PWV in childhood ($\beta=0.00663$, $p=0.037$; $\beta=0.0301$, $p=0.055$; and $\beta=0.0184$, $p=0.095$, respectively); BACH1 CpG6 methylation was inversely with childhood aortic PWV ($\beta=-0.0295$, $p=0.051$).

Conclusions

Variations in epigenetic state may alter vascular development in utero, changing arterial structure with long-term consequences for cardiovascular risk in later life.

EXAMPLE 4

Association of SLC6A4 DNA methylation in perinatal tissues with adiposity in early childhood and with obesity in adults.

SLC6A4 and associations with adiposity in early childhood

Materials and Methods

Southampton Women's Survey (SWS): Participants

Table 16 outlines the participant characteristics.

Body composition measurements in the offspring

At age 4 and 6 years a sub-set of children from the SWS study were assessed for body composition by dual-energy X-ray absorptiometry (Hologic Discovery, paediatric scan mode, Hologic Inc., Bedford, MA). At birth, 12 months, 2, 3, and 6 years of age the children were invited for triceps skinfold thickness measurements.

Instruments were calibrated daily; coefficients of variation were <1 %. Follow-up of the children and sample collection/analysis was carried out under Institutional Review Board approval (Southampton and SW Hampshire Research Ethics Committee) with written informed consent. Clinical investigations were conducted according to the principles expressed in the Declaration of Helsinki.

Skin fold measurement method

To access subcutaneous fat deposition as a measure of adiposity, skinfold thicknesses at standard sites were measured using a skinfold calliper (Holtain) (Inskip, 2006). Measurements were repeated three times and the mean used in analyses.

DNA methylation analysis by Pyrosequencing

Genomic DNA was prepared from umbilical cord and cord blood by a standard high salt method, and from adipose tissue using the QIAamp DNA mini kit (Qiagen, Germany). Bisulphite conversion was carried out on 1µg DNA (500 ng for adipose) using the EZ DNA Methylation-Gold kit (Zymo Research, USA) according to manufacturer's instructions. The Pyrosequencing reaction was carried out using primers listed in the following Table 15.

Table 15

Primer	Sequence 5'→3'	Amplicon Length	No. CpG's	SEQ ID NO.
SLC6A4 CpG 3-5 Forward	GTTTTTAGGAGAGTAGGGAGTATATAGTTT	151 bp	3	42
SLC6A4 CpG 3-5 Reverse	Biotin-ACCTACTTCTCACTCATCCATATTA			43
SLC6A4 CpG 3-5 Sequencing	GTGTGAAGTTATTGAGG			44

5' Biotin labelled primers are indicated. Amplicon length and number of CpG's within PCR amplicon are shown.

Bisulfite modified DNA was amplified using 1.25 U HotStar Taq Plus DNA polymerase (Qiagen). Pyrosequencing was carried out using the PyroMark Gold Q96 Reagent kit (Qiagen) on a PyroMark Q96 MD machine (Biotage, Sweden) according to manufacturer's instructions. % methylation was calculated using the Pyro Q CpG software (Biotage).

RT-PCR

Adipose tissue was processed using the RNeasy Lipid Tissue kit (Qiagen) according to manufacturer's instructions. 500 ng total RNA was used as a template for cDNA synthesis and incubated with 10 µm random nonamers (Sigma-Aldrich, USA), 0.5 µM dNTP's and 200 U M-MLV Reverse Transcriptase (Promega, UK) according to manufacturer's instructions. cDNA was amplified for quantification using 5µl QuantiTect SYBR Green PCR Mix (Qiagen), 4µl cDNA and 1µl SLC6A4 primer (Qiagen). Real time PCR was performed using the Roche LightCycler 480 Real Time PCR Detection System (Roche, Switzerland). Cycling conditions were as follows; initial denaturation of 94°C 15 minutes, followed by 40 cycles of 94°C 15s, 55°C 30s and 72°C 30sec.

Statistical methods

Statistical analysis was carried out using Stata (Statacorp) versions 13.1. Histograms of all continuous variables were plotted to check for normality. The distributions of some measures of adiposity and skinfold thickness were observed to be skewed and were therefore transformed using an appropriate transformation in order to satisfy the assumption of normality. DXA adiposity measurements were calculated without including the child's head in order to minimise the effect of head movement on the measurement. Regression models were built using either child's adiposity measurement (at either birth, 4 or 6 years), or child's

skinfold thickness (at either birth, 6 months, 12 months, 2, 3 or 6 years) as the outcome and CpG methylation as the predictor. Models were adjusted for child's sex and age at each time point (and additionally for gestational age at birth). Adiposity was measured as %fat and total fat in grams, converted to Z-scores to facilitate interpretation of the effect size, and methylation measurements as %; the regression coefficient can therefore be interpreted as the standard deviation change in % fat (or total fat) for each % change in methylation. The skinfold thickness measurements used at each age were calculated using the mean of three skinfold measures taken by trained research nurses and were measured in millimetres. The regression coefficients can therefore be interpreted as the change in (transformed) skinfold thickness (mm) for each percentage change in methylation. All results are presented as regression coefficients (β) for each 1% change in methylation with their associated P-values.

Statistical comparisons of gene expression, or CpG methylation with the lean and obese groups (based on BMI) from the BIOCLAIMS cohort was by Logistical regression, adjusting for age and sex.

Results

Cohort Characteristics

Genomic DNA was extracted from the umbilical cord of SWS cohort subjects (n=680) who had DXA (total and % fat mass) measurements at birth, 4 and 6 years of age and skinfold thickness measurements at birth, 6 months, 12 months, 2 years, 3 years and 6 years. The subjects had a median birth weight of 3.48 kg, median gestational age of 40.4 weeks, and 50% were female. The median maternal age at birth was 31.51 years with a pre-pregnancy median body mass index of 24.27. 48% were in their first pregnancy and 12% smoked in late pregnancy. 289 of these children also had DNA from cord blood available. These 289 children had a similar birth weight distribution to the full cohort with a median birth weight of 3.44kg, median gestational age of 40.06 with 47% being female. The median maternal age at birth was 30.67 years. They had a pre-pregnancy median BMI of 24.28 and 13% were smokers. DXA and skinfold thickness measurements showed a similar range in each group as illustrated by Table 16.

Table 16

Characteristic	% or Median (5th,95th percentile) for SWS 800 Cohort (Umbilical Cord n=680)	% or Median (5th, 95th percentile) for SWS 400 Cohort (Cord Blood n=289)
Educational qualifications (%):		
None	2.07 %	0.35 %
CSE	8.71 %	8.65 %
O levels	26.44 %	26.99 %
A levels	31.91 %	27.34 %
HND	7.83 %	8.30 %
Degree	23.04 %	28.37 %
Social class (%):		
Professional	6.01 %	7.02 %
Management and technical	40.09 %	43.16 %
Skilled non-manual	35.14 %	32.98 %
Skilled manual	6.76 %	7.02 %
Partly skilled	10.81 %	9.12 %
Unskilled	1.20 %	0.70 %
Age at birth, years	31.51 (24.6, 36.5)	30.67 (23.6, 35.5)
Smoking (during pregnancy)	12.98 %	13.19 %
BMI	24.27 (19.6, 34.8)	24.28 (19.6, 34.8)
<u>Child</u>		
Female (%)	50.15 %	47.40 %
Birth order (%)		
1st	47.94 %	48.10 %
2nd	38.24 %	37.37 %
3rd or higher	13.82 %	14.53 %
Birth weight, kg	3.48 (2.7, 4.4)	3.46 (2.7, 4.2)
Gestational age, weeks	40.14 (37.0, 41.9)	40.06 (37.0,41.9)
Child Fat mass Measurements:		
Infant total fat at birth, g	510.3 (264, 979)	508.5 (277, 923)
Infant total fat at 4 years, kg	4.11 (2.8, 6.7)	4.05 (2.8, 6.7)
Infant total fat at 6 years, kg	4.88 (3.0, 8.8)	4.85 (3.0, 8.4)
Infant percentage fat at birth	14.39 (8.8, 22.5)	14.37 (9.8, 21.7)
Infant percentage fat at 4 years	28.89 (21.7, 39.1)	28.38 (21.7, 37.7)
Infant percentage fat at 6 years	25.19 (17.7, 35.8)	24.69 (17.8, 34.7)
Child Skinfold Thickness Measurements:		
Infant triceps skinfold at birth, mm	4.67 (3.4, 6.3)	4.67 (3.6, 6.3)
Infant triceps skinfold at 6 months, mm	11.00 (7.8, 15.0)	10.97 (7.9, 15.1)
Infant triceps skinfold at 12 months, mm	10.92 (7.5, 15.2)	11.02 (7.0, 14.9)
Infant triceps skinfold at 2 years, mm	10.00 (6.9, 13.7)	10.10 (7.1, 14.0)
Infant triceps skinfold at 3 years, mm	9.95 (6.8, 13.9)	9.80 (6.8, 13.5)
Infant triceps skinfold at 6 years, mm	9.55 (6.3, 16.8)	9.87 (6.2, 16.8)

Perinatal SLC6A4 CpG Methylation is associated with measures of childhood adiposity

In umbilical cord, SLC6A4 methylation of the 5 measured CpG's ranged from 79.54 % for CpG 5 to 87.49 % for CpG 4, with the 5th-95th centiles of methylation between subjects ranging from 13.74% for CpG 4 to 17.59% for CpG 5. This is illustrated in Table 17.

Table 17

SLC6A4 CpG	Median (5th-95th Percentile)		
	SWS Umbilical Cord	SWS Cord Blood	BIOCLAIMS Adipose
3	84.47 (75.09-89.44)	76.18 (69.01-83.05)	84.54 (77.35-90.69)
4	87.49 (78.13-91.87)	75.97 (66.82-81.21)	87.38 (81.02-91.65)
5	79.54 (68.67-86.26)	54.16 (43.66-62.54)	70.17 (60.58-76.87)

Table 17 displays observed DNA methylation ranges for SLC6A4 CpG's within the different cohorts as measured by Pyrosequencing. Data is shown for the SWS cohorts (umbilical cord and cord blood) and the BIOCLAIMS cohort (adipose).

There was a negative association between SLC6A4 CpG5 methylation with % fat mass ($\beta=-0.005$ $p=0.013$) and total fat mass ($\beta=-0.012$ $p=0.0003$) at 6 years and with % fat mass ($\beta=-0.003$ $p=0.044$) and total fat mass ($\beta=-0.007$ $p=0.007$) at 4 years. This is outlined in Table 18 and illustrated in Figure 10. No associations were found between the methylation of CpG5 with total or % fat mass at birth.

To determine if the methylation of these CpG loci within SLC6A4 in cord tissue were associated with additional measures of adiposity in childhood, we investigated their relationship with triceps skinfold thickness in the children at birth through to 6 years. Here, higher methylation of CpG4 was associated with lower triceps skinfold thickness at birth ($\beta=-0.00$ $p=0.023$) but there was no association at later ages. Higher methylation of CpG 5 was significantly associated with lower triceps skinfold thickness at birth ($\beta=-0.003$ $p=0.013$), at 6 months ($\beta=-0.005$ $p=0.041$), at 2 years ($\beta=-0.006$ $p=0.0004$), at 3 years ($\beta=-0.007$ $p=0.0003$) and at 6 years ($\beta=-0.0001$ $p=0.013$) (Table 19 and Figure 11).

SLC6A4 CpG Methylation in cord blood is associated with measures of later adiposity in the SWS Cohort

The inventors sought to determine if an association would be found between SLC6A4 methylation and later adiposity in an alternative perinatal tissue such as cord blood. Pyrosequencing analysis of SLC6A4 CpGs 3 to 5 was performed on a subset of subjects from the SWS cohort where DNA from cord blood and measures of adiposity were available (n=289).

In cord blood, the median methylation ranged from 54.16 % for CpG 5 to 76.18 % for CpG 3. The 5th-95th centiles of methylation between subjects ranged from 14.04 % for CpG 3 to 18.88 % for CpG 5 (Table 20). The inventors also found associations between SLC6A4 methylation and measures of adiposity in cold blood. Cord blood methylation of CpG5 was negatively associated with total fat mass at age 4 years ($\beta=-0.003$, $p=0.037$) (Table 21), whilst weaker associations were also found between methylation of CpG 5 with total fat mass ($\beta=-0.004$, $p=0.097$) and % fat mass ($\beta=-0.003$, $p=0.08$) at birth. There were also associations between SLC6A4 methylation and skinfold thickness. Methylation of CpG3 was negatively associated with triceps skinfold thickness at birth ($\beta=-0.0064$, $p=0.004$), but positively associated at 2 years ($\beta=0.0645$, $p=0.034$) and 3 years ($\beta=0.007$, $p=0.036$). CpG4 ($\beta=-0.0043$, $p=0.045$) and CpG5 ($\beta=-0.0042$, $p=0.013$) were also negatively associated with triceps skinfold thickness at birth and CpG 5 weakly associated with triceps skinfold thickness at 6 years ($\beta=0.006$, $p=0.049$) (Table 22). For CpG 5, associations with triceps skinfold thickness at birth and total fat mass at 4 years were identified in both umbilical cord and cord blood (Figure 12A, 11B).

SWS 800	Ln transformed 6 year DXA: total fat, adjusted for sex and age			Ln transformed 6 year DXA: percentage fat, adjusted for sex and age			Ln transformed 4 year DXA: total fat, adjusted for sex			Ln transformed 4 year DXA: percentage fat, adjusted for sex			Ln transformed baby DXA: total fat, adjusted for sex, age, and gestational age			Ln transformed baby DXA: percentage fat, adjusted for sex, age, and gestational age		
	n	β	P-value	n	β	P-value	n	β	P-value	n	β	P-value	n	β	P-value	n	β	P-value
CpG 3	423	-0.002	0.659	421	-0.001	0.721	369	0.001	0.806	369	0.000	0.971	302	-0.009	0.086	302	-0.005	0.141
CpG 4 (FY)	421	0.002	0.924	419	0.005	0.781	366	0.005	0.787	366	0.006	0.661	299	-0.041	0.211	299	-0.018	0.46
CpG 5	420	-0.012	0.000**	418	-0.005	0.013*	363	-0.007	0.007**	363	-0.003	0.044*	299	-0.007	0.072	299	-0.004	0.177

Table 18: Associations between umbilical cord SLC6A4 CpG methylation levels with child's fatmass in the SWS cohort. Associations between percentage/total fatmass at birth, 4 & 6 years. * p 0.01-0.05, ** p $\leq$ 0.01

SWS 800	Ln transformed Triceps skinfold (mm) at birth, adjusted for gestational age, sex and age			Square root transformed Infant's triceps skinfold (mm) at 6 months, adjusted for age and sex			Infant's triceps skinfold (mm) at 12 months, adjusted for sex and age			Ln transformed child's triceps skinfold (mm) at 2 years, adjusted for sex and age			Ln transformed child's triceps skinfold (mm) at 3 years, adjusted for sex and age			Inverse square root transformed child's triceps skinfold (mm) at 6 years, adjusted for sex and age		
	n	β	P-value	n	β	P-value	n	β	P-value	n	β	P-value	n	β	P-value	n	β	P-value
CpG 3	637	-0.001	0.77	592	0.005	0.107	581	-0.005	0.831	566	-0.001	0.613	418	0.002	0.407	284	0.000	0.733
CpG 4	633	-0.005	0.023*	588	0.001	0.79	577	0.011	0.711	563	-0.004	0.197	415	-0.003	0.344	282	0.000	0.827
CpG 5	628	-0.003	0.013*	584	-0.005	0.041*	572	-0.033	0.062	557	-0.006	0.000**	413	-0.007	0.000**	281	0.001	0.013*

Table 19: Associations between umbilical cord CpG SLC6A4 methylation levels with child's skinfold thickness in the SWS cohort. Associations between triceps skinfold thickness at birth, 6 months, 12 months, 2, 3 & 6 years. * p 0.01-0.05, ** p $\leq$ 0.01

	Median (5th-95th Percentile)		
SLC6A4 CpG	SWS Umbilical Cord	SWS Cord Blood	BIOCLAIMS Adipose
3	84.47 (75.09-89.44)	76.18 (69.01-83.05)	84.54 (77.35-90.69)
4	87.49 (78.13-91.87)	75.97 (66.82-81.21)	87.38 (81.02-91.65)
5	79.54 (68.67-86.26)	54.16 (43.66-62.54)	70.17 (60.58-76.87)

Table 20: Observed DNA methylation ranges for SLC6A4 CpG's within the different cohorts as measured by Pyrosequencing. Data is shown for the SWS cohorts (umbilical cord and cord blood) and the BIOCLAIMS cohort (adipose).

SWS Cord Blood	Log transformed 6 year DXA: Total fat (g), without heads, adjusted for sex and age			Log transformed 6 year DXA: Percentage fat, without heads, adjusted for sex and age			Log transformed 4 year DXA: Total fat (g), without heads, adjusted for sex			4 year DXA: Percentage fat, without heads, adjusted for sex			Log transformed Baby DXA: Total fat (g), adjusted for gestational age, sex and age			Log transformed Baby DXA: Percentage fat, adjusted for gestational age, sex and age		
	n	b	p-value	n	b	p-value	n	b	p-value	n	b	p-value	n	b	p-value	n	b	p-value
3	197	0.001	0.679	197	0.001	0.529	212	-0.003	0.157	212	-0.034	0.702	111	-0.003	0.356	111	-0.002	0.51
4	196	0.003	0.281	196	0.002	0.213	211	-0.003	0.064	211	-0.095	0.251	111	-0.003	0.345	111	-0.003	0.265
5	196	0.001	0.504	196	0.001	0.326	211	-0.003	0.037*	211	-0.066	0.292	111	-0.004	0.097	111	-0.003	0.08

Table 21. Associations between cord blood SLC6A4 CpG methylation levels with child's fatmass in the SWS cohort. Associations between percentage/total fatmass at birth, 4 & 6 years. * p 0.01-0.05, ** p ≤ 0.01

SWS Cord Blood	Ln transformed Triceps skinfold (mm) at birth, adjusted for gestational age, sex and age			Infant's triceps skinfold (mm) at 6 months, adjusted for age and sex			Infant's triceps skinfold (mm) at 12 months, adjusted for sex and age			Child's triceps skinfold (mm) at 2 years, adjusted for sex and age			Ln transformed child's triceps skinfold (mm) at 3 years, adjusted for sex and age			Ln transformed child's triceps skinfold (mm) at 6 years, adjusted for sex and age		
	n	β	P-value	n	β	P-value	n	β	P-value	n	β	P-value	n	β	P-value	n	β	P-value
3	261	-0.0064	0.004**	260	0.0092	0.742	256	0.0075	0.822	240	0.0645	0.034*	250	0.007	0.036*	249	0.003	0.483
4	260	-0.0043	0.045*	259	0.0179	0.505	255	0.02	0.535	239	0.0455	0.112	249	0.004	0.2	248	0.005	0.213
5	260	-0.0042	0.013*	259	0.0037	0.862	255	0.0083	0.744	239	0.0135	0.552	249	7E-04	0.758	248	0.006	0.049*

Table 22. Associations between cord blood SLC6A4 CpG methylation levels with child's skinfold thickness in the SWS cohort.
Associations between triceps skinfold thickness at birth, 6 months, 12 months, 2, 3 & 6 years. * p 0.01-0.05, ** p $\leq$ 0.01

Lower SLC6A4 methylation is associated with obesity in adipose tissue from adults.

SLC6A4 is known to function not just on the appetite control circuitry within the CNS but also has been shown to have an important regulatory role in adipose tissue. The current inventors investigated whether methylation of the CpG loci within SLC6A4 were altered in adipose tissue of obese individuals compared to lean.

Genomic DNA was extracted from the adipose tissue of BIOCLAIMS subjects (n=65) who had BMI measurements. The subjects had a median BMI of 31.74, median height of 1.65 cm and median weight of 82.8 Kg. 76.92 % of participants were female. Subjects were determined as lean or obese according to their BMI: lean participants had a BMI between 18.5 and 25, whilst obese participants had a BMI from 30-40 (Table 23).

Table 23

Characteristic	% or Median (5th,95th percentile) for BIOCLAIMS Cohort (Adipose tissue) n=65
Female	76.92%
Age at visit	38.5 (19.53 to 61.19)
Lean	35.38%
Measurements:	
Weight (kg)	82.80 (49.52 to 124.37)
Height (M)	1.65 (1.53 to 1.85)
BMI	31.74 (19.50 to 39.91)
Waist, cm	94.40 (38.54 to 127.10)
Hip, cm	104.00 (85.06 to 133.42)
Percentage fat	34.00 (15.42 to 49.56)
Fatmass, kg	32.10 (9.68 to 55.40)
Total body water, kg	36.20 (27.08 to 59.54)
Lean mass, kg	49.50 (37.00 to 81.34)

Pyrosequencing analysis of the 3 CpG's within the SLC6A4 DMROI in adipose tissue of BIOCLAIMS subjects revealed that median methylation ranged from 70.17 % for CpG 5 to 87.38 % for CpG 4. Values for the spread (5th-95th centiles) of methylation between subjects

ranged from 10.63 % for CpG 4 to 16.29 % for CpG 5 (Table 20). Methylation of CpG 5 ($p=0.042$) was significantly reduced in adipose tissue from obese individuals compared to lean individuals (Table 24) (Figure 12A). There were no significant differences for CpG 3 or 4. The inventors performed real time PCR, in order to ascertain if the alteration in methylation in adipose tissue was accompanied by an alteration in gene expression of SLC6A4. Analysis of SLC6A4 expression by real time PCR demonstrated that its expression was significantly reduced in the obese individuals compared to lean ($p=0.0168$, $n=63$) (Table 24) (Figure 13B).

Table 24

BIOCLAIMS	P-value	Exp(B)	95% C.I. for EXP(B)	
			Lower	Upper
SLC6A4 Expression	0.026*	0.123	0.020	0.779
CpG 3	0.38	1.084	0.905	1.298
CpG 4	0.228	0.869	0.691	1.092
CpG 5	0.042*	0.568	0.330	0.979

Table 24 illustrates the associations between SLC6A4 gene expression or SLC6A4 CpG methylation and Obesity in the BIOCLAIMS cohort. * p 0.01-0.05, ** $p \leq 0.01$

Discussion

The above results demonstrate that altered methylation of specific CpGs 1, 2, 4 and 5 within the gene body of SLC6A4 in perinatal tissue is associated with several measures of adiposity in early childhood. These epigenetic marks are not only present in both umbilical cord and cord blood, but also present in adipose tissue from obese adults, alongside alterations in gene expression.

SLC6A4 is a serotonin transporter known to play an important role in the control of energy balance and appetite, and the current data indicates that SLC6A4 methylation in perinatal tissues may be an epigenetic biomarker for obesity risk. Furthermore, given that the biomarker is present in adipose tissue of a separate adult cohort with gene expression changes, it may also contribute to the development of adiposity.

The current inventors have shown that for a specific CpGs 1, 2, 4 and 5 within the first intron of the SLC6A4 gene, originally identified from an MBD array, that lower methylation in umbilical cord is associated with increased fat mass and triceps skinfold thickness at several time points during early childhood. This association was also found in cord blood. Next,

using adipose tissue from a separate adult cohort, the inventors also showed that SLC6A4 is hypomethylated at the CpG of interest, CpG5, in obese subjects compared to lean.

The inventors also confirmed reduced expression of SLC6A4 in adipose tissue in obese subjects. More specifically, the inventors have shown that SLC6A4 is expressed in human adipose tissue, and is down regulated in obese individuals compared to lean.

The inventors speculate that the downstream effects may result in increased peripheral serotonin availability in the obese state. During development, before its role as a neurotransmitter, serotonin has a trophic role. Therefore any alterations in its signalling in early life may have an impact on brain development. It is therefore plausible that alterations in central serotonin availability induced by altered SLC6A4 methylation in early life may induce programming which has long term effects on energy balance. Two distinct serotonin systems exist, one which functions in the CNS and has well known effects on mood, sleep and energy balance and one which functions in the periphery. Interestingly, 98% of serotonin is synthesised in the periphery. Peripheral serotonin is thought to act as a hormone involved in intestinal motility, the immune system and vasoconstriction. As the central serotonin system is involved in energy regulation, it is quite plausible that the peripheral serotonergic system has effects on energy balance also, for example by acting on adipocytes.

Most studies looking at the methylation of specific genes in relation to obesity measure methylation and obesity at the same time, i.e. after the onset of obesity (HIF ref). In these situations, it is consequently unknown if altered DNA methylation is a cause or consequence of the obesity. It is therefore very important to distinguish between those modifications which are a result of altered bodyweight and those which may be potential biomarkers, occurring prior to the onset of obesity. For this reason, studies which look in tissue at early stages of development are of utmost importance. The inventors is the first study to define biomarkers of fat mass in perinatal tissue and replicate this result in the adipose tissue of adults with associated alterations in gene expression. The inventors have shown that the epigenetic biomarker, i.e. altered SLC6A4 methylation, is present before the onset of obesity, is present in several perinatal tissues, and also found in adipose tissue of an adult cohort alongside alterations in gene expression. Therefore, not only is this epigenetics marker predictive of fat mass and measures of obesity thus a potential biomarker, it is also highly conceivable that it's altered methylation may have functional consequences on energy balance, making it an appealing target for intervention.

EXAMPLE 5

Expression of transcripts from the CDKN2A locus is also associated with measures of infant adiposity in the GUSTO cohort

Materials and Methods

Cohort

The GUSTO participants used in Example 1 and detailed in Table 3 were also used in Example 5

RNA Extraction

Umbilical cord tissue samples were collected at the time of delivery, flushed with saline to remove fetal blood and flash-frozen in liquid nitrogen within 30 min of collection. Umbilical cord tissue (300 mg) was first placed in a sterile Dispomix tube and homogenized for 55 s for 3 cycles in 3 ml of Trizol using the Dispomix (Medic Tools, AG, Zug, Switzerland). After spinning down the debris, the supernatants were divided equally into three 2 ml tubes. 200 µl of chloroform were added to each tube, vortexed vigorously and centrifuged for 15 min at 4°C. The aqueous phase was carefully transferred to a new tube containing 1 µl of linear acrylamide. An equal amount of isopropanol was added and mixed by inversion. After incubating at –20°C overnight to precipitate the RNA, the pellet was obtained by centrifuging at 13,200 rpm for 10 min at 4°C. The RNA pellet was washed twice in 70% (v/v) ethanol, air-dried and resuspended in RNase-free water. The isolated RNA was then purified using the RNeasy Mini Kit (Qiagen, Hilden, Germany). On-column DNase digestion was carried out before the first wash step according to the manufacturer's instructions. The purified RNA was then eluted in 30 µl of RNase-free water and stored at –80°C. RNA concentration and purity were measured using a nanodrop ND-8000 spectrophotometer (Nanodrop Technologies, Wilmington, DE, USA), and RNA integrity was determined using the Agilent 2100 Bioanalyzer and RNA 6000 Nano Labchips (Agilent Technologies, Santa Clara, CA, USA). RNA was reverse transcribed using a High Capacity cDNA Reverse Transcription Kit (Applied Biosystems Inc, ABI, CA, USA).

PCR

PCR reactions were prepared using 10 µl of Power SyBr Green PCR 2× Master mix (Applied Biosystems Inc, ABI, Foster City, CA, USA), 1 µl of each primer (2 µM), and 20 ng of cDNA in a total reaction volume of 20 µl. PCR for each sample was done in triplicates in 384-well

plates using the ABI 7900 HT Sequence Detection System. Expression was compared to two endogenous control genes (GAPDH and beta-actin).

Realtime PCR was carried out using Qiagen SYBR Green. Amplification conditions were: 15min 95°C, 40 cycles of: 15s 94°C, 30s 55°C, 30s 72°C.

Primer pairs:

p14ARF F:CCCTCGTGCTGATGCTACTGAG (SEQ ID NO: 34);

p14ARF R:GTGAGAGTGGCGGGGTCGGC (SEQ ID NO: 35);

P16INK4a F: AACGCACCGAATAGTTACGG (SEQ ID NO: 45);

p16INK4A R:ACCAGCGTGTCCAGGAAG (SEQ ID NO: 46).

Amplification of ANRIL transcripts was carried out for unspliced RNA using custom designed primers as follows.

ANRIL F: CAGTGGCTTCCTGTTCATGC (SEQ ID NO: 30);

ANRIL R: GGGCTTGACGTCTGATCTGT (SEQ ID NO: 31).

Amplification of ANRIL transcripts was carried out for ANRIL splice forms ANRIL14-5, 5-6, and 1-2 using primers from Burd et al 2010, as follows.

ANRIL14-5 F:GGAATGAGGAGCACAGTGATTA (SEQ ID NO: 47);

ANRIL14-5 R:CTGCTGTTGAATCAGAATGAGG (SEQ ID NO: 48);

ANRIL5-6 F:CCACCAGATATATGTTATCTGTGCTTA (SEQ ID NO: 49);

ANRIL5-6R:GTACTGACTCGGGAAAGGATTC (SEQ ID NO: 50);

ANRIL1-2F:GCCACGACATTTCAAAGGATTC (SEQ ID NO: 51);

ANRIL 1-2 R:GCACATACCACACCCTAACTAC (SEQ ID NO: 52).

Realtime PCR was carried out using Qiagen SYBR Green. Amplification conditions: 15min 95°C, 40 cycles of: 15s 94°C, 30s 60°C, 30s 72°C.

Results

An association was found between p14ARF expression and total fat mass ($p=0.027$) and percentage fat mass ($p=0.021$) at birth, and with ponderal index ($p=0.035$) on postnatal day 7. The inventors also found an association between p16INK4A expression and percentage fat mass ($p=0.024$) and total fat mass ($p=0.04$) on postnatal day 7, and with triceps ($p=0.031$) and subscapular ($p=0.019$) skinfold thicknesses and % fat mass ($p=0.035$) at 18 months. Expression of ANRIL was found to be inversely associated with triceps skinfold thickness ($p=0.045$) and subscapular skinfold thickness ($p=0.005$) at postnatal day 7.

Conclusion

The inventors have shown that expression of ANRIL transcript and the CDKN2A transcripts, p14ARF and p16INK4A, is associated with infant adiposity.

References

1. Frayling, T.M., et al., *A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity*. Science, 2007. **316**(5826): p. 889-94.
2. Loos, R.J., et al., *Common variants near MC4R are associated with fat mass, weight and risk of obesity*. Nat Genet, 2008. **40**(6): p. 768-75.
3. Thorleifsson, G., et al., *Genome-wide association yields new sequence variants at seven loci that associate with measures of obesity*. Nat Genet, 2009. **41**(1): p. 18-24.
4. Wellcome Trust Case Control, C., et al., *Genome-wide association study of CNVs in 16,000 cases of eight common diseases and 3,000 shared controls*. Nature, 2010. **464**(7289): p. 713-20.
5. Godfrey, K.M. and D.J. Barker, *Fetal programming and adult health*. Public Health Nutr., 2001. **4**(2B): p. 611-624.
6. Lillycrop, K.A., et al., *Induction of altered epigenetic regulation of the hepatic glucocorticoid receptor in the offspring of rats fed a protein-restricted diet during pregnancy suggests that reduced DNA methyltransferase-1 expression is involved in impaired DNA methylation and changes in histone modifications*. Br.J.Nutr., 2007. **97**(6): p. 1064-1073.
7. Gluckman, P.D., et al., *Metabolic plasticity during mammalian development is directionally dependent on early nutritional status*. Proc.Natl.Acad.Sci.U.S A, 2007. **104**(31): p. 12796-12800.
8. Heijmans, B.T., et al., *Persistent epigenetic differences associated with prenatal exposure to famine in humans*. Proc.Natl.Acad.Sci.U.S A, 2008. **105**(44): p. 17046-17049.
9. Godfrey, K.M., et al., *Epigenetic gene promoter methylation at birth is associated with child's later adiposity*. Diabetes, 2011. **60**(5): p. 1528-1534.
10. Congrains, A., et al., *Genetic variants at the 9p21 locus contribute to atherosclerosis through modulation of ANRIL and CDKN2A/B*. Atherosclerosis, 2012. **220**(2): p. 449-55.
11. Pomerantz, J., et al., *The Ink4a tumor suppressor gene product, p19Arf, interacts with MDM2 and neutralizes MDM2's inhibition of p53*. Cell, 1998. **92**(6): p. 713-23.
12. Abella, A., et al., *Cdk4 promotes adipogenesis through PPARgamma activation*. Cell Metab, 2005. **2**(4): p. 239-49.

13. Zhang, Y., Y. Xiong, and W.G. Yarbrough, *ARF promotes MDM2 degradation and stabilizes p53: ARF-INK4a locus deletion impairs both the Rb and p53 tumor suppression pathways*. Cell, 1998. **92**(6): p. 725-34.
14. Kim, W.Y. and N.E. Sharpless, *The regulation of INK4/ARF in cancer and aging*. Cell, 2006. **127**(2): p. 265-75.
15. Matheu, A., et al., *Anti-aging activity of the Ink4/Arf locus*. Aging Cell, 2009. **8**(2): p. 152-61.
16. Tzanetakou, I.P., et al., *"Is obesity linked to aging?": adipose tissue and the role of telomeres*. Ageing Res Rev, 2012. **11**(2): p. 220-9.
17. Melzer, D., et al., *A common variant of the p16(INK4a) genetic region is associated with physical function in older people*. Mech Ageing Dev, 2007. **128**(5-6): p. 370-7.
18. Cunningham, M.S., et al., *Chromosome 9p21 SNPs Associated with Multiple Disease Phenotypes Correlate with ANRIL Expression*. PLoS Genet, 2010. **6**(4): p. e1000899.
19. Grarup, N., et al., *Studies of association of variants near the HHEX, CDKN2A/B, and IGF2BP2 genes with type 2 diabetes and impaired insulin release in 10,705 Danish subjects: validation and extension of genome-wide association studies*. Diabetes, 2007. **56**(12): p. 3105-11.
20. Zhuang, J., et al., *Methylation of p15INK4b and expression of ANRIL on chromosome 9p21 are associated with coronary artery disease*. PLoS One, 2012. **7**(10): p. e47193.
21. Zeggini, E., et al., *Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes*. Science, 2007. **316**(5829): p. 1336-41.
22. Broadbent, H.M., et al., *Susceptibility to coronary artery disease and diabetes is encoded by distinct, tightly linked SNPs in the ANRIL locus on chromosome 9p*. Hum Mol Genet, 2008. **17**(6): p. 806-14.
23. Inskip, H.M., et al., *Cohort profile: The Southampton Women's Survey*. Int J Epidemiol, 2006. **35**(1): p. 42-8.
24. Crozier, S.R., et al., *Maternal vitamin D status in pregnancy is associated with adiposity in the offspring: findings from the Southampton Women's Survey*. Am J Clin Nutr, 2012. **96**(1): p. 57-63.
- 25.. Soh, S.E., et al., *Cohort Profile: Growing Up in Singapore Towards healthy Outcomes (GUSTO) birth cohort study*. Int J Epidemiol, 2013.

26. Palmke, N., D. Santacruz, and J. Walter, *Comprehensive analysis of DNA-methylation in mammalian tissues using MeDIP-chip*. Methods, 2011. **53**(2): p. 175-184.
27. Weber, M., et al., *Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells*. Nat Genet, 2005. **37**(8): p. 853-62.
28. Down, T.A., et al., *A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis*. Nat.Biotechnol., 2008. **26**(7): p. 779-785.
29. Jaffe, A.E., et al., *Bump hunting to identify differentially methylated regions in epigenetic epidemiology studies*. Int.J.Epidemiol., 2012. **41**(1): p. 200-209.
30. Barton, S.J., et al., *Correction of unexpected distributions of P values from analysis of whole genome arrays by rectifying violation of statistical assumptions*. BMC Genomics, 2013. **14**: p. 161.

CLAIMS:

1. A method of predicting the development of one or more phenotypic characteristics in an individual, wherein the phenotypic characteristic is selected from one or more of:

- (i) obesity and/or adiposity;
- (ii) bone health, and
- (iii) cardiovascular disease,

wherein the method comprises determining the methylation status of one or more CpG dinucleotides of a gene in a sample, wherein said gene is selected from the group comprising SLC6A4, CDKN2A, FEN1 and BACH1 or a combination thereof, and wherein the methylation status of the one or more CpG dinucleotides provides an indication of the development of said phenotypic characteristic, or characteristics, in said individual.

2. A method according to claim 1, wherein the phenotypic characteristic is one or more of obesity and adiposity, and

wherein the one or more CpG dinucleotide is selected from the group comprising CpG5, CpG1 and CpG2 of the SLC6A4 gene, CpG1, CpG2, CpG3, CpG4, CpG5, CpG6, CpG7, CpG8, and CpG9 of the CDKN2A gene, , CpG1 and 3 of the FEN1 gene and CpG1, CpG4, CpG5, CpG6 and CpG8 of the BACH1 gene, or a combination thereof.

3. A method according to claim 2, wherein the one or more CpG dinucleotide is selected from the group comprising CpG5, CpG1 and CpG2 of the SLC6A4 gene, or a combination thereof.

4. A method according to claim 2, wherein the one or more CpG dinucleotide is selected from the group comprising CpG1, CpG2, CpG3, CpG4, CpG5, CpG6, CpG7, CpG8, and CpG9 of the CDKN2A gene, or a combination thereof.

5. A method according to claim 2, wherein the one or more CpG dinucleotide is selected from the group comprising CpG1 and 3 of the FEN1 gene, or a combination thereof.

6. A method according to claim 2, wherein the one or more CpG dinucleotide is CpG1, CpG4, CpG5, CpG6 and CpG8 of the BACH1 gene, or a combination thereof.

7. A method according to claim 1, wherein the phenotypic characteristic is bone mineral content and/or bone health, and
wherein the one or more CpG dinucleotide is selected from the group comprising CpG4, CpG5, CpG6, CpG7, CpG8 and CpG9 of the CDKN2A gene, CpG1 of the SLC6A4 gene, and CpG1, CpG3, CpG4, CpG5, CpG7 and CpG8 of the BACH1 gene, or a combination thereof.
8. A method according to claim 7, wherein the one or more CpG dinucleotide is CpG4, CpG5, CpG6, CpG7, CpG8 and CpG9 of the CDKN2A gene, or a combination thereof.
9. A method according to claim 7, wherein the one or more CpG dinucleotide is CpG1 of the SLC6A4 gene.
10. A method according to claim 7, wherein the one or more CpG dinucleotide CpG1, CpG3, CpG4, CpG5, CpG7 and CpG8 of the BACH1 gene, or a combination thereof.
11. A method according to claim 1, wherein the phenotypic characteristic is cardiovascular disease, and
wherein the one or more CpG dinucleotide is selected from the group comprising CpG5 of the CDKN2A gene, CpG1 and CpG5 of the SLC6A4 gene and CpG6 of the BACH1 gene, or a combination thereof.
12. A method according to claim 11, wherein the one or more CpG dinucleotide is CpG5 of the CDKN2A gene.
13. A method according to claim 11, wherein the one or more CpG dinucleotide is CpG1 and CpG5 of the SLC6A4 gene, or a combination thereof.
14. A method according to claim 11, wherein the one or more CpG dinucleotide is CpG6 of the BACH1 gene.

15. A method according to any of preceding claim, wherein the level of methylation is an increase in methylation, a decrease in methylation, the presence of methylation or the absence of methylation.
16. A method according to any preceding claim, further comprising
(b) comparing the degree of methylation determined in (a), with the degree of methylation of said one or more CpG dinucleotides in an individual without said phenotypic characteristic at a selected older age.
17. A method according to claim 16, wherein the degree of methylation determined in (a) is an increase in methylation if the level of methylation is greater than that of the level in an individual without said phenotypic characteristic at a selected older age.
18. A method according to claim 16, wherein the degree of methylation determined in (a) is a decrease in methylation if the level of methylation is less than or equal to that of the level in an individual without said phenotypic characteristic at a selected older age.
19. A method according to any preceding claim, wherein determination of an increase in methylation is indicative of the development of said phenotypic characteristic in said individual.
20. A method according to any preceding claim, wherein the methylation status of a CpG site is determined by pyrosequencing, bisulfite sequencing, single molecular real time sequencing, methylation-sensitive PCR amplification, sequenom, bisulfite conversion followed by primer extension, or a combination thereof.
21. A method according to any preceding claim, wherein the at least one CpG site is selected from
SLC6A4 CpG1: Hg19 chr17: 28561601,
SLC6A4 CpG2: Hg19 chr17: 28561578,
SLC6A4 CpG3: Hg19 chr17: 28561505,
SLC6A4 CpG4: Hg19 chr17: 28561490,
SLC6A4 CpG5: Hg19 chr17: 28561468,
CDKN2A CpG1: Hg19 chr9: 21993721,

CDKN2A CpG2: Hg19 chr9: 21993697,
CDKN2A CpG3: Hg19 chr9: 21993694,
CDKN2A CpG4: Hg19 chr9: 21993654,
CDKN2A CpG5: Hg19 chr9: 21993645,
CDKN2A CpG6: Hg19 chr9: 21993638,
CDKN2A CpG7: Hg19 chr9: 21993629,
CDKN2A CpG8: Hg19 chr9: 21993603,
CDKN2A CpG9: Hg19 chr9: 21993583,
BACH1 CpG1: Hg19 chr21: 30668941,
BACH1 CpG2: Hg19 chr21: 30668995,
BACH1 CpG3: Hg19 chr21: 30669026,
BACH1 CpG4: Hg19 chr21: 30669196,
BACH1 CpG5: Hg 19 chr21: 30669217,
BACH1 CpG6: Hg19 chr21: 30669219,
BACH1 CpG7: Hg19 chr21: 30669245,
BACH1 CpG8: Hg19 chr21: 30669259,
FEN1 CpG1: Hg19 chr11: 61561170,
FEN1 CpG2: Hg19 chr11: 61561183,
FEN1 CpG3: Hg19 chr11: 61561220 and
FEN1 CpG4: Hg19 chr11: 61561291.

22. A method according to claim 21, wherein the at least one CpG site is selected from the CDKN2A CpG sites identified therein.

23. A method according to claim 21, wherein the at least one CpG site is selected from the SLC6A4 CpG sites identified therein.

24. A method according to claim 21, wherein the at least one CpG site is selected from the BACH1 CpG sites identified therein.

25. A method according to claim 21, wherein the at least one CpG site is selected from the FEN1 CpG sites identified therein.

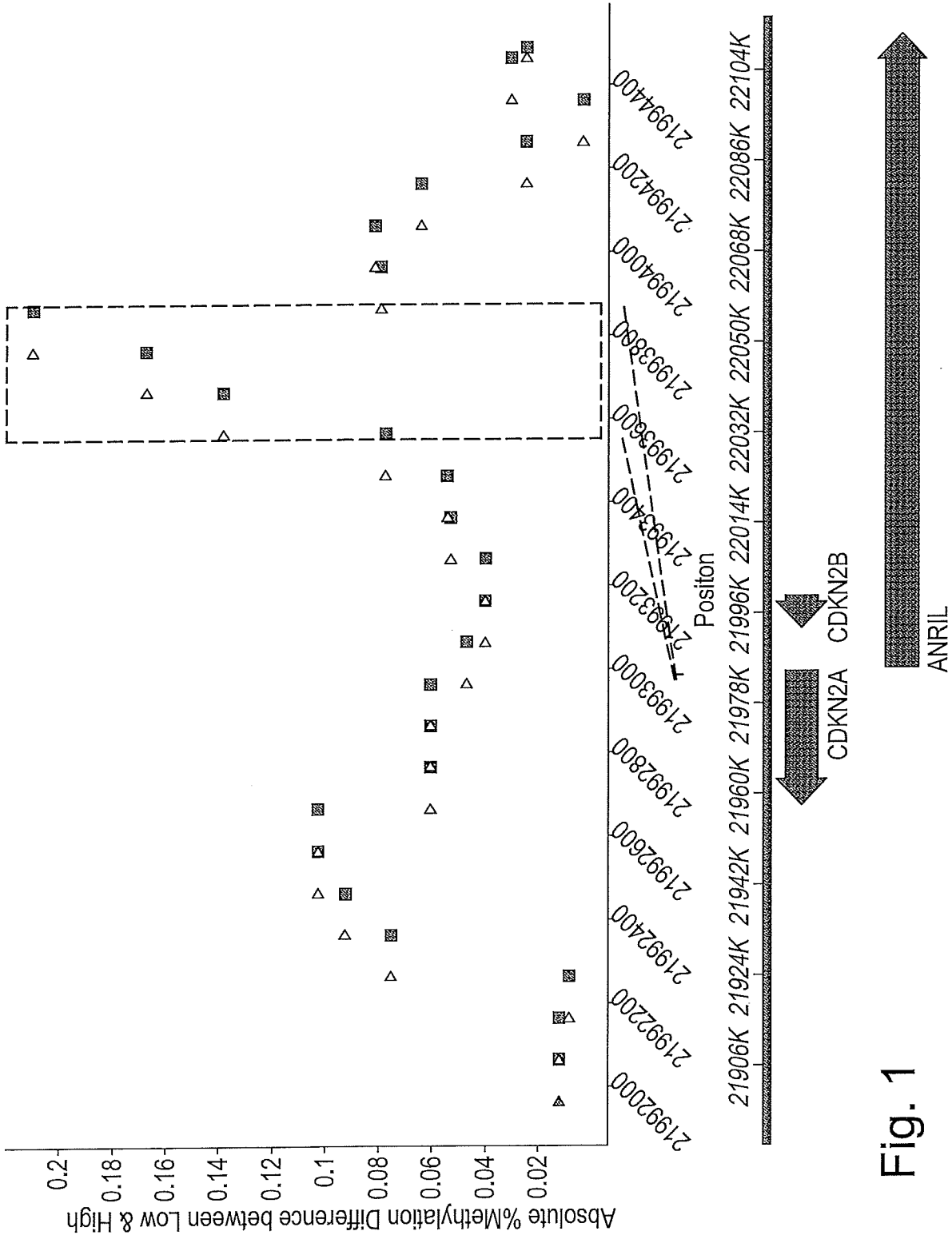


Fig. 1

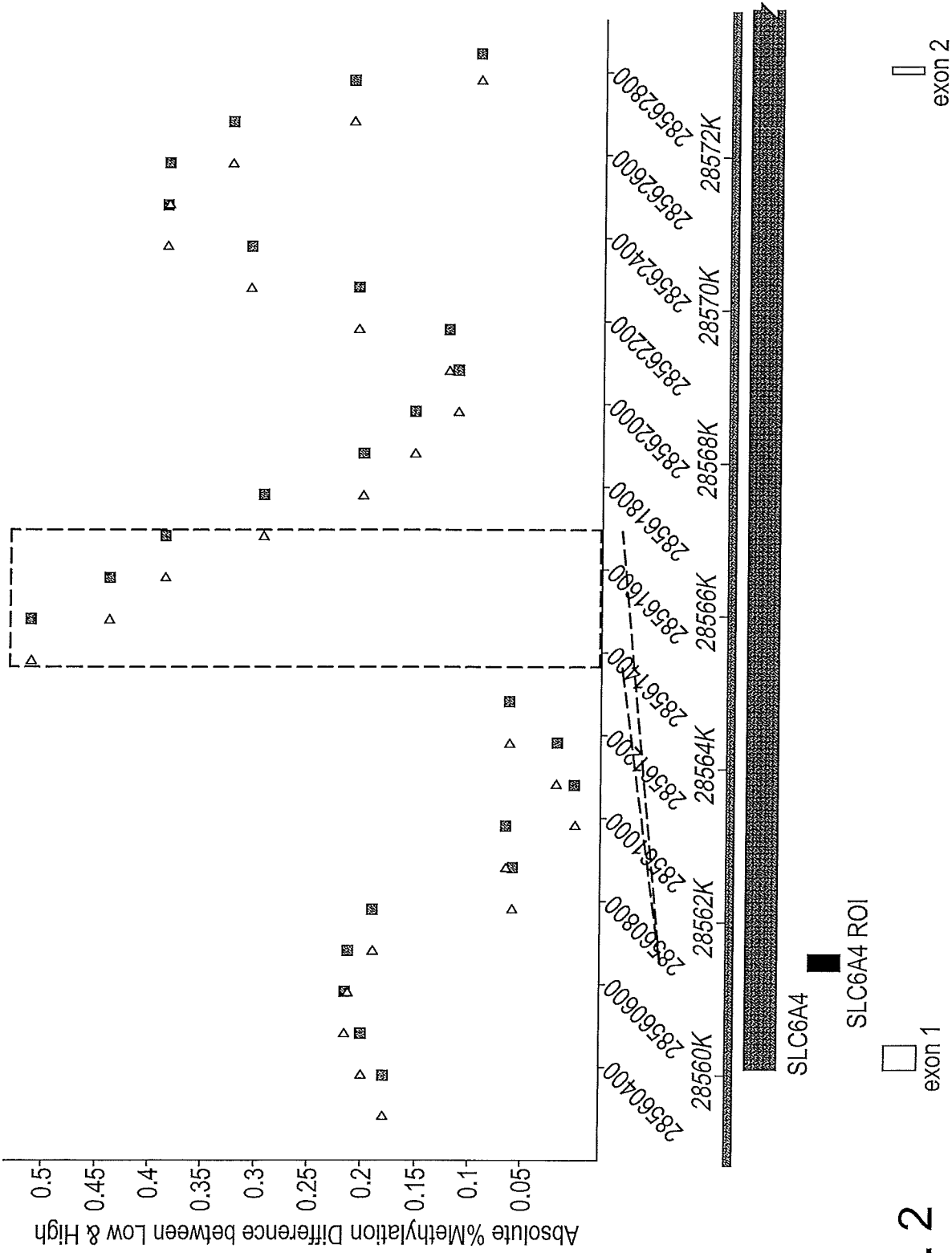


Fig. 2

3/23

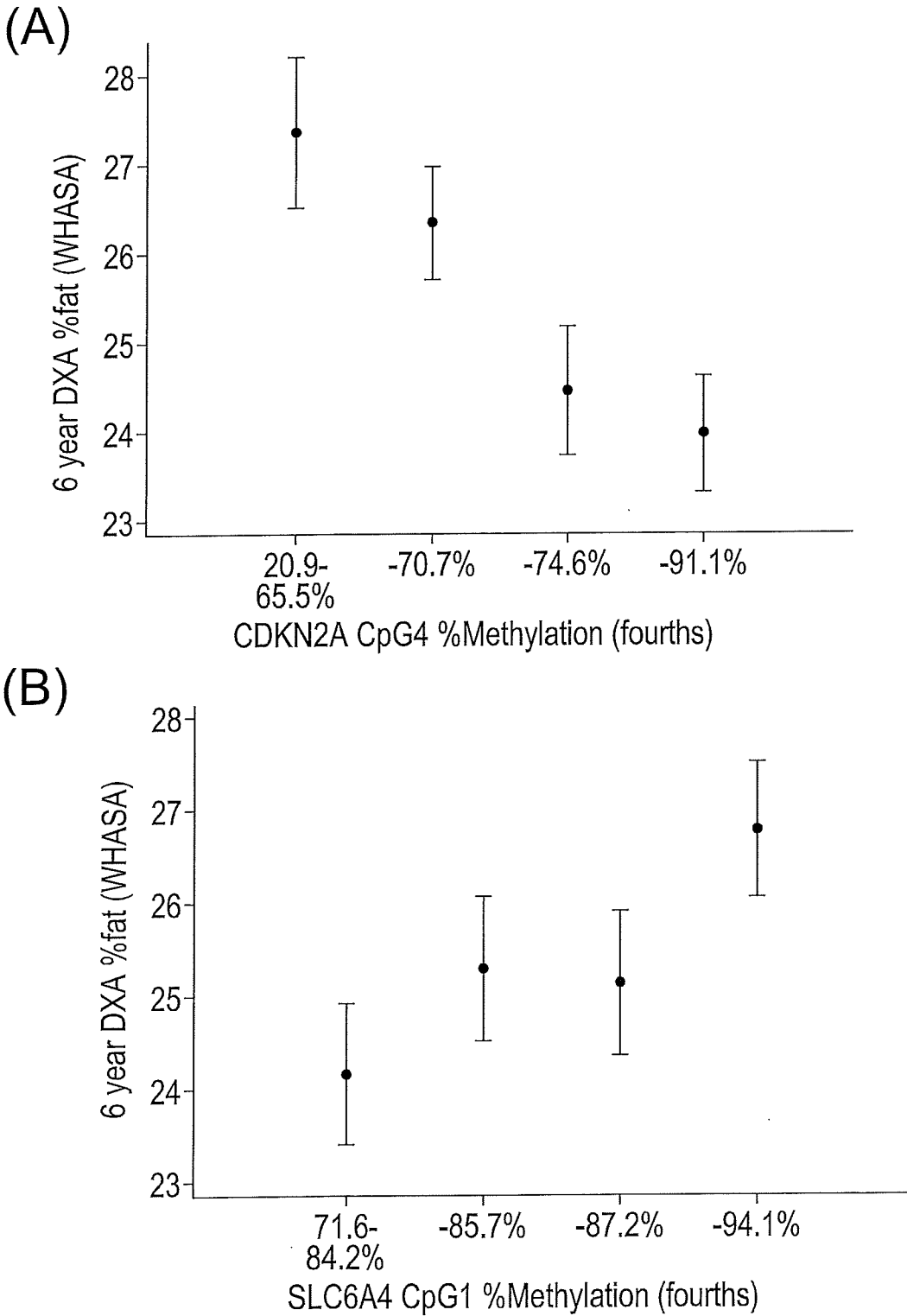
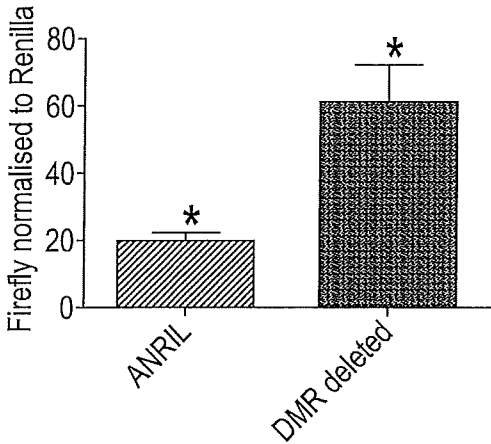
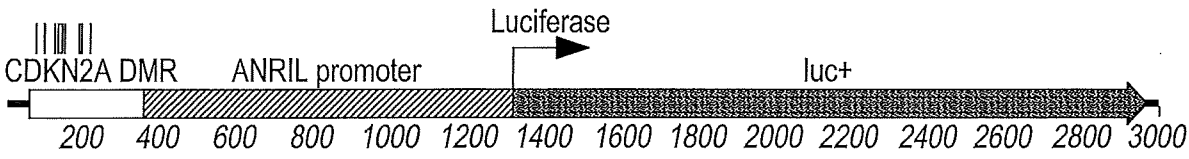


Fig. 3

(A)



(B)

Extract	-	+	+	+	+	+	+	+
CpG probe	+	+	+	+	+	+	+	+
Specific CpG 4-7 competitor	-	-	x50	x100	x500	-	-	-
Met CpG 4 competitor	-	-	-	-	-	x50	x100	x500

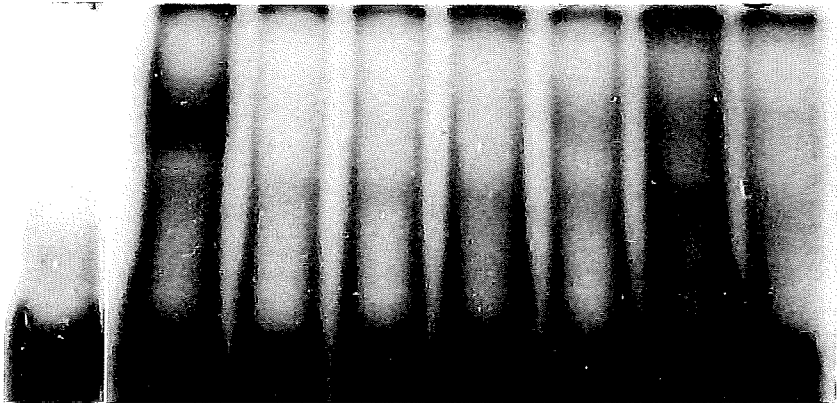


Fig. 4

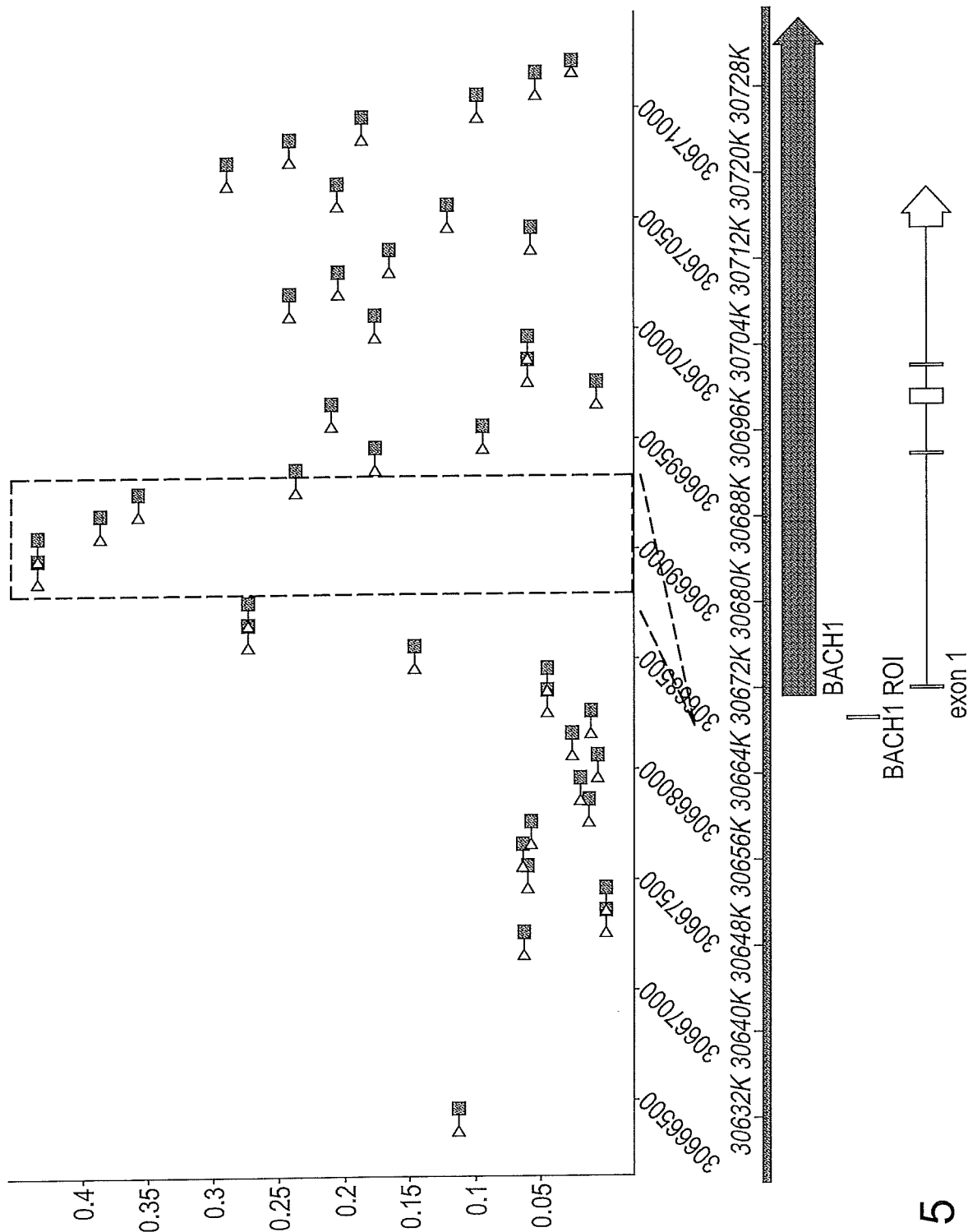


Fig. 5

6/23

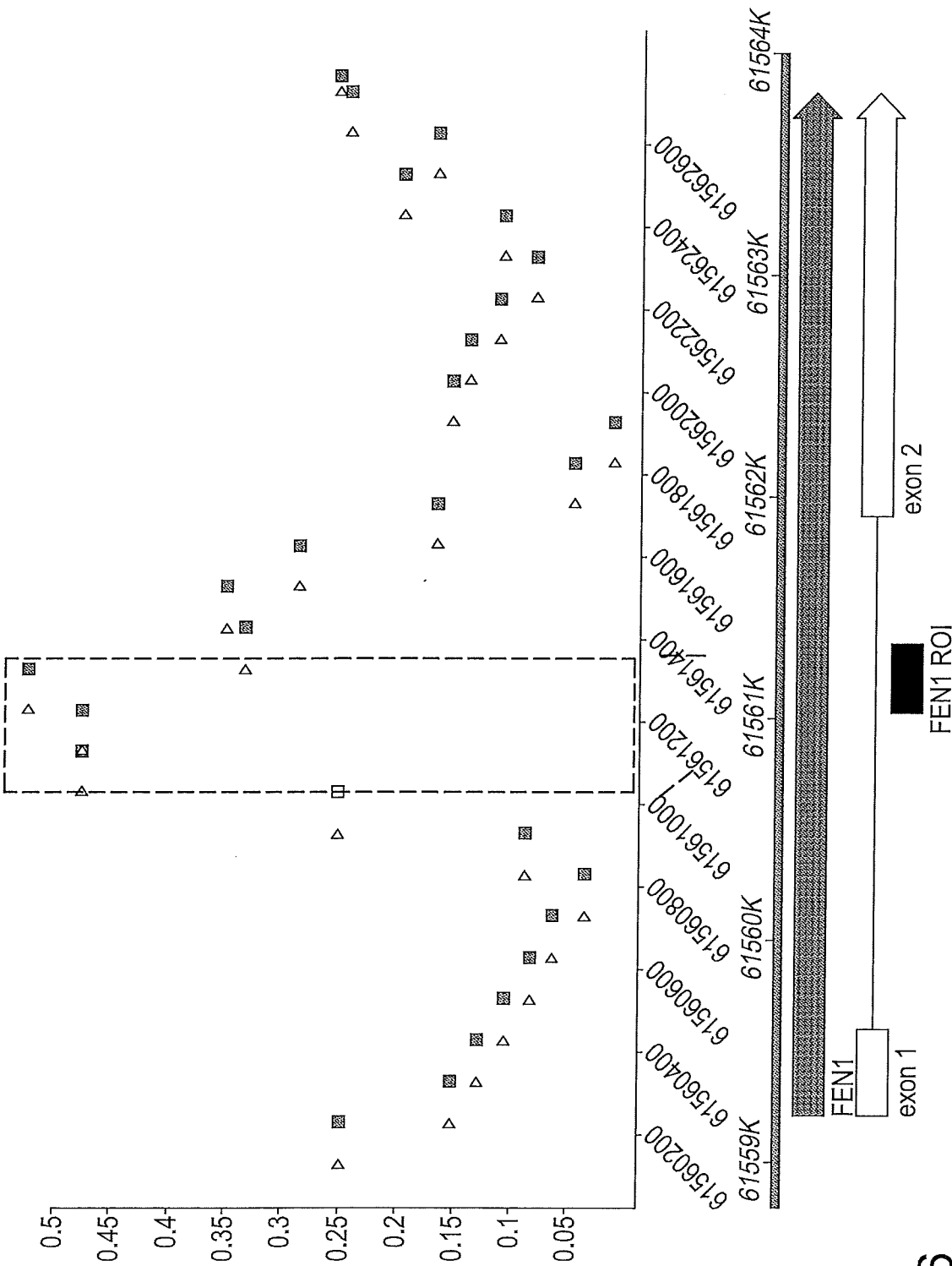


Fig. 6

7/23

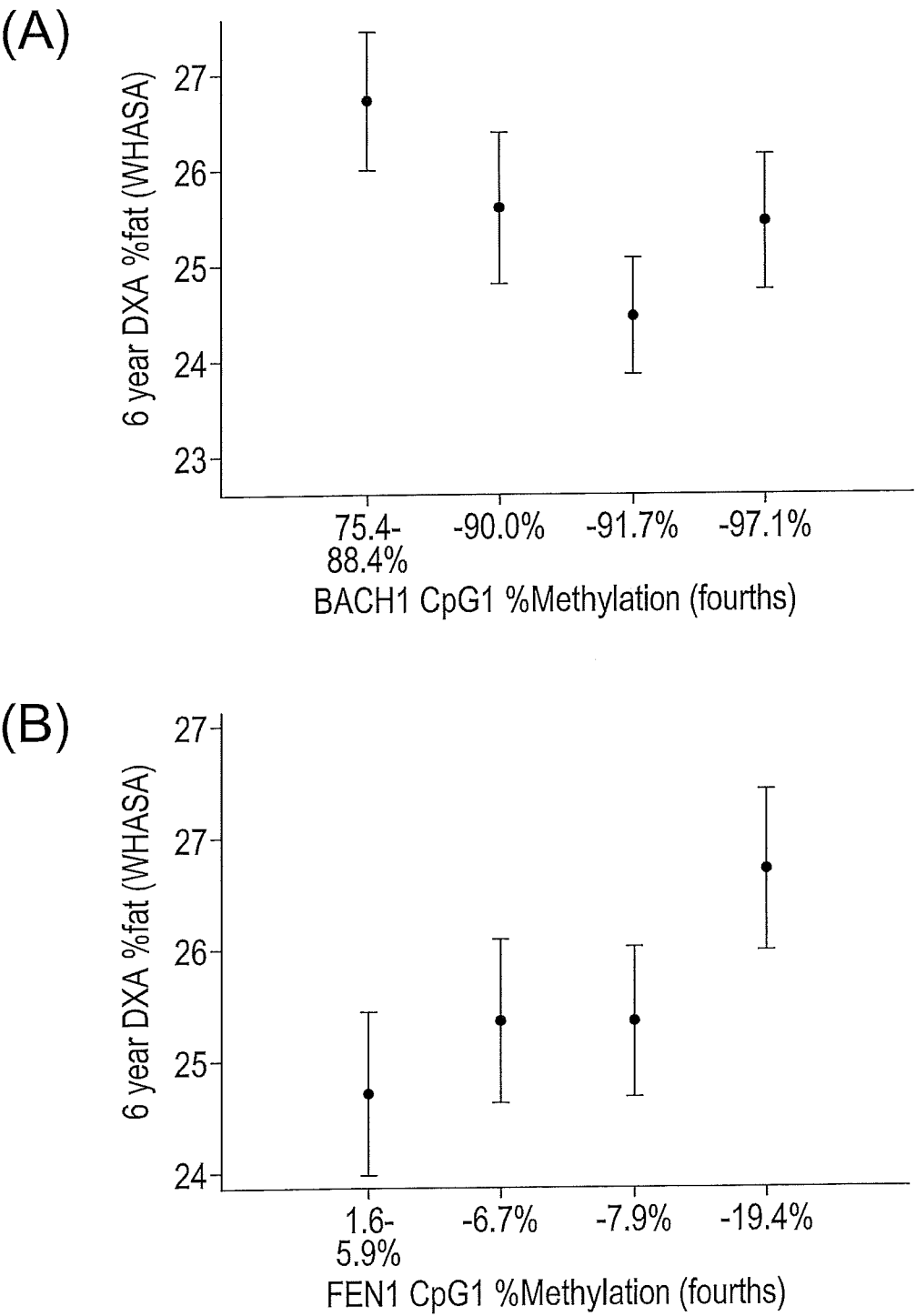


Fig. 7

8/23

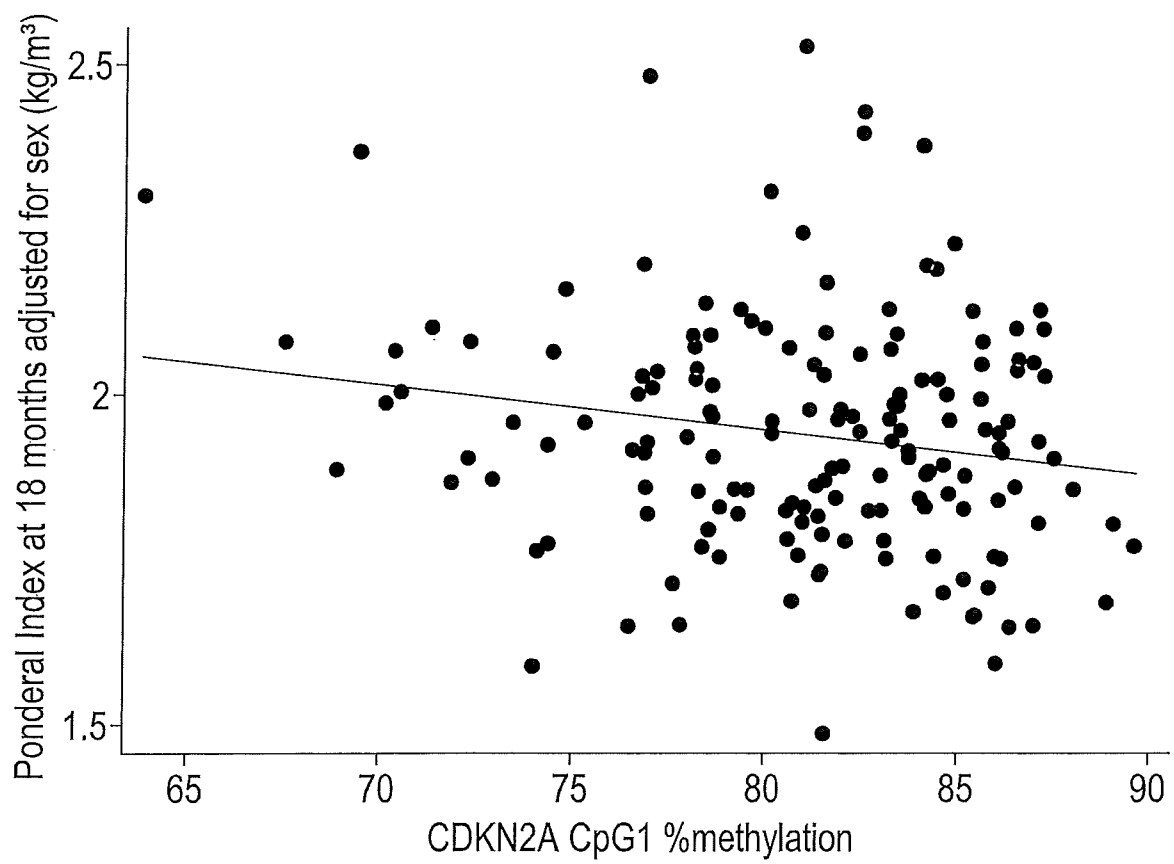


Fig. 8

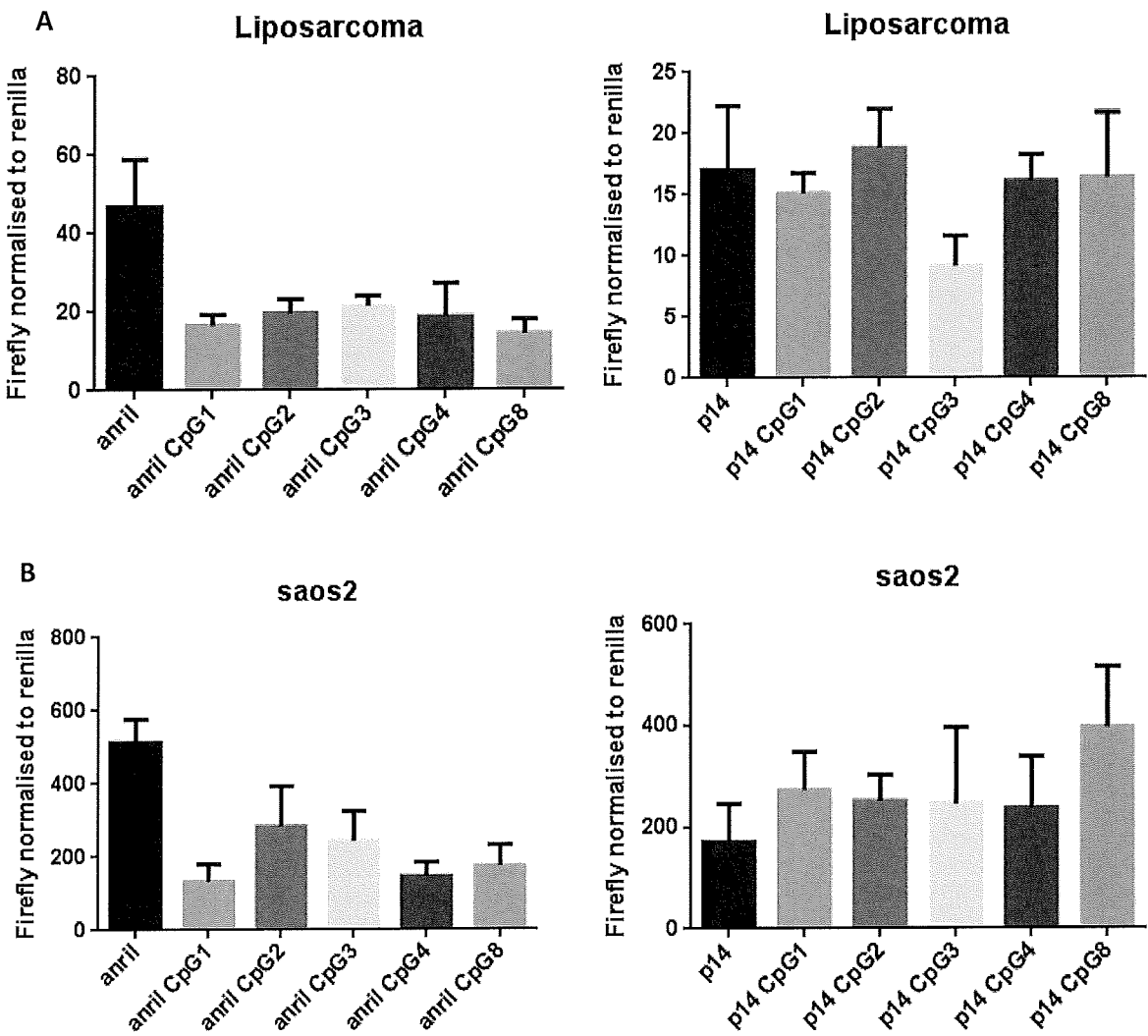


Fig. 9 [Continued]

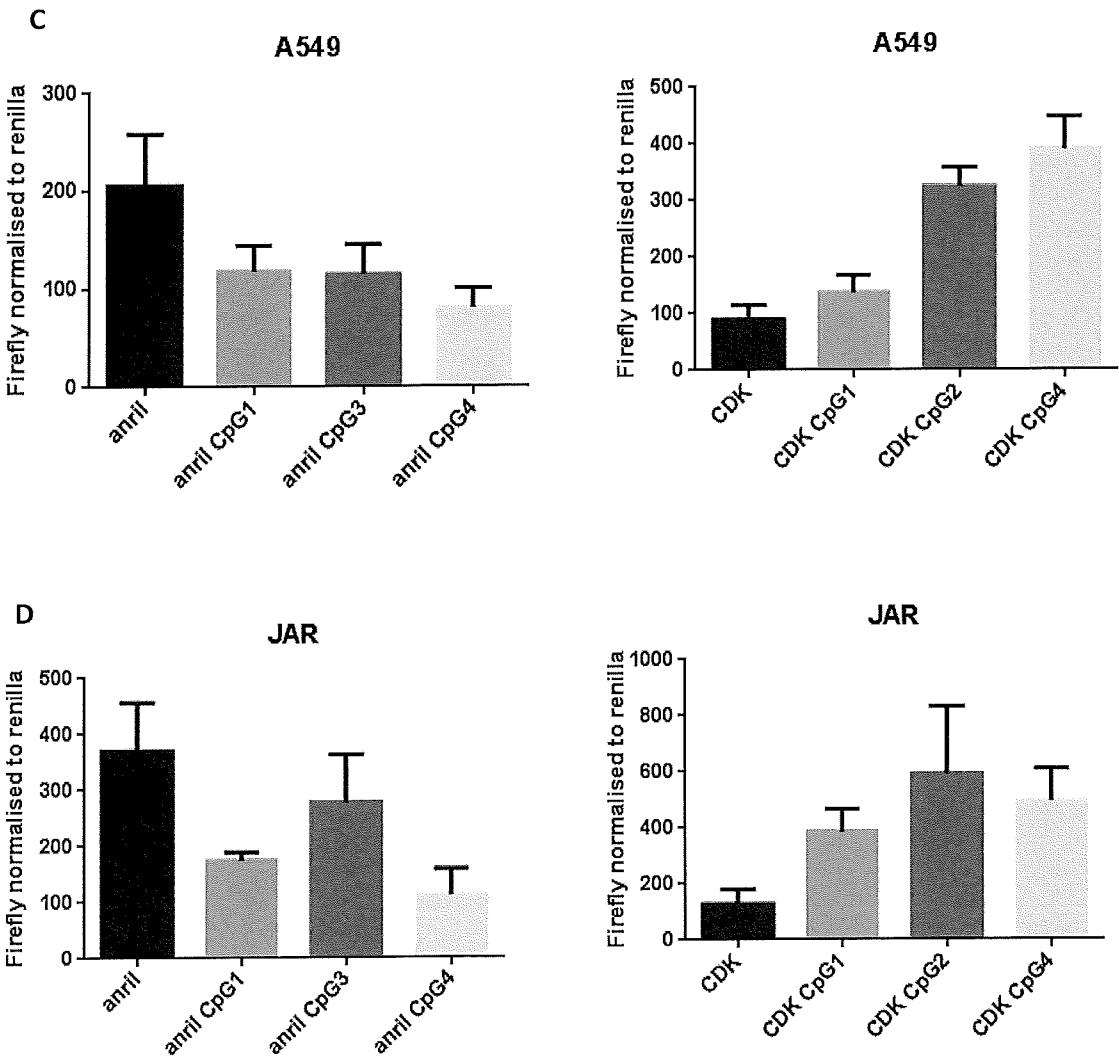


Fig. 9 [Continued]

E

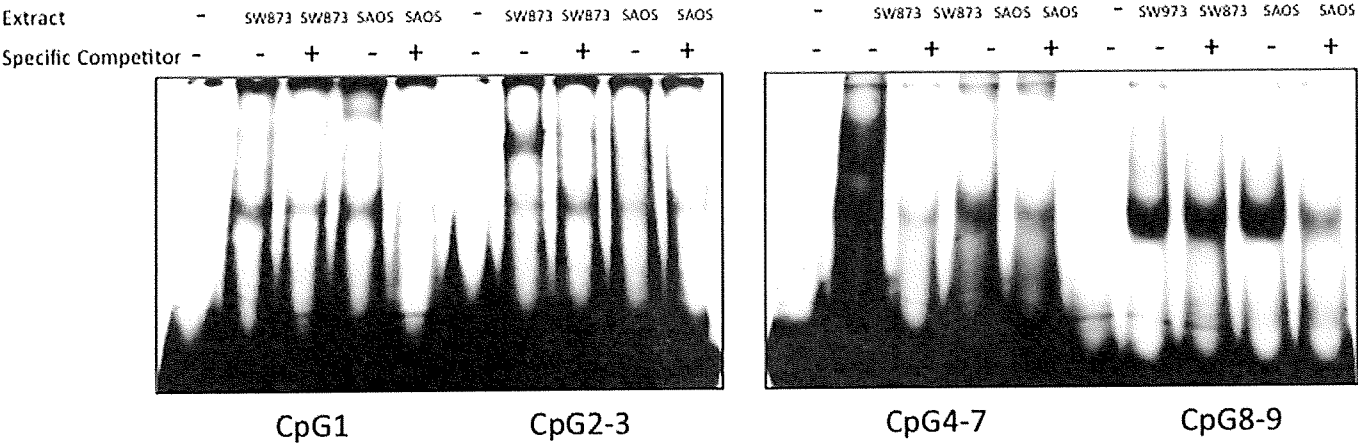


Fig. 9

12/23

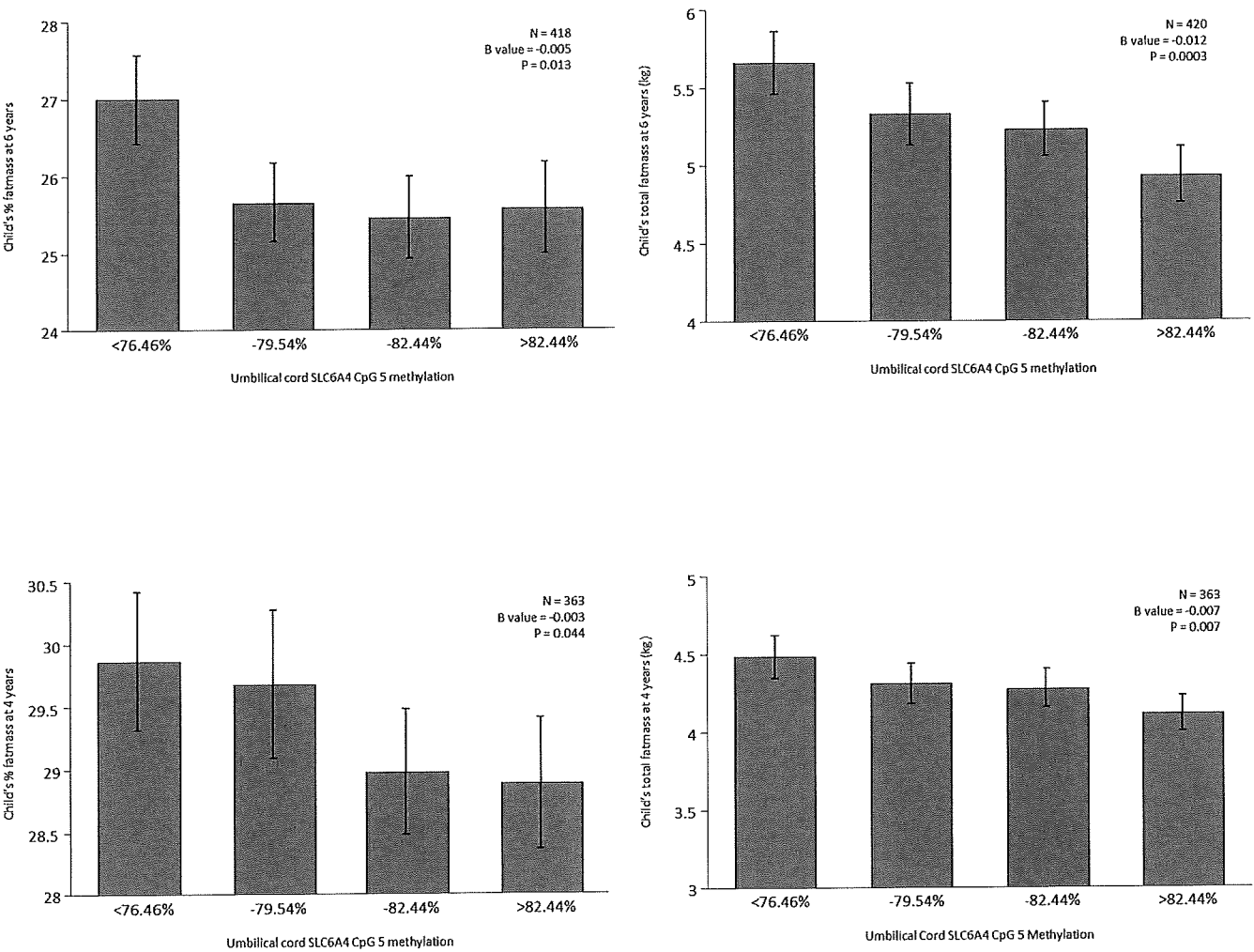


Fig. 10

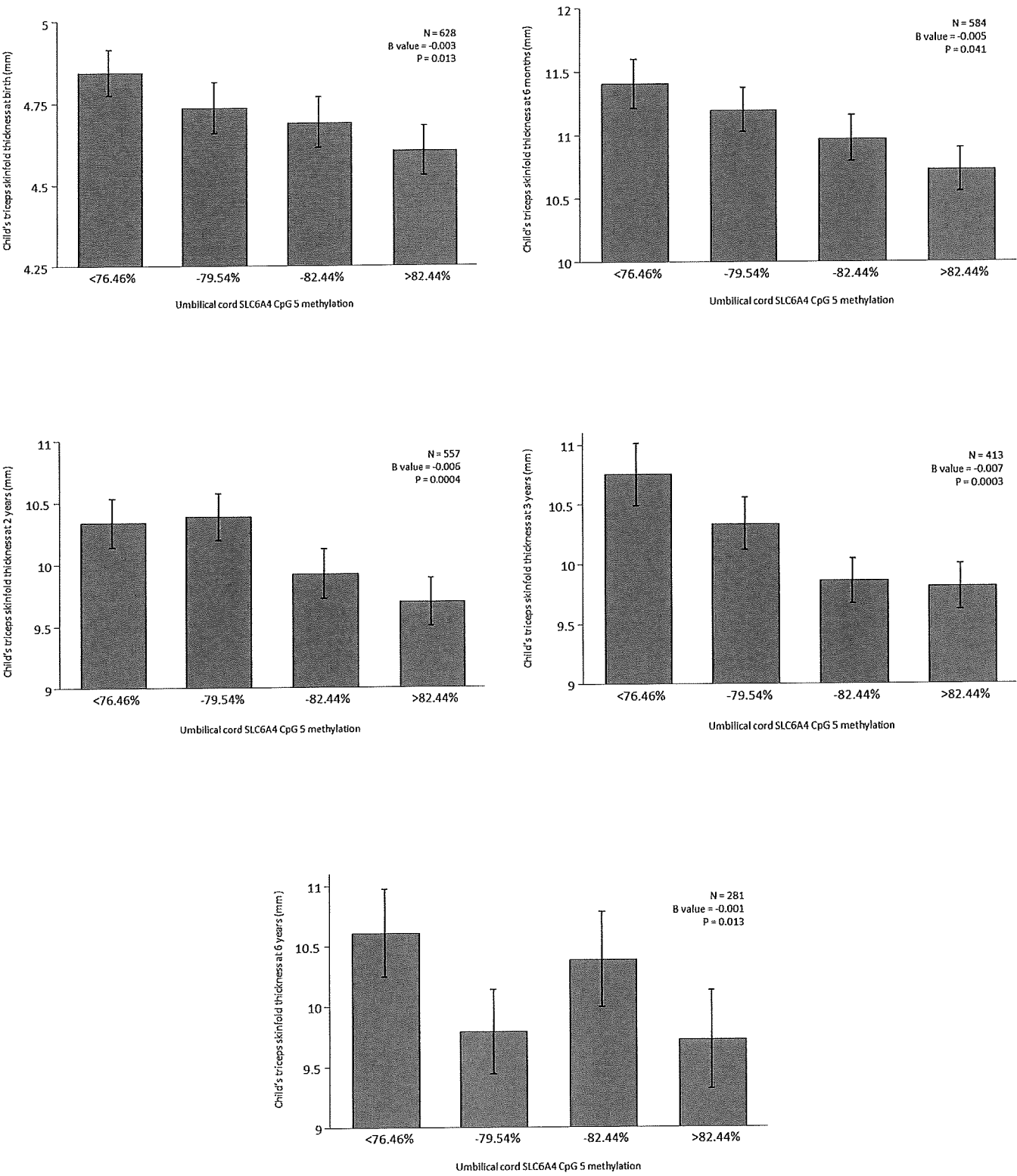


Fig. 11

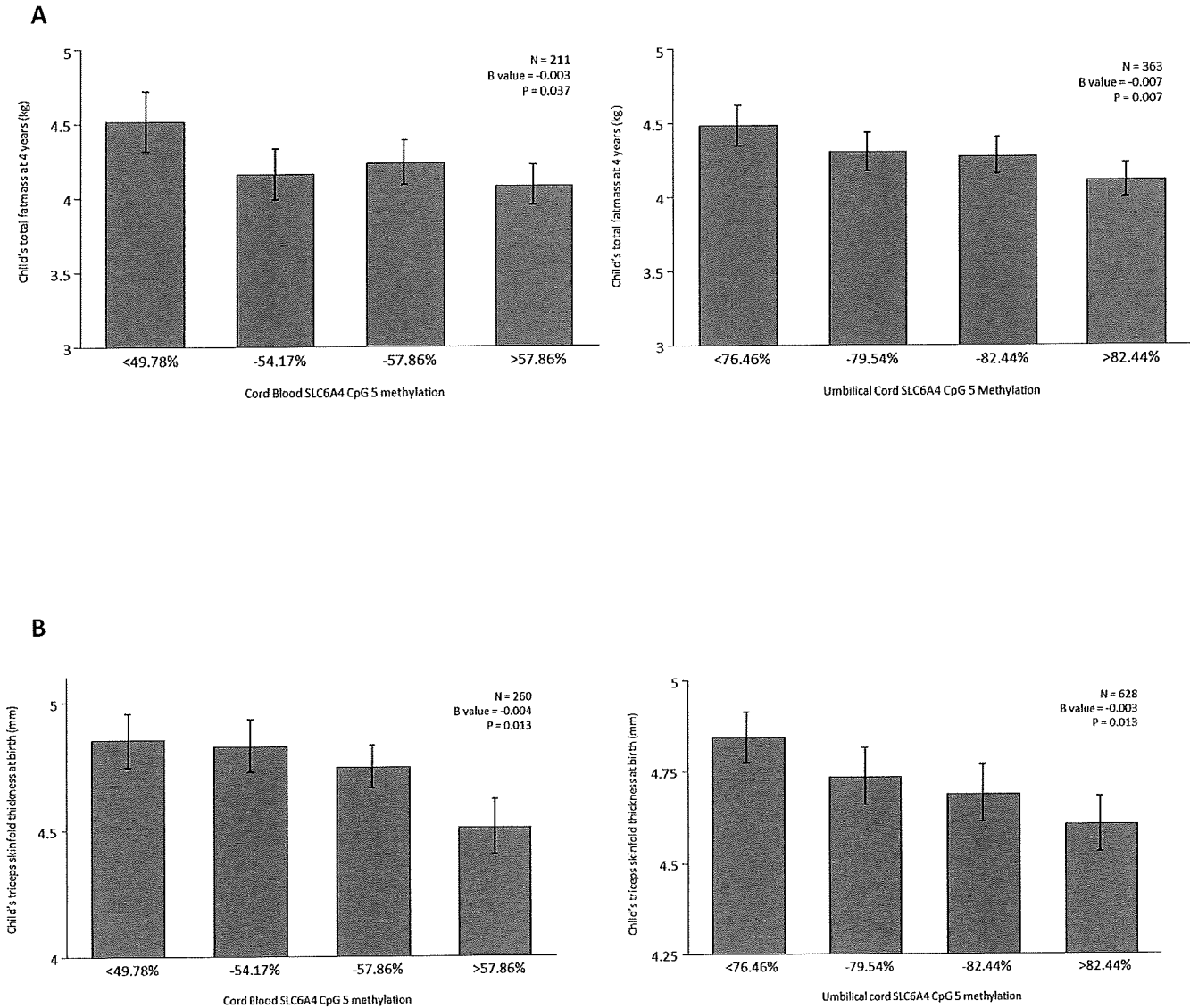


Fig. 12

15/23

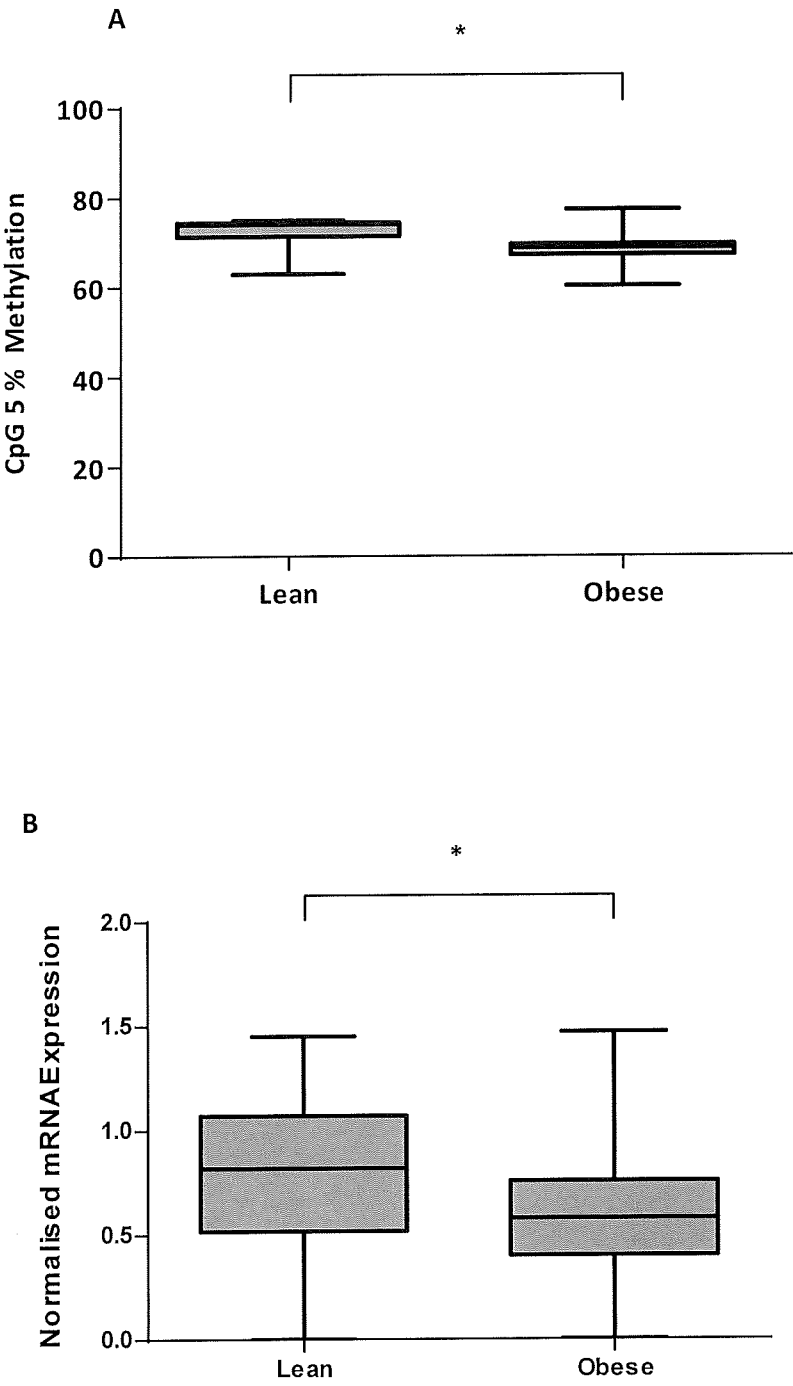


Fig. 13

16/23

A

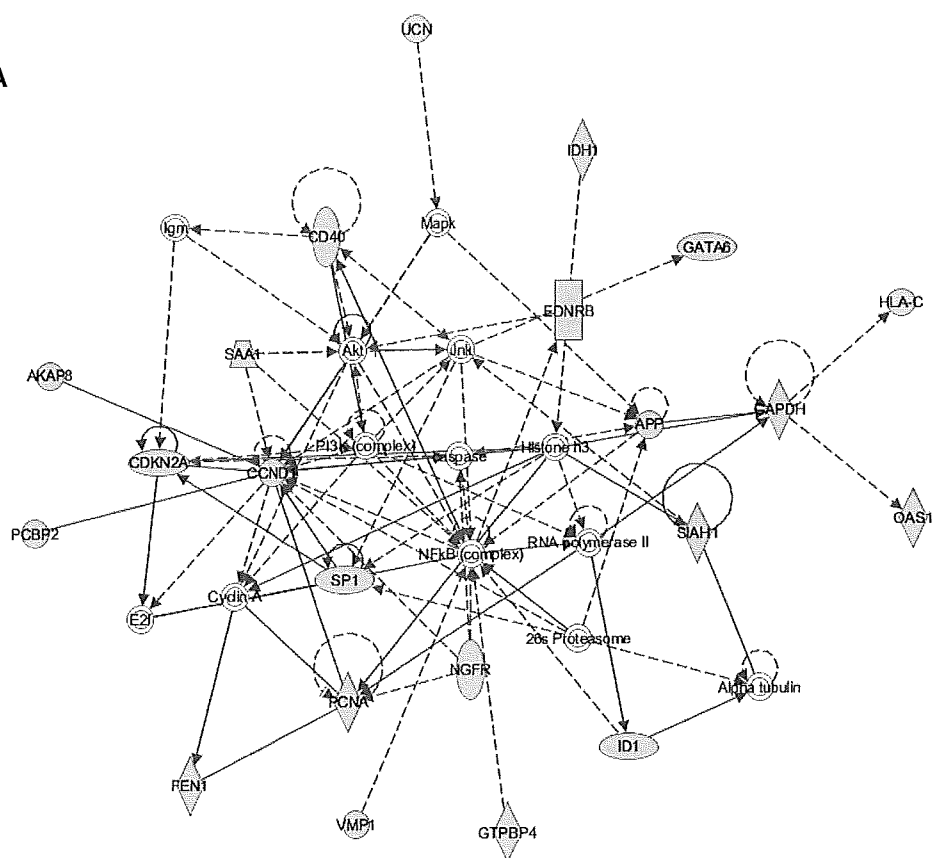
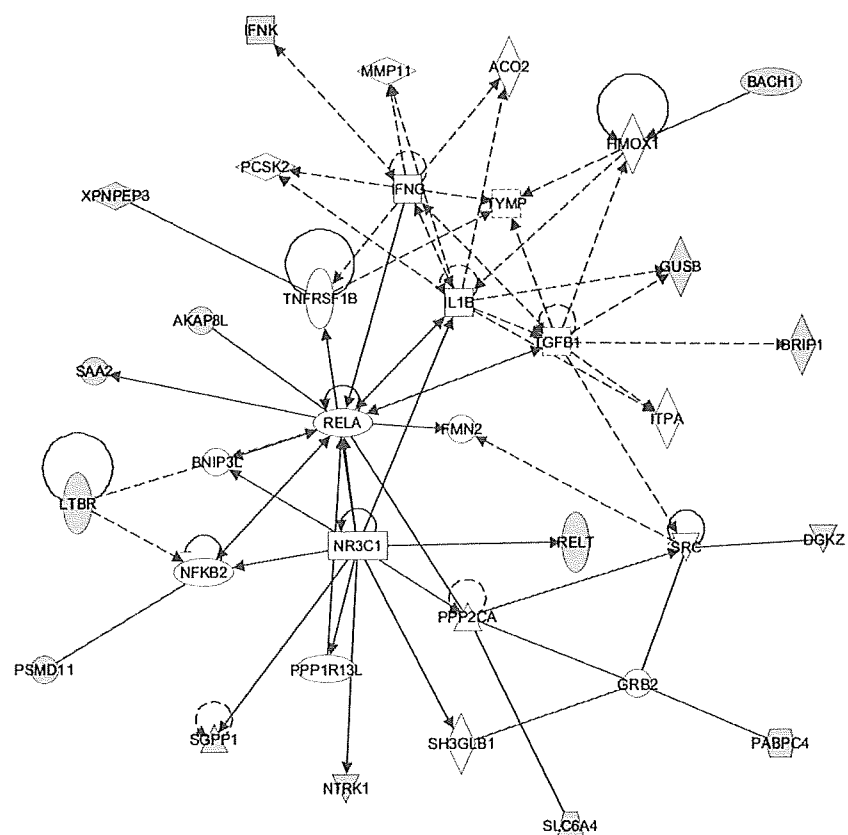
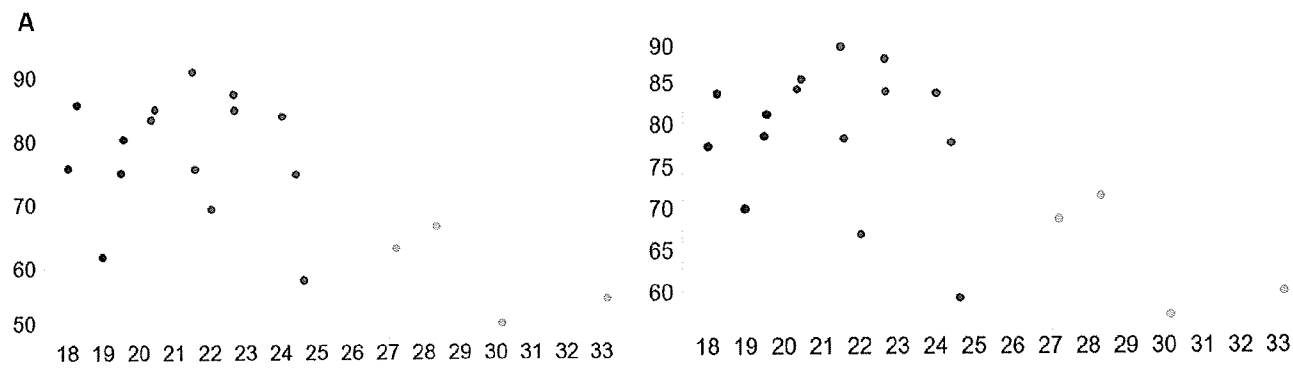
**B**

Fig. 14



B [Continued]

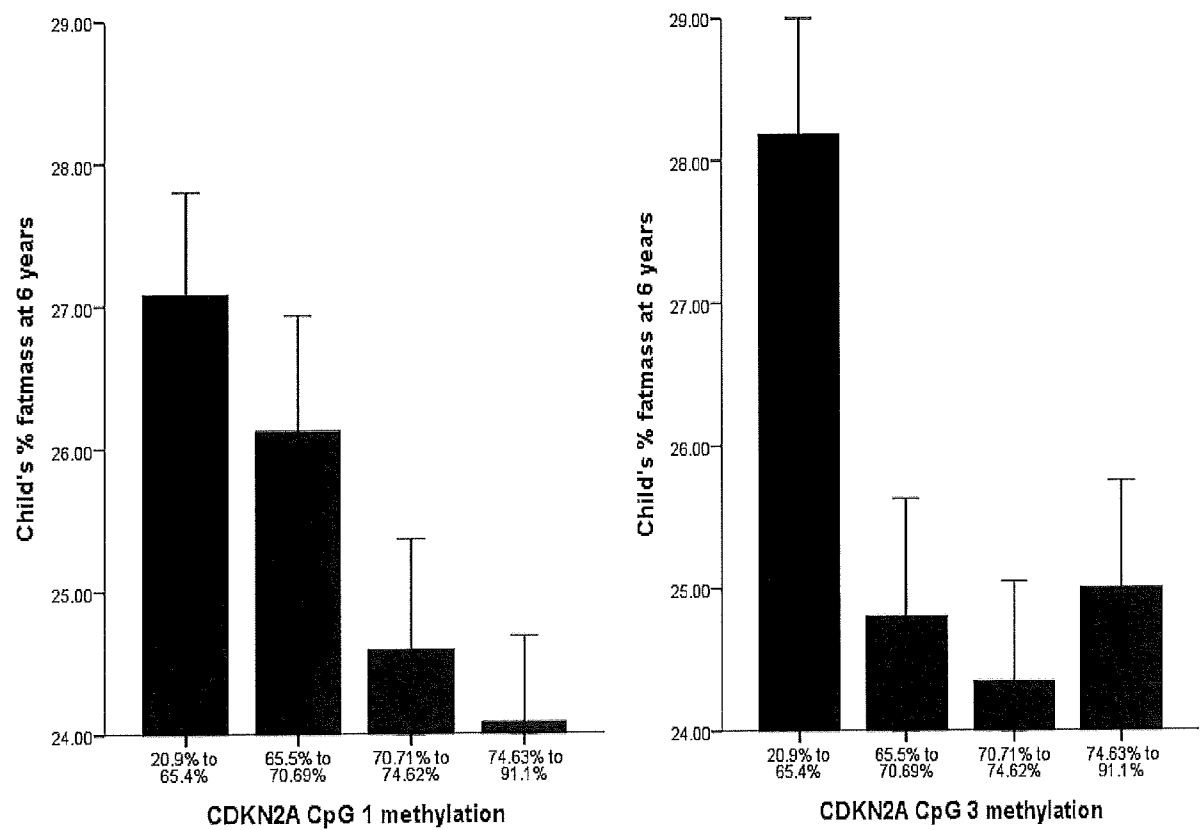


Fig. 15 [Continued]

B

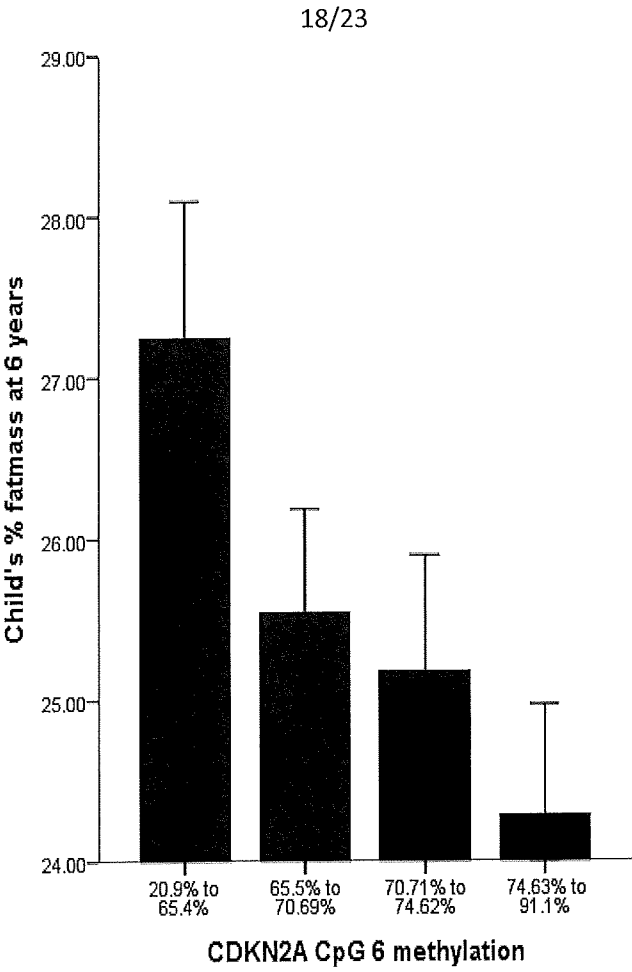


Fig. 15 [Continued]

19/23

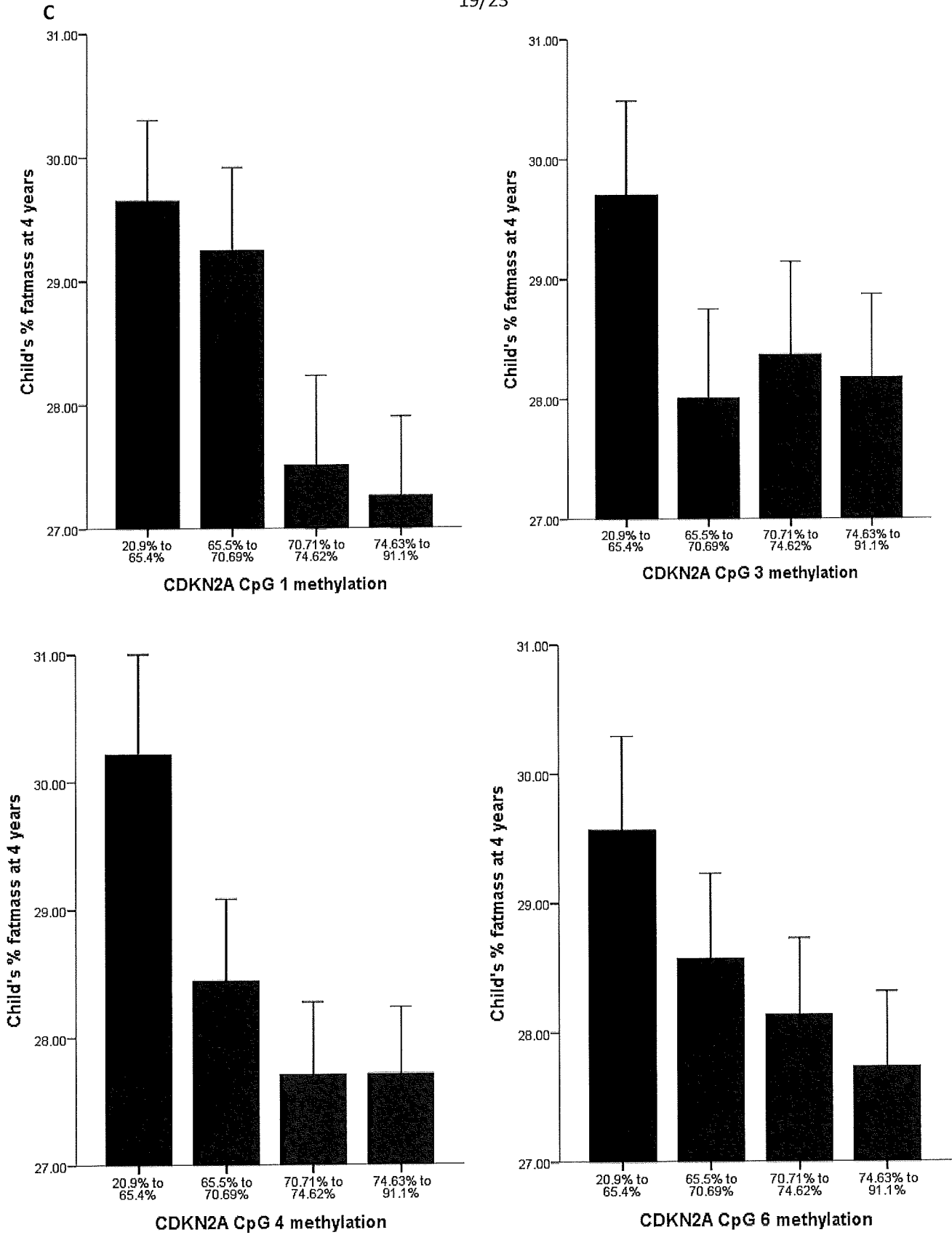


Fig. 15 [Continued]

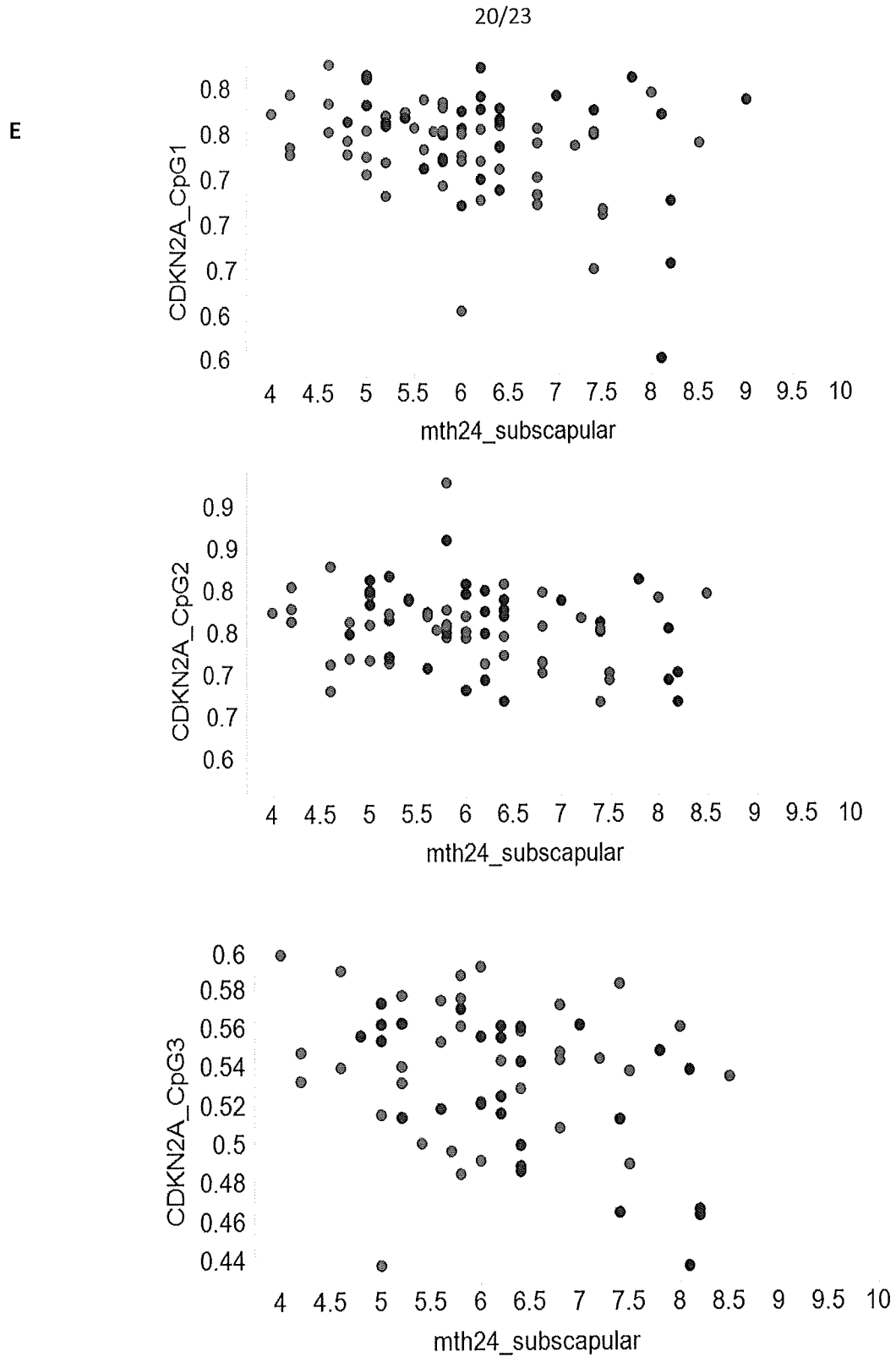


Fig. 15

A

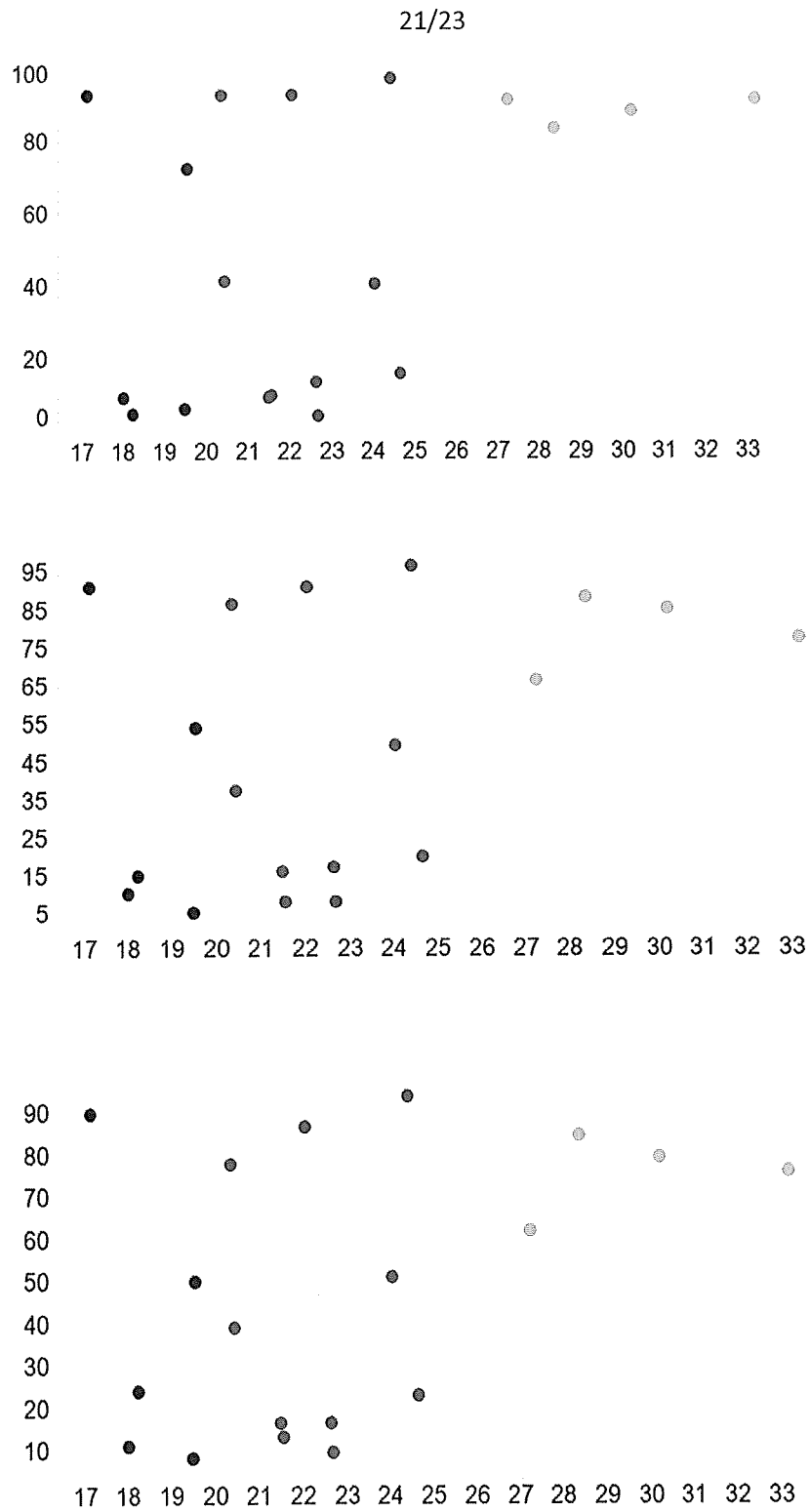


Fig. 16 [Continued]

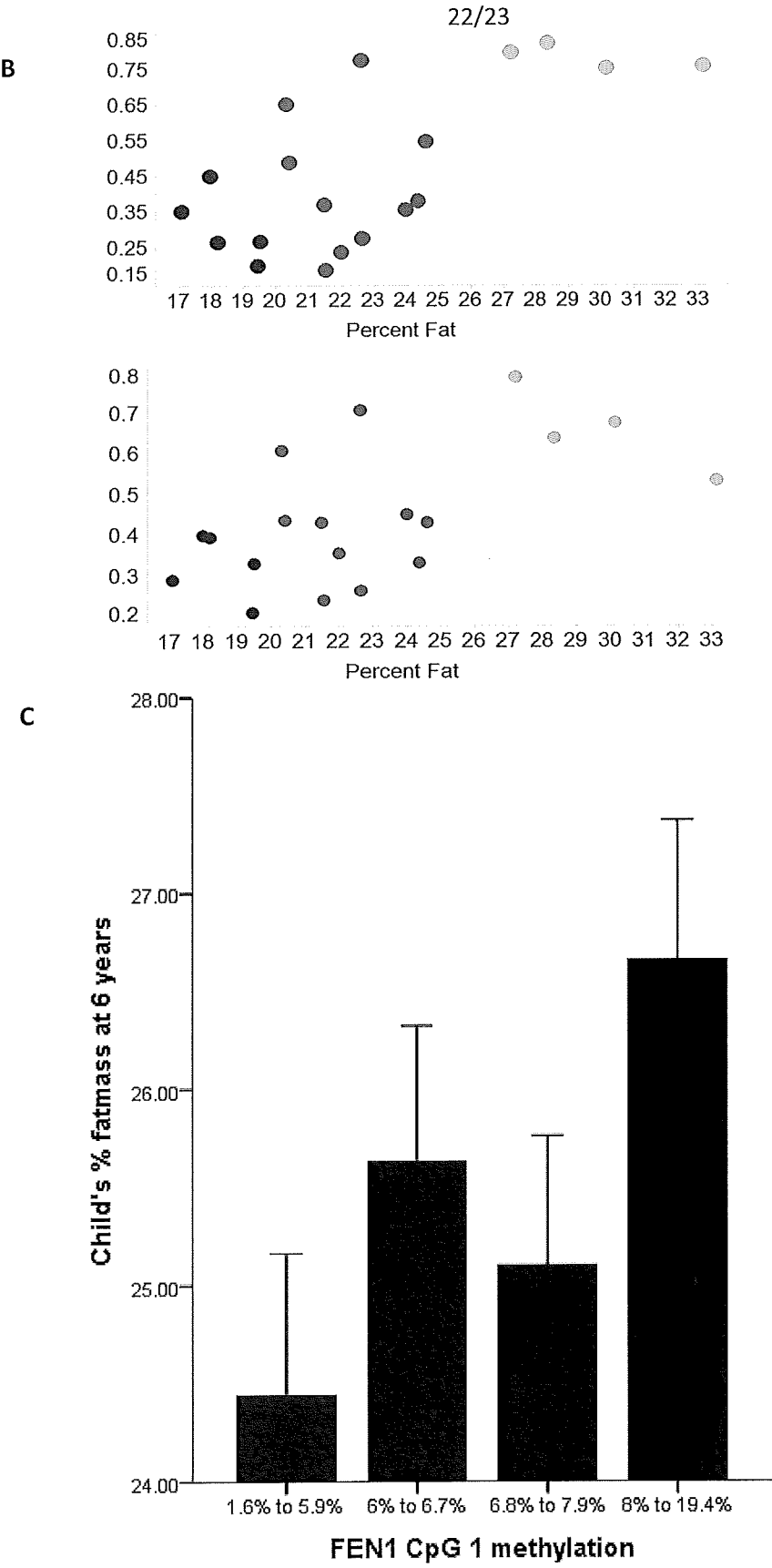


Fig. 16 [Continued]

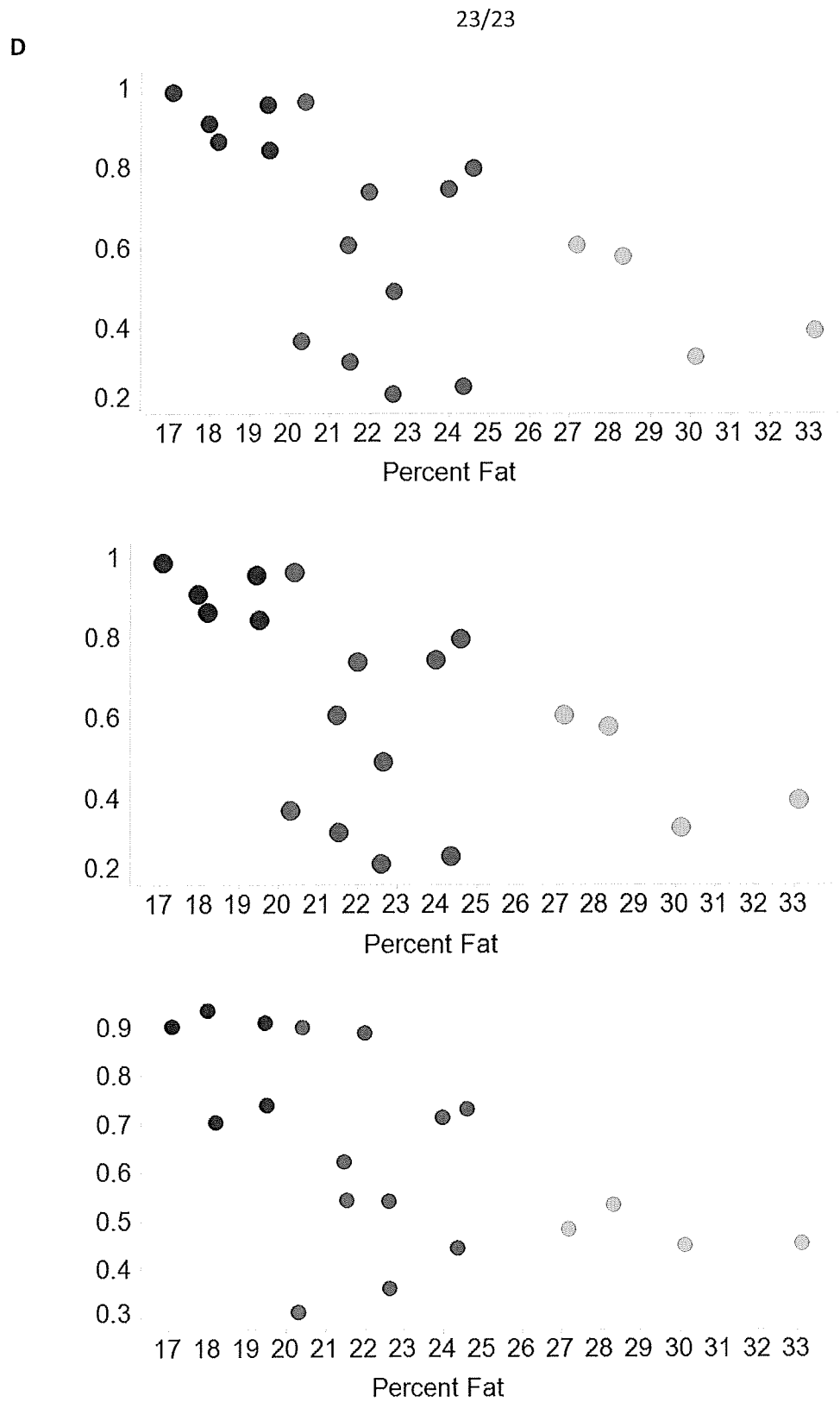


Fig. 16

International application No
PCT/GB2014/053308

According to International Patent Classification (IPC) or to both national classification and IPC

Minimum documentation searched (classification system followed by classification symbols)
C120

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, WPI Data

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2007/129113 A2 (UNIV SOUTHAMPTON [GB]; AUCKLAND UNISERVICES LTD [NZ]; GOODFREY KEITH M) 15 November 2007 (2007-11-15)	1,15-17, 19,20
Y	the whole document abstract, examples, claims ----- -/--	2,3,7, 11,13, 18,21,23

☒

Further documents are listed in the continuation of Box C.

☒

See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

30 March 2015

Date of mailing of the international search report

07/04/2015

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Sauer, Tinctura

INTERNATIONAL SEARCH REPORT

International application No

PCT/GB2014/053308

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	J ZHAO ET AL: "Promoter methylation of serotonin transporter gene is associated with obesity measures: a monozygotic twin study", INTERNATIONAL JOURNAL OF OBESITY, vol. 37, no. 1, 31 January 2012 (2012-01-31), pages 140-145, XP055161656, ISSN: 0307-0565, DOI: 10.1038/ijo.2012.8	1,15-17, 19,20
Y	the whole document table 2	2,3,7, 11,13, 18,21,23
X	----- WO 2009/036513 A1 (MURDOCH CHILDRENS RES INST [AU]; CRAIG JEFFREY MARK [AU]; SAFFERY RICH) 26 March 2009 (2009-03-26)	1,15-17, 19,20
Y	the whole document abstract, para. 7, 24, 42, 43, claims 1, 11	2,3,7, 11,13, 18,21,23
A	----- US 2012/040344 A1 (GODFREY KEITH MALCOLM [GB] ET AL) 16 February 2012 (2012-02-16) the whole document -----	1-3,7, 11,13, 15-21,23

INTERNATIONAL SEARCH REPORT

International application No.
PCT/GB2014/053308

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of Item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

3, 9, 13, 23(completely); 1, 2, 7, 11, 15-21(partially)

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 3, 9, 13, 23(completely); 1, 2, 7, 11, 15-21(partially)

A method of predicting the development of one or more phenotypic characteristics in an individual, wherein the method comprises determining the methylation status of one or more CpG dinucleotides of a gene in a sample, wherein said gene is SLC6A4;

2. claims: 4, 8, 12, 22(completely); 1, 2, 7, 11, 15-21(partially)

A method of predicting the development of one or more phenotypic characteristics in an individual, wherein the method comprises determining the methylation status of one or more CpG dinucleotides of a gene in a sample, wherein said gene is CDKN2A;

3. claims: 5, 25(completely); 1, 2, 15-21(partially)

A method of predicting the development of one or more phenotypic characteristics in an individual, wherein the method comprises determining the methylation status of one or more CpG dinucleotides of a gene in a sample, wherein said gene is FEN1;

4. claims: 6, 10, 14, 24(completely); 1, 2, 7, 11, 15-21(partially)

A method of predicting the development of one or more phenotypic characteristics in an individual, wherein the method comprises determining the methylation status of one or more CpG dinucleotides of a gene in a sample, wherein said gene is BACH1;

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/GB2014/053308

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2007129113 A2	15-11-2007	CA 2650443 A1 EP 2013364 A2 EP 2194143 A2 EP 2471951 A1 US 2009104615 A1 WO 2007129113 A2	15-11-2007 14-01-2009 09-06-2010 04-07-2012 23-04-2009 15-11-2007
WO 2009036513 A1	26-03-2009	NONE	
US 2012040344 A1	16-02-2012	CN 102439173 A EP 2391730 A2 JP 2012516148 A US 2012040344 A1 WO 2010086663 A2	02-05-2012 07-12-2011 19-07-2012 16-02-2012 05-08-2010