

RESEARCH

Open Access



Temporal associations between leukocytes DNA methylation and blood lipids: a longitudinal study

Zhiyu Wu^{1†}, Lu Chen^{1†}, Xuanming Hong¹, Jiahui Si², Weihua Cao¹, Canqing Yu^{1,3}, Tao Huang¹, Dianjianyi Sun¹, Chunxiao Liao¹, Yuanjie Pang¹, Zengchang Pang⁴, Liming Cong⁵, Hua Wang⁶, Xianping Wu⁷, Yu Liu⁸, Yu Guo⁹, Zhengming Chen¹⁰, Jun Lv^{1,3,11}, Wenjing Gao^{1*} and Liming Li^{1,3*}

Abstract

Background: The associations between blood lipids and DNA methylation have been investigated in epigenome-wide association studies mainly among European ancestry populations. Several studies have explored the direction of the association using cross-sectional data, while evidence of longitudinal data is still lacking.

Results: We tested the associations between peripheral blood leukocytes DNA methylation and four lipid measures from Illumina 450 K or EPIC arrays in 1084 participants from the Chinese National Twin Registry and replicated the result in 988 participants from the China Kadoorie Biobank. A total of 23 associations of 19 CpG sites were identified, with 4 CpG sites located in or adjacent to 3 genes (*TMEM49*, *SNX5/SNORD17* and *CCDC7*) being novel. Among the validated associations, we conducted a cross-lagged analysis to explore the temporal sequence and found temporal associations of methylation levels of 2 CpG sites with triglyceride and 2 CpG sites with high-density lipoprotein-cholesterol (HDL-C) in all twins. In addition, methylation levels of cg11024682 located in *SREBF1* at baseline were temporally associated with triglyceride at follow-up in only monozygotic twins. We then performed a mediation analysis with the longitudinal data and the result showed that the association between body mass index and HDL-C was partially mediated by the methylation level of cg06500161 (*ABCG1*), with a mediation proportion of 10.1%.

Conclusions: Our study indicated that the DNA methylation levels of *ABCG1*, *AKAP1* and *SREBF1* may be involved in lipid metabolism and provided evidence for elucidating the regulatory mechanism of lipid homeostasis.

Keywords: DNA methylation, Blood lipids, EWAS, Cross-lagged analysis, Mediation analysis

Background

An abnormal blood lipid profile, also known as dyslipidemia, is a major risk factor for atherosclerotic cardiovascular diseases such as myocardial infarction and stroke [1, 2], which mainly manifests as elevated plasma

triglyceride (TG), total cholesterol (TC), low-density lipoprotein-cholesterol (LDL-C) levels and decreased high-density lipoprotein-cholesterol (HDL-C) levels. The prevalence of dyslipidemia contributes to a great global burden of disease. For example, elevated LDL-C levels were one of the top 10 risk factors for all-cause mortality and disability-adjusted life-years worldwide in 2019 [3].

Blood lipid levels are a complex trait regulated by various genetic and environmental factors [4–9], with DNA methylation being a potential regulatory mechanism. DNA methylation is the "annotation system" of gene

[†]Zhiyu Wu and Lu Chen contributed equally to this work.

*Correspondence: pkuepigwj@126.com; lmlee@vip.163.com

¹ Department of Epidemiology and Biostatistics, School of Public Health, Peking University, Beijing 100191, China
Full list of author information is available at the end of the article



sequences and affects gene expression by dynamically modifying the methylation or demethylation state of CpG sites on DNA. The methylation status of the same individual and the same CpG site may change over time. DNA methylation modification is related to the physiological effects of environmental factors on the human body and is also involved in the pathological process of disease occurrence and progression [10]. Therefore, DNA methylation may play an important role in maintaining blood lipid homeostasis.

Several epigenome-wide association studies (EWAS) have been conducted in predominantly European populations to explore the association between DNA methylation and blood lipid levels and have identified CpG sites on genes such as *ABCG1*, *CPT1A*, and *SREBF1* [11–15]. However, most existing studies have mostly been based on cross-sectional data to analyze the association between DNA methylation and blood lipids, which makes it difficult to assess the causal relationship between them. Mendelian randomization (MR) is an effective method for causal inference with cross-sectional data [16]. Some studies have used MR to estimate the causal relationships between DNA methylation levels and blood lipid levels [17–19]. However, the application of MR is limited because of the sample size and strict assumptions. Due to the variability of DNA methylation, studies based on longitudinal data are the best approach for causal inference.

As one of the risk factors for dyslipidemia, body mass index (BMI) is also associated with DNA methylation, and several CpG sites are associated with both BMI and blood lipids [20]. Previous studies proved that BMI could influence the methylation level of certain CpG sites [21–23], but whether DNA methylation mediates the association between BMI and blood lipids remains unclear.

Thus, we performed an EWAS of blood lipids with a discovery stage and independent replication in participants from the Chinese National Twin Registry (CNTR) and China Kadoorie Biobank (CKB). We assessed the temporal association between replicated CpG sites and lipids using the cross-lagged panel model (CLPM) in 308 participants with follow-up data from CNTR. For the CpG sites influencing blood lipids, we further conducted a mediation analysis with the CLPM to examine the role of DNA methylation in the association between BMI and blood lipids.

Results

Demographic

After quality control, 1060 participants in CNTR and 948 participants in CKB were retained in the EWAS (see Additional file 1: Table S1 for further details). The characteristics of the EWAS participants were shown in Table 1. The mean age of the two populations was

around 50 years old (49.90 ± 12.15 and 50.56 ± 7.48). Males or never smokers accounted for more than half of the participants. In the discovery stage, approximately half of the people had never consumed alcohol, while in the replication stage, approximately 60% of the participants were current drinkers. Participants in the replication stage were more likely to have lower BMI, higher TG levels, and lower TC, HDL-C, and LDL-C levels than those in the discovery stage.

EWAS

For Model 1, the genomic inflation factors (λ) ranged between 1.009 and 1.049. We identified 26 CpG-lipid associations in the discovery stage, and methylation levels of 17, 3, 5, and 1 CpG sites were associated with TG, TC, HDL-C, and LDL-C, respectively (Additional file 1: Figure S1). A total of 23 associations of 19 CpG sites were replicated in the replication stage, and 15, 3, 5, and 0 CpG sites remained significant for the four lipid measures, respectively (Table 2). The direction of the effect was consistent in the two stages. cg06500151 (*ABCG1*), cg11024682 (*SREBF1*), cg19693031 (*TXNIP*) and cg27243685 (*ABCG1*) were associated with two lipid measures. In addition, we identified 4 novel CpG sites located in 2 genes (cg12054453 and cg18942579 in *TMEM49* and cg17507897 in *SNX5/SNORD17*) and 1 intergenic region (cg05176551 adjacent to *CCDC7*).

When we further adjusted for BMI in Model 2, only 1/3 (8/23) of the associations were robust. The number of significant CpG sites was reduced by half for TG (7/15) and TC (1/3), and no CpG site was associated with HDL-C (Additional file 1: Table S2). The effect estimates were attenuated except for cg05176551. Two of the four novel CpG sites (cg05176551 for TG and cg12054453 for TC) remained significant.

Enrichment analysis

The results of enrichment analyses based on the EWAS from model 1 were shown in Additional file 1: Table S3–S6. GO enrichment analysis for TG revealed lipid and cellular ketone metabolic pathways. Insulin and glucose metabolism, peptide synthesis, and regulatory processes were also enriched in the analysis, which appeared to be related to TG homeostasis. Sterol signaling pathway, as well as alcohol signaling pathway, were enriched in the GO enrichment analysis for HDL. KEGG and Reactome analysis suggested enrichment for pathways associated with liver disease, lipid metabolism, and antioxidants such as PPARalpha and HMOX1, although the pathways did not survive after multiple comparisons (FDR > 0.05).

Table 1 Characteristics of participants in the EWAS phase

	EWAS		Temporal association	
	Discovery stage	Replication stage	Baseline	Follow-up
<i>N</i>	1060	948	288	
Age, yrs	49.90 ± 12.15	50.56 ± 7.48	49.60 ± 10.05	54.26 ± 10.02
Sex, <i>n</i> (%)				
Female	337 (31.8)	416 (43.9)	110 (38.2)	
Male	723 (68.2)	532 (56.1)	178 (61.8)	
MZ, <i>n</i> (%)	748 (70.6)	–	174 (60.4)	
Fasting status, <i>n</i> (%)				
≥ 8 h	–	123 (13.0)	–	
< 8 h	–	825 (87.0)	–	
Smoking status, <i>n</i> (%)				
Never	575 (54.2)	514 (54.2)	175 (60.8)	162 (56.3)
Former	136 (12.8)	51 (5.4)	25 (8.7)	45 (15.6)
Current	349 (32.9)	383 (40.4)	88 (30.6)	81 (28.1)
Alcohol consumption, <i>n</i> (%)				
Never	539 (50.8)	365 (38.5)	134 (46.5)	125 (43.4)
Former	73 (6.9)	15 (1.6)	9 (3.1)	82 (28.5)
Current	448 (42.3)	568 (59.9)	145 (50.3)	81 (28.1)
BMI, kg/m ²	24.83 ± 3.91	23.57 ± 3.49	24.35 ± 3.56	24.46 ± 3.55
TG, mmol/L	1.85 ± 2.00	1.94 ± 1.18	1.86 ± 1.60	1.81 ± 1.34
TC, mmol/L	4.88 ± 1.04	4.58 ± 1.01	4.71 ± 0.98	4.79 ± 0.89
HDL-C, mmol/L	1.36 ± 0.38	1.21 ± 0.30	1.37 ± 0.32	1.29 ± 0.36
LDL-C, mmol/L	2.50 ± 0.77	2.28 ± 0.73	2.13 ± 0.61	2.72 ± 0.81

Continuous variables are expressed as mean ± SD and categorical variables are expressed as *n* (%)

EWAS, epigenome-wide association study; MZ, Monozygotic twins; BMI, body mass index; TG, triglyceride; TC, total cholesterol; HDL-C, high-density lipoprotein-cholesterol; LDL-C, low-density lipoprotein-cholesterol

Cross-lagged analysis

Our cross-lagged analysis showed that two CpG sites had significant temporal associations with TG and two had significant temporal associations with HDL-C in all twins (Table 3). The paths of TG at baseline to the methylation levels of cg27243685 and cg05778424 at follow-up (ρ_1) were significant, suggesting that the TG level had an effect on these CpG sites. We observed the same direction of association between HDL-C and the methylation level of cg11024682. The reverse path (ρ_2) was significant for cg06500161 and HDL-C, suggesting that this CpG site could influence HDL-C levels. We did not detect any temporal associations between CpG sites and TC.

We then performed a stratification analysis and revealed some new temporal associations. In monozygotic (MZ) twins, the methylation level of cg11024682 at baseline showed a significant association with TG at follow-up (ρ_2). In dizygotic (DZ) twins, cg17507897 also showed an association between HDL-C at baseline and methylation level at follow-up (ρ_1). The associations between cg05778424, cg06500161, and cg11024682

with TG or HDL-C could still be observed in only DZ twins (Additional file 1: Table S8).

Mediation analysis

According to the results of the cross-lagged analysis, cg06500161 and cg11024682, which had effects on HDL-C or TG, were selected as potential mediators in the mediation analysis. We first estimated the temporal association between BMI and lipid or CpG sites using the CLPM. In all twins, BMI at baseline was associated with TG, HDL-C, and the methylation level of the two CpG sites at follow-up (Table 4). We then performed a mediation analysis with the methylation level of CpG sites as the mediator, BMI as exposure, and lipid measures as the outcome, and only the mediating effect of cg06500161 was observed. The results are shown in Fig. 1. BMI was associated with HDL-C with a significant direct effect, indirect effect, and total effect (all $P < 0.05$). The methylation level of cg06500161 explained 10.1% ($P < 0.05$) of the association between BMI and HDL-C.

Table 2 Associations between DNA methylation and lipid measures (Model 1)

CpG	Discovery stage			Replication stage			Position	Gene
	β	SE	P_{adj}	β	SE	P_{adj}		
<i>TG</i>								
cg06500161	0.0130	0.0013	2.41E-14	0.0088	0.0013	1.51E-10	21:43,656,587	<i>ABCG1</i>
cg00574958	-0.0074	0.0008	3.63E-12	-0.0084	0.0010	1.20E-14	11:68,607,622	<i>CPT1A</i>
cg11024682	0.0085	0.0011	1.77E-09	0.0059	0.0011	9.33E-08	17:17,730,094	<i>SREBF1</i>
cg19693031	-0.0211	0.0027	2.41E-09	-0.0110	0.0022	1.15E-06	1:145,441,552	<i>TXNIP</i>
cg17058475	-0.0084	0.0011	4.34E-08	-0.0100	0.0012	5.70E-16	11:68,607,737	<i>CPT1A</i>
cg07504977	0.0102	0.0014	6.70E-07	0.0066	0.0016	3.47E-05	10:102,131,012	<i>OLMALINC</i>
cg27243685	0.0090	0.0013	3.88E-06	0.0084	0.0014	4.42E-09	21:43,642,366	<i>ABCG1</i>
cg01176028	0.0055	0.0009	5.82E-04	0.0047	0.0010	1.88E-06	21:43,653,234	<i>ABCG1</i>
cg05778424	0.0078	0.0014	7.36E-04	0.0057	0.0012	7.46E-06	17:55,169,508	<i>AKAP1</i>
cg08857797	0.0087	0.0016	3.16E-03	0.0073	0.0016	5.06E-06	17:40,927,699	<i>VPS25</i>
cg26403843	0.0108	0.0020	3.96E-03	0.0062	0.0024	1.11E-02	5:158,634,085	<i>RNF145</i>
cg06690548	-0.0089	0.0017	1.44E-02	-0.0037	0.0013	6.10E-03	4:139,162,808	<i>SLC7A11</i>
cg09737197	-0.0059	0.0012	2.13E-02	-0.0068	0.0012	7.32E-08	11:68,607,675	<i>CPT1A</i>
cg20544516	0.0054	0.0011	2.40E-02	0.0066	0.0011	7.27E-09	17:17,717,183	<i>MIR33B; SREBF1</i>
cg05176551	0.0079	0.0016	3.24E-02	0.0064	0.0014	5.83E-06	10:32,701,586	<i>CCDC7</i>
<i>TC</i>								
cg12054453	0.0395	0.0060	5.95E-05	0.0154	0.0059	1.22E-02	17:57,915,717	<i>TMEM49</i>
cg19693031	-0.0416	0.0078	2.76E-02	-0.0286	0.0062	1.62E-05	1:145,441,552	<i>TXNIP</i>
cg18942579	0.0278	0.0053	4.23E-02	0.0138	0.0055	1.22E-02	17:57,915,773	<i>TMEM49</i>
<i>HDL-C</i>								
cg06500161	-0.0218	0.0031	4.23E-06	-0.0204	0.0032	9.44E-10	21:43,656,587	<i>ABCG1</i>
cg17901584	0.0243	0.0043	3.01E-03	0.0305	0.0048	9.44E-10	1:55,353,706	<i>DHCR24</i>
cg11024682	-0.0138	0.0024	3.01E-03	-0.0122	0.0027	6.44E-06	17:17,730,094	<i>SREBF1</i>
cg27243685	-0.0167	0.0030	6.39E-03	-0.0166	0.0034	2.05E-06	21:43,642,366	<i>ABCG1</i>
cg17507897	-0.0179	0.0033	1.07E-02	-0.0093	0.0034	5.96E-03	20:17,943,694	<i>SNX5; SNORD17</i>

The novel CpG sites are in bold font. P -values are adjusted for FDR

β , regression coefficient; SE, standard error; TG, triglyceride; TC, total cholesterol; HDL-C, high-density lipoprotein-cholesterol

Table 3 Cross-lagged association between lipid measures and DNA methylation

CpG	Gene	Lipid _{baseline} → Methylation _{follow-up}			Methylation _{baseline} → Lipid _{follow-up}			Model fit	
		β	SE	P_{adj}	β	SE	P_{adj}	SRMR	CFI
<i>TG</i>									
cg27243685	<i>ABCG1</i>	0.3884	0.1426	0.0483	0.0264	0.0220	0.6682	<0.001	1
cg05778424	<i>AKAP1</i>	0.5423	0.1660	0.0163	0.0147	0.0209	0.9468	<0.001	1
<i>HDL-C</i>									
cg06500161	<i>ABCG1</i>	-0.5781	0.5812	0.3999	-0.0320	0.0088	0.0014	<0.001	1
cg11024682	<i>SREBF1</i>	-1.4046	0.4655	0.0127	-0.0193	0.0121	0.1837	<0.001	1

Adjusted P -values less than 0.05 are in bold font

β , regression coefficient; SE, standard error; TG, triglyceride; HDL-C, high-density lipoprotein-cholesterol; SRMR, standardized root mean squared residual; CFI, comparative fit index

Discussion

In the current study, we reported the associations between lipid measures and DNA methylation and inferred the potential causal direction of the associations.

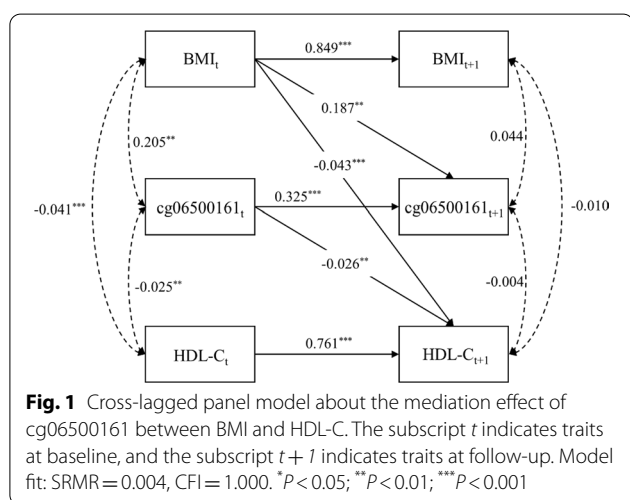
To our knowledge, this is the first EWAS of blood lipids reported in the Chinese population. We identified 23 associations of 19 CpG sites, and 4 CpG sites located in or adjacent to 3 genes were novel. With the cross-lagged

Table 4 Cross-lagged association between BMI and lipid measures or between BMI and DNA methylation

Lipid measures or CpG	Gene	BMI _{baseline} → Trait _{follow-up}			Trait _{baseline} → BMI _{follow-up}			Model fit	
		β	SE	P	β	SE	P	SRMR	CFI
<i>BMI and lipid measures</i>									
TG	–	0.0675	0.0199	0.0007	0.0698	0.1133	0.5376	<0.001	1
HDL	–	–0.0467	0.0113	<0.0001	–0.0009	0.2667	0.9972	<0.001	1
<i>BMI and CpG sites</i>									
cg06500161	<i>ABCG1</i>	0.1853	0.0558	0.0018	–0.0019	0.0387	0.9600	<0.001	1
cg11024682	<i>SREBF1</i>	0.1062	0.0512	0.0380	0.0237	0.0428	0.9600	<0.001	1

Adjusted P-values less than 0.05 are in bold font

β, regression coefficient; SE, standard error; TG, triglyceride; HDL-C, high-density lipoprotein-cholesterol; SRMR, standardized root mean squared residual; CFI, comparative fit index



panel model, we found potential causal relationships of 2 CpG sites with TG and 2 CpG sites with HDL-C among all the significant associations and a relationship of cg11024682 with TG in only MZ twins. In addition, we conducted a mediation analysis and found that BMI has an effect on HDL-C, and the methylation level of cg06500161 partially mediates this effect.

In our EWAS, we confirmed that the widely validated CpG sites in European populations, such as cg06500161 in *ABCG1*, cg00574958 in *CPT1A*, and cg11024682 in *SREBF1*, were also associated with lipid measures robustly in East Asian populations and the direction of the associations were consistent. This complemented the findings of Jhun et al. [24] that the associations between blood lipids and DNA methylation have consistency across ethnic groups, including Europeans, African Americans, and Hispanics. Another study compared differences in DNA methylation patterns of multiple traits related to non-communicable diseases between black South Africans and Europeans and found that 95% confidence intervals for effect estimates overlapped for more

than 85% of the shared lipid-related CpG sites [25]. However, we still identified 3 novel genes with TG (*CCDC7*), TC (*TMEM49*), and HDL-C (*SNX5/SNORD17*) with moderate sample size, and the latter two remained robust after adjustment for BMI. The differences in the methylation patterns may be associated with the genetic background of different ethnic groups [25], but the exact mechanisms need to be further explored.

Two CpG sites located in *TMEM49* were related to TC. *TMEM49*, also known as *VMPI*, encodes an endoplasmic reticulum (ER) transmembrane protein that regulates the formation of autophagosomes, lipid droplets, lipoproteins, and other ER-derived structures [26, 27]. *TMEM49* has been reported to affect the activity of lipoprotein-associated phospholipase A (2) (Lp-PLA2) [28], an inflammatory enzyme that is a risk factor for coronary heart disease and in which lipids may play a role [29, 30]. Lp-PLA2 was bound to serum cholesterol, and in vitro experiments showed that elevated LDL-C levels result in upregulation of Lp-PLA2 [31]. All these findings indicated a potential role for *TMEM49* in TC homeostasis. In addition, the methylation level of *TMEM49* was also associated with waist circumference [32], survival of cancer overall [33], and chemotherapy in breast cancer patients [34].

SNX5, short for sorting nexin 5, is a key regulator of endosomal trafficking [35], and *SNX5* has been identified as an LDL-C-associated gene [36]. Elevated or decreased expression of *SNX5* was found in multiple cancers [37, 38], and renal *SNX5* was reported to positively regulate insulin-degrading enzyme expression and function [39]. Noncoding small nucleolar RNA *SNORD17* is also related to carcinogenesis, including hepatocellular carcinoma [40], cervical cancer [41], and colon adenocarcinoma [42]. The association of *SNX5/SNORD17* with HDL-C needs to be further investigated.

Previous studies have reported an association between BMI and some lipid-related CpG sites [20, 43, 44], which

was validated to some extent in our Model 2 results. The magnitude of the association between lipid measures and methylation was attenuated when adjusting for BMI, indicating that BMI may partially influence the association. None of the HDL-C-related CpG sites remained significant. Given that these CpG sites were also BMI-related, we hypothesized that BMI might affect HDL-C through the CpG sites. This hypothesis was partially confirmed by the subsequent mediation analysis.

Based on the results of the cross-lagged analysis, we found that the methylation levels of *ABCG1* could influence HDL-C levels and be influenced by TG levels. Previous studies have explored the causal relationship between *ABCG1* methylation and TG and HDL-C using MR and gene expression analysis methods and obtained opposite directions [17, 45, 46]. Min et al. and Dekkers et al. [17] both found that the methylation level of *ABCG1* was influenced by TG or HDL-C using MR analysis, while Pfeiffer et al. [46] suggested that the methylation level of *ABCG1* could regulate TG and HDL-C levels through gene expression. Since our study used longitudinal data while the three studies above were based on cross-sectional data, we believe that our results are more credible. The ATP-binding cassette transporter (*ABCG1*) is the critical mediator of reverse cholesterol transport (RCT) and mediates cellular cholesterol efflux to HDL particles [47–50]. The methylation level of cg06500161 (*ABCG1*) was correlated with the *ABCG1* transcript level, which in turn had an impact on HDL-C [46].

Our results also showed that the methylation level of *AKAP1* might be influenced by TG. A-kinase anchoring protein 121 (*AKAP1*) binds protein kinase A and anchors it to the mitochondrial outer membrane to maintain mitochondrial function [51]. Animal experiments have demonstrated that *AKAP1* is involved in the regulation of endothelial cell behavior [52], oxidative stress, and apoptosis [53] and plays a role in multiple metabolism-related diseases. For instance, researchers observed that energy expenditure and thermogenesis were significantly enhanced in brown adipose tissue of *AKAP1* knockout obese mice, which could attenuate diet-induced obesity and insulin resistance [54]. DNA methylation may be an intermediate link in the association of *AKAP1* with these diseases, and TG, as one of the important metabolic indicators, may have a similar regulatory mechanism.

In the stratification analysis, a noteworthy result was that cg11024682 located in *SREBF1* had a potential influence on TG levels only in MZ twins. The direction of the association between TG and *SREBF1* in our results was not completely consistent with a previous study, which found that the methylation level of *SREBF1* was influenced by TG using MR analysis [17]. We hypothesize that the inconsistent results are due to genetic and early-life

environmental factors. MR analysis used genetic variants as instrumental variables and could not exclude confounding from genetic factors. Since MZ twins share almost 100% genetic background and early-life environment, our result, which was found only in MZ twins, suggests that genetic and early-life environmental factors may influence the effects of the CpG site on TG. *SREBF1* encodes sterol regulatory element-binding protein 1, which can activate and synthesize cholesterol and fatty acids [55]. MicroRNA (miR)-33b, located in the intron region of *SREBF1*, can also act as an important regulator of lipid metabolism [56, 57]. Antagonism of miR-33 inhibited the expression of genes involved in fatty acid synthesis, such as *SREBF1*, and thus reduced plasma very low-density lipoprotein triglyceride levels [58].

Our mediation analysis supported that cg06500161 (*ABCG1*) partially mediated the effect of BMI on HDL-C. Previous studies have reported that the methylation level of cg06500161 could be influenced by BMI [21–23] and mediate the effect of BMI on the expression of *ABCG1* [59]. Johansson and his colleagues found that *ABCG1* and *CETP* were the most upregulated genes that were differentially expressed in obese patients during weight loss and weight maintenance after weight loss, and the prevailing HDL concentration was correlated with the expression of *ABCG1* [60]. Similar results were observed in morbidly obese women [61]. However, the above two studies only analyzed *ABCG1* expression levels in adipose tissue rather than blood samples. Therefore, the mediating role of *ABCG1* and its methylation level between BMI and HDL-C remains to be further explored.

The strength of our study is that we explored the temporal association between CpG sites and blood lipids using longitudinal data and corresponding analysis, circumventing the limitations of cross-sectional data in causal inference. In addition, the analysis in twins, especially MZ twins, provides a natural matched design to present causal inference while controlling for genetic and early family environmental factors [62].

There are also some limitations of our study. In the EWAS phase, to maximize the sample size, only probes common to both the 450 K and EPIC methylation arrays were included in our study, while other probes present in only one methylation array were excluded. Some potential lipid-associated CpG sites may not have been identified. In addition, the CKB cohort did not require participants to fast prior to blood sample collection, which may have affected the validation results. Therefore, we adjusted the fasting status in our analysis and used SmartSVA to control for potential confounding, avoiding the interference of fasting time on the results to some extent. In the cross-lagged analysis phase, the power of the stratification analysis was limited due to the moderate

sample size. However, we still found some evidence that the association between TG and *SREBF1* may be influenced by genetic and early-life environmental factors.

Conclusions

We identified 4 novel CpG sites related to lipid measures in the Chinese population. Based on the longitudinal twin data, we found the temporal sequence of DNA methylation levels of *ABCG1*, *AKAP1*, and *SREBF1* with TG and HDL-C. In addition, we observed a potential mediation role of cg06500161 (*ABCG1*) in the temporal association between BMI and HDL-C. Our study provides evidence to elucidate the underlying biological mechanisms of DNA methylation in lipid metabolism, and future studies should continue to explore the biological role of DNA methylation in lipid metabolism and downstream effects on disease.

Method

Study population

The discovery stage of the EWAS phase and cross-lagged analysis phase was based on CNTR. Established in 2001, CNTR is a population-based twin registry, and the details of CNTR have been reported previously [63]. The current study included participants who participated in the follow-up surveys in 2013–2014 and 2017–2018. A total of 1392 blood specimens were collected for methylation measurements from 1084 participants, including 308 participants with repeated measurements. The replication stage of EWAS was based on CKB, a prospective cohort of more than 0.5 million adults in 10 geographically defined regions across China since 2004–2008 [64, 65]. DNA methylation data were available for 988 participants selected for a case–control study nested within CKB [66]. Information from the two cohorts was collected via standardized questionnaires, physical examinations, and blood biochemical examinations.

Data measurements

Lipid measures, including TG, TC, HDL-C, and LDL-C, were measured in blood samples. Participants were asked to fast for at least 8 h in CNTR, while fasting was not mandatory in CKB. The fasting time for each participant was recorded and adjusted in the subsequent analysis in CKB. Other covariates, such as age, sex, smoking status, and alcohol consumption, were collected using the interview-administered questionnaire. Smoking status and alcohol consumption were divided into three categories: never, former and current. Height and weight were collected in the physical examination to calculate BMI, which was defined as weight in kilograms divided by the square of height in meters. Medication history was also recorded, and participants taking lipid-lowering

medicine were excluded. Outliers that were three standard deviations from the mean of lipid measures in each cohort were removed. Blood samples with missing biochemical parameters or moderate to severe lipemia were also excluded. All lipid measures were natural log-transformed to approximate a normal distribution.

The zygosity of twins was determined based on the correlation of 59 SNPs in both the 450 K and EPIC methylation arrays. It was shown that the possible cutoff point was between 0.84 and 0.90 [67], which was set to 0.90 in the current study. Twin pairs with a correlation coefficient higher than 0.90 were considered MZ twins; otherwise, they were considered DZ twins.

DNA methylation and quality control

Genomic DNA was extracted from peripheral blood leukocytes and bisulfite-converted using the EZ DNA methylation kit (Zymo Research, Orange, CA, USA). Epigenome-wide DNA methylation levels were measured using the Infinium HumanMethylation450 BeadChip assay (Illumina, San Diego, CA, USA) or Infinium HumanMethylationEPIC BeadChip assay (Illumina, San Diego, CA, USA), and only the overlapping CpG sites of the two assays were included in the subsequent procedure.

DNA methylation measurement and preprocessing were conducted independently in the two cohorts. In CNTR, we applied the R package *minfi* [68] to process and combine the raw methylation data of two assays and obtained the β -value of each CpG site to report the methylation level. For quality control, we removed the probes if they (1) had detection p values > 0.05 in more than 1% of samples or had bead counts < 3 in more than 5% of samples; (2) were non-CpG or multi-hit probes; (3) were related to SNPs with MAF > 0.05 in the 1000 Genomes Project for the East Asian population; and (4) were located in sexual chromosomes. We excluded samples if they (1) were sex mismatched and (2) had a detection *P*-value greater than 0.01. The details of quality control of the raw methylation data in CKB can be found in a previous study [66]. The stratified quantile normalization method [69] in the *minfi* package was used for preprocessing and normalization. The cell proportions for each cell type were estimated using Houseman's method [70] and adjusted using the *champ.refbase* function in the R package ChAMP. To correct batch effects, we performed an optimized surrogate variable analysis (SVA) with the package SmartSVA [71] in the EWAS phase. SmartSVA provided a fast and robust method to remove potential confounding factors for epigenetic or other genomic studies and was developed based on the linear model. In the cross-lagged analysis phase, batch effects were corrected using the Combat method instead.

Finally, 378,654 CpG sites were retained in the discovery stage of EWAS. We annotated CpG sites to genes with the manifest file provided on the Illumina website. The CpG sites located in the intergenic region would be annotated to the nearest gene using the R package matchGenes or genome browser (<https://genome.ucsc.edu/>).

Statistical analysis

EWAS

We implemented a two-stage EWAS with participants in CNTR and CKB. For Model 1, linear mixed regression models were fitted with β -value at a CpG site as the dependent variable and each lipid measure (TG, TC, HDL-C, and LDL-C) as the independent variable using the R package nlme, adjusting for age, sex, smoking status, alcohol consumption, and all surrogate variables generated above as fixed effects. Due to the correlation within twin pairs, the unique ID of each twin pair was added to the model as a random effect in the discovery stage. The fasting status was further adjusted as a fixed effect (< 8 or ≥ 8 h) in the replication stage. BMI was further adjusted in Model 2.

Enrichment analysis

For all CpG sites associated with lipid measures, we then performed Gene Ontology (GO) term analyses, Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses, and Reactome pathway enrichment analysis with the R package "methylGSA" [72].

Cross-lagged analysis

We conducted a cross-lagged analysis based on participants with repeated measurements in CNTR, of whom the average follow-up duration was 4.67 ± 0.22 years. A residual analysis was performed first. Lipid measures were adjusted for age, sex, smoking status, and alcohol consumption as fixed effects at the corresponding time point and twin number as a random effect using the linear mixed regression model. The β -value of significant CpG sites was further adjusted with the blood cell proportions and batch effects. Residuals from the regression model were used in the subsequent cross-lagged analysis, and residuals of CpG sites were normalized using Z-transform. The CLPM simultaneously estimated the autoregressive and cross-lagged regressive effects of lipid measure and methylation data at two-time points, including (1) autoregression of lipid measure at follow-up on lipid measure at baseline, (2) autoregression of β value at follow-up on β value at baseline, (3) cross-lagged regression of β -value at follow-up on lipid measure at baseline (ρ_1), and (4) cross-lagged regression of lipid measure at follow-up on β -value at baseline (ρ_2). The significance and magnitude of ρ_1 and ρ_2 reflected

the temporal associations of the two variables. We fitted a structural equation model to estimate all the parameters and statistics above using the R package *lavaan* and set the cluster argument to adjust for the correlation of twins. We also reported model fit indexes to evaluate model fit, including standardized root mean squared residual (SRMR) and comparative fit index (CFI), and models with $SRMR < 0.08$ and $CFI > 0.95$ were considered good fits [73]. We repeated the steps above in only MZ or DZ, and only twin pairs with complete data at two-time points were included.

Mediation analysis

Since BMI was associated with DNA methylation and could affect blood lipids, we performed mediation analysis to assess whether CpG sites were the mediator of the effect of BMI on lipids. Only CpG sites that showed a potential effect on lipid measures in the cross-lagged analysis were included.

The temporal associations between BMI and lipid measures and between BMI and CpG sites were assessed with the CLPM first to deduce the potential mediator. The relationships among exposure X , mediator M and outcome Y at baseline (b) and follow-up (f) were as follows:

$$X_f = \beta_X X_b + \varepsilon_{Xf} \quad (1)$$

$$M_f = \beta_M M_b + aX_b + \varepsilon_{Mf} \quad (2)$$

$$Y_f = \beta_Y Y_b + bM_b + c'X_b + \varepsilon_{Yf} \quad (3)$$

where β is the autoregressive coefficient, ε is the residual item, ab is the direct effect, and c' is the indirect effect. The model fitting and parameter estimation method was similar to that of the cross-lagged analysis.

The false discovery rate (FDR) was used for multiple comparisons, and the threshold of significance was considered to be less than 0.05. Statistical analysis was performed using R version 4.0.2 or 4.0.3.

Abbreviations

ABCG1: ATP-binding cassette transporter; AKAP1: A-kinase anchoring protein 121; BMI: Body mass index; CFI: Comparative fit index; CKB: China Kadoorie Biobank; CLPM: Cross-lagged panel model; CNTR: Chinese national twin registry; DZ: Dizygotic twin; ER: Endoplasmic reticulum; EWAS: Epigenome-wide association study; FDR: False discovery rate; GO: Gene ontology; HDL-C: High-density lipoprotein-cholesterol; KEGG: Kyoto encyclopedia of genes and genomes; LDL-C: Low-density lipoprotein-cholesterol; Lp-PLA2: Lipoprotein-associated phospholipase A (2); MR: Mendelian randomization; MZ: Monozygotic twin; RCT: Reverse cholesterol transport; SNORD17: Small nucleolar RNA, C/D box 17; SNX5: Sorting nexin 5; SREBF1: Sterol regulatory element-binding protein 1; SRMR: Standardized root mean squared residual; SVA: Surrogate variable analysis; TC: Total cholesterol; TG: Triglyceride; TMEM49: Transmembrane protein 49.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13148-022-01356-x>.

Additional file 1: Table S1. Details of sample size and quality control. **Figure S1.** Manhattan plots of the epigenome-wide association study (Model 1). Manhattan plots for (a) triglyceride (TG), (b) total cholesterol (TC), (c) high-density lipoprotein-cholesterol (HDL-C), and (d) low-density lipoprotein-cholesterol (LDL-C) in Model 1. The red horizontal dashed lines represent the FDR-adjusted threshold of significance. **Table S2.** Associations between DNA methylation and lipid measures (Model 2). **Table S3.** Enriched GO terms based on the results of TG EWAS. **Table S4.** Enriched GO terms based on the Model 1 results of HDL EWAS. **Table S5.** KEGG enriched pathways based on the results of EWAS. **Table S6.** Reactome enriched pathways based on the results of EWAS. **Table S8.** Cross-lagged association between lipid measures and DNA methylation stratified by zygosity.

Acknowledgements

The most important acknowledgement is to the participants in the study and the members of the survey teams in each regional center, as well as to the project development and management teams based at Beijing, Oxford and each regional center.

Author contributions

ZW, WG, and LL designed the study. ZP, LC, HW, XW, and YL contributed to cohort design and data collection for CNTR, and YG, and ZC contributed to cohort design and data collection for CKB. ZW, LC, XH, and JS performed the analyses. ZW and WG drafted the manuscript. WC, JL, CY, TH, DS, CL, and YP were involved in critical review of manuscript. All authors read and approved the final manuscript.

Funding

The CNTR was supported by National Natural Science Foundation of China (82073633, 81973126, 81573223) and the Special Fund for Health scientific research in the public welfare (201502006, 201002007). This work for CKB was supported by National Natural Science Foundation of China (81390544, 81390540). The CKB baseline survey and the first re-survey were supported by a grant from the Kadoorie Charitable Foundation in Hong Kong. The long-term follow-up is supported by grants from the UK Wellcome Trust (212946/Z/18/Z, 202922/Z/16/Z, 104085/Z/14/Z, 088158/Z/09/Z), grants (2016YFC0900500) from the National Key R&D Program of China, National Natural Science Foundation of China (82192900, 82192901, 82192904, 91846303, 81941018), and Chinese Ministry of Science and Technology (2011BAI09B01). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The funders had no role in the study design, data collection, data analysis and interpretation, writing of the report, or the decision to submit the article for publication.

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

All participants provided written informed consent. The study protocol was approved by the Biomedical Ethics Committee of Peking University (Beijing, China: IRB00001052-13022, IRB00001052-14021), the Ethics Review Committee of the Chinese Center for Disease Control and Prevention (Beijing, China: 005/2004), and the Oxford Tropical Research Ethics Committee, University of Oxford (UK: 025-04).

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Epidemiology and Biostatistics, School of Public Health, Peking University, Beijing 100191, China. ²National Institute of Health Data Science at Peking University, Peking University, Beijing 100191, China. ³Peking University Center for Public Health and Epidemic Preparedness and Response, Beijing 100191, China. ⁴Qingdao Center for Disease Control and Prevention, Qingdao 266033, China. ⁵Zhejiang Center for Disease Control and Prevention, Hangzhou 310051, China. ⁶Jiangsu Center for Disease Control and Prevention, Nanjing 210008, China. ⁷Sichuan Center for Disease Control and Prevention, Chengdu 610041, China. ⁸Heilongjiang Center for Disease Control and Prevention, Harbin 150090, China. ⁹Fuwai hospital Chinese Academy of Medical Sciences, Beijing 100037, China. ¹⁰Clinical Trial Service Unit and Epidemiological Studies Unit (CTSUS), Nuffield Department of Population Health, University of Oxford, Oxford OX3 7LF, UK. ¹¹Key Laboratory of Molecular Cardiovascular Sciences (Peking University), Ministry of Education, Beijing 100191, China.

Received: 23 August 2022 Accepted: 13 October 2022

Published online: 23 October 2022

References

- Berberich AJ, Hegele RA. A modern approach to dyslipidemia. *Endocr Rev.* 2022;43(4):611–53.
- Crea F. Dyslipidaemias in stroke, chronic kidney disease, and aortic stenosis: the new frontiers for cholesterol lowering. *Eur Heart J.* 2021;42(22):2137–40.
- GBD 2019 Risk Factors Collaborators. Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the global burden of disease study 2019. *Lancet.* 2020;396(10258):1223–49.
- Gomez-Delgado F, Katsiki N, Lopez-Miranda J, Perez-Martinez P. Dietary habits, lipoprotein metabolism and cardiovascular disease: from individual foods to dietary patterns. *Crit Rev Food Sci Nutr.* 2021;61(10):1651–69.
- Lin WY. A large-scale observational study linking various kinds of physical exercise to lipoprotein-lipid profile. *J Int Soc Sports Nutr.* 2021;18(1):35.
- Ruiz-Ramie JJ, Barber JL, Sarzynski MA. Effects of exercise on HDL functionality. *Curr Opin Lipidol.* 2019;30(1):16–23.
- Aho S, Vuoristo MS, Raitanen J, Mansikkamäki K, Alanko J, Vähä-Ypää H, Luoto R, Kellokumpu-Lehtinen PL, Vasankari T. Higher number of steps and breaks during sedentary behaviour are associated with better lipid profiles. *BMC Public Health.* 2021;21(1):629.
- Oh SS, Jang JE, Lee DW, Park EC, Jang SI. Cigarette type or smoking history: which has a greater impact on the metabolic syndrome and its components? *Sci Rep.* 2020;10(1):10467.
- Klarin D, Damrauer SM, Cho K, Sun YV, Teslovich TM, Honerlaw J, Gagnon DR, DuVall SL, Li J, Peloso GM, et al. Genetics of blood lipids among ~300,000 multi-ethnic participants of the million veteran program. *Nat Genet.* 2018;50(11):1514–23.
- Dor Y, Cedar H. Principles of DNA methylation and their implications for biology and medicine. *Lancet.* 2018;392(10149):777–86.
- Braun KVE, Dhana K, de Vries PS, Voortman T, van Meurs JBJ, Uitterlinden AG, Hofman A, Hu FB, Franco OH, Dehghan A. Epigenome-wide association study (EWAS) on lipids: the Rotterdam study. *Clin Epigenetics.* 2017;9:15.
- Hedman ÅK, Mendelson MM, Marioni RE, Gustafsson S, Joehanes R, Irvin MR, Zhi D, Sandling JK, Yao C, Liu C, 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(1):e001487.
- Tobi EW, Slieker RC, Luijk R, Dekkers KF, Stein AD, Xu KM, Slagboom PE, van Zwet EW, Lumey LH, Heijmans BT. DNA methylation as a mediator of the association between prenatal adversity and risk factors for metabolic disease in adulthood. *Sci Adv.* 2018;4(1):eaao4364.
- Xie T, Gorenjak V, Stathopoulou MG, Dadé S, Marouli E, Masson C, Murray H, Lamont J, Fitzgerald P, Deloukas P, et al. Epigenome-wide association study (EWAS) of blood lipids in healthy population from STANISLAS family study (SFS). *Int J Mol Sci.* 2019;20(5):1014.
- Sayols-Baixeras S, Subirana I, Lluís-Ganella C, Civeira F, Roquer J, Do AN, Absher D, Cénarro A, Muñoz D, Soriano-Tárraga C, et al. Identification and validation of seven new loci showing differential DNA methylation

- related to serum lipid profile: an epigenome-wide approach the REGICOR study. *Hum Mol Genet.* 2016;25(20):4556–65.
16. Emdin CA, Khera AV, Kathiresan S. Mendelian randomization. *JAMA.* 2017;318(19):1925–6.
 17. Dekkers KF, van Iterson M, Sliker RC, Moed MH, Bonder MJ, van Galen M, Mei H, Zhernakova DV, van den Berg LH, Deelen J, et al. Blood lipids influence DNA methylation in circulating cells. *Genome Biol.* 2016;17(1):138.
 18. Sayols-Baixeras S, Tiwari HK, Aslibekyan SW. Disentangling associations between DNA methylation and blood lipids: a Mendelian randomization approach. *BMC Proc.* 2018;12(Suppl 9):23.
 19. Jiang L, Zhao K, Klein K, Canty AJ, Oualkacha K, Greenwood CMT. Investigating potential causal relationships between SNPs, DNA methylation and HDL. *BMC Proc.* 2018;12(Suppl 9):20.
 20. Wahl S, Drong A, Lehne B, Loh M, Scott WR, Kunze S, Tsai PC, Ried JS, Zhang W, Yang Y, et al. Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. *Nature.* 2017;541(7635):81–6.
 21. Sun D, Zhang T, Su S, Hao G, Chen T, Li QZ, Bazzano L, He J, Wang X, Li S, et al. Body mass index drives changes in DNA methylation: a longitudinal study. *Circ Res.* 2019;125(9):824–33.
 22. Chen Y, Kassam I, Lau SH, Kooner JS, Wilson R, Peters A, Winkelmann J, Chambers JC, Chow VT, Khor CC, et al. Impact of BMI and waist circumference on epigenome-wide DNA methylation and identification of epigenetic biomarkers in blood: an EWAS in multi-ethnic Asian individuals. *Clin Epigenet.* 2021;13(1):195.
 23. Reed ZE, Suderman MJ, Relton CL, Davis OSP, Hemani G. The association of DNA methylation with body mass index: distinguishing between predictors and biomarkers. *Clin Epigenet.* 2020;12(1):50.
 24. Jhun MA, Mendelson M, Wilson R, Gondalia R, Joehanes R, Salfati E, Zhao X, Braun KVE, Do AN, Hedman ÅK, et al. A multi-ethnic epigenome-wide association study of leukocyte DNA methylation and blood lipids. *Nat Commun.* 2021;12(1):3987.
 25. Cronjé HT, Elliott HR, Nienaber-Rousseau C, Pieters M. Replication and expansion of epigenome-wide association literature in a black South African population. *Clin Epigenet.* 2020;12(1):6.
 26. Jiang X, Fulte S, Deng F, Chen S, Xie Y, Chao X, He XC, Zhang Y, Li T, Li F, et al. Lack of VMP1 impairs hepatic lipoprotein secretion and promotes nonalcoholic steatohepatitis. *J Hepatol.* 2022;77(3):619–31.
 27. Hama Y, Morishita H, Mizushima N. Regulation of ER-derived membrane dynamics by the DedA domain-containing proteins VMP1 and TMEM41B. *EMBO Rep.* 2022;23(2):e53894.
 28. Chu AY, Guilianini F, Grallert H, Dupuis J, Ballantyne CM, Barratt BJ, Nyberg F, Chasman DI, Ridker PM. Genome-wide association study evaluating lipoprotein-associated phospholipase A2 mass and activity at baseline and after rosuvastatin therapy. *Circ Cardiovasc Genet.* 2012;5(6):676–85.
 29. Thompson A, Gao P, Orfei L, Watson S, Di Angelantonio E, Kaptoge S, Ballantyne C, Cannon CP, Criqui M, Cushman M, et al. Lipoprotein-associated phospholipase A(2) and risk of coronary disease, stroke, and mortality: collaborative analysis of 32 prospective studies. *Lancet.* 2010;375(9725):1536–44.
 30. Wilensky RL, Shi Y, Mohler ER 3rd, Hamamdzc D, Burgert ME, Li J, Postle A, Fenning RS, Bollinger JG, Hoffman BE, et al. Inhibition of lipoprotein-associated phospholipase A2 reduces complex coronary atherosclerotic plaque development. *Nat Med.* 2008;14(10):1059–66.
 31. Jackisch L, Kumsaiyai W, Moore JD, Al-Daghri N, Kyrou I, Barber TM, Randeve H, Kumar S, Tripathi G, McTernan PG. Differential expression of Lp-PLA2 in obesity and type 2 diabetes and the influence of lipids. *Diabetologia.* 2018;61(5):1155–66.
 32. Dhana K, Braun KVE, Nano J, Voortman T, Demerath EW, Guan W, Fornage M, van Meurs JBB, Uitterlinden AG, Hofman A, et al. An epigenome-wide association study of obesity-related traits. *Am J Epidemiol.* 2018;187(8):1662–9.
 33. Yu C, Wong EM, Joo JE, Hodge AM, Makalic E, Schmidt D, Buchanan DD, Severi G, Hopper JL, English DR, et al. Epigenetic drift association with cancer risk and survival, and modification by sex. *Cancers (Basel).* 2021;13(8):1881.
 34. Smith AK, Conneely KN, Pace TW, Mister D, Felger JC, Kilaru V, Akel MJ, Vertino PM, Miller AH, Torres MA. Epigenetic changes associated with inflammation in breast cancer patients treated with chemotherapy. *Brain Behav Immun.* 2014;38:227–36.
 35. Sun Q, Yong X, Sun X, Yang F, Dai Z, Gong Y, Zhou L, Zhang X, Niu D, Dai L, et al. Structural and functional insights into sorting nexin 5/6 interaction with bacterial effector IncE. *Signal Transduct Target Ther.* 2017;2:17030.
 36. Willer CJ, Schmidt EM, Sengupta S, Peloso GM, Gustafsson S, Kanoni S, Ganna A, Chen J, Buchkovich ML, Mora S, et al. Discovery and refinement of loci associated with lipid levels. *Nat Genet.* 2013;45(11):1274–83.
 37. Zhou Q, Huang T, Jiang Z, Ge C, Chen X, Zhang L, Zhao F, Zhu M, Chen T, Cui Y, et al. Upregulation of SNX5 predicts poor prognosis and promotes hepatocellular carcinoma progression by modulating the EGFR-ERK1/2 signaling pathway. *Oncogene.* 2020;39(10):2140–55.
 38. Jitsukawa S, Kamekura R, Kawata K, Ito F, Sato A, Matsumiya H, Nagaya T, Yamashita K, Kubo T, Kikuchi T, et al. Loss of sorting nexin 5 stabilizes internalized growth factor receptors to promote thyroid cancer progression. *J Pathol.* 2017;243(3):342–53.
 39. Li F, Yang J, Villar VAM, Asico LD, Ma X, Armando I, Sanada H, Yoneda M, Felder RA, Jose PA, et al. Loss of renal SNX5 results in impaired IDE activity and insulin resistance in mice. *Diabetologia.* 2018;61(3):727–37.
 40. Liang J, Li G, Liao J, Huang Z, Wen J, Wang Y, Chen Z, Cai G, Xu W, Ding Z, et al. Non-coding small nucleolar RNA SNORD17 promotes the progression of hepatocellular carcinoma through a positive feedback loop upon p53 inactivation. *Cell Death Differ.* 2022;29(5):988–1003.
 41. Cho O, Kim DW, Cheong JY. Screening plasma exosomal RNAs as diagnostic markers for cervical cancer: an analysis of patients who underwent primary chemoradiotherapy. *Biomolecules.* 2021;11(11):1691.
 42. Huang L, Liang XZ, Deng Y, Liang YB, Zhu X, Liang XY, Luo DZ, Chen G, Fang YY, Lan HH, et al. Prognostic value of small nucleolar RNAs (snoRNAs) for colon adenocarcinoma based on RNA sequencing data. *Pathol Res Pract.* 2020;216(6):152937.
 43. Do WL, Gohar J, McCullough LE, Galaviz KI, Conneely KN, Narayan KMV. Examining the association between adiposity and DNA methylation: a systematic review and meta-analysis. *Obes Rev.* 2021;22(10):e13319.
 44. Pavethynath S, Imai C, Jin X, Hichiwa N, Takimoto H, Okamitsu M, Tarui I, Aoyama T, Yago S, Fudono A, et al. Metabolic and immunological shifts during mid-to-late gestation influence maternal blood methylation of CPT1A and SREBF1. *Int J Mol Sci.* 2019;20(5):1066.
 45. Min JL, Hemani G, Hannon E, Dekkers KF, Castillo-Fernandez J, Luijk R, Carnero-Montoro E, Lawson DJ, Burrows K, Suderman M, et al. Genomic and phenotypic insights from an atlas of genetic effects on DNA methylation. *Nat Genet.* 2021;53(9):1311–21.
 46. Pfeiffer L, Wahl S, Pilling LC, Reischl E, Sandling JK, Kunze S, Holdt LM, Kretschmer A, Schramm K, Adamski J, et al. DNA methylation of lipid-related genes affects blood lipid levels. *Circ Cardiovasc Genet.* 2015;8(2):334–42.
 47. Frambach S, de Haas R, Smeitink JAM, Rongen GA, Russel FGM, Schirris TJJ. Brothers in arms: ABCA1- and ABCG1-mediated cholesterol efflux as promising targets in cardiovascular disease treatment. *Pharmacol Rev.* 2020;72(1):152–90.
 48. Xu D, Li Y, Yang F, Sun CR, Pan J, Wang L, Chen ZP, Fang SC, Yao X, Hou WT, et al. Structure and transport mechanism of the human cholesterol transporter ABCG1. *Cell Rep.* 2022;38(4):110298.
 49. Ouimet M, Barrett TJ, Fisher EA. HDL and reverse cholesterol transport. *Circ Res.* 2019;124(10):1505–18.
 50. Hoekstra M, Van Berkel TJ. Functionality of high-density lipoprotein as antiatherosclerotic therapeutic target. *Arterioscler Thromb Vasc Biol.* 2016;36(11):e87–94.
 51. Qi B, He L, Zhao Y, Zhang L, He Y, Li J, Li C, Zhang B, Huang Q, Xing J, et al. Akap1 deficiency exacerbates diabetic cardiomyopathy in mice by NDUFS1-mediated mitochondrial dysfunction and apoptosis. *Diabetologia.* 2020;63(5):1072–87.
 52. Schiattarella GG, Cattaneo F, Carrizzo A, Paolillo R, Boccia N, Ambrosio M, Damato A, Prionti G, Franzone A, Russo G, et al. Akap1 regulates vascular function and endothelial cells behavior. *Hypertension.* 2018;71(3):507–17.
 53. Perrino C, Feliciello A, Schiattarella GG, Esposito G, Guerriero R, Zaccaro L, Del Gatto A, Saviano M, Garbi C, Carangi R, et al. AKAP121 downregulation impairs protective cAMP signals, promotes mitochondrial dysfunction, and increases oxidative stress. *Cardiovasc Res.* 2010;88(1):101–10.
 54. Ji L, Zhao Y, He L, Zhao J, Gao T, Liu F, Qi B, Kang F, Wang G, Zhao Y, et al. AKAP1 deficiency attenuates diet-induced obesity and insulin resistance by promoting fatty acid oxidation and thermogenesis in brown adipocytes. *Adv Sci (Weinh).* 2021;8(6):2002794.

55. Braun KV, Voortman T, Dhana K, Troup J, Bramer WM, Troup J, Chowdhury R, Dehghan A, Muka T, Franco OH. The role of DNA methylation in dyslipidaemia: a systematic review. *Prog Lipid Res.* 2016;64:178–91.
56. Price NL, Singh AK, Rotllan N, Goedeke L, Wing A, Canfrán-Duque A, Diaz-Ruiz A, Araldí E, Baldán Á, Camporez JP, et al. Genetic ablation of miR-33 increases food intake, enhances adipose tissue expansion, and promotes obesity and insulin resistance. *Cell Rep.* 2018;22(8):2133–45.
57. Nishino T, Horie T, Baba O, Sowa N, Hanada R, Kuwabara Y, Nakao T, Nishiga M, Nishi H, Nakashima Y, et al. SREBF1/MicroRNA-33b axis exhibits potent effect on unstable atherosclerotic plaque formation in vivo. *Arterioscler Thromb Vasc Biol.* 2018;38(10):2460–73.
58. Rayner KJ, Esau CC, Hussain FN, McDaniel AL, Marshall SM, van Gils JM, Ray TD, Sheedy FJ, Goedeke L, Liu X, et al. Inhibition of miR-33a/b in non-human primates raises plasma HDL and lowers VLDL triglycerides. *Nature.* 2011;478(7369):404–7.
59. Ding J, Reynolds LM, Zeller T, Müller C, Lohman K, Nicklas BJ, Kritchevsky SB, Huang Z, de la Fuente A, Soranzo N, et al. Alterations of a cellular cholesterol metabolism network are a molecular feature of obesity-related type 2 diabetes and cardiovascular disease. *Diabetes.* 2015;64(10):3464–74.
60. Johansson LE, Danielsson AP, Parikh H, Klintonberg M, Norström F, Groop L, Ridderstråle M. Differential gene expression in adipose tissue from obese human subjects during weight loss and weight maintenance. *Am J Clin Nutr.* 2012;96(1):196–207.
61. Choromanska B, Mysliwiec P, Hady HR, Dadan J, Mysliwiec H, Bonda T, Chabowski A, Miklosz A. The implication of adipocyte ATP-binding cassette A1 and G1 transporters in metabolic complications of obesity. *J Physiol Pharmacol.* 2019;70(1):143–52.
62. McGue M, Osler M, Christensen K. Causal inference and observational research: the utility of twins. *Perspect Psychol Sci.* 2010;5(5):546–56.
63. Gao W, Cao W, Lv J, Yu C, Wu T, Wang S, Meng L, Wang D, Wang Z, Pang Z, et al. The Chinese national twin registry: a “gold mine” for scientific research. *J Intern Med.* 2019;286(3):299–308.
64. Chen Z, Lee L, Chen J, Collins R, Wu F, Guo Y, Linksted P, Peto R. Cohort profile: the Kadoorie study of chronic disease in China (KSCDC). *Int J Epidemiol.* 2005;34(6):1243–9.
65. Chen Z, Chen J, Collins R, Guo Y, Peto R, Wu F, Li L. China Kadoorie Biobank of 0.5 million people: survey methods, baseline characteristics and long-term follow-up. *Int J Epidemiol.* 2011;40(6):1652–66.
66. Si J, Yang S, Sun D, Yu C, Guo Y, Lin Y, Millwood IY, Walters RG, Yang L, Chen Y, et al. Epigenome-wide analysis of DNA methylation and coronary heart disease: a nested case-control study. *Elife.* 2021;10:e68671.
67. Wang B, Gao W, Yu C, Cao W, Lv J, Wang S, Pang Z, Cong L, Wang H, Wu X, et al. Determination of zygosity in adult Chinese twins using the 450 K methylation array versus questionnaire data. *PLoS ONE.* 2015;10(4):e0123992.
68. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, Irizarry RA. Minfi: a flexible and comprehensive bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics.* 2014;30(10):1363–9.
69. Touleimat N, Tost J. Complete pipeline for Infinium® human methylation 450 K BeadChip data processing using subset quantile normalization for accurate DNA methylation estimation. *Epigenomics.* 2012;4(3):325–41.
70. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, Wiencke JK, Kelsey KT. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinf.* 2012;13:86.
71. Chen J, Behnam E, Huang J, Moffatt MF, Schaid DJ, Liang L, Lin X. Fast and robust adjustment of cell mixtures in epigenome-wide association studies with SmartSVA. *BMC Genomics.* 2017;18(1):413.
72. Ren X, Kuan PF. methylGSA: a bioconductor package and shiny app for DNA methylation data length bias adjustment in gene set testing. *Bioinformatics.* 2019;35(11):1958–9.
73. Li H, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Model.* 1999;6(1):1–55.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

