BMC Genomics

Role of DNA methylation on the association between physical activity and cardiovascular diseases: results from the longitudinal multi-ethnic study of atherosclerosis (MESA) cohort

Hangchuan Shi[1,2], Deborah J. Ossip[2], Nicole L. Mayo[2], Daniel A. Lopez[2], Robert C. Block[2], Wendy S. Post[3], Alain G. Bertoni[4], Jingzhong Ding[5], Si Chen[6,7], Chen Yan[6,7], Zidian Xie[1], Ina Hoeschele[8], Yongmei Liu[9*] and Dongmei Li[1*]

Abstract

Background: The complexity of physical activity (PA) and DNA methylation interaction in the development of cardiovascular disease (CVD) is rarely simultaneously investigated in one study. We examined the role of DNA methylation on the association between PA and CVD.

Results: The Multi-Ethnic Study of Atherosclerosis (MESA) cohort Exam 5 data with 1065 participants free of CVD were used for final analysis. The quartile categorical total PA variable was created by activity intensity (METs/week). During a median follow-up of 4.0 years, 69 participants developed CVD. Illumina HumanMethylation450 BeadChip was used to provide genome-wide DNA methylation profiles in purified human monocytes (CD14+). We identified 23 candidate DNA methylation loci to be associated with both PA and CVD. We used the structural equation modeling (SEM) approach to test the complex relationships among multiple variables and the roles of mediators. Three of the 23 identified loci (corresponding to genes *VPS13D*, *PIK3CD* and *VPS45*) remained as significant mediators in the final SEM model along with other covariates. Bridged by the three genes, the 2nd PA quartile (β = − 0.959; 95%CI: − 1.554 to − 0.449) and the 3rd PA quartile (β = − 0.944; 95%CI: − 1.628 to − 0.413) showed the greatest inverse associations with CVD development, while the 4th PA quartile had a relatively weaker inverse association (β = − 0.355; 95%CI: − 0.713 to − 0.124).

* Correspondence: yongmei.liu@duke.edu; Dongmei_Li@urmc.rochester.edu
[9]Department of Medicine, Division of Cardiology, Duke Molecular Physiology Institute, Duke University Medical Center, Durham, NC 27701, USA
[1]Department of Clinical and Translational Research, University of Rochester Medical Center, Rochester, NY 14642-0708, USA
Full list of author information is available at the end of the article

Conclusions: The current study is among the first to simultaneously examine the relationships among PA, DNA methylation, and CVD in a large cohort with long-term exposure. We identified three DNA methylation loci bridged the association between PA and CVD. The function of the identified genes warrants further investigation in the pathogenesis of CVD.

Keywords: Cardiovascular disease, Physical activity, DNA methylation, MESA, Structural equation modeling

Background

Cardiovascular disease (CVD) is the leading cause of mortality in the United States [1]. Understanding contributors to the etiology of CVD will help us better control the disease at the level of primary prevention [2]. The etiology of CVD involves complex multifactorial mechanisms including environmental and genetic factors [3]. Among the environmental or behavioral exposures, physical activity (PA) has consistently been shown in epidemiological studies to be associated with reduced risk of CVD [4]. Although multiple mediating mechanisms have been proposed (e.g., an anti-inflammatory response, insulin-sensitivity) [4, 5], our knowledge about the underlying mechanisms of the benefits of PA on cardiovascular health is still sparse. Owing to the evolutionary development of genomic study methods, recent studies suggest that epigenetic modifications are also involved in the association between PA and CVD [6].

DNA methylation is the best-studied epigenetic modification. It is a dynamic epigenetic mechanism in several cellular process such as regulating gene expression, genomic imprinting, and X-chromosome inactivation [7]. In eukaryotes, the methylated cytosine (5-methylcytosine, 5mC) occurs predominantly at C followed by guanine (G) forming the so-called CpG site. Although the result of DNA methylation is usually linked to the silencing of genes [8], growing evidences showed that DNA methylation can also increase gene expression in some instances [9]. Epigenome-wide association studies (EWAS) have shown that many DNA methylation loci are associated with CVD traits [10, 11] and influenced by physical activity [12]. However, because the associations between PA or CVD and DNA methylation were studied separately in previous studies, the role of DNA methylation on PA's benefits on cardiovascular health was only deduced from the biological functions of identified genes [13]. There is a lack of evidence directly showing the complexity of PA-DNA methylation interaction in the development of CVD [14].

To investigate the complex relationships among PA, DNA methylation, and CVD, we applied the structural equation modeling (SEM) approach to a cohort. Compared with common statistical methodologies, SEM allows for testing of complex relationships among multiple variables including the roles of mediators [15].

In epigenetic epidemiology, it is useful to investigate whether DNA methylation is a component of the pathway linking the exposure and the outcome [14]. In the current study, we evaluated the role of DNA methylation on the longitudinal association between PA and CVD, along with other covariates, using data from the Multi-Ethnic Study of Atherosclerosis (MESA).

Methods

Data source and study cohort

The MESA study is a prospective cohort study designed to investigate the prevalence, associations, and progression of subclinical CVD in 6814 men and women aged 45–84 years and free of clinical CVD at the first examination (Exam 1, 2000–2002). Following Exam 1, five follow-up examinations (Exam 2–6) have occurred. Written informed consent was obtained from all participating subjects under research protocols approved by the review boards of the six filed centers. A detailed description of the MESA study be found elsewhere [16]. Corresponding to Exam 5 (2010–2012), the MESA Epigenomics and Transcriptomics Study (2010–2012) was conducted to investigate the association between DNA methylation and gene expression in purified human monocytes (CD14$^+$) [17]. Asian participants were not included since the study was not available for participants from two of six MESA sites [18]. The DNA methylation data are publicly accessible through the GEO website with accession ID GSE56046.

In this study, we linked the MESA DNA methylation data to Exam 5 data through unique MESA participant IDs. As a result, we used exposure (physical activity), covariate, and DNA methylation data of 1264 participants collected in 2010–2012 (Exam 5). Data were analyzed and compared for CVD outcomes occurring through December 31, 2015. We excluded participants with prevalent CVD at baseline ($n = 155$) and those with missing exposure or covariate data ($n = 44$). The final sample size of the analyzed cohort was 1065.

Variables and analysis

Physical activity

Physical activity and sedentary behaviors were assessed through an interviewer-administered MESA Typical Week Physical Activity Survey (TWPAS) which was

designed to estimate participants' habitual activities as of their last interview. The survey contains 28 question item categories asking participants about the duration of time when doing each category of activities. The unit of the continuous activity variables was measured as both minutes per week and metabolic equivalence of tasks (METs) per week. According to the MET value of each activity, the intensity was classified according to MET value as sedentary behavior (METs ≤1.5), light PA (1.5 < METs < 3.0), moderate PA (3 ≤ METs < 6), and vigorous PA (METs ≥6) [19, 20]. In the screening stage of this study, the continuous total PA variable was defined as the sum of total METs per week spent doing light, moderate, and vigorous PAs. The quartile categorical total PA variable (METs/week) was created from the continuous total PA variable used both in screening stage and in the SEM model to examine potential non-linear pattern. In addition to total PA, each PA components (i.e., light, moderate, and vigorous PA), two combinations of PA intensity (i.e., moderate to vigorous PA [MVPA] and exercise PA), and leisure sedentary behavior were also used for screening PA related DNA methylation loci to increase the screening power. The detailed descriptions of all PA variables are listed in Supplementary Table S1.

CVD outcomes
Each MESA participant was contacted every 9–12 months to collect information on new CVD conditions, which were classified using medical records and death certificates [16]. MESA assessed 11 individual cardiovascular events. The detailed descriptions of all CVD outcomes are listed in Supplementary Table S2. In this study, CVD (all types), which was the combination of all 11 individual cardiovascular events, was used as the CVD outcome in the SEM model. The endpoint outcome was the binary response (Yes/No) to CVD (all types) occurring through December 31, 2015. However, in the screening stage of DNA methylation loci identification, all individual CVD outcomes were used separately to increase the sensitivity of screening.

Covariates
Atherosclerotic CVD (ASCVD) 10-year risk score is developed by the American College of Cardiology (ACC) and American Heart Association (AHA) in collaboration with the National Heart Lung and Blood Institute (NHLBI). The score is in a percentage form to estimate the risk of a hard ASCVD event within the next ten years [21, 22]. The risk estimator is calculated based on an individual's sex, race, age, total cholesterol (mg/dL), high-density lipoprotein cholesterol (HDL, mg/dL), systolic blood pressure, use of blood pressure medication, smoking status, and diabetes status. The calculation

additionally takes into account the potential interactions between these variables.

Some potential confounders associated with CVD in the literature [23–25], but absent in the ASCVD score, included ethnicity (Hispanic or not), body mass index (BMI), waist circumference (WC, cm), and coronary artery calcium (CAC) score. BMI and WC were categorized into three categories based on well-established definitions [26]. CAC score is the quantification of the coronary artery calcification by computed tomography (CT) [27]. In this study, we used the Agatston CAC score and categorized it into four levels (0, 1–100, 101–400, > 400) according to one of the most widely used classification systems [28].

Identification of DNA methylation locus
Illumina HumanMethylation450 (450 K) BeadChip was used to provide genome-wide DNA methylation profiles with over 485,000 DNA methylation loci. Quality control was described in detail in a previous study [18]. Prior to identification, CpG sites were filtered according to beta values (cut-off = 0.2) to exclude background changes [29]. A two-stage strategy was then used. In the first stage (screening stage), EWAS were performed using *limma* and *preprocessCore* packages in R [30, 31]. Separate univariate logistic regression models were fitted using the normalized M-value for each CpG site as the dependent variable. Each of the PA related variables (continuous or categorical) or each of the CVD outcome variables (binary) was used as the independent variable. *P*-values were adjusted for multiple testing using the Benjamini-Hochberg method for controlling the false discovery rate (FDR), which was discussed in our previous publications [32, 33]. CpG sites with adjusted *p*-value < 0.05 were selected as potential candidates. All candidate CpG sites screened from regression models with PA related variables were pooled as Set A, and all candidate CpG sites from models with CVD outcome variables were pooled as Set B. In the second stage (identification stage), CpG sites overlapped by Set A and Set B were finally identified as the final candidate DNA methylation loci for the further SEM model. The corresponding genes were determined using Illumina annotation. Volcano plots and heatmaps were created using *ggplot2* and *pheatmap* packages in R [34, 35] to show the differentially methylated level on the identified DNA loci.

SEM model
SEM analyses were performed to examine the proposed conceptual model, namely, the role of DNA methylation on the association between PA and CVD. Model fits were conducted using Mplus version 8.5 (Muthen and Muthen, 2017). All the parameters in the SEM model

Shi *et al. BMC Genomics*   (2021) 22:790

Page 4 of 13

were estimated by a maximum likelihood method [36]. The estimated coefficient (a.k.a. coefficient load, β) generated from the SEM model was used to measure the strength of association between two variables. Prior to model fitting, the correlation matrix of all identified DNA methylation loci was examined based on M-values to avoid multicollinearity. We did SEM model selection by including the categorical total PA variable (MET-mins/week), CVD (all), and all identified DNA methylation loci, along with all covariates. Association was considered non-significant if the coefficient load β had a *p*-value > 0.1. Variables lacking significant associations with other variables were removed from the SEM model. The final SEM model was fitted using variables with at least one significant direct or indirect path from PA to CVD. Sensitivity analyses were conducted by fitting SEM models and substituting the continuous PA variable for the categorical PA variable or excluding mediators from the final SEM model.

Results

Demographic and clinical characteristics of the studied samples

During a median of 4.0 years' follow-up, 69 of 1065 participants had a cardiovascular disease event. The cumulative incidence is 6.5%. The baseline characteristics of the 1065 participants revealed an average age of 69.2 years (range: 54–92 years) with a slightly higher proportion of females (52.3%) than males (Table 1). The largest group of the participants were non-Hispanic White (46.9%), followed by Hispanics (31.3%) and non-Hispanic Blacks (21.8%). Compared with participants who did not develop any CVD event (non-incident cases), incident cases were more likely to be older people (72.7 years vs 69.0 years), to have hypertension (71.0% vs 57.0%), and undergoing hypertension treatment (69.6% vs 52.4%). Notably, compared with non-incident cases, incident cases were more likely to be identified in the lowest PA quartile (39.1% vs 24.4%), but less likely to be in the second (17.4% vs 25.5%) or third PA quartile (15.9% vs 25.5%). At baseline, the mean ASCVD (atherosclerotic cardiovascular disease) score of the future incident cases was higher (28.5%) than that of non-incident cases (19.3%). Also, incident cases were more likely to have a CAC (coronary artery calcium) score greater than 100 (55.0%) than non-incident cases (33.6%).

As shown in Table 2, several baseline summary measures were significantly different across the PA quartiles (PA Q1–4). From the PA Q1 to Q4, the mean ages of participants decreased from 72.6 years to 66.0 years, which indicated that older people were more likely to be physically inactive. A higher proportion of males (55.1%) than females (44.9%) were observed in PA Q1. Females were dominant (> 54.0%) in all other three higher

quartiles (PA Q2–3), indicating that females were more likely to be physically active than males. Regarding the clinical characteristics, hypertension prevalence at baseline was higher in lower PA quartiles (i.e., PA Q1 and Q2; > 62.0%) than that in upper PA quartiles (i.e., PA Q3 and Q4; < 55%). Notably, the mean ASCVD score was highest in PA Q1 (25.9%) and lowest in the PA Q4 (15.8%). Similarly, the proportion of the participants with a high CAC score (i.e., greater than 400) decreased from 20.6% (PA Q1) to 14.4% (PA Q4).

Table 3 displays the cumulative incidence of each CVD outcome across the PA quartiles. Besides overall CVD (6.5%), the most frequently identified outcome during the follow-up was hard CVD (3.9%), followed by stroke (2.3%). Across the PA quartiles, it is noticeable that more CVD (all) incident cases were observed in PA Q1 and PA Q4 with the cumulative incidences of about 10.0%. The middle two quartiles (i.e., PA Q2 and Q3) had less CVD (all) incident cases, with no more than 4.5% cumulative incidence. A similar "U" shape pattern was also observed for the CVD (hard) outcome. In contrast, stroke incidence did not have this pattern. The highest stroke cumulative incidence (5.2%) was observed in PA Q1. All the other three quartiles (PA Q2–4) had lower incidences (≤ 1.5%). Although the interesting "U" shape pattern was observed in this study, the similar pattern was not appreciable in cross-sectional studies using MESA dataset [37].

DNA methylation loci identification

Representative volcano plots (Fig. 1a and b), Manhattan plots (Fig. S1), and Q-Q plots (Fig. S2) show the differentially methylated loci identified at the screening stage. The total number of candidate loci identified from each variable is presented in Supplementary Tables S3 and S4. As shown in the Venn diagram (Fig. 1c), a total number of 1048 loci were selected from the EWAS on PA variables (Set A), and 124 loci were from the EWAS on all the CVD variables (Set B). The overlapped 23 loci were then identified from the two sets. The detailed information of the overlapped 23 loci and corresponding gene names are listed in Supplementary Table S5. The heatmap shows the differential degree of methylation of the 23 loci grouped by four PA quartiles (PA Q1–4) and CVD (all) outcome (Fig. 1d).

SEM model

The final SEM model was fitted among PA, 3 identified loci and CVD, along with factors such as waist circumference (WC), ASCVD score, CAC score, and ethnicity. 20 loci were excluded during the model selection due to non-significant correlations with CVD outcome. Similarly, sedentary time and BMI were also excluded due to a lack of correlation with other variables. The remaining

Table 1 Baseline demographic, behavioral, and risk factor characteristics by CVD outcome during follow-up (mean, SD or N, %)

| | | CVD during follow-up | | |
| | Total sample | Yes | No | |
Characteristics	(N = 1,065)	(n = 69)	(n = 996)	p-value*
Age (years)	69.2 (9.3)	72.7 (9.5)	69.0 (9.3)	**0.002**
Gender				0.164
Female	557 (52.3)	30 (43.5)	527 (52.9)	
Male	508 (47.7)	39 (56.5)	469 (47.1)	
Race/ethnicity				0.136
Non-Hispanic White	500 (46.9)	27 (39.1)	473 (47.5)	
Non-Hispanic Black	232 (21.8)	13 (18.8)	219 (22.0)	
Hispanics	333 (31.3)	29 (42.0)	304 (30.5)	
Smoking status				0.250
Never	440 (41.3)	22 (10.1)	418 (42.0)	
Former	530 (49.8)	40 (58.0)	490 (49.2)	
Current	92 (8.6)	7 (31.9)	85 (8.5)	
Total cholesterol (mg/dl)	183.4 (36.2)	176.6 (34.9)	183.9 (36.3)	0.100
HDL cholesterol (mg/dl)	54.7 (16.2)	55.0 (17.0)	54.7 (16.1)	0.873
Hypertension	617 (57.9)	49 (71.0)	568 (57.0)	**0.033**
Hypertension treatment	570 (53.5)	48 (69.6)	522 (52.4)	**0.008**
Diabetes mellitus	192 (18.0)	16 (23.2)	176 (17.7)	0.322
Body mass index (BMI, kg/m^2)	29.5 (5.5)	28.8 (5.0)	29.5 (5.5)	0.249
BMI categories (kg/m^2)				0.605
Normal (≤ 24.9)	215 (20.2)	17 (24.6)	198 (19.9)	
Overweight (25 - 29.9)	411 (38.6)	24 (34.8)	387 (38.9)	
Obese (≥ 30.0)	438 (41.1)	28 (40.6)	410 (41.2)	
Waist circumference (WC, cm)	101.1 (14.4)	99.6 (13.0)	101.2 (14.5)	0.341
WC categories (C1 - C3, cm)				0.590
C1: < 90 (female); < 100 (male)	386 (36.2)	29 (42.0)	357 (35.8)	
C2: < 105 (female); < 110 (male)	375 (35.2)	22 (31.9)	353 (35.4)	
C3: ≥ 105 (female); ≥ 110 (male)	303 (28.5)	18 (26.1)	285 (28.6)	
ASCVD score (%)	19.9 (17.9)	28.5 (20.3)	19.3 (17.6)	**< 0.001**
Physical activity (PA, MET-mins/week)	7705 (5,478.0)	7,229 (5,918.6)	7,738 (5,447.8)	0.204
PA quartiles (MET-mins/week)				**0.017**
Q1: ≤ 3,555	267 (25.1)	27 (39.1)	243 (24.4)	
Q2: ≤ 6,555	267 (25.1)	12 (17.4)	254 (25.5)	
Q3: ≤ 10,650	266 (25.0)	11 (15.9)	254 (25.5)	
Q4: > 10,650	265 (24.9)	19 (27.5)	245 (24.6)	
Sedentary time (MET-mins/week)	1,624 (1,194.5)	1,576 (1,085.1)	1,627 (1,202.2)	0.860
Sedentary time quartiles (MET-mins/week)				0.695
Q1: ≤ 840	316 (29.7)	22 (31.9)	294 (29.5)	
Q2: ≤ 1,260	222 (20.8)	15 (21.7)	207 (20.8)	
Q3: ≤ 2,160	264 (24.8)	13 (18.8)	251 (25.2)	
Q4: > 2,160	263 (24.7)	19 (27.5)	244 (24.5)	
CAC score categories (C1 - C4)				**0.001**
C1: CAC = 0	339 (31.8)	12 (17.4)	327 (32.8)	
C2: 0 < CAC ≤ 100	307 (28.8)	15 (21.7)	292 (29.3)	
C3: 100 < CAC ≤ 400	204 (19.2)	19 (27.5)	185 (18.6)	
C4: CAC > 400	168 (15.8)	19 (27.5)	149 (15.0)	

* Chi-square test, Student's t-test, or Mann-Whitney U test;

ASCVD, atherosclerotic CVD; CAC score, coronary artery calcium score; MET, metabolic equivalence of tasks.

Table 2 Baseline demographic, behavioral, and risk factor characteristics by physical activity intensity (mean, SD or N, %)

Characteristics	PA Quartile 1 (PA Q1) ≤ 3,555 (n = 267)	PA Quartile 2 (PA Q2) ≤ 6,555 (n = 267)	PA Quartile 3 (PA Q3) ≤ 10,650 (n = 266)	PA Quartile 4 (PA Q4) > 10,650 (n = 265)	p-value*
Age (years)	72.6 (9.3)	70.0 (9.5)	68.2 (8.6)	66.0 (8.7)	**< 0.001**
Gender					**0.048**
Female	120 (44.9)	145 (54.3)	144 (54.1)	148 (55.8)	
Male	147 (55.1)	122 (45.7)	122 (45.9)	117 (44.2)	
Race/ethnicity					0.106
Non-Hispanic White	114 (42.7)	145 (54.3)	128 (48.1)	113 (42.6)	
Non-Hispanic Black	61 (22.8)	53 (19.9)	54 (20.3)	64 (24.2)	
Hispanics	92 (34.5)	69 (25.8)	84 (31.6)	88 (33.2)	
Smoking status					0.348
Never	107 (40.1)	110 (41.2)	113 (42.5)	110 (41.5)	
Former	141 (52.8)	132 (49.4)	135 (50.8)	122 (46.0)	
Current	18 (6.7)	25 (9.4)	18 (6.8)	31 (11.7)	
Total cholesterol (mg/dl)	178.3 (38.8)	185.3 (35.9)	184.5 (35.4)	185.6 (34.3)	0.066
HDL cholesterol (mg/dl)	55.3 (16.8)	55.9 (16.3)	54.9 (17.5)	52.6 (13.6)	0.091
Hypertension	166 (62.2)	167 (62.5)	139 (52.3)	145 (54.7)	**0.029**
Hypertension treatment	155 (58.1)	151 (56.6)	128 (48.1)	136 (51.3)	0.078
Diabetes mellitus	54 (20.2)	52 (19.5)	39 (14.7)	47 (17.7)	0.347
Body mass index (BMI, kg/m^2)	29.7 (5.6)	29.4 (5.4)	29.0 (5.5)	29.7 (5.4)	0.769
BMI categories (kg/m^2)					0.706
Normal (≤ 24.9)	51 (19.1)	55 (20.7)	61 (22.9)	48 (18.1)	
Overweight (25 - 29.9)	102 (38.2)	105 (39.5)	105 (39.5)	99 (37.4)	
Obese (≥ 30.0)	114 (42.7)	106 (39.8)	100 (37.6)	118 (44.5)	
Waist circumference (WC, cm)	102.8 (14.4)	101.5 (15.2)	99.8 (14.1)	100.3 (13.6)	0.063
WC categories (C1 - C3, cm)					0.306
C1: < 90 (female); < 100 (male)	88 (33.1)	98 (36.7)	104 (39.1)	96 (36.2)	
C2: < 105 (female); < 110 (male)	90 (33.8)	88 (33.0)	94 (35.3)	103 (38.9)	
C3: ≥ 105 (female); ≥ 110 (male)	88 (33.1)	81 (30.3)	68 (25.6)	66 (24.9)	
ASCVD score (%)	25.9 (21.2)	20.8 (18.0)	16.9 (14.7)	15.8 (15.4)	**< 0.001**
Sedentary time (MET-mins/week)	1,562 (1,090.9)	1,731 (1,340.2)	1,640 (1,117.3)	1,562 (1,212.2)	0.389
Sedentary time quartiles (MET-mins/week)					0.187
Q1: ≤ 840	74 (27.7)	79 (29.6)	75 (28.2)	88 (33.2)	
Q2: ≤ 1,260	68 (25.5)	50 (18.7)	55 (20.7)	49 (18.5)	
Q3: ≤ 2,160	64 (24.0)	63 (23.6)	62 (23.3)	75 (28.3)	
Q4: > 2,160	61 (22.8)	75 (28.1)	74 (27.8)	53 (20.0)	
CAC score categories (C1 - C4)					**0.036**
C1: CAC = 0	70 (27.7)	69 (27.5)	100 (38.9)	100 (38.9)	
C2: 0 < CAC ≤ 100	78 (30.8)	79 (31.5)	78 (30.4)	72 (28.0)	
C3: 100 < CAC ≤ 400	53 (20.9)	58 (23.1)	45 (17.5)	48 (18.7)	
C4: CAC > 400	52 (20.6)	45 (17.9)	34 (13.2)	37 (14.4)	

* Chi-square test, ANOVA or Kruskal-Wallis test;
ASCVD, atherosclerotic CVD; CAC score, coronary artery calcium score; MET, metabolic equivalence of tasks.

Table 3 Total number of cases (N) and crude cumulative incidences (%) of CVD outcomes during follow-up by baseline physical activity (PA) intensity

	Overall	PA Q1[††]	PA Q2[††]	PA Q3[††]	PA Q4[††]	
	(N = 1,065)	(n = 267)	(n = 267)	(n = 266)	(n = 265)	p-value*
CVD, all[†]	69 (6.5)	27 (10.1)	12 (4.5)	11 (4.1)	27 (10.2)	**0.017**
CVD, hard[‡]	42 (3.9)	18 (6.7)	6 (2.2)	6 (2.3)	18 (6.8)	**0.020**
CHD, all[§]	30 (2.8)	9 (3.4)	6 (2.2)	3 (1.1)	9 (3.4)	0.092
CHD, hard[||]	21 (2.0)	7 (2.6)	3 (1.1)	2 (0.8)	7 (2.6)	0.091
CVD death[¶]	12 (1.1)	5 (1.9)	3 (1.1)	2 (0.8)	5 (1.9)	0.573
Revascularization[#]	22 (2.1)	8 (3.0)	3 (1.1)	4 (1.5)	8 (3.0)	0.382
PTCA	19 (1.8)	6 (2.2)	3 (1.1)	3 (1.1)	7 (2.6)	0.469
CABG	1 (0.1)	0 (0.0)	0 (0.0)	1 (0.4)	0 (0.0)	0.499
MI	17 (1.6)	6 (2.2)	2 (0.7)	2 (0.8)	6 (2.3)	0.157
Angina	14 (1.3)	5 (1.9)	3 (1.1)	2 (0.8)	5 (1.9)	0.710
RCA	2 (0.2)	0 (0.0)	2 (0.7)	0 (0.0)	0 (0.0)	0.249
PVD	4 (0.4)	3 (1.1)	0 (0.0)	1 (0.4)	3 (1.1)	0.170
CHF	15 (1.4)	6 (2.2)	3 (1.1)	1 (0.4)	6 (2.3)	0.218
Stroke	25 (2.3)	14 (5.2)	3 (1.1)	4 (1.5)	4 (1.5)	**0.011**
TIA	3 (0.3)	0 (0.0)	1 (0.4)	1 (0.4)	1 (0.4)	0.811

* Chi-square test or Fisher exact test;

¶ CVD death is death due to CHD, stroke or other CVD;

Revascularization (REV) is CABG + PTCA + other revascularizations;

† CVD (all) is PVD + CHF + TIA + MI + RCA + REV + Stroke + Angina + CVD death;

‡ CVD (hard) is MI + RCA + Stroke + CVD death (due to CHD or Stoke);

§ CHD (all) is MI + RCA + Angina + CVD death (due to CHD);

|| CHD (hard) is MI + RCA + CVD death (due to CHD);

†† PA Q1 to Q4 denotes PA Quartile 1 to Quartile 4 based on the unit of MET-min/week: Quartile 1 (≤ 3,555), Quartile 2 (≤ 6,555), Quartile 3 (≤ 10,650), Quartile 4 (> 10,650).

Abbreviations: CHD, coronary heart disease; PTCA, percutaneous transluminal coronary angioplasty; CABG, coronary artery bypass grafting; MI, myocardial infarction; RCA, resuscitated cardiac arrest; PVD, peripheral vascular disease; CHF, congestive heart failure; TIA, transient ischemic attack.
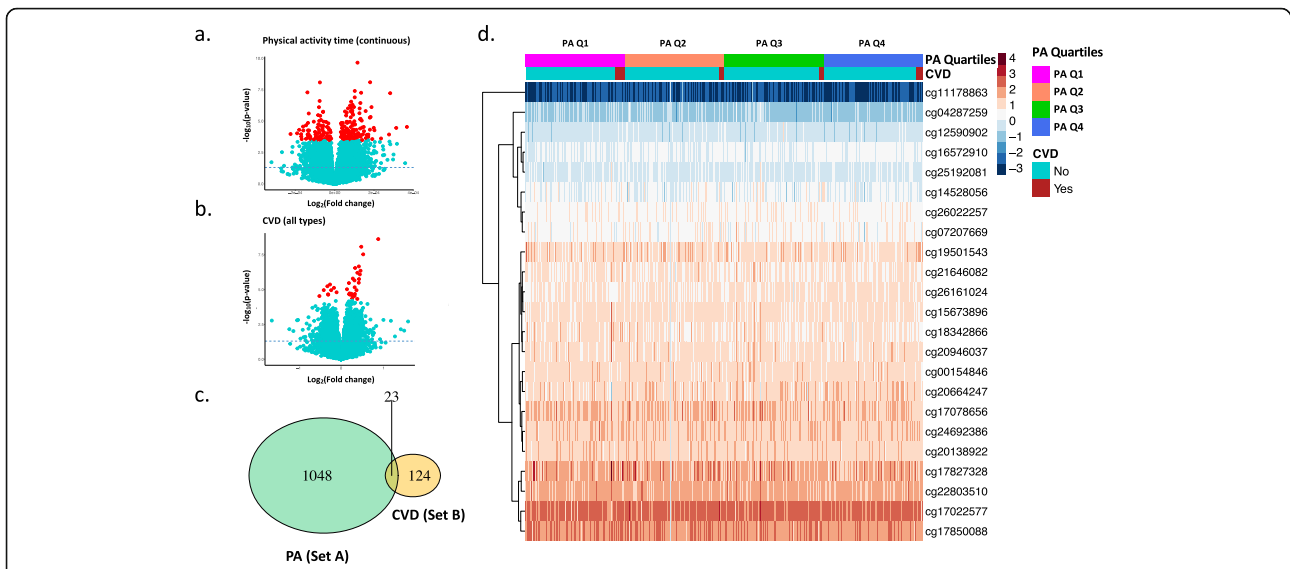
Fig. 1 DNA methylation loci associated with PA and CVD. Representative volcano plots show the significant DNA methylation loci associated with continuous PA variable (**a**) and binary CVD outcome (**b**). DNA methylation loci were denoted as dots in red (significant loci) or in cyan (non-significant loci) according to the FDR cut-off value (0.05). The horizontal dashed line represents where *p*-value is 0.05. The Venn diagram shows the overlapped 23 DNA methylation loci from two sets (**c**). The green and orange circles represent the set of DNA methylation loci associated with all PA variables (Set A) and all CVD variables (Set B), respectively. Heatmap (with M-value) shows the differential methylation level of the overlapped 23 loci in 1065 participants across PA quartiles and CVD status (**d**). Numbers beginning with "cg" are IlmnIDs

three DNA methylation loci were cg17850088, cg04287259, and cg17078656, which corresponded to gene *VPS45*, *PIK3CD*, and *VPS13D*, respectively.

Figure 2 shows the path diagram of the final SEM with estimated standardized regression coefficients (β). A positive load suggested a positive association with CVD incidence (i.e., a "stimulation" path), while a negative load suggested an inverse association (i.e., an "inhibition" path). Focusing on the correlations among PA, DNA methylation, and CVD, direct loads on CVD were found from PA quartile 2 (PA Q2 ⇒ CVD, β – 0.782), PA quartile 3 (PA Q3 ⇒ CVD, β – 0.805), and the three genes
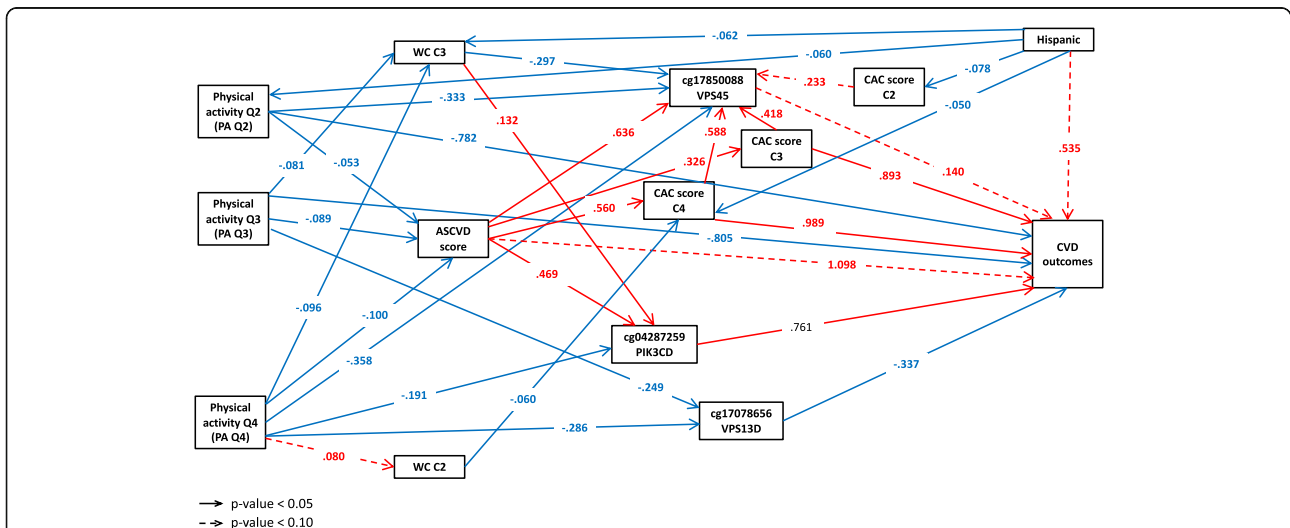

Fig. 2 SEM model with standardized coefficient loads among PA, DNA methylation loci, CVD, and covariates. Each variable in the box represents a continuous variable or a quartile/category of a categorical variable. Numbers beginning with "cg" are IlmnIDs. *VPS45*, *PIK3CD*, and *VPS13D* are the corresponding gene names of the IlmnIDs. A red arrow represents a positive load ("stimulation") from one variable to another, while a blue arrow represents a negative load ("inhibition"). Solid lines indicate the associations with a *p*-value < 0.05, whereas dashed lines for a *p*-value between 0.05 and 0.10. Abbreviations: ASCVD, atherosclerotic CVD; CAC, coronary artery calcium score; WC, waist circumference; Q2-Q4, quartile 2 to 4; C2-C4, category 2 to 4

Shi *et al. BMC Genomics* (2021) 22:790

Page 9 of 13

(*VPS45* ⇒ CVD [β 0.140], *PIK3CD* ⇒ CVD [β 0.761], and *VPS13D* ⇒ CVD [β – 0.337]). PA quartile 2 (PA Q2) and PA quartile 3 (PA Q3) showed negative loads on CVD (i.e., inhibition on CVD incidence), with the greatest absolute value in the association from PA quartile 3 (|β| = 0.805). Among the three genes, *VPS45* and *PIK3CD* had positive loads ("stimulation") on CVD, indicating they were positively associated with CVD incidence. In contrast, gene *VPS13D* was inversely associated with CVD incidence ("inhibition") due to a significant negative load. Located in the central part of the path diagram, the three genes, along with ASCVD score, CAC score, and WC, engaged as mediators bridging other variables to CVD through indirect paths. Counted from the total number of path arrows pointed to the mediators, gene *VPS45* (seven in total) was found to have the largest number of indirect paths mediating to CVD, followed by mediators such as gene *PIK3CD*, ASCVD score, and gene *VPS13D*.

The standardized total effect, direct path effect, and indirect path effects are summarized in Table 4. We calculated the total effects of PA on CVD from the sum of all direct and indirect path effects. For example, the total effect of PA Q2 → CVD (β = – 0.959 = [– 0.782] + [– 0.047] + [– 0.130]) can be decomposed into direct effect (PA Q2 ⇒ CVD, β – 0.782) and indirect paths through *VPS45* (PA Q2 ⇒ *VPS45* → CVD, β – 0.047) and ASCVD score (PA Q2 ⇒ ASCVD→CVD, β – 0.130). Similarly, PA Q3 and Q4 (vs PA Q1) each had multiple paths involving a significant indirect path through reducing the effect of ASCVD score (the variable with the largest total effect [|β| = 2.454] on CVD). However, different from PA Q2, PA Q3 and Q4 started to have indirect "stimulation" path to CVD incidence. The total effect of PA Q3 on CVD consisted of three "inhibition" paths, and one "stimulation" path through *VPS13D* (β = 0.084). The total effect of PA Q4 on CVD was the summation of five "inhibition" paths, and one "stimulation" indirect path through *VPS13D* (β = 0.096). Despite the existence of the "stimulation" paths, the total effects of all PA quartiles were negative, indicating inverse associations ("inhibition") with CVD incidence. However, the overall association of PA Q4 with CVD was much weaker (β = – 0.355) than PA Q2 (β = – 0.959) and Q3 (β = – 0.944). In sensitivity analyses, the coefficients and significant paths between CVD and other variables were robust in all models described in the Methods section.

Discussion

The current study is the first to simultaneously examine the relationships among PA, DNA methylation, and CVD in a large multiethnic cohort with long-term exposure. The major finding of the present work is the identification of three DNA methylation loci

Table 4 Overall coefficient loads on CVD in SEM model

Variables	Coefficient ß	95%CI
VPS45 → CVD	**0.140**	(0.054, 0.226)
PIK3CD → CVD	**0.761**	(0.467, 1.055)
VPS13D → CVD	**-0.337**	(-0.478, -0.196)
PA Q2 → CVD	**-0.959**	(-1.554, -0.449)
PA Q2 ⇒ CVD	-0.782	(-1.165, -0.339)
PA Q2 ⇒ VPS45 → CVD	-0.047	(-0.106, -0.011)
PA Q2 ⇒ ASCVD score → CVD	-0.130	(-0.283, -0.039)
PA Q3 → CVD	**-0.944**	(-1.628, -0.413)
PA Q3 ⇒ CVD	-0.805	(-1.205, -0.405)
PA Q3 ⇒ WC C3 → CVD	-0.005	(-0.019, -0.001)
PA Q3 ⇒ VPS13D → CVD	0.084	(0.029, 0.068)
PA Q3 ⇒ ASCVD score → CVD	-0.218	(-0.433, -0.076)
PA Q4 → CVD	**-0.355**	(-0.713, -0.124)
PA Q4 ⇒ WC C2 ⇒ CAC score C4 → CVD	-0.005	(-0.017, -0.001)
PA Q4 ⇒ WC C3 → CVD	-0.006	(-0.021, -0.001)
PA Q4 ⇒ VPS13D → CVD	0.096	(0.036, 0.185)
PA Q4 ⇒ PIK3CD → CVD	-0.145	(-0.269, -0.059)
PA Q4 ⇒ VPS45 → CVD	-0.050	(-0.111, -0.012)
PA Q4 ⇒ ASCVD score → CVD	-0.245	(-0.479, -0.087)
ASCVD score → CVD	**2.454**	(1.027, 4.165)
ASCVD score ⇒ CVD	1.098	(0.431, 1.765)
ASCVD score ⇒ CAC score C3 → CVD	0.310	(0.129, 0.566)
ASCVD score ⇒ CAC score C4 → CVD	0.600	(0.294, 0.984)
ASCVD score ⇒ PIK3CD → CVD	0.357	(0.155, 0.640)
ASCVD score ⇒ VPS45 → CVD	0.089	(0.019, 0.210)
CAC score C2 → CVD	**0.033**	(0.006, 0.080)
CAC score C2 ⇒ VPS45 → CVD	0.033	(0.006, 0.080)
CAC score C3 → CVD	**0.952**	(0.508, 1.418)
CAC score C3 ⇒ CVD	0.893	(0.493, 1.293)
CAC score C3 ⇒ VPS45 → CVD	0.059	(0.015, 0.125)
CAC score C4 → CVD	**1.071**	(0.594, 1.575)
CAC score C4 ⇒ CVD	0.989	(0.570, 1.408)
CAC score C4 ⇒ VPS45 → CVD	0.082	(0.024, 0.167)
WC C2 → CVD	**-0.064**	(-0.137, -0.020)
WC C2 ⇒ CAC score C4 → CVD	-0.064	(-0.137, -0.020)
WC C3 → CVD	**0.059**	(0.013, 0.158)
WC C3 ⇒ PIK3CD → CVD	0.100	(0.035, 0.198)
WC C3 ⇒ VPS45 → CVD	-0.042	(-0.041, -0.022)
Hispanic → CVD	**0.533**	(0.230, 0.897)
Hispanic ⇒ CVD	0.535	(0.262, 0.808)
Hispanic ⇒ PA Q2 → CVD	0.057	(0.014, 0.138)
Hispanic ⇒ CAC score C2 → CVD	-0.003	(-0.004, -0.001)
Hispanic ⇒ CAC score C4 → CVD	-0.054	(-0.117, -0.015)
Hispanic ⇒ WC C3 → CVD	-0.004	(-0.005, -0.001)

→ denotes the overall effect of the left-sided variable on CVD;
⇒ denotes the direct effect of the left-sided variable on the right-sided variable;
PA Q1 to Q4 denote PA Quartile 1 to Quartile 4.
PA (MET-mins/week): Q1 ≤ 3,555, Q2 ≤ 6,555, Q3 ≤ 10,650, Q4 > 10,650.
WC (cm): C1 < 90 (female) or < 100 (male), C2 < 105 (female) or < 110 (male), C3 ≥ 105 (female) or ≥ 110 (male). CAC score: C1 CAC = 0, C2 0 < CAC ≤ 100, C3 100 < CAC ≤ 400, C4 > 400. PA, physical activity; ASCVD, atherosclerotic CVD; WC, waist circumference; CAC score, coronary artery calcium score.

Shi *et al. BMC Genomics*   (2021) 22:790

Page 10 of 13

(corresponding to genes *PIK3CD*, *VPS45*, and *VPS13D*) playing as mediators linking the association between PA and CVD. Although we used the screened significant loci to fit the SEM model, the *p*-values distribution of the screened loci in SEM complied uniform pattern. Results of the SEM depicted a network with both direct and indirect pathways showing how PA was associated with the reduced CVD incidence through different genes and/or other risk factors.

Consistent with most previous studies [38], we showed that the increase of habitual PA time was associated with decreased incidence of CVD, although not totally linear (strongest with PA Q2 and Q3, but relatively weaker with Q4). The magnitude of the association between CVD and PA Q2 ($\beta = -0.959$) or Q3 ($\beta = -0.944$) was equivalent to > 30% of the association between CVD and ASCVD score ($\beta = 2.454$). Such strong quantified associations highlighted the importance of being physically active in the reduction of CVD incidence. However, the highest level of PA quartile (PA Q4) was found to have the weakest protective effect on the cardiovascular system due to the greatest contribution of the "stimulation" path components (Fig. 2 and Table 4). The reduction of the inhibition effect of PA as the intensity approaches vigorous activity has been previously observed [39, 40]. Additionally, growing evidence has shown that vigorous activity can even acutely and transiently increase the risk of CVD in susceptible persons with congenital anomalies or existing cardiac diseases [41]. It is noted that in the present study we fitted models using total PA, which may be different from the studies focusing on vigorous PA or excessive exercise [20, 37]. The difference may explain why we observed harmful components (with overall protective effect) of heigh level PA, but not overall harmful effect reported in some previous studies [42]. Although the mechanism by which vigorous PA has harmful components or effect on the cardiovascular system is not well understood [43], we provided one promising biological mechanism – the epigenetic modification of gene *VPS13D* – based on the role as a mediator in the SEM model.

The biological functions of the identified genes in the SEM model were strongly related to the development of atherosclerosis. As the dominant cause of CVD, atherosclerosis is in fact a chronic inflammatory condition involving immune cells (e.g., macrophages, etc.) and pro-inflammatory cytokines [44]. The monocyte-derived macrophage migrates to lesions in the vessel and contributes to every phase of atherogenesis [45]. The initial atherosclerotic plaque is characterized by the accumulation of lipid-laden macrophages. The activated macrophages then secrete cytokines (e.g., TNF-α and IL-1), enzymes, and growth factors (e.g., PDGF), and contribute to the further development of atherosclerotic plaque [46]. As the key mediators in the SEM model, genes *VPS13D* and *PIK3CD* are both related to monocyte-macrophage differentiation [47]. The *VPS13D* (vacuolar protein sorting 13 homolog D) gene encodes a protein that plays an important role in autophagy [48]. The induction of autophagy is found to be essential for the survival and differentiation of monocytes to macrophages [49]. *PIK3CD* (phosphatidylinositol-4,5-bisphosphate 3-kinase [PIK3] catalytic subunit δ) encodes a subunit of the enzyme PIK3 which phosphorylates inositol and is involved in the immune response through the PIK3/AKT pathway [50]. The PIK3/AKT/mTOR pathway has been demonstrated to regulate monocyte-derived macrophage activation and polarization [47]. Furthermore, PIK3 was also found to limit the production of TNF-α in activated monocytes [51] and to regulate the activation of platelet [52]. Notably, the polymorphism of the genes *VPS13D* and *PIK3* were also identified to be associated with CVD in other cohorts [52, 53]. Gene *VPS45* (vacuolar protein sorting 45 homolog) is currently not well studied. The knowledge of *VPS45*-related disease in humans is limited to its association with congenital neutropenia [54]. Considering the important effect of *VPS45* identified in our SEM model, future biological studies on *VPS45* and its role in the development of CVD are promising.

Several limitations in this study are noted. First, the habitual PA time was collected through surveys at Exam 5 along with the collection of samples for epigenetics. Although the interviewers were well trained, recall bias may have been introduced because of self-reporting. Also, although the formation of PA habitus occurred prior to the collection of epigenetics information, the time order of PA habitus and epigenetic pattern formation was still not clear. Second, the time-varying variables are difficult to control in an observational study without a follow-up visit between baseline and endpoint. Third, a survivorship bias was introduced when we used MESA Exam 5 as a baseline and excluded the CVD prevalent cases. By Exam 5, the MESA study has continued for nearly 10 years since Exam 1. People who did not develop CVD during the last four visits were possibly those who were less likely to have CVD due to genetic or environmental factors. However, since the DNA methylation data were collected during Exam 5, the introduction of a survivorship bias is inevitable. Fourth, although the MESA study included Asian-American participants, this group was not included in the current study because they were not represented in the MESA Epigenomics and Transcriptomics Study. Thus, the

Shi *et al. BMC Genomics*   (2021) 22:790

Page 11 of 13

interpretation of the results may not be generalizable to Asian-American populations. In addition, because of the potential heterogeneity [55] and misclassification [56] in the Hispanic population of the MESA cohort. The interpretation of the direct effect of Hispanics on CVD had limitation. Fifth, although two of the three identified genes in this study were also identified by other GWAS cohorts, the current study was the only study reported the association between CVD and the methylation of the genes *PIK3CD*, *VPS45*, and *VPS13D*. The possibility of false positive results should be noted due to the lack of replication. Finally, the effects of genes in this study were only indicative due to the lack of causal inference regarding both statistics and biology. The term "mediator" used in this study only defined the linking variable in SEM, without demonstration of causal mediation effects. Although it is reasonable to assume that DNA methylations mediate the effects of an exposure on a disease since DNA methylation is malleable to environmental changes [57, 58], causation assumptions in high-dimensional omics data are usually not justified without the application of approaches such as global test [59] and Mendelian randomization [60]. It is noteworthy that we interpreted the "stimulation" or "inhibition" effects of genes based on the positive or negative signs of the coefficients, and the assumption that DNA methylation commonly represses gene expression. However, growing evidences showed that DNA methylation can also increase gene expression in some instances [9]. Thus, the actual biological roles of identified genes should be further studied by mechanism research.

Interpreted with a few limitations in mind, the current study is still among the first to simultaneously examine the relationships among PA, DNA methylation, and CVD in a large multiethnic cohort with long-term exposure. The identified genes might have the potential to be therapeutic targets for future CVD prevention or to be involved in the modification of CVD risk score.

Conclusion

Our study was among the first to simultaneously examine the complexity of physical activity (PA) and DNA methylation interaction in the development of (CVD) in a large cohort. The major finding of the present work is the identification of three DNA methylation loci (corresponding to genes *PIK3CD*, *VPS45*, and *VPS13D*) bridging the association between PA and CVD. The function of those three genes warrants further investigation in the pathogenesis of CVD, with potentially therapeutic or predictive implications for clinical care.

Abbreviations
CVD: Cardiovascular disease; PA: Physical activity; MESA: Multi-Ethic Study of Atherosclerosis; SEM: Structural equation modeling; BMI: Body mass index; WC: Waist circumference; ASCVD: Atherosclerotic cardiovascular disease; CAC: Coronary artery calcium; MET: Metabolic equivalence of tasks

Supplementary Information

Acknowledgements

Authors' contributions

Funding

Availability of data and materials

Declarations

Ethics approval and consent to participate
This study was conducted in accordance with the Declaration of Helsinki. Written informed consent was obtained from all participating subjects under research protocols approved by the institutional review boards of the six filed centers (Baltimore City and Baltimore County, Maryland; Chicago, Illinois; Forsyth County, North Carolina; Los Angeles County, California; New York, New York; and St. Paul, Minnesota). Information is available at www.mesa-nhlbi.org. The Multi-Ethnic Study of Atherosclerosis (MESA) Genetics P&P Committee has reviewed and approved our manuscript proposal titled "MESA Genetics Manuscript Proposal: G 602, Mediation of DNA Methylation and Gene Expression on Physical Activity and Sedentary Behavior effects on Cardiovascular Disease: The Multi-Ethnic Study of Atherosclerosis" to access the data used for the present study. The manuscript was also reviewed and approved by the MESA Genetics P&P Committee before the submission. In addition, our current study was not research involving human subjects as defined by U.S. Department of Health and Human Services (DHHS) and U.S. Food and Drug Administration (FDA) regulations determined by the Research Subjects Review Board (RSRB) office at the University of Rochester on May 28, 2019 (Study ID: Study00003812).

Shi *et al. BMC Genomics*   (2021) 22:790

Page 12 of 13

Author details
[1]Department of Clinical and Translational Research, University of Rochester Medical Center, Rochester, NY 14642-0708, USA. [2]Department of Public Health Sciences, University of Rochester Medical Center, Rochester, NY 14642, USA. [3]Division of Cardiology, Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, MD 21287, USA. [4]Division of Public Health Sciences, Wake Forest School of Medicine, Winston-Salem, NC 27157, USA. [5]Department of Medicine, Wake Forest Baptist Medical Center, Winston-Salem, NC 27157, USA. [6]Aab Cardiovascular Research Institute, University of Rochester, School of Medicine and Dentistry, Rochester, NY 14642, USA. [7]Department of Pharmacology and Physiology, University of Rochester, School of Medicine and Dentistry, Rochester, NY 14642, USA. [8]Department of Statistics, Fralin Life Sciences Institute at Virginia Tech, Blacksburg, VA 24061, USA. [9]Department of Medicine, Division of Cardiology, Duke Molecular Physiology Institute, Duke University Medical Center, Durham, NC 27701, USA.

References
1. Roth GA, Mensah GA, Johnson CO, Addolorato G, Ammirati E, Baddour LM, et al. Global burden of cardiovascular diseases and risk factors, 1990-2019: update from the GBD 2019 study. J Am Coll Cardiol. 2020;76(25):2982–3021.
2. Stewart J, Manmathan G, Wilkinson P. Primary prevention of cardiovascular disease: a review of contemporary guidance and literature. JRSM Cardiovasc Dis. 2017;6:2048004016687211.
3. Kathiresan S, Srivastava D. Genetics of human cardiovascular disease. Cell. 2012;148(6):1242–57.
4. Eijsvogels TM, Molossi S, Lee DC, Emery MS, Thompson PD. Exercise at the extremes: the amount of exercise to reduce cardiovascular events. J Am Coll Cardiol. 2016;67(3):316–29.
5. Gleeson M, Bishop NC, Stensel DJ, Lindley MR, Mastana SS, Nimmo MA. The anti-inflammatory effects of exercise: mechanisms and implications for the prevention and treatment of disease. Nat Rev Immunol. 2011;11(9):607–15.
6. Grazioli E, Dimauro I, Mercatelli N, Wang G, Pitsiladis Y, Di Luigi L, et al. Physical activity in the prevention of human diseases: role of epigenetic modifications. BMC Genomics. 2017;18(Suppl 8):802.
7. Devaux Y, Robinson EL. Epigenetics in cardiovascular disease: Elsevier science; 2021.
8. van der Harst P, de Windt LJ, Chambers JC. Translational perspective on epigenetics in cardiovascular disease. J Am Coll Cardiol. 2017;70(5):590–606.
9. Moore LD, Le T, Fan G. DNA methylation and its basic function. Neuropsychopharmacology. 2013;38(1):23–38.
10. Goulle JP, Guerbet M. Recreational use of cannabis: from effects to harm. Epidemiological data. Bull Acad Natl Med. 2020;204(6):543–50.
11. Meder B, Haas J, Sedaghat-Hamedani F, Kayvanpour E, Frese K, Lai A, et al. Epigenome-wide association study identifies cardiac gene patterning and a novel class of biomarkers for heart failure. Circulation. 2017;136(16):1528–44.
12. Fernandez-Sanles A, Sayols-Baixeras S, Castro DEMM, Esteller M, Subirana I, Torres-Cuevas S, et al. Physical activity and genome-wide DNA methylation: the REgistre GIroni del COR study. Med Sci Sports Exerc. 2020;52(3):589–97.
13. Recchioni R, Marcheselli F, Antonicelli R, Mensa E, Lazzarini R, Procopio AD, et al. Epigenetic effects of physical activity in elderly patients with cardiovascular disease. Exp Gerontol. 2017;100:17–27.
14. Zhong J, Agha G, Baccarelli AA. The role of DNA methylation in cardiovascular risk and disease: methodological aspects, study design, and data analysis for epidemiological studies. Circ Res. 2016;118(1):119–31.
15. Antonogeorgos G, Panagiotakos DB, Pitsavos C, Papageorgiou C, Chrysohoou C, Papadimitriou GN, et al. Understanding the role of depression and anxiety on cardiovascular disease risk, using structural equation modeling; the mediating effect of the Mediterranean diet and physical activity: the ATTICA study. Ann Epidemiol. 2012;22(9):630–7.
16. Bild DE, Bluemke DA, Burke GL, Detrano R, Diez Roux AV, Folsom AR, et al. Multi-ethnic study of atherosclerosis: objectives and design. Am J Epidemiol. 2002;156(9):871–81.
17. Liu Y, Ding J, Reynolds LM, Lohman K, Register TC, De La Fuente A, et al. Methylomics of gene expression in human monocytes. Hum Mol Genet. 2013;22(24):5065–74.
18. Chi GC, Liu Y, MacDonald JW, Barr RG, Donohue KM, Hensley MD, et al. Long-term outdoor air pollution and DNA methylation in circulating monocytes: results from the multi-ethnic study of atherosclerosis (MESA). Environ Health. 2016;15(1):119.
19. Ainsworth BE, Haskell WL, Herrmann SD, Meckes N, Bassett DR Jr, Tudor-Locke C, et al. 2011 compendium of physical activities: a second update of codes and MET values. Med Sci Sports Exerc. 2011;43(8):1575–81.
20. Joseph JJ, Echouffo-Tcheugui JB, Golden SH, Chen H, Jenny NS, Carnethon MR, et al. Physical activity, sedentary behaviors and the incidence of type 2 diabetes mellitus: the multi-ethnic study of atherosclerosis (MESA). BMJ Open Diabetes Res Care. 2016;4(1):e000185.
21. Goff DC Jr, Lloyd-Jones DM, Bennett G, Coady S, D'Agostino RB, Gibbons R, et al. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association task force on practice guidelines. Circulation. 2014;129(25 Suppl 2):S49–73.
22. Grundy SM, Stone NJ, Bailey AL, Beam C, Birtcher KK, Blumenthal RS, et al. 2018 AHA/ACC/AACVPR/AAPA/ABC/ACPM/ADA/AGS/APhA/ASPC/NLA/PCNA guideline on the Management of Blood Cholesterol: a report of the American College of Cardiology/American Heart Association task force on clinical practice guidelines. Circulation. 2019;139(25):e1082–143.
23. Greenland P, Blaha MJ, Budoff MJ, Erbel R, Watson KE. Coronary calcium score and cardiovascular risk. J Am Coll Cardiol. 2018;72(4):434–47.
24. Zhu S, Heymsfield SB, Toyoshima H, Wang Z, Pietrobelli A, Heshka S. Race-ethnicity-specific waist circumference cutoffs for identifying cardiovascular disease risk factors. Am J Clin Nutr. 2005;81(2):409–15.
25. McClelland RL, Chung H, Detrano R, Post W, Kronmal RA. Distribution of coronary artery calcium by race, gender, and age: results from the multi-ethnic study of atherosclerosis (MESA). Circulation. 2006;113(1):30–7.
26. Ross R, Neeland IJ, Yamashita S, Shai I, Seidell J, Magni P, et al. Waist circumference as a vital sign in clinical practice: a consensus statement from the IAS and ICCR working group on visceral obesity. Nat Rev Endocrinol. 2020;16(3):177–89.
27. Neves PO, Andrade J, Moncao H. Coronary artery calcium score: current status. Radiol Bras. 2017;50(3):182–9.
28. Budoff MJ, Nasir K, McClelland RL, Detrano R, Wong N, Blumenthal RS, et al. Coronary calcium predicts events better with absolute calcium scores than age-sex-race/ethnicity percentiles: MESA (multi-ethnic study of atherosclerosis). J Am Coll Cardiol. 2009;53(4):345–52.
29. Sundar IK, Yin Q, Baier BS, Yan L, Mazur W, Li D, et al. DNA methylation profiling in peripheral lung tissues of smokers and patients with COPD. Clin Epigenetics. 2017;9:38.
30. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W. Smyth GK: limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43(7):e47.
31. Bolstad B. preprocessCore: A collection of pre-processing functions; 2019.
32. Li D, Xie Z, Pape ML, Dye T. An evaluation of statistical methods for DNA methylation microarray data analysis. BMC Bioinformatics. 2015;16:217.
33. Li D, Xie Z, Zand M, Fogg T, Dye T. Bon-EV: an improved multiple testing procedure for controlling false discovery rates. BMC Bioinformatics. 2017;18(1):1.
34. Wickham H. ggplot2: elegant graphics for data analysis. New York: Springer-Verlag; 2016.
35. Kolde R. pheatmap: Pretty Heatmaps; 2019.
36. Muthén B. A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. Psychometrika. 1984;49(1):115–32.
37. Bertoni AG, Whitt-Glover MC, Chung H, Le KY, Barr RG, Mahesh M, et al. The association between physical activity and subclinical atherosclerosis: the multi-ethnic study of atherosclerosis. Am J Epidemiol. 2009;169(4):444–54.
38. Shiroma EJ, Lee IM. Physical activity and cardiovascular health: lessons learned from epidemiological studies across age, gender, and race/ethnicity. Circulation. 2010;122(7):743–52.
39. Lee DC, Pate RR, Lavie CJ, Sui X, Church TS, Blair SN. Leisure-time running reduces all-cause and cardiovascular mortality risk. J Am Coll Cardiol. 2014;64(5):472–81.

40. Mohamadzade B, Hashmi RM, Simorangkir R, Gharaei R, Ur Rehman S, Abbasi QH. Recent Advances in Fabrication Methods for Flexible Antennas in Wearable Devices: State of the Art. Sensors (Basel). 2019;19(10).

41. Buttar HS, Li T, Ravi N. Prevention of cardiovascular diseases: role of exercise, dietary interventions, obesity and smoking cessation. Exp Clin Cardiol. 2005; 10(4):229–49.

42. O'Keefe JH, Patil HR, Lavie CJ, Magalski A, Vogel RA, McCullough PA. Potential adverse cardiovascular effects from excessive endurance exercise. Mayo Clin Proc. 2012;87(6):587–95.

43. Thompson PD, Franklin BA, Balady GJ, Blair SN, Corrado D, Estes NA 3rd, et al. Exercise and acute cardiovascular events placing the risks into perspective: a scientific statement from the American Heart Association Council on nutrition, physical activity, and metabolism and the council on clinical cardiology. Circulation. 2007;115(17):2358–68.

44. Frostegard J. Immunity, atherosclerosis and cardiovascular disease. BMC Med. 2013;11:117.

45. Ross R. Atherosclerosis--an inflammatory disease. N Engl J Med. 1999;340(2): 115–26.

46. Alexander MR, Owens GK. Epigenetic control of smooth muscle cell differentiation and phenotypic switching in vascular development and disease. Annu Rev Physiol. 2012;74:13–40.

47. Vergadi E, Ieronymaki E, Lyroni K, Vaporidi K, Tsatsanis C. Akt signaling pathway in macrophage activation and M1/M2 polarization. J Immunol. 2017;198(3):1006–14.

48. Wang Z, Zhang H. Mitophagy: Vps13D couples mitochondrial fission and Autophagic clearance. Curr Biol. 2018;28(2):R66–8.

49. Zhang Y, Morgan MJ, Chen K, Choksi S, Liu ZG. Induction of autophagy is essential for monocyte-macrophage differentiation. Blood. 2012;119(12): 2895–905.

50. Weichhart T, Saemann MD. The PI3K/Akt/mTOR pathway in innate immune cells: emerging therapeutic applications. Ann Rheum Dis. 2008;67(Suppl 3): iii70–4.

51. Kramer PR, Winger V, Reuben J. PI3K limits TNF-alpha production in CD16-activated monocytes. Eur J Immunol. 2009;39(2):561–70.

52. Durrant TN, Hers I. PI3K inhibitors in thrombosis and cardiovascular disease. Clin Transl Med. 2020;9(1):8.

53. Koriyama H, Nakagami H, Katsuya T, Sugimoto K, Yamashita H, Takami Y, et al. Identification of evidence suggestive of an association with peripheral arterial disease at the OSBPL10 locus by genome-wide investigation in the Japanese population. J Atheroscler Thromb. 2010;17(10):1054–62.

54. Stepensky P, Saada A, Cowan M, Tabib A, Fischer U, Berkun Y, et al. The Thr224Asn mutation in the VPS45 gene is associated with the congenital neutropenia and primary myelofibrosis of infancy. Blood. 2013;121(25):5078–87.

55. Manichaikul A, Palmas W, Rodriguez CJ, Peralta CA, Divers J, Guo X, et al. Population structure of Hispanics in the United States: the multi-ethnic study of atherosclerosis. PLoS Genet. 2012;8(4):e1002640.

56. Divers J, Redden DT, Rice KM, Vaughan LK, Padilla MA, Allison DB, et al. Comparing self-reported ethnicity to genetic background measures in the context of the multi-ethnic study of atherosclerosis (MESA). BMC Genet. 2011;12:28.

57. Fasanelli F, Baglietto L, Ponzi E, Guida F, Campanella G, Johansson M, et al. Hypomethylation of smoking-related genes is associated with future lung cancer in four prospective cohorts. Nat Commun. 2015;6:10192.

58. Sandanger TM, Nost TH, Guida F, Rylander C, Campanella G, Muller DC, et al. DNA methylation and associated gene expression in blood prior to lung cancer diagnosis in the Norwegian women and Cancer cohort. Sci Rep. 2018;8(1):16714.

59. Djordjilovic V, Page CM, Gran JM, Nost TH, Sandanger TM, Veierod MB, et al. Global test for high-dimensional mediation: testing groups of potential mediators. Stat Med. 2019;38(18):3346–60.

60. Richmond RC, Hemani G, Tilling K, Davey Smith G, Relton CL. Challenges and novel approaches for investigating molecular mediation. Hum Mol Genet. 2016;25(R2):R149–56.

Publisher's Note