OXFORD

Human Molecular Genetics, 2016, Vol. 25, No. 20 4556–4565

doi: 10.1093/hmg/ddw285
Advance Access Publication Date: 15 September 2016
Original Article

ORIGINAL ARTICLE

Identification and validation of seven new loci showing differential DNA methylation related to serum lipid profile: an epigenome-wide approach. The REGICOR study

S. Sayols-Baixeras[1,2], I. Subirana[1,3], C. Lluis-Ganella[1], F. Civeira[4], J. Roquer[5], AN Do[6], D. Absher[7], A. Cenarro[4], D. Muñoz[8], C. Soriano-Tárraga[5], J. Jiménez-Conde[5], J.M. Ordovas[9], M. Senti[10], S. Aslibekyan[6], J. Marrugat[1], D.K. Arnett[11] and R. Elosua[1,*]

[1]Cardiovascular Epidemiology and Genetics Research Group, IMIM (Hospital del Mar Medical Research Institute), Barcelona, Catalonia, Spain, [2]Universitat Pompeu Fabra (UPF), Barcelona, Catalonia, Spain, [3]CIBER Epidemiology and Public Health, Barcelona, Catalonia, Spain, [4]Unidad Clínica y de Investigación en Lípidos y Arteriosclerosis, Hospital Universitario Miguel Servet, Instituto de Investigación Sanitaria Aragón, Zaragoza, Spain, [5]Department of Neurology, Neurovascular Research Group, IMIM (Hospital del Mar Medical Research Institute), Barcelona 08003, Catalonia, Spain, [6]Department of Epidemiology, University of Alabama at Birmingham, Birmingham, AL, USA, [7]Hudson Alpha Institute for Biotechnology, Huntsville, AL, USA, [8]Cardiovascular Risk and Nutrition Research Group, IMIM (Hospital del Mar Medical Research Institute), Barcelona, Catalonia, Spain, [9]Jean Mayer USDA Human Nutrition Research Center on Aging, Tufts University, Boston, MA, USA, [10]Department of Experimental and Health Sciences, Pompeu Fabra University, Barcelona, Catalonia, Spain and [11]Dean's Office, College of Public Health, University of Kentucky, Lexington, KY, USA

*To whom correspondence should be addressed at: Roberto Elosua, MD, PhD, IMIM, Hospital del Mar Medical Research Institute, Dr Aiguader 80, 08003 Barcelona, Catalonia, Spain. Tel: (+34) 933 160800; Fax: (+34) 933 160796; Email: relosua@imim.es

Abstract

Lipid traits (total, low-density and high-density lipoprotein cholesterol, and triglycerides) are risk factors for cardiovascular disease. DNA methylation is not only an inherited but also modifiable epigenetic mark that has been related to cardiovascular risk factors. Our aim was to identify loci showing differential DNA methylation related to serum lipid levels. Blood DNA methylation was assessed using the Illumina Human Methylation 450 BeadChip. A two-stage epigenome-wide association study was performed, with a discovery sample in the REGICOR study ($n = 645$) and validation in the Framingham Offspring Study ($n = 2,542$). Fourteen CpG sites located in nine genes (*SREBF1*, *SREBF2, PHOSPHO1, SYNGAP1, ABCG1, CPT1A, MYLIP, TXNIP* and *SLC7A11*) and 2 intergenic regions showed differential methylation in association with lipid traits. Six of these genes and 1 intergenic region were new discoveries showing differential methylation related to total cholesterol (*SREBF2*), HDL-cholesterol (*PHOSPHO1, SYNGAP1* and an intergenic region in chromosome 2) and triglycerides (*MYLIP*, *TXNIP* and *SLC7A11*). These CpGs explained 0.7%, 9.5% and 18.9% of the variability of total cholesterol, HDL cholesterol and triglycerides in the

Received: July 13, 2016. **Revised:** July 13, 2016. **Accepted:** August 19, 2016

Framingham Offspring Study, respectively. The expression of the genes *SREBF2* and *SREBF1* was inversely associated with methylation of their corresponding CpGs (*P*-value = 0.0042 and 0.0045, respectively) in participants of the GOLDN study ($n = 98$). In turn, *SREBF1* expression was directly associated with HDL cholesterol (*P*-value = 0.0429). Genetic variants in *SREBF1*, *PHOSPHO1*, *ABCG1* and *CPT1A* were also associated with lipid profile. Further research is warranted to functionally validate these new loci and assess the causality of new and established associations between these differentially methylated loci and lipid metabolism.

Introduction

Total cholesterol (TC), low-density lipoprotein cholesterol (LDL-c), high-density lipoprotein cholesterol (HDL-c), and triglycerides (TG) are among the most important risk factors for cardiovascular disease, the leading cause of death worldwide (1,2). Serum lipid levels are determined by a complex interplay between environmental, lifestyle and genetic factors (3).

In recent years, new technological advances have allowed the study of epigenetic mechanisms that affect chromatin structure and influence transcription levels. In the field of cardiovascular diseases, several studies have analysed the association between DNA methylation and lipid traits (4–7). DNA methylation is a heritable but reversible addition of a methyl group to a nucleotide. This process usually occurs in cytosine nucleotides in the presence of cytosine-phosphate-guanine (CpG), and can now be analysed by epigenome-wide association studies (EWAS) using arrays that cover thousands of CpGs across the genome.

The aim of the present study was to identify and validate new individual CpGs showing differential methylation related to serum lipid traits (TC, LDL-c, HDL-c and TG) in a population-based survey using an EWAS design. Additionally, we used shotgun transcriptomics (RNASeq) data to assess the associations between methylation at each validated CpG site and gene transcript level, as well as between gene transcript level and lipid phenotypes.

Results

Discovery stage of the EWAS analysis

After quality control steps, the discovery analysis included 645 individuals (99.5%; three individuals were excluded) and 427,948 probes (88.7%; 57,629 probes were excluded). A descriptive analysis of the main sociodemographic and clinical characteristics of the included individuals is shown in Table 1.

The relationship between observed and predicted *P*-values for all the associated tests between methylation levels and each lipid trait is shown in QQ and Manhattan plots (Supplementary Material, Fig. S1). In the discovery sample, we identified 113 CpGs (27 using model 1 –adjusted for covariates but not for surrogates variables– and 105 using model 2–adjusted for covariates and surrogates variables–) showing a suggestive association between methylation levels and lipid traits with an arbitrary *P*-value threshold $<1x10^{-5}$ (Supplementary Material, Table S1 and Fig. 1A). These CpG sites were located in 88 genes, 3 lncRNA and 17 intergenic regions. In model 1 (Fig. 1B), 4 CpG sites were associated with TC, 4 with HDL-c, 18 with TG and 1 with both TG and HDL-c; in model 2 (Fig. 1C), 16 CpGs were associated with TC, 7 with LDL-c, 39 with HDL-c, 35 with TG, 5 with both TC and LDL-c and 3 with HDL-c and TG.

We also performed sensitivity analyses excluding those participants taking cholesterol-lowering drugs and adjusting for diabetes and body mass index (BMI). The results remained consistent (Supplementary Material, Table S2).

Validation stage of the EWAS analysis

After applying the same quality control steps as in the discovery analysis, the validation study included 2,542 individuals (99.0%; 26 individuals excluded) and all the CpG sites identified in the discovery stage. The characteristics of the population included in the validation stage are shown in Table 1. The associations observed between the 113 CpGs selected and the lipid traits in the Framingham Offspring study are shown in Supplementary Material, Table S1.

Meta-analyses

After meta-analysis of the results obtained in the REGICOR and Framingham Offspring studies, we validated 12 CpG sites using model 1, and 2 additional CpGs using model 2 (Table 2). Of the 14 hits, one CpG was associated with TC, 3 CpGs with HDL-c, 7 with TG and 3 CpGs with both TG and HDL-c (Table 2 and Fig. 1D). Validated CpG sites were located in 9 genes and 2 intergenic regions. One gene (*SREBF2*) was associated with TC, 4 (*SYNGAP1*, *SREBF1*, *PHOSPHO1* and *ABCG1*) with HDL-c and 6 (*TXNIP*, *SLC7A11*, *MYLIP*, *CPT1A*, *SREBF1* and *ABCG1*) with TG.

These 14 CpGs explained 3.84%, 10.33% and 16.07% of the variability of TC, HDL-c and TG, in REGICOR, and 0.74%, 9.51% and 18.89% in Framingham, respectively (Table 3, Supplementary Material, Fig. S2).

Gene expression analyses

The results of the Genetics of Lipid-lowering Drugs and Diet Network (GOLDN) study showed that the methylation of cg16000331 and cg11024682 was inversely associated with the expression of the genes *SREBF2 and SREBF1*, respectively (Table 4). In turn, *SREBF1* expression was directly associated with higher levels of HDL cholesterol (Table 4).

Genetic variation and lipid profile

The aggregated results of the Global Lipids Genetics Consortium showed some genetic variants in *SREBF1*, *PHOSPHO1*, *ABCG1* and *CPT1A* associated with the lipid trait of interest (Table 5, Supplementary Material, Table S4).

Discussion

In this EWAS we have identified and validated 14 CpG sites located in 9 genes (*SREBF1*, *SREBF2*, *PHOSPHO1*, *SYNGAP1*, *ABCG1*, *CPT1A*, *MYLIP*, *TXNIP* and *SLC7A11*) and 2 intergenic regions that showed differential methylation in association with 3 lipid traits: TC, HDL-c and TG. This differential methylation in the validated CpGs explained around 10% of the variability of HDL-c and around 16-18% of the variability of TG. These results replicated previous findings showing an association between lipid traits and DNA methylation in three genes (*ABCG1*, *SREBF1* and

Table 1. Descriptive characteristics of the participants in the discovery (REGICOR study) and validation (Framingham Offspring Study) stage.

	REGICOR $n = 645$	Framingham Offspring Study $n = 2,542$	P-value
Age	63.2 (11.7)	66.3 (8.9)	<0.001
Sex, female, *n* (%)	329 (51.0)	1,378 (54.2)	0.158
Total cholesterol, mg/dl*	208.0 (36.5)	186.0 (37.2)	<0.001
LDL cholesterol, mg/dl*‡	135 (32.3)	105 (31.3)	<0.001
HDL cholesterol, mg/dl*‡	53.0 (12.3)	57.3 (18.2)	<0.001
Triglycerides, mg/dl†	89.0 [67.0; 121.0]	102.0 [73.0; 142.0]	<0.001
SBP, mmHg*‡	131.0 (18.5)	126.0 (17.1)	<0.001
DBP, mmHg*‡	76.0 (9.9)	71.6 (10.2)	<0.001
Hypertension, *n*(%)	302 (47.0)	1442 (57.0)	<0.001
Diabetes, *n* (%)	63 (9.8)	–	–
Glucose, mg/dl‡	97.7 (20.4)	107 (23.7)	<0.001
BMI, kg/m^{2}‡	26.9 (4.1)	28.2 (5.4)	<0.001
Waist (cm)*	94.2 (11.5)	101 (14.6)	<0.001
Obesity, *n* (%)	128 (19.9)	792 (31.3)	<0.001
Smoking status, *n* (%)			<0.001
Current smokers	107 (16.6)	252 (10.0)	
Former 1–5 years	31 (4.8)	56 (2.2)	
Former >5 years	165 (25.6)	10 (0.4)	
Never smokers	342 (53.0)	2212 (87.4)	
Cholesterol treatment, *n* (%)	154 (24.0)	1098 (43.3)	<0.001
Diabetes treatment, *n* (%)	44 (6.9)	–	–
Blood pressure treatment, *n* (%)	200 (31.2)	1236 (48.8)	<0.001

*Mean (Standard deviation).

†Median (Interquartile range).

‡LDL, Low-density lipoprotein; HDL, High-density lipoprotein; SBP, Systolic blood pressure; DBP, Diastolic blood pressure; Hypertension, defined as previous treatment or SBP ≥ 140 mmHg or DBP ≥ 90 mmHg; Diabetes, defined as previous treatment or glycaemia ≥126 mg/dl; BMI, Body mass index; Obesity, defined as BMI ≥30 kg/m^{2}.

CPT1A) and 1 intergenic region (5,6,8,9). In addition, we identified 6 new genes and 1 intergenic region showing differential methylation in relation to TC (*SREBF2*), HDL-c (*PHOSPHO1, SYNGAP1* and an intergenic region located in chromosome 2) and TG (*MYLIP*, *TXNIP* and *SLC7A11*). The expression of the genes *SREBF2* and *SREBF1* was inversely associated with methylation of their corresponding CpGs. In turn, the expression of *SREBF1* was associated with higher levels of HDL cholesterol. Finally, genetic variants in *SREBF1*, *PHOSPHO1*, *ABCG1* and *CPT1A* were also associated with the lipid profile.

Previous findings associated *ABCG1* methylation with TG and HDL-c levels (6,8,9). A direct relationship of *ABCG1* methylation with coronary artery disease and aging has also been reported (6,9,10). Reduced levels of DNA methylation in the *ABCG1* promoter region have also been observed in patients under aspirin therapy (11). Our results on *SREBF1* methylation and TG are also consistent with previous findings, validating the direction and the magnitude of the association (5,6). Furthermore, we observed a statistically significant association between methylation levels of *SREBF1* and HDL-c in the opposite direction to that observed with TG. Previous studies also found an association between BMI and methylation of CpG sites located in *CPT1A* and *SREBF1* (12,13). In a sensitivity analysis, we further adjusted our models including BMI as a potential intermediate variable; the magnitude of the association between methylation in these CpGs and TG or HDL-c was slightly reduced (16–28%; Supplementary Material, Table S3), suggesting that this association is only partially BMI-dependent. In this study, we also provide data supporting that *SREBF1* methylation is related to *SREBF1* expression, which in turn is associated with HDL cholesterol. Moreover, genetic variants in this gene are also associated with HDL cholesterol. Globally, these results support the causal association between *SREBF1* and HDL cholesterol and suggest a potential functional mechanism.

Novel findings in the present study include an association between lipid traits and different CpG methylation levels in *SREBF2*, *PHOSPHO1*, *SYNGAP1*, *TXNIP*, *MYLIP* and *SLC7A11*. Four of these genes, *SREBF2*, *PHOSPHO1*, *TXNIP and MYLIP*, merit detailed comment (See Supplementary Material for further discussion).

Sterol regulatory element binding factor 2 (*SREBF2*) is similar to *SREBF1* and encodes membrane-bound transcription factors that control the metabolism of cholesterol and fatty acids. A common genetic variant of *SREBF2* has been associated with intima-media thickness (14). In the present study, we report a positive association between *SREBF2* methylation and TC, in the same direction as the association between *SREBF1* and TG. We also report an inverse association between *SREBF2* methylation and expression. These results suggest a potential role of increased *SRBEF2* methylation in atherosclerosis progression through lipid metabolism regulation.

Phosphatase orphan 1 (*PHOSPHO1*) plays an important role at early steps of inorganic phosphate-induced and matrix vesicle-mediated calcification. Published studies show an up-regulation of *PHOSPHO1* in vascular smooth muscle cells; administration of *PHOSPHO1* inhibitors reduced vascular smooth calcification by 2.5-fold. Therefore, a procalcific role of *PHOSPHO1* in pathologic vascular ossification is plausible (15,16). The present study found a direct association between the CpG methylation at this locus and HDL-cholesterol levels; other recent studies found an association with diabetes (17,18). The complex interrelations between lipids, diabetes and arterial calcifications merit additional studies (19–22), our data suggest that the low methylation level at this locus could be associated with higher PHOSPHO1 levels. This could induce calcification due to low HDL cholesterol levels and altered glucose homeostasis or other potential mechanisms.

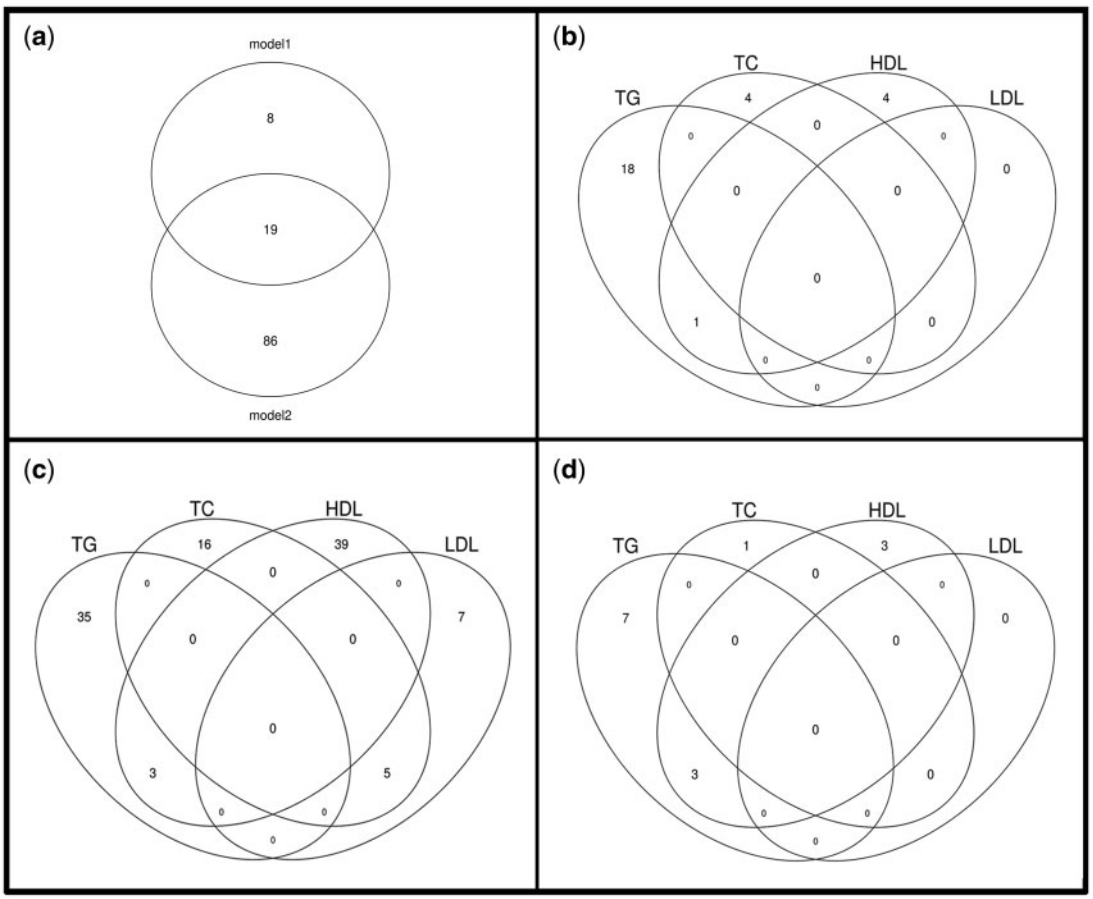

Figure 1. (**A**) Venn diagrams of discovery hits in REGICOR study using model 1 and model 2. (**B**) Venn diagrams of discovery hits in REGICOR study using model 1. (**C**) Venn diagrams of discovery hits in REGICOR study using model 2. (**D**) Venn diagrams of meta-analysis hits in REGICOR study and Framingham Offspring study using model 2. Model 1 is adjusted for age, sex, smoking exposure, batch effect and estimated cell count, whereas model 2 is adjusted for age, sex, smoking exposure, batch effect, estimated cell count and surrogate variables. TC= total cholesterol, LDL= low-density lipoprotein, HDL=high-density lipoprotein and TG= triglycerides.

Thioredoxin-interacting protein (*TXNIP*) is involved in redox homeostasis. It is induced by glucose but suppressed by insulin, and acts as a regulator of glucose homeostasis in humans. Recent studies have shown different methylation levels of *TXNIP* gene in relation to type 2 diabetes (17,18,23–26). *TXNIP* has also been related to regulation of TG levels. It was proposed as a candidate gene for familial combined hyperlipidemia; however, other studies have not replicated this finding (27–30). A *TXNIP* genetic variant was associated with TG levels in diabetic participants (31). In addition, the same *TXNIP* methylation CpG site was reported to be associated with TG levels, even though the association did not reach the statistical significance threshold in the meta-analyses (6). Additional adjustment of our models including diabetes as a potential confounder variable did not modify the magnitude of the association between methylation in this CpG and TG levels (Supplementary Material, Table S2), suggesting that the association is independent of diabetes. In summary, low *TXNIP* methylation is associated with high triglycerides and altered glucose homeostasis and could increase the risk of atherosclerosis-related diseases.

Finally, we report a direct association between methylation in myosin regulatory light chain interacting protein (*MYLIP*), a ubiquitin ligase, and TG. *MYLIP* was recently identified as a novel regulator of the LDL receptor in a genome-wide association study, so now it is also called *MYLIP/IDOL* (inducible degrader of the LDL receptor) (32,33). The induction of *MYLIP* in macrophages and liver leads to the ubiquitination of LDL receptors, whereas *MYLIP* downregulation by RNA silencing increases LDL receptor levels and enhances LDL uptake (34). This mechanism could be similar to that of PCSK9 inhibitors, making it a new therapeutic target if causality between *MYLIP* methylation and increased lipid levels can be demonstrated (35). Our findings and previously published results support the hypothesis that high *MYLIP* methylation could decrease *MYLIP* expression and increase triglycerides by decreasing LDL receptor availability, increasing the risk of atherosclerosis-related diseases.

A major strength of the study was the use of standardized protocols to remove non-biological sources of variation. Moreover, we used a powerful statistical method that reduces outlier effects (robust linear regression) and adjusted for residual confounding factors (surrogate variables). We replicated and meta-analysed data in a very large external population in order to improve the quality of our evidence. We also have analysed the association between DNA methylation and gene expression, and between gene expression and lipid traits.

Some limitations of the study should be considered. First, some heterogeneity exists between the REGICOR and Framingham Offspring studies. To address the differences, we carried out a meta-analysis of CpG sites with significant coefficient differences between studies, using the random effects method, and the results did not change. Second, the design of the study was cross-sectional and therefore we cannot infer causality

Table 2. Significant CpG sites differentially methylated in relation to different lipids traits (total cholesterol, low-density lipoprotein, high-density lipoprotein and triglycerides) in the discovery, replication and meta-analysis phase using model 2. CpG id, associated gene or transcript, chromosome (chr) location, genomic position, and observed coefficient (coeff) and standard error (SE) of M-stand and P-value.

Total Cholesterol (TC)													Previously associated with:
Features				Discovery Phase: REGICOR			Replication Phase: Framingham Offspring Study			Meta-analysis (Phase1 + Phase2)			
CpG	Chr.	Pos.	Gene	Coeff	SE	P-value	Coeff	SE	P-value	Coeff	SE	P-value	
cg16000331	22	42230138	SREBF2	9.00	2.03	8.96E-06	3.99	0.96	3.08E-05	4.91	0.87	1.46E-08	–
Low-density lipoprotein (LDL) cholesterol													**Previously associated with:**
Features				**Discovery Phase: REGICOR**			**Replication Phase: Framingham Offspring Study**			**Meta-analysis (Phase1 + Phase2)**			
CpG	**Chr.**	**Pos.**	**Gene**	**Coeff**	**SE**	**P-value**	**Coeff**	**SE**	**P-value**	**Coeff**	**SE**	**P-value**	
–	–	–	–	–	–	–	–	–	–	–	–	–	–
High-density lipoprotein (HDL) cholesterol													**Previously associated with:**
Features				**Discovery Phase: REGICOR**			**Replication Phase: Framingham Offspring Study**			**Meta-analysis (Phase1 + Phase2)**			
CpG	**Chr.**	**Pos.**	**Gene**	**Coeff**	**SE**	**P-value**	**Coeff**	**SE**	**P-value**	**Coeff**	**SE**	**P-value**	
ch.2.101965435R	2	102599003	NA	2.61	0.57	4.77E-06	2.06	0.44	2.33E-06	2.26	0.35	6.57E-11	–
cg09572125	6	33400477	SYNGAP1	−4.22	0.90	2.96E-06	−3.72	0.73	3.41E-07	−3.92	0.57	5.08E-12	Age(51)
cg11024682	17	17730094	SREBF1	−2.72	0.60	5.81E-06	−3.76	0.41	3.10E-20	−3.43	0.34	0.00E+00	TG(6), BMI(12) and diabetes(17,18)
cg02650017	17	47301614	PHOSPHO1	2.83	0.51	2.81E-08	2.23	0.41	4.49E-08	2.47	0.32	9.77E-15	Diabetes(17,18)
cg27243685	21	43642366	ABCG1	−2.44	0.54	5.72E-06	−4.54	0.40	2.81E-30	−3.80	0.32	0.00E+00	TG(6) and BMI(12)
cg06500161	21	43656587	ABCG1	−3.47	0.63	3.08E-08	−6.16	0.41	4.00E-52	−5.37	0.34	0.00E+00	TG(6), HDL-c(6), BMI(12) and diabetes(17,18)
Triglycerides													**Previously associated with:**
Features				**Discovery Phase: REGICOR**			**Replication Phase: Framingham Offspring Study**			**Meta-analysis (Phase1 + Phase2)**			
CpG	**Chr.**	**Pos.**	**Gene**	**Coeff**	**SE**	**P-value**	**Coeff**	**SE**	**P-value**	**Coeff**	**SE**	**P-value**	
cg19693031	1	145441552	TXNIP	−0.04	0.01	1.68E-06	−0.05	0.00	5.72E-26	−0.05	0.00	0.00E+00	Diabetes(17,18,25,26)
cg06690548	4	139162808	SLC7A11	−0.04	0.01	4.86E-07	−0.04	0.00	5.38E-22	−0.04	0.00	0.00E+00	Diabetes(17)
cg03717755	6	16136539	MYLIP	0.06	0.01	7.31E-09	0.05	0.01	2.45E-14	0.05	0.01	0.00E+00	–
cg07504977	10	102131012	NA	0.04	0.01	7.54E-07	0.04	0.01	1.98E-15	0.04	0.00	0.00E+00	Age (51) and BMI(13,52)
cg00574958	11	68607622	CPT1A	−0.05	0.01	4.68E-06	−0.08	0.01	4.21E-56	−0.07	0.00	0.00E+00	TG(5,6), BMI(12,13) and diabetes(17)
cg11024682	17	17730094	SREBF1	0.07	0.01	5.63E-09	0.06	0.01	9.81E-26	0.06	0.01	0.00E+00	TG(6), BMI(12) and diabetes(17,18)
cg27243685	21	43642366	ABCG1	0.06	0.01	2.46E-10	0.07	0.01	2.59E-38	0.06	0.00	0.00E+00	TG(6) and BMI(12)
cg01881899	21	43652704	ABCG1	0.05	0.01	7.41E-07	0.05	0.01	9.56E-19	0.05	0.00	0.00E+00	–
cg02370100	21	43655256	ABCG1	0.05	0.01	4.74E-06	0.04	0.01	6.91E-13	0.04	0.00	0.00E+00	BMI(52)
cg06500161	21	43656587	ABCG1	0.08	0.01	3.34E-19	0.09	0.01	2.28E-67	0.09	0.00	0.00E+00	TG(6), HDL-c(6), BMI(12) and diabetes(17,18)

NA: Non-annotated gene; All P values were calculated using standardized M values.

Table 3. Variability of total and HDL cholesterol, and triglycerides explained by differential methylation of the validated CpG

Phenotype	CpG	Explained variability (%)	
		REGICOR	Framingham Offspring Study
Total cholesterol	cg16000331	3.84	0.74
LDL cholesterol	–	–	–
HDL cholesterol	cg06500161	4.05	5.53
	cg27243685	2.16	2.98
	cg02650017	2.13	0.64
	cg11024682	1.56	1.41
	ch.2.101965435R	1.16	0.82
	cg09572125	0.02	0.00
	cg06500161+ cg27243685+ cg02650017+ cg11024682+ ch.2.101965435R + cg09572125	10.33	9.51
Triglycerides	cg06500161	5.00	5.76
	cg27243685	3.66	3.36
	cg19693031	3.18	3.22
	cg00574958	2.96	7.49
	cg06690548	2.77	3.34
	cg11024682	2.44	1.91
	cg01881899	2.37	1.49
	cg03717755	1.95	1.17
	cg07504977	1.71	2.18
	cg02370100	0.69	0.92
	cg06500161+ cg27243685+ cg19693031+ cg00574958+ cg06690548+ cg11024682+ cg01881899+ cg03717755+ cg07504977+ cg02370100	16.07	18.89

in the reported association between lipid traits and DNA methylation levels. Proper Mendelian randomization studies or prospective data are warranted to assess this issue (36).

In summary, this study replicated 4 loci (*ABCG1*, *SREBF1*, *CPT1A* and 1 intergenic region at chromosome 10) and discovered and validated 7 new loci (*SREBF2*, *PHOSPHO1, SYNGAP1, TXNIP, MYLIP*, *SLC7A11*, and one intergenic region at chromosome 2) showing differential methylation related to serum lipid profile. The findings point out some genes and pathways that may be related to lipid metabolism regulation and atherosclerosis. We also report a potential functional explanation for the association between *SREBF1* methylation and HDL cholesterol that could be mediated by gene expression. Further research is needed to functionally validate the new loci and assess the causality of these associations.

Materials and Methods

Study design and participants

A cross-sectional two-stage epigenome-wide association study was designed.

Discovery stage

We used the REGICOR (REgistre GIroní del COR) study, a population-based cohort that includes participants from different towns representing the urban and rural diversity of Girona Province in Catalonia (Spain). For this study, we considered those individuals initially enrolled in a survey performed during 2003–2005 ($n = 6{,}352$; response rate 71.5%) who were still residing in these towns and attended a follow-up visit during 2009–2013 ($n = 4{,}980$ participants; response rate 78.4%). From those participants, we randomly selected a subsample of 648 participants, all of whom reported that they were of European descent. All the participants gave written informed consent before their participation. This study follows the principles expressed in the Declaration of Helsinki and the legislation in Spain, and was approved by the local ethics committee.

Validation stage

We obtained methylation and phenotype data from the Framingham Offspring Study through the Database of Genotypes and Phenotypes (dbGAP; http://dbgap.ncbi.nlm.nih.gov; project number #9047). The Framingham Offspring Study started in 1971 and included 5,209 offspring from the participants in the original Framingham Heart Study. Those participants attending exam 8 and with available DNA methylation data were included in the present analysis.

Serum lipid profile

In the REGICOR study, a team of trained nurses collected fasting blood samples, which were centrifuged between 15 and 30 min after extraction. Both serum and plasma samples were aliquoted and frozen at -80°C until analysis. Serum concentrations of TC and TG were determined enzymatically (Horiba ABX, Montpellier, France). Serum HDL-c concentrations were measured as soluble cholesterol determined by an accelerator selective detergent method (Horiba ABX). Analyses were performed in an ABX Pentra 400 (Horiba ABX, Montpellier, France). External quality assessment was performed with External Quality Control BIORAD (Bio-Rad, Hercules, California, USA) and Assessment-SEQC (Sociedad Española Química Clínica, Barcelona, Spain). LDL-c was calculated by the Friedewald equation when TG levels were lower than 300 mg/dl.

In the Framingham Offspring Study, lipid traits of interest were obtained from examination 8 data through dbGaP. Fasting TC, HDL-c and TG were determined directly and LDL-c was estimated by the Friedewald equation.

Other covariates

In the REGICOR study, the same group of trained nurses examined and administered all validated and standardized questionnaires and methods used to collect sociodemographic, lifestyle and cardiovascular risk factors information. In the Framingham Offspring study, the covariates of interest were obtained from examination 8 data through dbGaP.

Smoking status was self-reported and then categorized in four different groups: current smokers (smoked on average ≥ 1 cigarette/day at the time of the visit or gave up smoking < 1 year before the visit); former smokers, between 1 and 5 years (gave up smoking up to 5 years before the visit); former smokers, more than 5 years; and never smokers (never smoked regularly, defined as an average ≥ 1 cigarette/day). Other phenotypes of interest are available in Supplementary Material.

Table 4. Results of the GOLDN study corresponding to the association between methylation at the CpGs of interest and gene expression, and between gene expression and the corresponding lipid trait

			Methylation-Gene Expression				Gene Expression-Lipid trait		
Lipid trait	CpG	Gen	Read count	Beta	SE	P-value	Beta	SE	P-value
Total cholesterol									
	cg16000331	*SREBF2*	ENSG00000198911	−3182	1080	0.0042	−0.397	0.733	0.5897
HDL cholesterol									
	cg09572125	*SYNGAP1*	ENSG00000197283	−812	454	0.0775	1.047	1.98	0.5985
	cg11024682	*SREBF1*	ENSG00000072310	−3329	1139	0.0045	6.706	3.261	0.0429
	cg02650017	*PHOSPHO1*	ENSG00000173868	−55600	42230	0.1918	−44.39	33.01	0.1825
	cg06500161	*ABCG1*	ENSG00000160179	364	542	0.5036	1.921	1.966	0.3314
	cg27243685	*ABCG1*	ENSG00000160179	−318	908	0.7274			
	cg01881899	*ABCG1*	ENSG00000160179	−391	2017	0.8468			
	cg02370100	*ABCG1*	ENSG00000160179	−1891	1058	0.0778			
Triglycerides									
	cg19693031	*TXNIP*	ENSG00000117289	−59270	30320	0.0542	25.34	20.25	0.2142
	cg06690548	*SLC7A11*	ENSG00000151012	−38	36	0.2950	−0.01	0.018	0.5616
	cg03717755	*MYLIP*	ENSG00000007944	−318	598	0.5968	−0.726	0.37	0.0533
	cg00574958	*CPT1A*	ENSG00000110090	−7725	3048	0.0133	0.1147	0.87	0.8955
	cg11024682	*SREBF1*	ENSG00000072310	−3329	1139	0.0045	−0.541	0.484	0.2670
	cg06500161	*ABCG1*	ENSG00000160179	364	542	0.5036	−0.332	0.286	0.2499

Beta: linear regression coefficient; SE: standard error of the regression coefficient.

Table 5. Aggregated results of the Global Lipids Genetic Consortium: selection of the SNP showing the lowest P-value for the associations between genetic variants in the loci of interest and the corresponding lipid trait. (Adapted from Reference 33)

Gene	rsID	chr	pos	A1	A2	A1 allele frequency	Regression coefficient	Standard Error	P-value
Total cholesterol									
SREBF2	rs2267443	22	42287454	G	A	0.5989	0.0115	0.0054	0.0828
HDL-Cholesterol									
SYNGAP1	rs211456	6	33389381	T	G	0.3654	0.0062	0.0036	0.0618
SREBF1	rs11653007	17	17716662	G	C	1	0.0388	0.0112	0.0013
PHOSPHO1	rs648980	17	47303456	C	T	0.4617	0.0079	0.0050	0.0120
ABCG1	rs482303	21	43691469	C	G	NA	0.0261	0.0089	0.0051
Triglycerides									
TXNIP	rs7212	1	145442254	G	C	1.0000	0.0281	0.0120	0.0608
SLC7A11	rs13141329	4	139125405	A	T	NA	0.0058	0.0047	0.3340
MYLIP	rs2072781	6	16147349	T	C	0.9195	0.0139	0.0075	0.1270
CPT1A	rs7938117	11	68598054	A	G	0.3259	0.0155	0.0035	0.0004
SREBF1	rs11653007	17	17716662	C	G	NA	0.0301	0.0108	0.0087
ABCG1	rs450808	21	43706944	T	C	0.2098	0.0192	0.0090	0.0836

Infinium methylation 450K technology

In the REGICOR study, DNA was extracted from whole peripheral blood collected in 10 mL EDTA tubes using a standardized method (Puregen TM; Gentra Systems). In the Framingham Offspring Study, DNA was extracted from buffy coat using the same method. DNA methylation was assessed using the Illumina HumanMethylation450 BeadChip (Illumina), which interrogates 485,577 CpGs, following the Illumina Infinium HD Methylation protocol in both studies (37,38). Infinium Methylation 450K is an array-based technology allowing a maximum of 12 samples for each Beadchip. Each analysis runs onto a plate where it is possible to accommodate 8 Beadchips, permitting processing of 96 samples in a batch (39).

The 645 REGICOR samples were analysed in seven batches in two laboratories of the Spanish National Genotyping Centre: the Centre for Genomic Regulation in Barcelona ($n = 187$) and the Centro Nacional de Investigaciones Oncológicas in Madrid (n= 458). Two duplicated samples were included in all the batches as an internal quality control. In the Framingham Offspring Study, 2,542 samples were analysed distributed in 30 batches in two different laboratories ($n = 499$ and 2,043) (5,40). Duplicated samples were also included.

M-values were calculated according to Equation 1 and β-values were calculated using Equation 2:

$$M_{value} = \log_2\left(\frac{M_i + \alpha}{U_i + \alpha}\right) \quad (1)$$

$$\beta_{value} = \left(\frac{M_i}{M_i + U_i + \alpha}\right) \quad (2)$$

,

Where: M_i is the intensity of methylated probes, U_i is the intensity of unmethylated probes and α is a constant offset that takes values of 1 and 100 in the respective equations.

An M-value close to 0 means the CpG site is about half-methylated. Positive M-values mean that there are more methylated than unmethylated cytosines; negative M-values indicate the opposite ratio. β-value ranges between 0 (completely unmethylated) and 1 (completely methylated). M-value was the main outcome variable due to its good statistical properties (41,42).

We assessed the quality control of the methylation data using a well-defined pipeline (see Supplementary Material) and standardized the M-values for batches following Equation 3:

$$Z = \frac{(X - \bar{X})}{\sqrt{\frac{\sum (X-\bar{X})^2}{(n-1)}}}; \quad (3)$$

Where: Z = standardized M-value (M-stand), X = M-value for a specific individual, $\bar{X}$ = mean of M-value for a specific batch and n = sample size.

Statistical analysis

All the analyses were carried out by robust multivariable linear regression. For each association test, methylation was considered as the independent variable and lipid traits (TC, LDL-c, HDL-c and logarithm of TG) were considered as the outcome or dependent variables. We defined two models for each lipid trait: Model 1, adjusted for a pre-defined set of variables including sex, age, smoking status, batch effect and estimated cell count, and Model 2, additionally adjusted for surrogate variables to control for unmeasured potential confounding variables, including technical, environmental or demographic variables (43). Cell count was estimated using Houseman's algorithm implemented in *R::minfi* and surrogate variables were estimated using the *R::sva* R package (44–46).

We passed on to the validation stage those CpG sites associated with lipid traits which exceeded an arbitrary *P*-value threshold of 1×10^{-05} for either of the two models.

Finally, we performed a fixed-effects meta-analysis using data from the REGICOR study and Framingham Offspring Study for both models. We declared as statistically significant any association that fulfilled the Bonferroni criteria applied according to the number of probes that passed the quality control (*P*-value $< 1.17 \times 10^{-07}$; 0.05/427,948 probes that passed the quality control). We also performed sensitivity analyses to detect for possible confounding factors, excluding those participants receiving treatment for dyslipidemia and adjusting for diabetes and BMI.

We calculated the statistical power of the meta-analysis to replicate those CpGs identified in the discovery stage: the magnitudes of the associations for the 113 CpGs included in the validation stage that could be considered as statistically significant, accepting an alpha risk of 1.17×10^{-07}, in a two-sided test and an 80% power are shown in Supplementary Material, Table S3.

Additional functional and genetic validation

Methylation, gene expression and lipid profile analyses: the GOLDN study

The GOLDN study included the families of self-reported European descent with at least two siblings, recruited from two centres in the National Heart, Lung, and Blood Institute Family Heart Study: Minneapolis, Minnesota, and Salt Lake City, Utah (47). An extended description of the study is provided in the Supplementary Material.

Briefly, we used data from the baseline GOLDN visit for the present analysis. The Infinium HumanMethylation450 BeadChip (Illumina, San Diego, CA, USA) was used to quantify DNA methylation in CD4+ T-cells (5,48). For transcriptional profiling, we selected 102 unrelated GOLDN participants from the extremes of the BMI distribution. We extracted RNA from buffy coats using the TRIzol method (ThermoFisher Scientific, Waltham, MA, USA) and evaluated the quality of each RNA using Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) to ensure integrity. We fragmented purified poly A+ RNA by controlled hydrolysis with a weak base, followed by conversion to cDNA with random-primed reverse transcriptase. We then incorporated the cDNA into the Illumina libraries and sequenced on the HiSeq2000 platform. We used the STAR-2 pass method (49) to align pair-end read sequences. We obtained the aligned read count information for each gene transcript from the STAR output using the HTSeq-count program (50). Two samples appeared as outliers and another two individuals were missing covariate data; the final sample in the analysis was 98 GOLDN participants.

We fitted linear mixed models to test for associations between methylation scores at each CpG site and gene transcript level, as well as between gene transcript level and lipid phenotypes. We only used CpG sites that could be mapped to a known gene in the methylation vs. expression analyses. We set the statistical significance level according to the Bonferroni principle, $\alpha = 0.05/9$ genes $= 0.0056$.

Genetic variation and lipid profile: aggregated results of the global lipid research consortium

We accessed the publicly available and aggregated summary data of the Global Lipid Genetics Consortium to explore the additive association between genetic variants in the loci of interest and their corresponding lipid trait (33).

Supplementary Material

Supplementary Material is available at *HMG* online.

Acknowledgements

Elaine M. Lilly, PhD, Writers First Aid, for revision of the English text.

Conflict of Interest statement. None declared.

Funding

This work was supported by the following sources: Agència de Gestió Ajuts Universitaris de Recerca [2014 SGR 240]; the Spanish Ministry of Economy through the Carlos III Health Institute [ISCIII-FIS-FEDER-ERDF PI12-00232, PI12-01238, PI11-01801, PI08-1327, PI05-1251, PI05-1297, PI02-0471, FIS99/0013-01, FIS96/0026-01, FIS93/0568, FIS92/0009-05], and the Red de Investigación Cardiovascular [RD12/0042/0013, RD12/0042/0020, RD12/0042/0055, RD12/0042/0061]. S.S-B. was funded by a contract from Instituto de Salud Carlos III FEDER [IFI14/00007] and Daniel Bravo Andreu Private Foundation.

GOLDN: The GOLDN study (AND, DA, JO, SA, DKA) was funded by the US National Institute of Health (NIH)/National Heart, Lung and Blood Institutes (http://www.nhlbi.nih.gov) grants R01HL104135 and U01HL72524.

References

1. O'Donnell, C.J. and Elosua, R. (2008) Cardiovascular risk factors. Insights from Framingham Heart Study. *Rev. Esp. Cardiol.*, **61**, 299–310.
2. Mozaffarian, D., Benjamin, E.J., Go, A.S., Arnett, D.K., Blaha, M.J., Cushman, M., de Ferranti, S., Despres, J.P., Fullerton, H.J., Howard, V.J., *et al.* (2014) Heart Disease and Stroke Statistics - 2015 Update: A Report From the American Heart Association. *Circulation*, **131**, e29–e294.
3. Sayols-Baixeras, S., Lluis-Ganella, C., Lucas, G. and Elosua, R. (2014) Pathogenesis of coronary artery disease: focus on genetic risk factors and identification of genetic variants. *Appl. Clin. Genet.*, **7**, 15–32.
4. Martinez-Jimenez, C.P. and Sandoval, J. (2015) Epigenetic crosstalk: a molecular language in human metabolic disorders. *Front. Biosci. (Schol. Ed.)*, **7**, 46–57.
5. Irvin, M.R., Zhi, D., Joehanes, R., Mendelson, M., Aslibekyan, S., Claas, S.A., Thibeault, K.S., Patel, N., Day, K., Jones, L.W., *et al.* (2014) Epigenome-wide association study of fasting blood lipids in the genetics of lipid-lowering drugs and diet network study. *Circulation*, **130**, 565–572.
6. Pfeiffer, L., Wahl, S., Pilling, L.C., Reischl, E., Sandling, J.K., Kunze, S., Holdt, L.M., Kretschmer, A., Schramm, K., Adamski, J., *et al.* (2015) DNA methylation of lipid-related genes affects blood lipid levels. *Circ. Cardiovasc. Genet.*, **8**, 334–342.
7. Guay, S.P., Voisin, G., Brisson, D., Munger, J., Lamarche, B., Gaudet, D. and Bouchard, L. (2012) Epigenome-wide analysis in familial hypercholesterolemia identified new loci associated with high-density lipoprotein cholesterol concentration. *Epigenomics*, **4**, 623–639.
8. Guay, S.P., Brisson, D., Lamarche, B., Gaudet, D. and Bouchard, L. (2014) Epipolymorphisms within lipoprotein genes contribute independently to plasma lipid levels in familial hypercholesterolemia. *Epigenetics*, **9**, 718–729.
9. Guay, S.P., Brisson, D., Munger, J., Lamarche, B., Gaudet, D. and Bouchard, L. (2012) ABCA1 gene promoter DNA methylation is associated with HDL particle profile and coronary artery disease in familial hypercholesterolemia. *Epigenetics*, **7**, 464–472.
10. Peng, P., Wang, L., Yang, X., Huang, X., Ba, Y., Chen, X., Guo, J., Lian, J. and Zhou, J. (2014) A preliminary study of the relationship between promoter methylation of the ABCG1, GALNT2 and HMGCR genes and coronary heart disease. *PLoS One*, **9**, 1–8.
11. Guay, S.P., Légaré, C., Houde, A.A., Mathieu, P., Bossé, Y. and Bouchard, L. (2014) Acetylsalicylic acid, aging and coronary artery disease are associated with ABCA1 DNA methylation in men. *Clin. Epigenetics*, **6**, 1–7.
12. Demerath, E.W., Guan, W., Grove, M.L., Aslibekyan, S., Mendelson, M., Zhou, Y.H., Hedman, ÅK., Sandling, J.K., Li, L.A., Irvin, M.R., *et al.* (2015) Epigenome-wide Association Study (EWAS) of BMI, BMI Change, and Waist Circumference in African American Adults Identifies Multiple Replicated Loci. *Hum. Mol. Genet.*, **24**, 4464–4479.
13. Aslibekyan, S., Demerath, E.W., Mendelson, M., Zhi, D., Guan, W., Liang, L., Sha, J., Pankow, J.S., Liu, C., Irvin, M.R., *et al.* (2015) Epigenome-wide study identifies novel methylation loci associated with body mass index and waist circumference. *Obesity*, **23**, 1493–1501.
14. Robinet, P., Védie, B., Chironi, G. and Gariépy, J. (2003) Characterization of polymorphic structure of SREBP-2 gene: role in atherosclerosis. *Atherosclerosis*, **168**, 381–387.
15. Kiffer-Moreira, T., Yadav, M.C., Zhu, D., Narisawa, S., Sheen, C., Stec, B., Cosford, N.D., Dahl, R., Farquharson, C., Hoylaerts, M.F., *et al.* (2013) Pharmacological inhibition of PHOSPHO1 suppresses vascular smooth muscle cell calcification. *J. Bone Miner. Res.*, **28**, 81–91.
16. Bobryshev, Y.V., Orekhov, A.N., Sobenin, I. and Chistiakov, D.A. (2014) Role of bone-type tissue-nonspecific alkaline phosphatase and PHOSPO1 in vascular calcification. *Curr. Pharm. Des.*, **20**, 5821–5828.
17. Kulkarni, H., Kos, M.Z., Neary, J., Dyer, T.D., Göring, H.H.H., Cole, S.A., Comuzzie, A.G., Almasy, L., Mahaney, M.C., Curran, J.E., *et al.* (2015) Novel epigenetic determinants of type 2 diabetes in mexican american families. *Hum. Mol. Genet.*, **24**, 5330–5344.
18. Chambers, J.C., Loh, M., Lehne, B., Drong, A., Kriebel, J., Motta, V., Wahl, S., Elliott, H.R., Rota, F., Scott, W.R., *et al.* (2015) Epigenome-wide association of DNA methylation markers in peripheral blood from Indian Asians and Europeans with incident type 2 diabetes: a nested case-control study. *Lancet Diabetes Endocrinol.*, **3**, 526–534.
19. Bjornstad, P., Eckel, R.H., Pyle, L., Rewers, M., Maahs, D.M. and Snell-Bergeon, J.K. (2015) Relation of combined non-high-density lipoprotein cholesterol and apolipoprotein B with atherosclerosis in adults with type 1 diabetes mellitus. *Am. J. Cardiol.*, **116**, 1057–1062.
20. Eren, E., Ellidag, H.Y., Aydin, O. and Yılmaz, N. (2014) HDL-Associated paraoxonase 1 as a bridge between postmenopausal osteoporosis and cardiovascular disease. *Chonnam. Med. J.*, **50**, 75–81.
21. Rattazzi, M., Faggin, E., Buso, R., Di Virgilio, R., Puato, M., Plebani, M., Zaniotto, M., Palmosi, T., Bertacco, E., Fadini, G.P., *et al.* (2016) Atorvastatin reduces circulating osteoprogenitor cells and T-Cell RANKL expression in osteoporotic women: implications for the bone-vascular axis. *Cardiovasc. Ther.*, **34**, 13–20.
22. García-Sánchez, C., Posadas-Romero, C., Posadas-Sánchez, R., Carreón-Torres, E., Rodríguez-Pérez, J.M., Juárez-Rojas, J.G., Martínez-Sánchez, C., Fragoso, J.M., González-Pacheco, H., Vargas-Alarcón, G., *et al.* (2015) Low concentrations of phospholipids and plasma HDL cholesterol subclasses in asymptomatic subjects with high coronary calcium scores. *Atherosclerosis*, **238**, 250–255.
23. Parikh, H., Carlsson, E., Chutkow, W.A., Johansson, L.E., Storgaard, H., Poulsen, P., Saxena, R., Ladd, C., Schulze, P.C., Mazzini, M.J., *et al.* (2007) TXNIP regulates peripheral glucose metabolism in humans. *PLoS Med*, **4**, 0868–0879.
24. Zhou, J. and Chang, W.J. (2013) Roles of thioredoxin binding protein (TXNIP) in oxidative stress, apoptosis and cancer. *Mitochondrion*, **13**, 163–169.
25. Florath, I., Butterbach, K., Heiss, J., Bewerunge-Hudler, M., Zhang, Y., Schöttker, B. and Brenner, H. (2016) Type 2 diabetes and leucocyte DNA methylation: an epigenome-wide association study in over 1,500 older adults. *Diabetologia*, **59**, 130–138.
26. Soriano-Tárraga, C., Jiménez-Conde, J., Giralt-Steinhauer, E., Mola-Caminal, M., Vivanco-Hidalgo, R.M., Ois, A., Rodríguez-Campello, A., Cuadrado-Godia, E., Sayols-Baixeras, S., Elosua, R., *et al.* (2015) Epigenome-wide association study identifies TXNIP gene associated with type 2 diabetes mellitus and sustained hyperglycemia. *Hum. Mol. Genet.*, **25**, 609–619.
27. Bodnar, J.S., Chatterjee, A., Castellani, L.W., Ross, D.A., Ohmen, J., Cavalcoli, J., Wu, C., Dains, K.M., Catanese, J., Chu, M., *et al.* (2002) Positional cloning of the combined hyperlipidemia gene Hyplip1. *Nat. Genet.*, **30**, 110–116.

28. Coon, H., Singh, N., Dunn, D., Eckfeldt, J.H., Province, M.A., Hopkins, P.N., Weiss, R., Hunt, S.C. and Leppert, M.F. (2004) TXNIP gene not associated with familial combined hyperlipidemia in the NHLBI Family Heart Study. *Atherosclerosis*, **174**, 357–362.
29. Pajukanta, P., Lilja, H.E., Sinsheimer, J.S., Cantor, R.M., Lusis, A.J., Gentile, M., Duan, X.J., Soro-Paavonen, A., Naukkarinen, J., Saarela, J., *et al.* (2004) Familial combined hyperlipidemia is associated with upstream transcription factor 1 (USF1). *Nat. Genet.*, **36**, 371–376.
30. Van der Vleuten, G.M., Hijmans, A., Kluijtmans, L.A., Blom, H.J., Utalenhoef, A.F. and de Graaf, J. (2004) Thioredoxin interacting protein in Dutch families with familial combined hyperlipidemia. *Am. J. Med. Genet. A.*, **130**, 73–75.
31. Van Greevenbroek, M.M.J., Vermeulen, V.M.M.J., Feskens, E.J.M., Evelo, C.T., Kruijshoop, M., Hoebee, B., van der Kallen, C.J.H. and de Bruin, TW. a. (2007) Genetic variation in thioredoxin interacting protein (TXNIP) is associated with hypertriglyceridaemia and blood pressure in diabetes mellitus. *Diabet. Med.*, **24**, 498–504.
32. Teslovich, T.M., Musunuru, K., Smith, A.V., Edmondson, A.C., Stylianou, I.M., Koseki, M., Pirruccello, J.P., Ripatti, S., Chasman, D.I., Willer, C.J., *et al.* (2010) Biological, clinical and population relevance of 95 loci for blood lipids. *Nature*, **466**, 707–713.
33. Global Lipids Genetics Consortium(2013) Discovery and refinement of loci associated with lipid levels. *Nat. Genet.*, **45**, 1274–1283.
34. Lindholm, D., Bornhauser, B.C. and Korhonen, L. (2009) Mylip makes an Idol turn into regulation of LDL receptor. *Cell. Mol. Life Sci.*, **66**, 3399–3402.
35. Raal, F.J., Stein, E.A., Dufour, R., Turner, T., Civeira, F., Burgess, L., Langslet, G., Scott, R., Olsson, A.G., Sullivan, D., *et al.* (2015) PCSK9 inhibition with evolocumab (AMG 145) in heterozygous familial hypercholesterolaemia (RUTHERFORD-2): a randomised, double-blind, placebo-controlled trial. *Lancet*, **385**, 331–340.
36. Relton, C.L. and Davey Smith, G. (2015) Mendelian randomization: applications and limitations in epigenetic studies. *Epigenomics*, **7**, 1239–1243.
37. Bibikova, M., Barnes, B., Tsan, C., Ho, V., Klotzle, B., Le, J.M., Delano, D., Zhang, L., Schroth, G.P., Gunderson, K.L., *et al.* (2011) High density DNA methylation array with single CpG site resolution. *Genomics*, **98**, 288–295.
38. Sandoval, J., Heyn, H.A., Moran, S., Serra-Musach, J., Pujana, M.A., Bibikova, M. and Esteller, M. (2011) Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. *Epigenetics*, **6**, 692–702.
39. Yousefi, P., Huen, K., Schall, R.A., Decker, A., Elboudwarej, E., Quach, H., Barcellos, L. and Holland, N. (2013) Considerations for normalization of DNA methylation data by Illumina 450K BeadChip assay in population studies. *Epigenetics*, **8**, 1141–1152.
40. Kannel, W.B., Feinleib, M., McNamara, P.M., Garrison, R.J. and Castelli, W.P. (1979) An investigation of coronary heart disease in families. The Framingham offspring study. *Am. J. Epidemiol.*, **110**, 281–290.
41. Du, P., Zhang, X., Huang, C.C., Jafari, N., Kibbe, W. a., Hou, L. and Lin, S.M. (2010) Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics*, **11**, 1–9.
42. Dedeurwaerder, S., Defrance, M., Bizet, M., Calonne, E., Bontempi, G. and Fuks, F. (2013) A comprehensive overview of Infinium HumanMethylation450 data processing. *Brief. Bioinform.*, **15**, 929–941.
43. Teschendorff, A.E., Zhuang, J. and Widschwendter, M. (2011) Independent surrogate variable analysis to deconvolve confounding factors in large-scale microarray profiling studies. *Bioinformatics*, **27**, 1496–1505.
44. Houseman, E.A., Accomando, W.P., Koestler, D.C., Christensen, B.C., Marsit, C.J., Nelson, H.H., Wiencke, J.K. and Kelsey, K.T. (2012) DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics*, **13**, 1–16.
45. Aryee, M.J., Jaffe, A.E., Corrada-Bravo, H., Ladd-Acosta, C., Feinberg, A.P., Hansen, K.D. and Irizarry, R.A. (2014) Minfi: A flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics*, **30**, 1363–1369.
46. Leek, J.T., Johnson, W.E., Parker, H.S., Fertig, E.J., Jaffe, A.E. and Storey, J.D. sva: Surrogate Variable Analysis. R package version 3.10.0.
47. Corella, D., Arnett, D.K., Tsai, M.Y., Kabagambe, E.K., Peacock, J.M., Hixson, J.E., Straka, R.J., Province, M., Lai, C.Q., Parnell, L.D., *et al.* (2007) The -256T>C polymorphism in the apolipoprotein A-II gene promoter is associated with body mass index and food intake in the genetics of lipid lowering drugs and diet network study. *Clin. Chem.*, **53**, 1144–1152.
48. Absher, D.M., Li, X., Waite, L.L., Gibson, A., Roberts, K., Edberg, J., Chatham, W.W. and Kimberly, R.P. (2013) Genome-wide DNA methylation analysis of systemic lupus erythematosus reveals persistent hypomethylation of interferon genes and compositional changes to CD4+ T-cell populations. *PLoS Genet.*, **9**, e1003678.
49. Engstrom, P.G., Steijger, T., Sipos, B., Grant, G.R. and Kahles, A., The RGASP Consortium. (2013) Systematic evaluation of spliced alignment programs for RNA-Seq data. *Nat. Methods*, **10**, 1185–1191.
50. Anders, S., Pyl, P.T. and Huber, W. (2015) HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics*, **31**, 166–169.
51. Ali, O., Cerjak, D., Kent, J.W.J., James, R., Blangero, J., Carless, M.A. and Zhang, Y. (2015) An epigenetic map of age-associated autosomal loci in northern European families at high risk for the metabolic syndrome. *Clin. Epigenetics*, **7**, 12.
52. Arner, P., Sinha, I., Thorell, A., Rydén, M., Dahlman-Wright, K. and Dahlman, I. (2015) The epigenetic signature of subcutaneous fat cells is linked to altered expression of genes implicated in lipid metabolism in obese women. *Clin. Epigenetics*, **7**, 1–13.