


# Obesity susceptible novel DNA methylation marker on regulatory region of inflammation gene: results from the Korea Epigenome Study (KES)

In-Uk Koh,<sup>1</sup> Nak-Hyeon Choi,<sup>1</sup> Kibaick Lee,<sup>1</sup> Ho-Yeong Yu,<sup>1</sup> Jun Ho Yun,<sup>1</sup> Jin-Hwa Kong,<sup>1</sup> Hyo Jin Kim,<sup>2</sup> Song Lee,<sup>3</sup> Song Cheol Kim,<sup>3</sup> Bong-Jo Kim,<sup>1</sup> Sanghoon Moon <sup>1</sup>

**To cite:** Koh I-U, Choi N-H, Lee K, *et al.* Obesity susceptible novel DNA methylation marker on regulatory region of inflammation gene: results from the Korea Epigenome Study (KES). *BMJ Open Diab Res Care* 2020;**8**:e001338. doi:10.1136/bmjdr-2020-001338

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/bmjdr-2020-001338>).

I-UK and N-HC contributed equally.

Received 12 March 2020  
Revised 3 June 2020  
Accepted 2 July 2020



© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

**Correspondence to**  
Dr Sanghoon Moon;  
[moon.sanghoon@daum.net](mailto:moon.sanghoon@daum.net)  
and Dr Bong-Jo Kim;  
[kbj6181@korea.kr](mailto:kbj6181@korea.kr)

## ABSTRACT

**Introduction** Obesity is growing global health concern and highly associated with increased risk of metabolic diseases including type 2 diabetes. We aimed to discover new differential DNA methylation patterns predisposing obesity and prioritize surrogate epigenetic markers in Koreans.

**Research design and methods** We performed multistage epigenome-wide analyses to identify differentially expressed CpGs in obesity using the Illumina HumanMethylationEPIC array (EPIC). Forty-eight CpGs showed significant differences across three phases: 902 whole blood DNAs from two cohorts (phase 1: n=450, phase 2: n=377) and a hospital-based sample (phase 3: n=75). Samples from phase III participants were used to examine whether the 48 CpGs are significant in the fat tissue and influenced gene expression. Furthermore, we investigated the epigenetic effect of CpG loci in childhood obesity (n=94).

**Results** Seven of the 48 CpGs exhibited similar changes in the fat tissue along with gene expression changes. In particular, hypomethylated CpG (cg13424229) on the GATA1 transcription factor cluster of CPA3 promoter was related to its increased gene expression and showed consistent effect in childhood obesity. Interestingly, subsequent analysis using RNA sequencing data from 21 preadipocytes and 26 adipocytes suggested CPA3 as a potential obesity-related gene. Moreover, expression patterns from RNA sequencing and public Gene Expression Omnibus showed the correlation between CPA3 and type 2 diabetes (T2D) and asthma.

**Conclusions** Our finding prioritizes influential genes in obesity and provides new evidence for the role of CPA3 linking obesity, T2D, and asthma.

## INTRODUCTION

Obesity is a complex trait determined by multiple factors, including the genome, epigenome, environment, and personal behavior. It is a significant risk factor of type 2 diabetes (T2D), cardiovascular disease, and related metabolic and inflammatory disturbances.<sup>1</sup> The changes in lifestyle over the past decades have caused a global obesity epidemic, with

## Significance of this study

### What is already known about this subject?

- DNA methylation can be altered on changes in genetic, environmental, and lifestyle factors.
- In the pathogenesis of obesity, DNA methylation patterns in human samples have been analyzed to reveal novel genes and pathways.

### What are the new findings?

- In multiphase study of Korean subjects, obesity specific differential methylation patterns were prioritized with intertissue and interage consistencies.
- CPA3, well-known inflammation marker, was hypomethylated (cg13424229) and showed decreased expression in obesity.

### How might these results change the focus of research or clinical practice?

- Prioritized novel obesity-related DNA methylation changes in this study could be surrogates in blood for adiposity in fat tissue and treatment targets in the concept of obesity as an inflammatory disease.
- Differential methylation of inflammation marker CPA3 in obesity may be a mechanism linking obesity to clinical conditions of asthma and type 2 diabetes.

an estimate of over 700 million affected people in 2015.<sup>2</sup> In South Korea, according to the guideline for obesity definition from the Korean Society for the Study of Obesity, the prevalence of adult obesity (Class I obesity is defined as body mass index (BMI) 25 kg/m<sup>2</sup> to less than 30 kg/m<sup>2</sup>) increased from 29.7% in 2009 to 32.4% in 2015.<sup>3</sup> A recent report revealed that the estimated obesity prevalence in 2030 would continue to increase by 1.5-fold and 1.4-fold in men and women, respectively, compared with that in 2015.<sup>4</sup> Moreover, as the incidence of T2D, hypertension, and dyslipidemia gradually increases with increase in the

BMI, the socioeconomic costs associated with obesity in Korea are also expected rise.<sup>3</sup>

In the past decade, several genome-wide association studies (GWAS) identified genetic loci predisposing to obesity. Over 500 loci were reported to be associated with obesity including, BMI and waist-hip ratio.<sup>5</sup> However, genome-wide estimates suggested that common variants account only for ~20% of the obesity to date.<sup>6</sup> Most known obesity-related genetic associations account only for a small portion of the disease or trait heritability. Therefore, considering the complex etiology of obesity, possible associations from nongenetic factors including epigenetic and environmental factors need to be explored.

DNA methylation is the reversible and heritable addition of a methyl group to a nucleotide<sup>7</sup> and a key regulator of gene expression and phenotype.<sup>1</sup> Genetic, environmental, and lifestyle factors can alter the extent of DNA methylation<sup>7</sup> and thus contribute to the pathogenesis of various diseases. Studying DNA methylation patterns may reveal novel genes and pathways involved in the pathogenesis of obesity that might be missed by GWAS.<sup>8</sup> It has been reported that blood lipids affect DNA methylation in the circulating cells.<sup>9</sup> Moreover, obesity has also been associated with genome-wide DNA methylation changes in the peripheral blood.<sup>10</sup>

In this study, we performed epigenome-wide analyses of DNA methylation and gene expression to prioritize reliable genes influenced by epigenetic marks and identify new biological targets in obesity. We also investigated whether genes affected by the newly identified hypomethylated marker are related to obesity, T2D, and asthma.

## RESEARCH DESIGN AND METHODS

### Study subjects and phenotypic measurements

The Korea Epigenome Study (KES) subjects for blood methylome analyses were recruited from the Health Examinees Study and Ansan and Ansung Study in the Korean Genome and Epidemiology Study population cohorts.<sup>11 12</sup> A total of 450 males (average age: 47.4, SD=5.9) were included in our discovery analysis (phase

I) of differentially methylated CpG sites in obesity. An additional 373 males (average age: 59.5, SD=8.3) were included in phase II. Seventy-five hospital-based individuals (average age: 57.3, SD=13.2) were recruited in phase III from whom blood and fat tissue were collected for the analysis of differential methylation and correlation with gene expression (table 1). Following the multistage epigenome-wide analyses in phases I–III (online supplementary figure S1), 94 individuals were selected from the Korean Child-Adolescent Cohort Study<sup>13</sup> to study differential methylation in childhood obesity.

Tissue samples were obtained for extracting single-type cells for studying the role of methylome and (regulation of) transcriptome in adipocyte differentiation from hospital-based subjects recruited to the KES, a study in collaboration with the International Human Epigenome Consortium (IHEC).<sup>14</sup> Along with the analysis performed in multistage discovery and replication for prioritization of DNA methylation markers (online supplementary figure S1), all datasets generated and used in this study are listed in online supplementary table 1. In the study, obesity was defined by BMI; BMI of both  $\geq 30$  and  $\geq 27$  kg/m<sup>2</sup> was considered obese based on the Asia-Pacific standards of WHO<sup>15</sup> and the newly proposed standards for the Korean population.<sup>16</sup>

### Genome-wide DNA methylation profiling

DNA samples for population-based DNA methylation profiling were obtained from the National Biobank of Korea, with written informed consent, following the Korea National Institute for Bioethics Policy and the corporation's Institutional Review Board (IRB) (approval number: P01-201703-31-004 and LASIRB-201180222-001/002). For hospital-based DNA methylation profiling, whole blood and fat tissues were obtained from the Asan Medical Center (AMC) in Seoul, South Korea. The IRB of AMC approved this study (2013–0619), and all subjects provided informed consent. The fat tissue was purified into mature adipocytes and preadipocytes using collagenase I digestion followed by centrifugation and MACS/FACS (Magnetic activated cell sorting/Fluorescence

**Table 1** General characteristics of the subjects for analysis on obesity-related differentially methylated CpGs

	Phase I: Discovery/ population		Phase II: Replication 1/ population		Phase III: Replication 2/ hospital		Childhood obesity	
	Lean	Obese*	Lean	Obese†	Lean	Obese†	Lean	Obese*
N	250	200	240	133	46	29	46	48
Gender (male)	250 (100%)	200 (100%)	240 (100%)	133 (100%)	18 (39%)	11 (38%)	46 (100%)	48 (100%)
Age (years)	47.2±5.9	47.7±5.9	61.0±8.8	56.9±7.5	58.2±14.4	55.9±11.2	14.0±0.8	13.9±0.8
BMI (kg/m <sup>2</sup> )	21.4±1.1	31.6±1.6	21.4±1.1	28.5±1.4	20.8±1.4	30.3±3.0	19.7±1.4	34.8±3.7

\*BMI>30 kg/m<sup>2</sup>.

†BMI>27 kg/m<sup>2</sup>.

BMI, body mass index.

activated cell sorting).<sup>17</sup> Genomic DNA was extracted from hospital-based blood samples and purified cells using the QIAamp DNA mini kit (Qiagen, Hilden, Germany).

The Infinium EPIC chip (Illumina, California, USA), composing >800 000 methylation marks, was used for methylome profiling. Illumina intensity data (IDAT) files were imported to obtain detailed bead-level information. Focusing on CpG sites and eliminating potential artifacts, non-CpG sites, single nucleotide polymorphisms (SNPs), and probes on sex chromosomes were removed during the filtering process. Probes with a detection  $p > 0.01$  in all samples and low SD were also excluded. Raw IDAT files were parsed into the analysis pipeline for DNA methylation adapted basically from Bioconductor Minfi (<https://github.com/hansenlab/minfi>) and Enmix R packages (<https://bioconductor.org/packages/release/bioc/html/ENmix.html>). Preprocessing of background correction and dye-bias adjustment was performed, and normalizations of between-sample bias (quantile) with probe-type bias (BMIQ) were performed. In addition, to find reliable differentially methylated positions, we estimated the impact of technical batch (ie, chip slides and arrays) by pcrplot function in the Enmix package, and blood cell counts using estimateCellCounts2 function in the Minfi package.

The beta-value was calculated as the intensity of the methylated channel divided by the sum of the methylated and nonmethylated probe intensities, reflecting the methylation level at each CpG site (0–1.0 to signify per cent methylations of 0%–100%). Corrected and normalized methylation beta-values were transversed into M-values ( $\log_2(\text{beta-value}/(1 - \text{beta-value}))$ ), and linear regression analysis was conducted with adjustment of cell types (blood DNA only) and technical batch effects, using the limma package (<http://bioinf.wehi.edu.au/limma>).

### Gene expression assays

Total RNA from fat tissues and purified cells (adipocytes and preadipocytes) were extracted using an RNeasy Lipid tissue kit (Qiagen, Hilden, Germany) and RNeasy Micro kit (Qiagen, Hilden, Germany) following the manufacturer's recommendations. Gene expression profiling was performed on mRNA samples of fat tissues ( $n=75$ ) using the HT-12 v4 expression array (Illumina, San Diego, California, USA) to identify differentially expressed genes (DEGs). Probes with a detection  $p > 0.01$  were removed, and a quantile normalization method was applied to reduce potential sources of nonbiological variation among samples using GenomeStudio software (Illumina, California, USA). A two-sample *t*-test was performed to compare the obese and lean groups (29:46), and 2166 DEGs with  $p < 0.05$  were identified.

RNA samples from purified adipocytes/preadipocytes were prepared into cDNA libraries, using TruSeq (Stranded) mRNA Sample Preparation Kit, and sequenced with the HiSeq platform (Illumina, San Diego, California, USA). The nonaligned raw

sequencing reads of the samples were preprocessed to eliminate low-quality reads and library adapter residue. Preprocessed reads were aligned on the human reference genome hg19 using STAR aligner with default parameters. Gene expression estimates of the read counts were calculated and compared between the obese and lean samples using RSEM 1.3.0 (<https://deweylab.github.io/RSEM>).

### Correlations between DNA methylation and gene expression

Spearman's correlation coefficients were calculated between gene expression and DNA methylation in fat tissues using the 'cor' function and visualized using the corrplot package (<https://github.com/taiyun/corrplot>) in R.

### ChIP-seq and chromatin state annotation

Libraries for ChIP-sequencing were constructed according to the Illumina's standard protocol and sequenced on Illumina HiSeq 2500 platform, with 100 bp paired-end reads. For each sample, six histone modification marks and input (control) were sequenced. All ChIP-seq reads were mapped to the human reference genome hg19, and unmapped or duplicated reads were marked and eliminated by Picard (V.2.5.0, <https://broadinstitute.github.io/picard/>) and samtools (V.0.1.19, <http://www.htslib.org>).

To investigate the chromatin states, we performed ChromHMM (V.1.14, <http://compbio.mit.edu/ChromHMM>) analysis with 12 samples obtained from three different tissue types. Chromatin states for the samples were annotated using the 'Expanded 18-state model' proposed by the Roadmap Epigenomics Consortium.<sup>18</sup>

### Association of CPA3 on obesity, T2D, and asthma

To examine the association of newly identified genes and other diseases, we conducted additional orthogonal analyses. First, we performed an epigenome-wide association study (EWAS) meta-analysis of novel marker on T2D using three different datasets with >2300 cohort participants (online supplementary table S1). Two different adjusting algorithms for cell-type heterogeneity were used: ReFACTor (<https://github.com/cozygene/refactor>) for reference-free and RefbaseEWAS in ChAMP package (<https://bioconductor.org/packages/release/bioc/html/ChAMP.html>) for reference-based (online supplementary table S3). Second, reanalyses of CPA3 gene expression on obesity, T2D, and asthma were performed on publicly available Gene Expression Omnibus (GEO) (<http://www.ncbi.nlm.nih.gov/geo>) datasets on human and mice adipocytes (GDS3601 and GDS2743, respectively), human pancreatic islets (GDS3782), and human PBMC (GDS3615) (table 2). The raw data (.CEL file) were visualized using ggplot2 (<https://ggplot2.tidyverse.org>) in R.

**Table 2** EWAS and/or GWAS (within 200 kb of CpG) replication of 48 differentially methylated CpG sites

CpG_ID	Chromosome:Position	Gene	Gene_Group	Obesity(-related) associations in EWAS catalog ( $p < 1 \times 10^{-3}$ )	Obesity(-related) association in EWAS Atlas (excluded $p = NR$ )	Obesity(-related) association in GWAS catalog (within 200 Kb of CpG, $p < 5 \times 10^{-8}$ )
<i>Obesity(-related) CpGs known in EWAS*</i>						
cg06500161	21:43 656 587	<b>ABCG1</b>	Body	BMI, TG, WC, T2D, Fasting Glucose	T2D, BMI, Blood TG, MetS	–
cg27243685	21:43 642 366	ABCG1	5'UTR	BMI, TG, WC, Total lipids	BMI	–
cg01082498	11:68 608 225	CPT1A	5'UTR	TG, LDL cholesterol	LDL cholesterol, Blood TG	BMI
cg14465376	3:81 766 440	GBE1	Body	TG, Cholesterol, HC	–	BMI, WC
cg13001289	17:1 415 550	INPP5K	Body	Total cholesterol	Obesity	–
cg02335306	11:2 559 880	KCNQ1	Body	Free cholesterol to total lipids ratio in small VLDL	–	T2D
<b>cg16744911</b>	10:102 108 974	<b>SCD</b>	Body	–	Birth weight	Fatty acid desaturase activity (fat tissue), Total cholesterol
cg11024682	17:17 730 094	SREBF1	Body	BMI, TG, CRP, WC, T2D	BMI, WC, Blood TG, T2D	BMI, T2D
<i>Novel CpGs but known gene in GWAS†</i>						
cg20655350	2:219 124 263	GPBAR1	TSS1500	–	–	WHR
cg04321608	22:50 689 962	HDAC10	TSS200	–	–	BMI
cg17827949	2:207 974 984	MIR2355;KLF7	TSS200;Body	–	–	BMI
<b>cg18706511</b>	6:46 350 774	<b>RCAN2</b>	Body	–	–	BMI
cg11884741	11:134 449 995	–	–	–	–	BMI
cg23386509	16:80 843 728	–	–	–	–	WHR
cg23891633	12:11 695 118	–	–	–	–	Subcutaneous adipose tissue
cg27413950	14:60 789 354	–	–	–	–	Hip circumference adjusted for BMI
<i>Novel CpGs but known gene in GWAS† related to complex trait</i>						
<b>cg01269795</b>	6:26 440 101	<b>BTN3A3</b>	TSS1500	–	–	TG
<b>cg13424229</b>	3:148 581 837	<b>CPA3</b>	TSS1500	–	–	Birth weight
cg00857282	6:16 130 727	MYLIP	Body	–	–	Cholesterol, LDL cholesterol, Total cholesterol
cg13980454	14:36 985 905	NKX2-1	3'UTR	–	–	T2D
cg14546197	17:38 137 043	PSMD3	TSS200	–	–	Subcutaneous adipose tissue
cg03904213	21:40 457 407	–	–	–	–	TG
<i>Novel CpGs in both GWAS† and EWAS*</i>						
cg16740586	21:43 655 919	ABCG1	Body	–	–	–
<b>cg14234394</b>	1:212 787 866	<b>ATF3</b>	5'UTR;TSS1500	–	–	–
cg23762306	16:3 817 973	CREBBP	Body	–	–	–
cg05779336	16:23 682 007	DCTN5	Body	–	–	–
cg01579884	7:37 004 390	ELMO1	5'UTR	–	–	–
cg21739530	4:174 186 068	GALNT7	Body	–	–	–

Continued



**Table 2** Continued

CpG_ID	Chromosome:Position	Gene	Gene_Group	Obesity(-related) associations in EWAS catalog ( $p < 1 \times 10^{-3}$ )	Obesity(-related) association in EWAS Atlas (excluded $p = NR$ )	Obesity(-related) association in GWAS catalog (within 200 Kb of CpG, $p < 5 \times 10^{-8}$ )
cg03134285	1:67 836 673	IL12RB2	Body	–	–	–
cg05935283	3:190 231 667	IL1RAP	TSS200	–	–	–
cg17909892	10:135 009 388	KNDC1	Body	–	–	–
<b>cg13279894</b>	4:151 600 922	<b>LRBA</b>	Body	–	–	–
cg22796849	19:17 530 340	MVB12A	TSS1500	–	–	–
cg19554457	12:93 774 772	NUDT4	Body	–	–	–
cg23610075	12:14 722 220	PLBD1	TSS1500	–	–	–
cg13689699	9:137 225 311	RXRA	Body	–	–	–
cg03001463	2:200 323 278	SATB2	5'UTR	–	–	–
cg19842683	10:102 795 320	SFXN3	Body	–	–	–
cg15833511	20:35 256 798	SLA2	Body	–	–	–
cg26257411	6:33 397 122	SYNGAP1	Body	–	–	–
cg26547928	12:33 589 385	SYT10	Body	–	–	–
cg27465654	2:71 221 575	TEX261	Body	–	–	–
cg10851763	21:43 545 792	UMODL1	Body	–	–	–
cg23487560	2:10 832 689	–	–	–	–	–
cg02006035	9:118 859 925	–	–	–	–	–
cg15730234	1:101 759 418	–	–	–	–	–
cg21098652	11:121 129 115	–	–	–	–	–
cg24271501	6:47 042 976	–	–	–	–	–

IDs of CpG sites and genes in bold indicate that differential methylation was correlated with differences of gene expression, and each obesity (-related) CpGs were differentially methylated in both blood and fat tissue.

\*EWAS: EWAS catalog (<http://www.ewascatalog.org/>) and EWAS atlas (<https://bigd.big.ac.cn/ewas>).

†GWAS: GWAS catalog (<https://www.ebi.ac.uk/gwas/>).

BMI, body mass index; CRP, C-reactive Protein; EWAS, epigenome-wide association study; GWAS, genome-wide association studies; HC, Hepatic Cholesterol; LDL, Low-density lipoprotein; T2D, type 2 diabetes; TG, Triglyceride; WC, Waist Circumference; WHR, waist-hip ratio.

## RESULTS

### Discovery and replication phase analyses

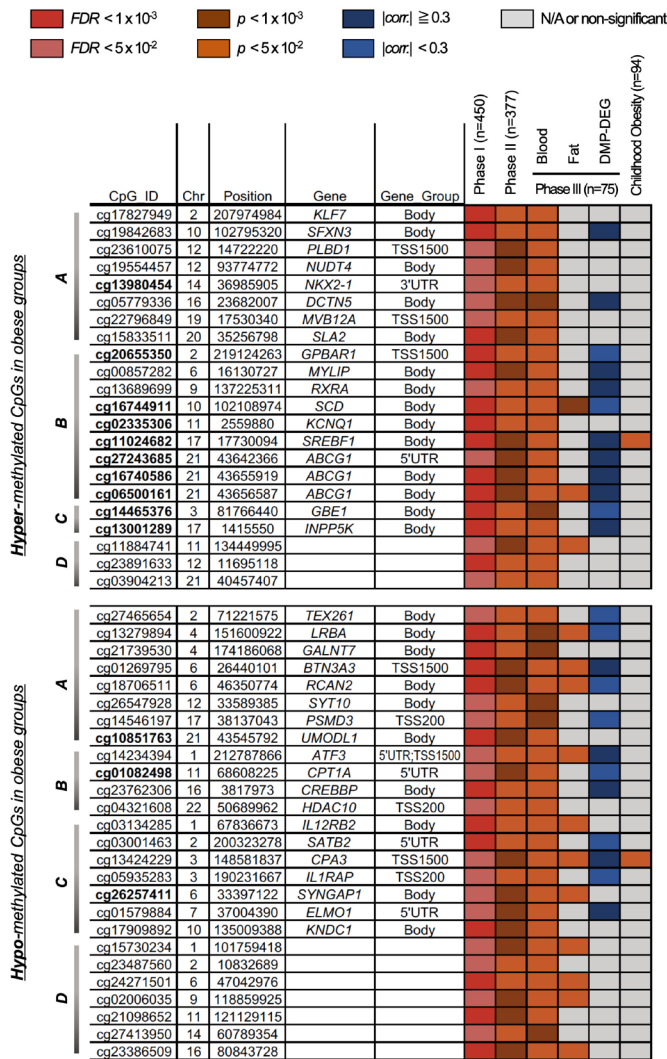
**Table 1** lists the general characteristics of the subjects from the discovery cohort ( $n=450$ , phase I), and the replication cohorts ( $n=373$  and  $75$ , phase II and III, respectively). We identified 142 CpG sites with significant differential methylation patterns (DMPs, false discovery rate (FDR) $<0.05$ ) between obese and lean subjects of phase I, and filtered with significance ( $p < 0.05$ ) across phases II and III (online supplementary figure S1). In addition, among the replicated 142 DMPs, 48 sites (of 38 genes) showed unidirectional changes in three phases (22 hypermethylation and 26 hypomethylation, online supplementary table S2, **figure 1** and online supplementary figure S2). The most significant CpG site (cg06500161,  $FDR=4.38 \times 10^{-9}$ ) was located in the gene body of ABCG1 (ATP-binding cassette subfamily G member 1). Other CpG sites (cg16740586,  $FDR=0.0055$ ; cg27243685,  $FDR=0.0413$ ) for the same gene also passed the threshold for significance. CpG sites cg06500161 and cg27243685 have been reported to be associated with BMI in previous studies;<sup>19</sup> however, cg16740586 was newly

identified as obesity-related ABCG1 methylation located 668 bp and 13 553 bp away from cg06500161 and cg27243685, respectively. Differential methylation of cg00857282 was also a significant hit ( $FDR=1.55 \times 10^{-8}$ ) and novel for a previously reported obesity-related gene, MYLIP.<sup>20</sup>

Among the CpG sites with methylation differences in phases I and II, only five CpG sites (in three genes) showed significance ( $FDR < 0.05$ ) in both phases (online supplementary table S4). Cg00574958 and cg17058475 of CPT1A, and cg06500161 of ABCG1 were consistent with previous reports for obesity-related methylation changes,<sup>19 21</sup> with an additional CpG site (cg05325763) for CPT1A. Cg02988288 was identified as a significant CpG site (Phase I,  $FDR=6.01 \times 10^{-4}$ ; Phase II,  $FDR=1.79 \times 10^{-3}$ ) located in TXNIP, previously known to be linked to obesity and metabolic diseases.<sup>22</sup>

### Significant DMPs annotated with genes of obesity-related mechanisms

Of the 142 differentially methylated sites (of 111 genes) with significance in obese groups across the phases, 67



**Figure 1** Forty-eight DMPs in obese groups in multiple phases of analyses (Phase I: Discovery and Phase II and III: Replication). A total of 22 hypermethylated and 26 hypomethylated CpGs and annotated genes in obesity were categorized into four gene-groups from IPA results—Gene-group A (lipid metabolism, small molecule biochemistry, cellular compromise), B (lipid metabolism, small molecule biochemistry, vitamin and mineral metabolism), C (cell-mediated immune response, cellular development, cellular function and maintenance), and D (uncategorized) from IPA results. DMPs were further prioritized with intertissue (fat and blood) and/or interage (childhood and adult) consistency and methylation-expression correlation in fat tissue. DEG, differentially expressed genes; DMPs, differentially methylated CpGs; FDR, false discovery rate.

CpG sites (45 genes) were hypomethylated and 75 sites (66 genes) hypermethylated in phase I discovery analysis (online supplementary table S2). DMPs located at gene body or the 3'UTR were hypomethylated (28/59) and hypermethylated (31/59) in obesity. In contrast, the majority of DMPs (35 out of 52) located at the upstream regulatory regions (TSS or 5'UTR) were hypermethylated in obesity. Among them, 23 genes, including ATF3, CPT1A, IL1RAP, HDAC10, SYNGAP1, and UMODL1, were the most hypomethylated genes in obesity, and 26

genes, including FAM59B, HNRNPA1, MSH2, MVB12A, and RNF182, were the most hypermethylated ones in obesity (|difference in methylation|>2%).

We performed gene-level pathway analyses, using IPA, for the 111 genes identified in our multistage differential DNA methylation analyses of obese versus control and observed significant enrichment of 50 specific gene networks. The top ones were 'Loss of adipose tissue' (p=6.53×10<sup>-6</sup>), 'Dysglycemia' (p=2.08×10<sup>-5</sup>), 'Homeostasis of cholesterol' (p=4.09×10<sup>-5</sup>), and 'Oxidation of fatty acids' (p=5.68×10<sup>-4</sup>) (online supplementary table S5).

### Identification of correlations between differential fat DNA methylation and gene expression

In a correlation analysis between gene expression and DNA methylation in fat tissue of same individuals in phase III (n=75), differences in DNA methylation in 111 CpG sites were correlated with expression of 113 genes (|spearman's correlation|>0.2; 57 positive and 59 negative correlations) (online supplementary table S6). Among these, 17 CpG sites showed significant and consistent differences in blood DNA methylation of case-control comparison in Phase I–III. Most CpG sites (12 out of 17) were located on the body and 5'UTR of annotated genes. However, CpG methylations of five genes, FAM125A, CPA3, BTN3A3, ATF3, and IL1RAP were located at the genomic regions of TSS1500 and TSS200 for transcriptional regulation (Gene\_group in online supplementary table S7).

### Prioritization of differentially methylated sites discovered and replicated

Obesity-related DMPs with unidirectional changes in multiphased case-control analyses were prioritized using a combination of the consistency of methylation changes in fat tissues as in blood samples and correlations between differential methylation and expression in fat tissues (figure 1). Among the 48 DMPs discovered and replicated in blood samples, methylation values of 24 CpG sites showed correlation with expression of the annotated genes, and 14 were consistently significant in methylation differences both in fat and blood (n=75, p<0.05). Consequently, seven obesity-related DMPs—cg16744911, cg06500161, cg01269795, cg14234394, cg18706511, cg13279894, and cg13424229—annotated to genes SCD, ABCG1, BTN3A3, ATF3, RCAN2, LRBA, and CPA3, respectively, were overlapped between combinations of additional analyses in fat tissue (online supplementary figures S2 and S3). Most of prioritized DMPs (five of seven) were novel CpGs that have not been previously reported as obesity (-related) CpGs in EWA-studies (bold in table 2). Alongside ABCG1 for novel site (cg16740586), CpGs of ATF3 (cg14234394) and LRBA (cg13279894) were novel in both EWA- and GWA-studies (table 2, 'Novel CpGs in both GWAS and EWAS'). BTN3A3 (cg01269795), CPA3 (cg13424229), and RCAN2 (cg18706511) were also novel CpGs in EWAS

with evidences in GWAS catalog, but the latter one only had associations with obesity directly (table 2, 'Novel CpGs but known gene in GWAS').

### Two obesity-related DMPs (cg13424229 and cg17730094) were also significant in childhood obesity

We performed additional analysis for differential methylation in blood samples of childhood obesity (n=94, table 1) and identified 32 CpG sites (16 hypermethylated and 16 hypomethylated) with significance ( $p < 0.05$ ; online supplementary table S8). Among these childhood obesity-specific DMPs, only cg13424229 ( $p = 0.04$ , CPA3) and cg17730094 ( $p = 0.02$ , SREBF1) were found in 48 DMPs of multiphased obesity-specific analyses and showed consistent directions in methylation changes in obese groups (figure 1 and online supplementary figure S2). Cg17730094 located within the gene body of SREBF1, previously reported as obesity-specific DNA methylation of an obesity-related gene, was hyper-methylated in obese groups of childhood and adult cohorts (Phase I–III). Contrastingly, cg13424229 was hypomethylated in obesity and located in the promoter region (TSS1500) of CPA3 (online supplementary figure S4), which is a well-known gene for inflammation and immune pathways in several tissues/cells but has not been reported as an obesity-specific CpG site.<sup>23</sup>

### Association of CPA3 gene with obesity, T2D, and asthma

To assess the potential regulation and differences in expression of CPA3 in fat tissue, we analyzed RNA sequencing data of cells isolated from the fat tissue (n=47, 21 preadipocytes, and 26 adipocytes) of Korean hospital-based subjects collected for the KES in collaboration with IHEC.<sup>14</sup> Gene expression of CPA3 was significantly higher in adipocytes than in preadipocytes (figure 2A) with BMI-correlated patterns (data not shown). In particular, seven subjects (online supplementary table S9; BMI range

23.8–43.4) with both types of cells showed increases in CPA3 expression correlated with BMI (figure 2B;  $r = 0.59$  in preadipocytes and 0.39 in adipocytes).

The EWAS meta-analysis showed that cg13424229 was significantly associated with T2D considering both adjusting algorithms ReFACTor ( $p = 1.17 \times 10^{-4}$ ) and RefbaseEWAS ( $p = 7.00 \times 10^{-11}$ ) (online supplementary table S3).

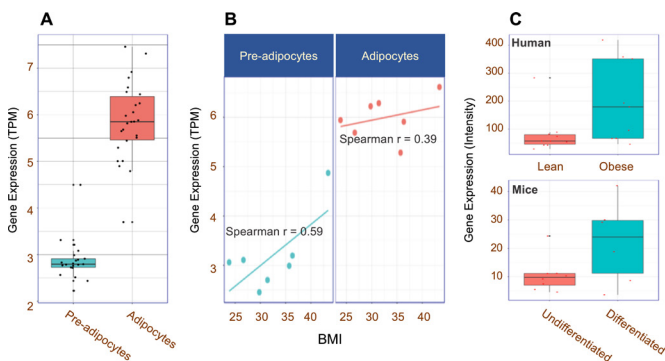
In the reanalysis of gene expression data of human and mice from public GEO (GDS2743 and GDS3601), we confirmed the overexpression of CPA3 in the adipocytes of obese subjects and in more differentiated cells (figure 2C). The relative expressions of CPA3 in T2D (GDS3782) and asthma (GDS3615) was also elevated compared with that in healthy individuals (online supplementary figure S6).

### DISCUSSION

In this study, we identified several obesity-specific CpG sites with targeted genes using multistage analyses of the blood DNA methylome. In addition, we prioritized obesity-susceptible CpG sites and genes, in children and older individuals, using differential methylation analyses and correlations with gene expression levels in fat tissue. Even in our relatively small sample size (n=902 in three phases), we found and replicated DNA methylation changes in CpG sites and/or annotated genes, many of which were previously reported in the studies of BMI-methylation associations in larger samples (n=7983 or more).<sup>1–24</sup> In addition to ABCG1 (cg06500161), GPBAR1 (cg20655350), SCD (cg16744911), and SREBF1 (cg11024682), of which CpG sites were consistently reported as obesity-specific DNA methylation changes,<sup>25</sup> we identified novel loci designating obesity-related genes such as CPT1A (cg01082498), MYLIP (cg00867282), SYNGAP1 (cg26257411), and KCNQ1 (cg02335306).<sup>26–28</sup>

By comparing the DNA methylome of the adipose tissue and blood in the same individual, we identified 14 potential surrogate CpG sites in blood for adipose tissue (figure 1, DMPs at Phase III Fat). The 14 DMPs identified in blood and fat were mostly hypomethylated (11 out of 14), and the hypomethylated sites were on genes related to inflammatory/stress response and immune mechanism (ATF3, IL12RB2, CPA3, LRBA, and RCAN2), whereas the hypermethylated sites were on genes related to lipid metabolism (ABCG1 and SCD). In addition, the CpG sites and genes identified as obesity-specific DMPs in the blood samples showed concordant methylation changes and correlations both in methylation changes and gene expression in the fat tissue (online supplementary table S6).

Since DNA methylation is dynamic and influenced by multiple environmental factors, even in chronological manners, we analyzed differential methylation in childhood obesity in addition to the DNA methylome from cohorts of aged subjects. In this analysis, only 32 CpG sites were differentially methylated ( $p < 0.05$ ) in obese childhood subjects, with the replication of two CpGs



**Figure 2** Expression of CPA3 in fat cells. (A) Boxplot showing relative expression of CPA3 in preadipocytes (n=21) and adipocytes (n=26). (B) Relationships between relative expression of CPA3 and BMI in preadipocytes and adipocytes from same subjects (n=7). (C) Relative CPA3 expression differences in human (lean and obese, n=19) and mice (undifferentiated and differentiating, n=14) models. The original data were downloaded from GEO (GDS3601, GDS2743). BMI, body mass index; GEO, Gene Expression Omnibus.



(SREBF1 and CPA3) observed in aged subjects (Phase I–III). The relatively smaller number of DMPs and replication might be due to the difference in the platform used for DNA methylome analysis (450K vs EPIC). However, cg11024682 in the gene body of SREBF1 is known to be obesity-related in other populations,<sup>25</sup> and SREBF1 encodes sterol regulatory element-binding transcription factor 1, known to regulate sterol biosynthesis in interaction with molecules of glucose and fatty-acid/lipid metabolism,<sup>29</sup> including CREBBP, which showed differential methylation (cg23762306,  $FDR=1.73 \times 10^{-4}$ ) in multiphased blood samples also (online supplementary table S2). Interestingly, one of the two CpG sites with obesity-specific methylation changes in both childhood and aged subjects, cg13424229 in the promoter region (–1206 bp) of CPA3, was also observed in the fat tissue and showed correlation with gene expression changes ( $r=-0.55$ ).

Carboxypeptidase A3, encoded by CPA3, is an enzyme located in the secretory granule of several types of cells, mostly mast cells and mast-cell-like lines.<sup>30</sup> CPA3 is reported to be involved in defense mechanisms with inflammation, and its upregulation is a useful biomarker of allergic inflammation.<sup>31</sup> The expression of CPA3 was higher in fat tissue of obese group; it was also higher in the differentiated fat cells (adipocytes) compared with that in the precursors (preadipocytes) (figure 2A). On comparing CPA3 expression in two different cell types from the same subject ( $n=7$ ), we observed that the expression in fat cells correlated with the BMI (figure 2B,  $r=0.59$  in preadipocytes and 0.39 in adipocytes). These BMI-related and cell-type-specific increments in CPA3 expression implicate the possible role of CPA3 in adipocyte differentiation and/or adipogenesis mechanism.<sup>32</sup> Similarly, in a transcriptome analysis of brown and white adipocyte differentiation in a mouse model, increase in CPA3 was followed by fat cell differentiation.<sup>33</sup> Our findings of increased expression of CPA3 with decreased methylation of cg13424229 in obesity and mature adipocytes may imply its epigenetic role in differentiation or proliferation of preadipocytes. In particular, the obesity-related hypomethylated CpG site, cg13424229, resides within the GATA binding cluster in the promoter region of CPA3 (online supplementary figure S4), which was reported as one of the specific target genes of GATA-1,<sup>34</sup> a transcription repressor with cell-type specificity (online supplementary figure S5). In addition, as one of the genetic variants (rs10935733) of CPA3 was reported to be associated with birth weight,<sup>35</sup> and as maternal obesity could affect CPA3 expression in the fetus,<sup>36</sup> the differential methylation of CPA3 might be contributing to the pathogenesis of obesity.

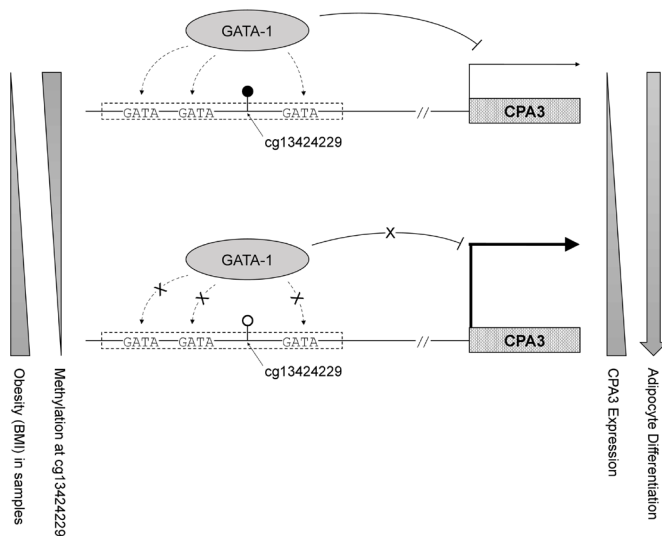
Since obesity is considered an inflammatory disease and is associated with increase in several inflammatory mediators in the diseased tissues, including the adipose tissue,<sup>37</sup> the mast cells, which play critical roles in the release of inflammatory mediators and CPA3, as one of the designated markers might have an implication in

obesity.<sup>38</sup> Interestingly, besides inflammation, CPA3 has also been reported as a marker for adipocyte differentiation under both obese and nonobese conditions;<sup>32</sup> and CPA3 protein expression was downregulated with concomitant differences in the immune system, complement activation, and renin-angiotensin system.<sup>39</sup> In human (Pima Indians) profiles of gene expression in abdominal subcutaneous adipocytes, levels (TPM) of CPA3 were higher in the obese group with an overrepresentation of genes in the inflammation/immune response categories.<sup>40</sup> In particular, methylation changes of cg13424229 in the regulatory promoter region of CPA3 was significant in multi-stage analysis of obesity (figure 1) and meta-analysis of type 2 diabetes (online supplementary table S3), as the gene expression was also increased in conditions of obesity/adipogenesis (figure 2), T2D, and asthma (online supplementary figure S6). Since epigenetics is one of the mechanisms linking environmental factors to altered gene activity, differential methylation of CPA3 in obesity may also be a mechanism linking obesity to clinical conditions of other metabolic severities such as asthma and T2D.

Our study of differential methylation in Koreans confirmed previous reports and revealed novel CpG sites in relation to obesity. Furthermore, we reported obesity-specific DMPs, prioritized with intertissue (fat and blood) and interage (young and middle-aged) consistencies in changes of DNA methylation and correlations to gene expression in fat tissue. Among the 48 obesity-specific DMPs, most DMPs (26/48) were also SNPs, and there were known genomic associations (GWAS catalog  $p<5 \times 10^{-8}$ ) in neighbors. However, almost none of the CpG sites have been reported to be associated with obesity or related diseases/traits in (epi-)genome-wide analyses (online supplementary table S10). Therefore, with the evidences that differential methylation was correlated with gene expressions in obesity and related mechanisms, the regulation of gene expression via DNA methylation may explain an additional component of interindividual variation in lipid levels beyond genetic sequence variants, reinforcing the epigenetic complementation of molecular mechanisms associated with genetic variations.

However, further studies are needed to clarify these relationships. Studies on the methylome of obesity performed in larger cohort might replicate our prioritized obesity CpGs with relatively low significance in this study. To validate the differential methylation of cg13424229 in blood as a causal candidate, tissue-level expression of CPA3 and the epigenetic regulation of the promoter methylation in obesity and related diseases (ie, asthma or T2D) need to be studied in multiple models. In addition, GATA binding cluster of CPA3 promoter that cg13424229 resides in should further be assessed for the possibilities of other GATA proteins with zinc-finger domain, since GATA family members share consensus DNA binding motifs and some are known to be required in differentiation of mesoderm-derived mesenchymal stem cells into adipocytes.<sup>41–43</sup>





**Figure 3** Hypothetical epigenetic regulation model that hypomethylation of cg13424229 on the cluster of GATA binding motifs diminishes recruitment of GATA-1, transcriptional repressor of CPA3, in obesity and/or differentiating adipocytes. BMI, body mass index.

In conclusion, we show that blood DNA methylation (eg, cg13424229) status could differ in obese subjects. Further, the blood DNA methylation could be concordant to the DNA methylation in the fat tissues and correlated with the expression of obesity-related genes (eg, CPA3) (figure 3). Our study sheds light on how DNA methylation could influence to regulate gene expression in obesity and inflammation-related diseases.

#### Author affiliations

<sup>1</sup>Division of Genome Research, Center for Genome Science, Korea National Institute of Health, Cheongju-si, Chungcheongbuk-do, South Korea

<sup>2</sup>Division of Endocrine and Metabolic Diseases, Center for Biomedical Sciences, Korea National Institute of Health, Cheongju-si, Chungcheongbuk-do, South Korea

<sup>3</sup>Division of Hepato-Biliary and Pancreatic Surgery, Department of Surgery, Asan Medical Center, AMIST, University of Ulsan College of Medicine, Asan Medical Center, Songpa-gu, Seoul, South Korea

**Contributors** I-UK, B-JK, and SM conceptualized and supervised the study. N-HC, KL, and J-HK performed the data analysis and prepared figures and tables. H-YY, JHY, and SL performed experiments. HJK and SCK provided samples and clinical information. I-UK and SM coordinated the research, interpreted the data, searched the literature, and wrote the original draft of manuscript. I-UK, B-JK, and SM reviewed and edited the manuscript. B-JK is the guarantor/supervisor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

**Funding** This research was supported by intramural grants from the Korea National Institute of Health (2017-NI73002-02 and 2018-NI001-02). Data were provided by Reference Data Production of Regulation of Gene Expression (4848–308) and Korean Genome Analysis Project (4845–301) that were supported by the Korea Centers for Disease Control and Prevention, Republic of Korea.

**Competing interests** None declared.

**Patient consent for publication** Not required.

**Ethics approval** The Institutional review boards at Korea National Institute of Health approved this study (2017-03-02-2C-A).

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Data are available on reasonable request. The deidentified data generated and/or analyzed in this study are available from the

Division of Genome Research in Korea National Institute of Health ([www.nih.go.kr](http://www.nih.go.kr)) on reasonable requests.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

#### ORCID iD

Sanghoon Moon <http://orcid.org/0000-0001-5987-9505>

#### REFERENCES

- Wahl S, Drong A, Lehne B, *et al*. Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. *Nature* 2017;541:81–6.
- Toubal A, Treuter E, Clément K, *et al*. Genomic and epigenomic regulation of adipose tissue inflammation in obesity. *Trends Endocrinol Metab* 2013;24:625–34.
- Seo MH, Lee W-Y, Kim SS, *et al*. 2018 Korean Society for the study of obesity guideline for the management of obesity in Korea. *J Obes Metab Syndr* 2019;28:40–5.
- Baik I. Forecasting obesity prevalence in Korean adults for the years 2020 and 2030 by the analysis of contributing factors. *Nutr Res Pract* 2018;12:251–7.
- Loos RJ. The genetics of adiposity. *Curr Opin Genet Dev* 2018;50:86–95.
- Locke AE, Kahali B, Berndt SI, *et al*. Genetic studies of body mass index yield new insights for obesity biology. *Nature* 2015;518:197–206.
- Dick KJ, Nelson CP, Tsaprouni L, *et al*. DNA methylation and body-mass index: a genome-wide analysis. *Lancet* 2014;383:1990–8.
- Hedman Åsa K, Mendelson MM, Marioni RE, *et al*. Epigenetic patterns in blood associated with lipid traits predict incident coronary heart disease events and are enriched for results from genome-wide association studies. *Circ Cardiovasc Genet* 2017;10.
- Dekkers KF, van Iterson M, Sliker RC, *et al*. Blood lipids influence DNA methylation in circulating cells. *Genome Biol* 2016;17:138.
- Benton MC, Johnstone A, Eccles D, *et al*. An analysis of DNA methylation in human adipose tissue reveals differential modification of obesity genes before and after gastric bypass and weight loss. *Genome Biol* 2015;16:8.
- Health Examinees Study Group. The health examinees (HEXA) study: rationale, study design and baseline characteristics. *Asian Pac J Cancer Prev* 2015;16:1591–7.
- Kim Y, Han B-G, KoGES group. Cohort profile: the Korean genome and epidemiology study (KoGES) Consortium. *Int J Epidemiol* 2017;46:1350.
- Yoo S, Lee S-Y, Kim K-N, *et al*. Obesity in Korean pre-adolescent school children: comparison of various anthropometric measurements based on bioelectrical impedance analysis. *Int J Obes* 2006;30:1086–90.
- International Human Epigenome Consortium. The International human epigenome Consortium: a blueprint for scientific collaboration and discovery. *Cell* 2016;167:1145–9.
- WHO Expert Consultation. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. *Lancet* 2004;363:157–63.
- Yoon JL, Cho JJ, Park KM, *et al*. Diagnostic performance of body mass index using the Western Pacific regional office of World Health organization reference standards for body fat percentage. *J Korean Med Sci* 2015;30:162–6.
- Acosta JR, Douagi I, Andersson DP, *et al*. Increased fat cell size: a major phenotype of subcutaneous white adipose tissue in non-obese individuals with type 2 diabetes. *Diabetologia* 2016;59:560–70.
- Roadmap Epigenomics Consortium, Kundaje A, Meuleman W, *et al*. Integrative analysis of 111 reference human epigenomes. *Nature* 2015;518:317–30.
- Mendelson MM, Marioni RE, Joeannes R, *et al*. Association of body mass index with DNA methylation and gene expression in blood cells and relations to cardiometabolic disease: a Mendelian randomization approach. *PLoS Med* 2017;14:e1002215.
- Berisha SZ, Serre D, Schauer P, *et al*. Changes in whole blood gene expression in obese subjects with type 2 diabetes following bariatric surgery: a pilot study. *PLoS One* 2011;6:e16729.
- Demerath EW, Guan W, Grove ML, *et al*. Epigenome-wide association study (EWAS) of BMI, BMI change and waist

- circumference in African American adults identifies multiple replicated loci. *Hum Mol Genet* 2015;24:4464–79.
- 22 Gateva AT, Assyov YS, Velikova T, *et al*. Higher levels of thioredoxin interacting protein (TXNIP) in patients with prediabetes compared to obese normoglycemic subjects. *Diabetes Metab Syndr* 2019;13:734–7.
- 23 Liang L, Willis-Owen SAG, Laprise C, *et al*. An epigenome-wide association study of total serum immunoglobulin E concentration. *Nature* 2015;520:670–4.
- 24 Dhana K, Braun KVE, Nano J, *et al*. An epigenome-wide association study of obesity-related traits. *Am J Epidemiol* 2018;187:1662–9.
- 25 Akinyemiju T, Do AN, Patki A, *et al*. Epigenome-wide association study of metabolic syndrome in African-American adults. *Clin Epigenetics* 2018;10:49.
- 26 Abete I, Gómez-Uríz AM, Mansego ML, *et al*. Epigenetic changes in the methylation patterns of KCNQ1 and WT1 after a weight loss intervention program in obese stroke patients. *Curr Neurovasc Res* 2015;12:321–33.
- 27 Sayols-Baixeras S, Subirana I, Fernández-Sanlés A, *et al*. DNA methylation and obesity traits: an epigenome-wide association study. The REGICOR study. *Epigenetics* 2017;12:909–16.
- 28 Sawamura T. New Idol for cholesterol reduction? *Clin Chem* 2009;55:2082–4.
- 29 Eberlé D, Hegarty B, Bossard P, *et al*. SREBP transcription factors: master regulators of lipid homeostasis. *Biochimie* 2004;86:839–48.
- 30 Reynolds DS, Gurley DS, Austen KF. Cloning and characterization of the novel gene for mast cell carboxypeptidase A. *J Clin Invest* 1992;89:273–82.
- 31 Takabayashi T, Kato A, Peters AT, *et al*. Glandular mast cells with distinct phenotype are highly elevated in chronic rhinosinusitis with nasal polyps. *J Allergy Clin Immunol* 2012;130:410–20.
- 32 Ishijima Y, Ohmori Shin'ya, Ohneda K. Mast cell deficiency results in the accumulation of preadipocytes in adipose tissue in both obese and non-obese mice. *FEBS Open Bio* 2013;4:18–24.
- 33 Timmons JA, Wennmalm K, Larsson O, *et al*. Myogenic gene expression signature establishes that brown and white adipocytes originate from distinct cell lineages. *Proc Natl Acad Sci U S A* 2007;104:4401–6.
- 34 DeVilbiss AW, Boyer ME, Bresnick EH. Establishing a hematopoietic genetic network through locus-specific integration of chromatin regulators. *Proc Natl Acad Sci U S A* 2013;110:E3398–407.
- 35 Horikoshi M, Beaumont RN, Day FR, *et al*. Genome-wide associations for birth weight and correlations with adult disease. *Nature* 2016;538:248–52.
- 36 Edlow AG, Hui L, Wick HC, *et al*. Assessing the fetal effects of maternal obesity via transcriptomic analysis of cord blood: a prospective case-control study. *BJOG* 2016;123:180–9.
- 37 Ellulu MS, Patimah I, Khaza'ai H, *et al*. Obesity and inflammation: the linking mechanism and the complications. *Arch Med Sci* 2017;13:851–63.
- 38 Finlin BS, Zhu B, Confides AL, *et al*. Mast cells promote seasonal white adipose Beiging in humans. *Diabetes* 2017;66:1237–46.
- 39 Plubell DL, Wilmarth PA, Zhao Y, *et al*. Extended multiplexing of tandem mass tags (TMT) labeling reveals age and high fat diet specific proteome changes in mouse epididymal adipose tissue. *Mol Cell Proteomics* 2017;16:873–90.
- 40 Lee YH, Nair S, Rousseau E, *et al*. Microarray profiling of isolated abdominal subcutaneous adipocytes from obese vs non-obese Pima Indians: increased expression of inflammation-related genes. *Diabetologia* 2005;48:1776–83.
- 41 Li X, Huynh H, Zuo H, *et al*. Gata2 is a rheostat for mesenchymal stem cell fate in male mice. *Endocrinology* 2016;157:1021–8.
- 42 Tong Q, Dalgin G, Xu H, *et al*. Function of GATA transcription factors in preadipocyte-adipocyte transition. *Science* 2000;290:134–8.
- 43 Tsai J, Tong Q, Tan G, *et al*. The transcription factor GATA2 regulates differentiation of brown adipocytes. *EMBO Rep* 2005;6:879–84.