



Contents lists available at ScienceDirect

Nutrition

journal homepage: www.nutritionjrn.com

Applied nutritional investigation

Nutritional habits, lifestyle, and genetic predisposition in cardiovascular and metabolic traits in Turkish population

Sefayet Karaca PhD^{a,b,*}, Sema Erge PhD^{b,c}, Tomris Cesuroglu MD^{b,d}, Renato Polimanti PhD^e

^aSchool of Health Science, Aksaray University, Aksaray, Turkey^bGENAR Institute for Public Health and Genomics Research, Ankara, Turkey^cDepartment of Nutrition and Dietetics, Zirve University, Gaziantep, Turkey^dDepartment of Social Medicine, Maastricht University, Maastricht, The Netherlands^eDepartment of Psychiatry, Yale University School of Medicine and VA CT Healthcare Center, West Haven, CT, USA

ARTICLE INFO

Article history:

Received 28 August 2015

Accepted 17 December 2015

Keywords:

Cardiometabolic traits

Genetic predisposition

Diet

Interactive mechanisms

Turkey

ABSTRACT

Objectives: Cardiovascular and metabolic traits (CMT) are influenced by complex interactive processes including diet, lifestyle, and genetic predisposition. The present study investigated the interactions of these risk factors in relation to CMTs in the Turkish population.

Methods: We applied bootstrap agglomerative hierarchical clustering and Bayesian network learning algorithms to identify the causative relationships among genes involved in different biological mechanisms (i.e., lipid metabolism, hormone metabolism, cellular detoxification, aging, and energy metabolism), lifestyle (i.e., physical activity, smoking behavior, and metropolitan residency), anthropometric traits (i.e., body mass index, body fat ratio, and waist-to-hip ratio), and dietary habits (i.e., daily intakes of macro- and micronutrients) in relation to CMTs (i.e., health conditions and blood parameters).

Results: We identified significant correlations between dietary habits (soybean and vitamin B12 intakes) and different cardiometabolic diseases that were confirmed by the Bayesian network-learning algorithm. Genetic factors contributed to these disease risks also through the pleiotropy of some genetic variants (i.e., *F5* rs6025 and *MTR* rs180508). However, we also observed that certain genetic associations are indirect since they are due to the causative relationships among the CMTs (e.g., *APOC3* rs5128 is associated with low-density lipoproteins cholesterol and, by extension, total cholesterol).

Conclusions: Our study applied a novel approach to integrate various sources of information and dissect the complex interactive processes related to CMTs. Our data indicated that complex causative networks are present: causative relationships exist among CMTs and are affected by genetic factors (with pleiotropic and non-pleiotropic effects) and dietary habits.

© 2016 Elsevier Inc. All rights reserved.

Introduction

S.K and S.E. contributed equally to this work. This research was a part of the GENTEST model and partially funded by the project OBESIS-Development of a Genetic and Information Technologies Assisted System on Obesity Treatment and Control (Funded by The TEYDEP/TUBITAK; Project no: 3060480). We are greatly indebted to Dr. Serdar Savas for his great contribution to the development of the GENTEST Model. GENTEST was identified as a "Best Practice Model in Europe for Public Health Genomics" by the Public Health Genomics European Network (PHGEN). We would also like to thank the staff members of the GENAR Molecular Biology and Genetics Research and Application Laboratories for their assistance. The authors declare no conflict of interest.

* Corresponding author. Tel.: +90-53-2506-8202; fax: +90-38-2288-2545.

E-mail address: skaraca@aksaray.edu.tr (S. Karaca).

<http://dx.doi.org/10.1016/j.nut.2015.12.027>

0899-9007/© 2016 Elsevier Inc. All rights reserved.

Cardiometabolic diseases are the leading cause of death worldwide. They include a wide range of disorders, such as obesity, insulin resistance, metabolic dyslipidemia, atherosclerosis, type II diabetes, and cardiovascular diseases (CVD), various combinations of which frequently occur simultaneously [1]. Anthropometric measurements and blood parameters are phenotypic traits used to estimate their risk, as well as to monitor and diagnose these disorders [2]. Studies of cardiometabolic traits (CMT) demonstrated how dietary habits, lifestyle, and genetic defects converge to alter the same molecular mechanisms (e.g.,

lipid metabolism, hormone metabolism, cellular detoxification, aging, and energy metabolism) that increase the risk of cardiometabolic diseases [3–5]. These findings indicate that CMTs, dietary habits, lifestyle, and genetic predisposition are all included in the same intricate network of causative relationships. Although this knowledge suggests that investigations utilizing multiple types of data are necessary to dissect the complex pathogenesis of these traits, most studies still focus on approaches based on a single type of data. To our knowledge, few studies have applied network-based approaches to investigate CMTs [6,7].

Here, we used hierarchical clustering analysis and Bayesian networks to analyze the causative relationships among genetic loci involved in relevant biological mechanisms (i.e., lipid metabolism, hormone metabolism, cellular detoxification, aging, and energy metabolism), lifestyle (i.e., physical activity, smoking behavior, and metropolitan residency), anthropometric traits (i.e., body mass index [BMI], body fat ratio [BFR], and waist-to-hip ratio [WHR]), and dietary habits (i.e., daily intakes of macro- and micronutrients) in relation to health conditions and blood parameters related to CMTs. Specifically, we investigated a representative sample of the Turkish population. To our knowledge, genetic and epidemiologic data are not widely available regarding the pathogenesis of CMTs in the Turkish population [8,9]. The genetic structure of the Turkish population is complex due to its demographic history and location near Central Asia, Europe, and the Middle East [10]. Relevant genetic differences are present between Turkish and European populations, partially explaining the health disparities of Turkish communities in Northern Europe [11] and also contributing to the pharmacogenetics differences observed [12]. Moreover, comprehensive information regarding the lifestyle and dietary habits of the Turkish population could enhance our understanding of the relationship between environmental factors and the risk of cardiovascular diseases. Accordingly, the present study provided what is to our knowledge the first comprehensive report of the dietary habits in the Turkish population and proposed an innovative approach to integrate genetic and non-genetic data in the investigation of CMTs.

Material and methods

Samples

All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975, as revised in 2000. Informed consent was obtained from all subjects included in the study. Participants were recruited from the seven largest cities (Istanbul, Ankara, Izmir, Bursa, Adana, Antalya, and Samsun) within five regions of Turkey. The demographic characteristics of these cities are similar to the overall structure of population in Turkey since these large cities harbor 40% of the population in Turkey. All of the participants voluntarily applied to the GENAR Biotechnology and Molecular Genetics, Research and Application Laboratories to enroll in a preventive health care intervention program [13]. The study population included unrelated Turkish participants, who were at least 18 y of age. Personal and health information collection, genetic and biochemical tests, and nutritional assessments were performed after obtaining informed consent. All collected data were selected according to their involvement in CMTs. The screened genetic and biochemical parameters were grouped according to their contributions to different biological mechanisms. Accordingly, different disease-related test packages, which contain relevant combinations of genetic and non-genetic analyses, were generated. The volunteers could choose one or more of those packages for genetic and biochemical testing according to their health priorities. These packages had different genetic and blood parameters and therefore the number of screened individuals per parameter varied.

Collection of personal and health information, nutritional assessment, and blood-related analyses

A standard questionnaire was used to collect the following: personal information, health status, living and working conditions, lifestyle information

including nutritional habits, food consumption, physical activity and exercise, and smoking status (Table 1). The questionnaire was completed by the participants with the assistance of trained personnel. The most appropriate test package(s) were chosen with the assistance of the health professionals (i.e., physicians authorized to be practitioners after an appropriate training). Collected health information refers to the known health conditions, based on the volunteers' past medical records. Blood samples were collected during fasting for a number of blood parameters. Biochemical tests were carried out in licensed medical laboratories. In accordance with the study design, the health conditions and biochemical parameters are independent data.

The food consumption information was collected using a food frequency questionnaire (FFQ) and a detailed photographic food portion size atlas, both developed for this program according to nutrition habits in Turkey. The nutritional content of the FFQ was calculated using the professional nutrition database software BeBis version 5 [14]. Anthropometric measurements (height, weight, waist circumference, and BFR) were made by dietitians. Further information about data collection is available in our previous article [13].

Genotyping

Genomic DNAs were isolated from buccal swabs or whole blood samples using an MN DNA isolation kit (Macherey Nagel-Nucleospin, Düren, Germany). SequenomRealSNP software was used to design sequence-specific amplification primers (Metabion, Martinsreid, Germany) for the multiplex level (details regarding the PCR primers, extended probes, and multiplex combinations are available upon the request). The investigated variants were selected according to their involvement in molecular pathways related to CMTs (e.g. lipid metabolism, cardio-vascular diseases, hormone metabolism, cellular detoxification, aging, and energy metabolism) [11], developing the Genteset practice model to screen genetic predisposition to complex diseases of healthy Turkish volunteers [13]. All laboratory procedures were performed according to the manufacturer's protocol of the MassArray platform (Sequenom Inc., San Diego, CA, USA). Detailed information about the genotyping pipeline is reported in our previous study [15].

Statistical analysis

Statistical analysis was performed using the computing environment R. Non-genetic variables were regressed with respect to age and sex to compute the residuals and correct for the strong confounding effect of age and sex following the analysis. Genotype data were converted in dosage variables on the basis of their minor allele frequency (MAF). Spearman's correlation test was used to test the correlation among the variables collected and we used the rho coefficient to analyze the hierarchical clustering among the tested variables. Specifically, we calculated dissimilarity as $1 - \text{Abs}(\text{Correlation})$ and used it as distance. Distance matrices were used for the hierarchical clustering analysis and the R package pvclust [16] to calculate *P* values for each cluster after 10,000 bootstrap replications. We considered clusters with approximately unbiased *P* value > 95% as significant. We used PANTHER Classification Systems and the PANTHER Pathways annotation data set to detect the molecular pathways related to the genetic clusters identified [17]. To investigate the causative relationship of genetic factors with CMTs, we performed a causative model analysis based on Bayesian networks. Specifically, we used the R package bnlearn [18], applying the hill-climbing learning algorithm, and determined the best causative network for each cluster considering the BIC (Bayesian information criterion) score.

Results

Table 1 shows the characteristics of the study population in terms of personal information, anthropometric traits, diet, lifestyle, health conditions, aspirin use, blood parameters, vital signs, and genetic polymorphisms. In our sample, we observed 39% of overweight subjects (BMI 25–30) and 32% of obese subjects (BMI > 30) in accordance with prevalence reported for the Turkish general population [19]. To identify the relationships between non-genetic factors (i.e., anthropometric traits, diet, lifestyle, and aspirin use) and the CMTs (i.e., blood parameters and health conditions), we estimated the correlation matrix among these variables (Supplemental Table 1) and performed a bootstrap agglomerative hierarchical clustering using the dissimilarity matrix (Fig. 1). We observed 11 significant clusters, suggesting relationships among CMTs, nutritional factors, and other environmental factors. Most of the significant clusters consisted of parameters included in the same phenotypic

Table 1
Characteristics of study populations

Characteristic	Category	Unit	N	Measurements	
Sex	Personal information	Females (%)	247	87 (35)	
Age		y	247	45.2 ± 11.8	
Body mass index (BMI)	Anthropometric traits	kg/m ²	244	28.5 ± 5.5	
Body fat ratio (BFR)		%	230	29.4 ± 9.1	
Waist hip ratio (WHR)		–	170	0.89 ± 0.08	
Sodium intake	Nutritional information	mg	171	4060 ± 1374	
Optimal energy intake ratio (OEIR)		daily	247	1.18 ± 0.26	
Daily fat intake		g	247	92 ± 29	
Fiber intake		g	247	34 ± 10	
Alcohol intake		g	247	2 (0–176)	
Vitamin A intake		IU	247	3201 (1135–17 447)	
Vitamin E intake		IU	247	26.9 (7–107)	
Vitamin B6 intake		mg	247	2 (0.8–253)	
Folate intake		µg	247	440 (50–990)	
Vitamin B12 intake		µg	247	5.4 (0.4–73.6)	
Vitamin C intake		mg	247	210 (51–714)	
Potassium intake		mg	247	3830 (1270–9204)	
Magnesium intake		mg	247	448 (191–924)	
Zinc intake		mg	246	15.2 (7.8–28.6)	
Sucrose intake		g	246	44.8 ± 21.3	
Saturated fat acid (SFA) intake		g	247	30.9 ± 12	
Monounsaturated fatty acids (MUFA) intake		g	247	36.4 ± 12.6	
Polyunsaturated fatty acids (PUFA) intake		g	245	24 (9–127)	
Cholesterol intake		mg	247	269 ± 120	
Caffeine intake		mg	247	270 ± 159	
Ratio of omega-6 to omega-3 essential fatty acids	–	244	9.8 ± 2.9		
Vegetable consumption	Life-style	serving	247	2 (1–6)	
Fruit consumption		serving	247	2 (0–8)	
Soybean consumption		g	247	0 (0–23)	
Omega-3 fatty acids intake		g	247	2.5 ± 0.9	
Physical activity		Daily (1/2/3) (Sedentary/Medium/Heavy)	244	1 (1–3)	
Tobacco consumption		0/1/2 (No/Past/Active)	245	1 (0–2)	
Passive smokers		Y (%)	190	83 (43)	
Living in metropolitan area for at least 10 y		Y (%)	205	197 (96)	
Impaired glucose tolerance; insulin resistance; hyperglycemia		Health information	Y (%)	245	21 (9)
Diabetes – Type I (T1D)		Y (%)	245	3 (1)	
Diabetes – Type II (T2D)	Y (%)	245	20 (8)		
Heart attack or stroke	Y (%)	245	3 (1)		
Hypertension	Y (%)	246	55 (22)		
Hypercholesterolemia	Y (%)	245	93 (38)		
Hypertriglyceridemia	Y (%)	244	46 (19)		
Cardiovascular diseases (CVD)	Y (%)	242	19 (8)		
Thrombosis	Y (%)	242	3 (1)		
Aspirin use	Drug use	Y (%)	247	67 (27)	
Fasting plasma glucose (FPG)	Blood parameters	mg/dL	244	91 (61–281)	
Glycated hemoglobin (HbA1 c)		%	119	5.7 (4.8–10.2)	
Fasting insulin (FI)		µIU/mL	102	13.5 ± 11.6	
Homeostatic model assessment - insulin resistance (HOMA-IR)		–	102	3.7 ± 3.0	
Total cholesterol		mg/dL	245	198 ± 39	
High-density lipoproteins (HDL) cholesterol		mg/dL	244	51 ± 15	
Total/HDL cholesterol ratio		–	245	4.2 ± 1.3	
Low-density lipoproteins (LDL) cholesterol		mg/dL	245	121 ± 34	
Triacylglycerols		mg/dL	244	120 (28–1193)	
Lipoprotein a (LPA)		mg/dL	85	24 ± 22	
High-sensitivity C-reactive protein (hs-CRP)		mg/L	244	1.5 (0–38.8)	
Erythrocyte sedimentation rate (ESR)		mm/hour	160	12 ± 9	
Homocysteine		µmol/L	243	11 ± 4	
Thyroid-stimulating hormone (TSH)		mIU/L	145	1.8 ± 1.5	
Uric acid		mg/dL	157	5.5 ± 1.4	
Systolic blood pressure (SBP)		Vital signs	mmHg	228	119 ± 16
Diastolic blood pressure (DBP)			mmHg	228	77 ± 11
Pulse pressure	mmHg		228	42 ± 11	

(Continued)

Table 1 (Continued)

Characteristic	Category	Unit	N	Measurements
APOC3 rs5128	Genetic polymorphisms	MAF	243	0.187
LPL rs328		MAF	240	0.085
CETP rs708272		MAF	218	0.452
F5 rs6025		MAF	236	0.047
F2 rs1799963		MAF	238	0.027
MTHFR rs1801133		MAF	241	0.336
MTHFR rs1801131		MAF	246	0.358
MTRR rs1801394		MAF	245	0.480
NOS3 rs1799983		MAF	245	0.318
NOS3 rs2070744		MAF	172	0.357
IL6 rs1800795		MAF	247	0.273
IL6 rs1800796		MAF	245	0.082
TNF rs1800629		MAF	247	0.101
VDR rs1544410		MAF	242	0.393
VDR rs731236		MAF	246	0.360
VDR rs2228570		MAF	246	0.307
COL1 A1 rs1800012		MAF	244	0.117
ACE Ins/Del (rs4646994/rs1799752)		MAF	182	0.431
PPARG rs1801282		MAF	228	0.083
SOD2 rs1799725		MAF	245	0.420
CYP1 A1 rs4646903		MAF	231	0.143
MTR rs1805087		MAF	144	0.181
LIPC rs1800588		MAF	206	0.201
PON1 rs662		MAF	201	0.294
APOA1 rs670		MAF	196	0.130
CYP1 B1 rs1056836		MAF	146	0.360
CYP1 B1 rs1800440		MAF	125	0.228
SERPINE1 rs1799889	MAF	237	0.222	
COMT rs4680	MAF	193	0.484	
AGT rs699	MAF	200	0.490	
GNB3 rs5443	MAF	246	0.329	
ITGB3 rs5918	MAF	188	0.149	
CYP19 A1 rs10046	MAF	187	0.500	
CYP17 A1 rs743572	MAF	180	0.333	
ADRB1 rs1801253	MAF	186	0.293	
ADRB2 rs1042713	MAF	187	0.398	
ADRB2 rs1042714	MAF	186	0.309	
ADRB3 rs4994	MAF	186	0.056	
ELAC2 rs34152967	MAF	218	0.062	
IL10 rs1800896	MAF	136	0.460	
APOE rs429358	MAF	166	0.069	
APOE rs7412	MAF	163	0.046	
PLIN1 rs894160	MAF	113	0.376	

Absolute frequency (%), mean (\pm standard deviation), median (minimum-maximum), and minor allele frequency (MAF) are reported for dichotomous, normal distributed, non-normal distributed variables, and genotypes, respectively

categories. In some cases, we observed potential relationships between different categories. The largest cluster that included multiple cardio-metabolic diseases (i.e., CVD, T2D, HA-stroke, IGT-IR-hyperglycemia, T1D, and thrombosis) also comprised two “dietary habits” variables: soybean intake and vitamin B12 intake. Soybean intake is positively correlated with T1 D (Spearman’s rho = 0.46) and negatively correlated with the other cardio-metabolic diseases present in the cluster (Spearman’s rho: -0.57 [CVD], -0.61 [T2 D], 0.76 [HA-stroke], -0.40 [IGT-IR-hyperglycemia], and -0.53 [thrombosis]). The opposite correlations are observed for vitamin B12 intake: Negative correlation with T1 D (Spearman’s rho = -0.66) and positive correlations with the other cardio-metabolic diseases (Spearman’s rho: 0.50 [CVD], 0.53 [T2 D], 0.56 [HA-stroke], 0.63 [IGT-IR-hyperglycemia], and 0.68 [thrombosis]). The cluster with HOMA-IR (homeostatic model assessment–insulin resistance) and fasting insulin (FI) also included the anthropometric traits BMI and BFR, which are positively correlated with FI and HOMA-IR values (BMI Spearman’s rho: 0.55 [FI] and 0.51 [HOMA-IR], BFR Spearman’s rho: 0.53 [FI] and 0.52 [HOMA-IR]).

We tested the associations of genetic polymorphisms screened with CMTs (Supplemental Table 2). On the basis of

these genetic associations, we performed hierarchical clustering analysis for the phenotypic traits and genetic loci, to group these parameters in accordance with their similarities (i.e., phenotypic traits clustered by their genetic similarity and genetic loci clustered by their phenotype-association similarity). Figure 2 reports clustering trees for phenotypic traits (Fig. 2A) and genetic loci (Fig. 2B) and the heatmap of the genetic associations (Fig. 2C). Through a hierarchical clustering analysis of genetic loci, we identified three significant clusters: 1) a cluster that includes VDR variants (i.e., rs1544410 and rs731236); 2) a cluster related to PON1 rs662 and IL10 rs1800896; and 3) a third cluster related to CETP rs708272 and F2 rs1799963. The VDR cluster is due to the high linkage disequilibrium (LD) between VDR variants, whereas the other gene clusters may be due to biological mechanisms shared by the included genes. No high LD was observed among the other variants investigated. The PANTHER Classification System detected enriched molecular pathways for both of these gene clusters: the PON1-IL10 cluster for the interleukin signaling pathway ($P = 0.009$), and the CETP-F2 cluster for blood coagulation ($P = 0.005$) and CCKR signaling ($P = 0.016$).

Considering the hierarchical clustering analysis of CMTs based on their genetic similarity (Fig. 2A), we identified four

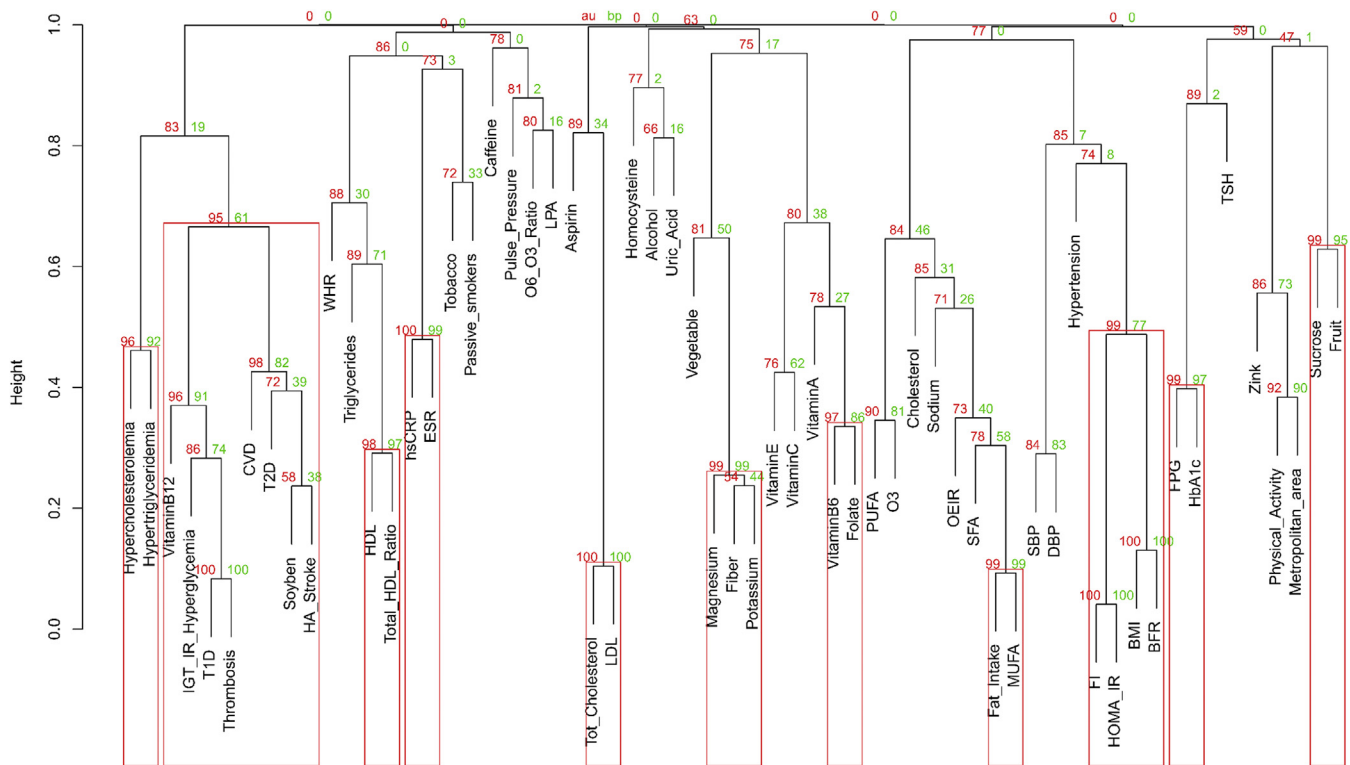


Fig. 1. Hierarchical clustering tree considering lifestyle, diet, personal information, and cardiovascular and metabolic traits. For each edge, it is reported the approximately unbiased P -value (left) and the bootstrap probability (right). Significant clusters (approximately unbiased P -value > 95%) are highlighted in red. The acronym descriptions are reported in Table 1.

significant clusters. A large cluster included the cardiometabolic diseases (i.e., CVD, T2D, HA-stroke, IGT-IR-hyperglycemia, T1D, thrombosis, hypercholesterolemia, and hyperglycemia). Three small clusters included: 1) HOMA-IR and FI; 2) BMI and BFR; and 3) total cholesterol and LDL. Comparing the hierarchical clustering analyses based on their genetic similarity and on their correlations, we observed that significant clusters overlapped. To understand whether the genetic similarity observed was due to causative relationships among the CMTs or shared genetic architecture, we performed a causative model analysis based on Bayesian networks. Specifically, for each significant phenotypic clusters generated by genetic-similarity analysis we selected the genetic variants associated with the CMTs ($P < 0.05$) and the non-genetic factors that clustered with the same CMTs in the first clustering analysis. Then, we tested the causative relationships between CMTs regard to genetic and non-genetic factors. On this basis, we determined the best causative model for: A) CVD, T2D, HA-stroke, IGT-IR-hyperglycemia, T1D, thrombosis, hypercholesterolemia, hyperglycemia, soybean intake, vitamin B12 intake, *IL10* rs1800896, *COL1 A1* rs1800012, *VDR* rs1544410, *ITGB3* rs5918, *APOC3* rs5128, *IL6* rs1800795, *F2* rs1799963, *F5* rs6025, and *MTR* rs1805087; B) BMI, BFR, *COL1 A1* rs1800012, and *NOS3* rs1799983; and C) total cholesterol, LDL, *MTRR* rs1801394, *LIPC* rs1800588, *COL1 A1* rs1800012, *APOE* rs7412, *CYP1 A1* rs4646903, *ITGB3* rs5918, *APOC3* rs5128, *IL6* rs1800795, and *CETP* rs708272 (Fig. 3).

Discussion

Since investigating multiple ethnic groups has been demonstrated to be an effective way to dissect the pathogenesis of complex traits [20], we performed an investigation on a Turkish

study population regarding the interactions among gene variations, lifestyle and dietary habits as they pertain to health conditions and blood parameters related to CMTs.

Initially, we focused on the relationships among CMTs (i.e., health conditions and blood parameters), nutritional (i.e., daily intakes of macro- and micronutrients), and other environmental factors (i.e., lifestyle and personal characteristics). In our bootstrap agglomerative hierarchical clustering, we observed 11 significant clusters that highlighted the complex relationship among CMTs, nutritional and other environmental factors. Most of them included parameters that belong to the same phenotypic categories. Regarding CMTs, several health conditions and blood parameters clustered together in agreement with current knowledge about the cooccurrence of cardiovascular and metabolic alterations [21]. We also identified clusters that included parameters related to dietary habits, a correlation which agrees with previous epidemiologic studies that demonstrated how different dietary habits tend to increase or reduce the intakes of healthy and unhealthy nutrients together [22]. Further, we identified a correlation between variables included in different phenotypic categories. Soybean and vitamin B12 intakes correlated with multiple cardiometabolic disorders, suggesting potential effects on the disease risk. Epidemiologic studies indicated that soybean intake reduced the risk of cardiovascular and metabolic diseases [23], and molecular experiments proved its antiinflammatory and anti-oxidant effects [24]. Accordingly, the observed correlation between soybean intake and cardiometabolic diseases reinforces the beneficial effects related to its consumption. Vitamin B12 is an important micronutrient and its deficiency is associated with hyperhomocysteinemia and comorbidities related to cardiovascular and metabolic diseases [25]. However, its role in the

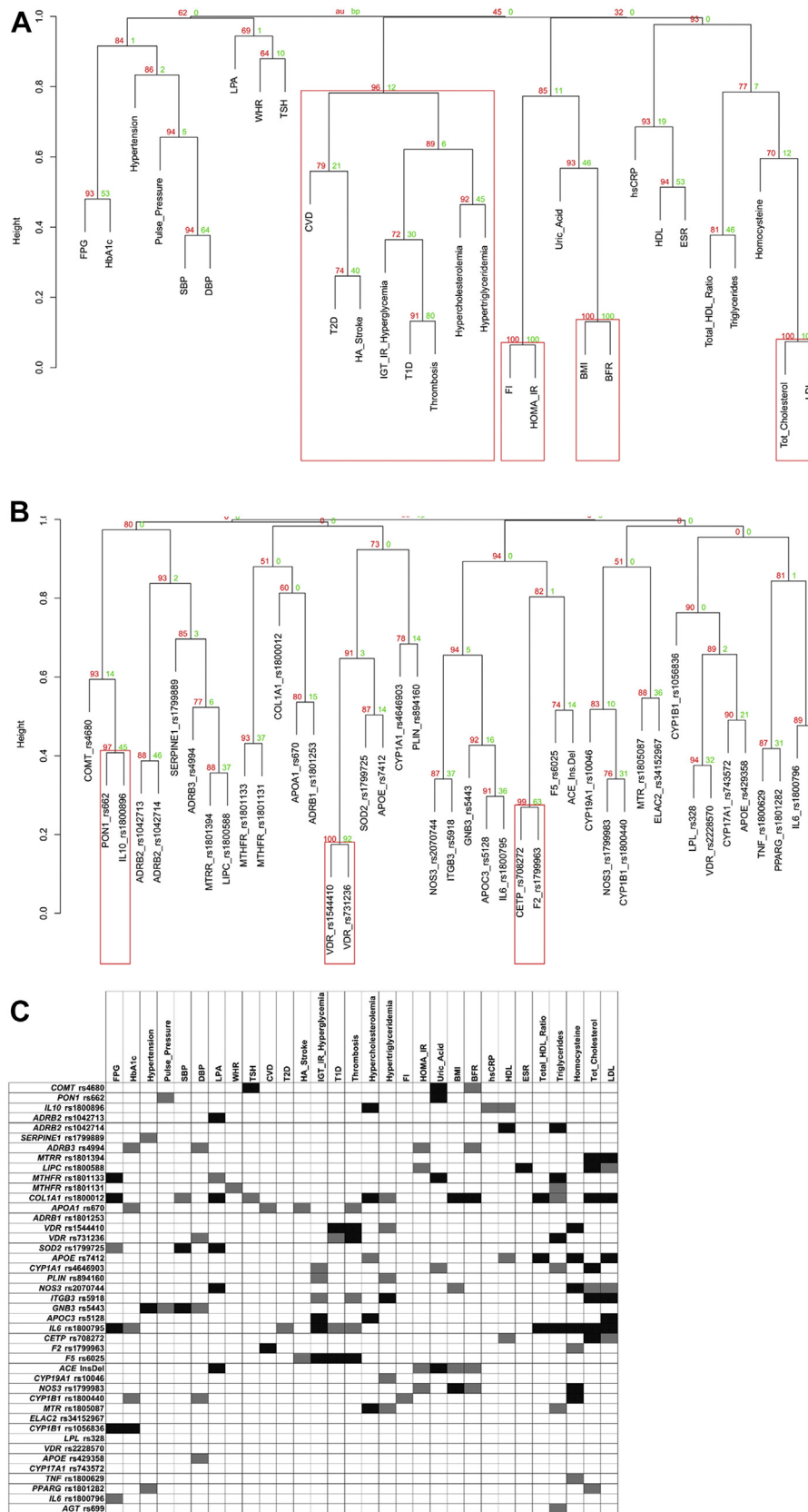


Fig. 2. Hierarchical clustering trees related cardiovascular and metabolic traits (A) and genetic loci (B) and the heatmap of their associations (C). In heatmap, black cells represent significant associations ($P < 0.05$) and gray cells represent trend associations ($P < 0.1$). The description of the hierarchical clustering trees is reported in Figure 1.

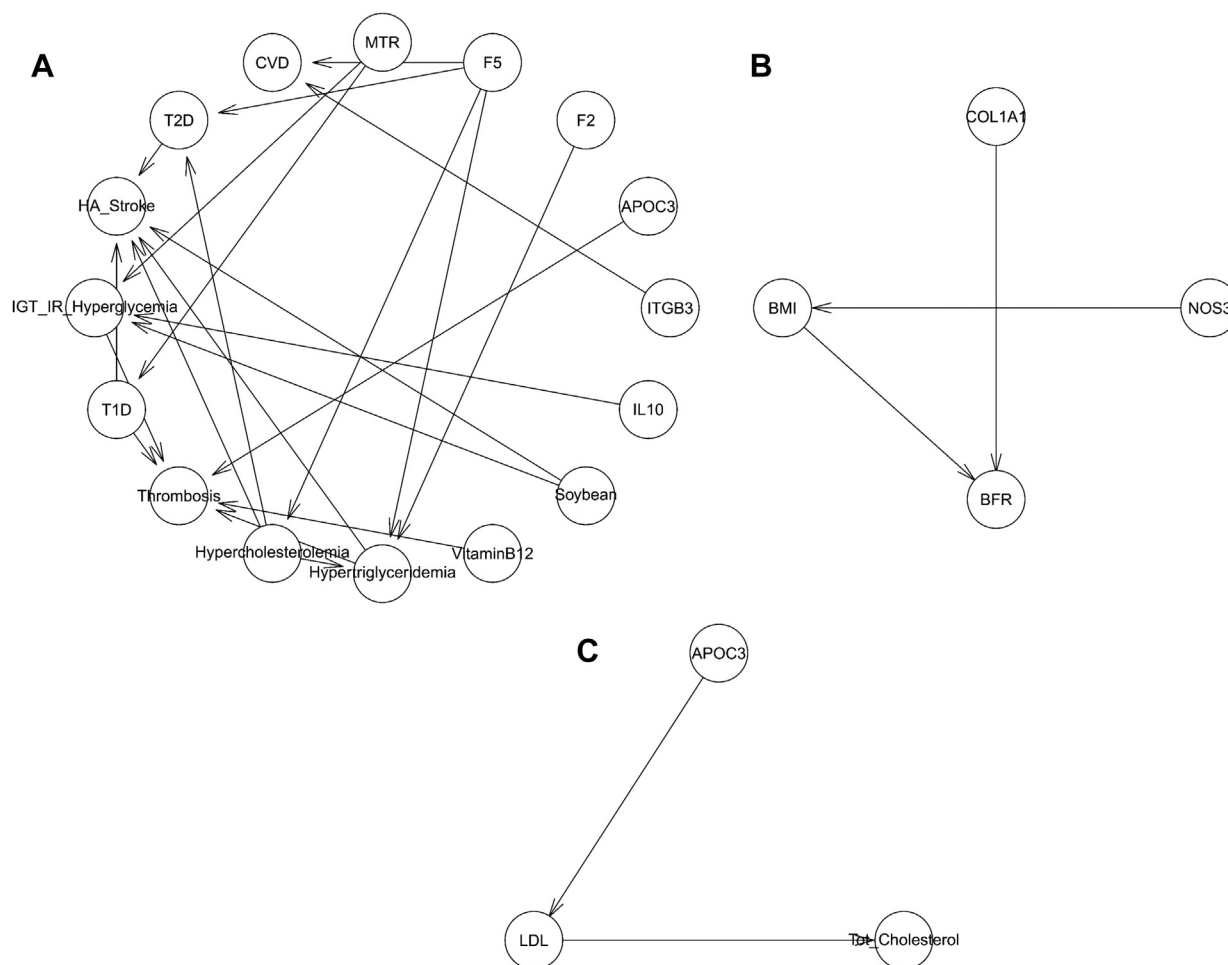


Fig. 3. Best causative models of the networks identified: (A) CVD, T2D, HA-stroke, IGT-IR-hyperglycemia, T1D, thrombosis, hypercholesterolemia, hyperglycemia, soybean intake, vitamin B12 intake, IL10 rs1800896, COL1 A1 rs1800012, VDR rs1544410, ITGB3 rs5918, APOC3 rs5128, IL6 rs1800795, F2 rs1799963, F5 rs6025, and MTR rs1805087; (B) BMI, BFR, COL1 A1 rs1800012, and NOS3 rs1799983; (C) total cholesterol, LDL, MTR rs1801394, LIPC rs1800588, COL1 A1 rs1800012, APOE rs7412, CYP1 A1 rs4646903, ITGB3 rs5918, APOC3 rs5128, IL6 rs1800795, and CETP rs708272.

risk of CMTs in adults is still unclear [26]. The main sources of vitamin B12 are foods of ruminant origin, so dairy and meat products contribute to its intake [27]. Accordingly, vitamin B12 intake can also be a marker of the consumption of dairy and meat products, foods strongly associated with cardiometabolic risk. Thus, the correlation observed may be linked to an over-consumption of animal products.

Considering the correlations between CMTs and the genetic variants tested, we performed hierarchical clustering analyses among genetic loci (i.e., variants grouped in accordance with their association with phenotypic traits), and among the CMTs (i.e., the CMTs are clustered in accordance with the associations with genetic variants). Analyzing the genetic loci, we identified a significant cluster due to the LD between VDR variants and two potentially related molecular pathways. The *PON1* and *IL10* genes are both involved in the interleukin signaling pathway and have been previously associated with cardiovascular and immunologic phenotypes [28,29]. The interleukin signaling pathway plays a relevant role in the cross talk between inflammatory processes and cardiometabolic dysregulation [30], supporting the phenotypic convergence observed between *PON1* and *IL10* in relation to CMTs. The *CETP* and *F2* genes showed enrichments for the blood coagulation and CCKR signaling pathways. Previous studies reported the involvement of both these genes in

coagulation, homocysteine, and lipoprotein metabolism [31]. Furthermore, their involvement in CCKR signaling also indicates a potential association in digestion, appetite control and body weight regulation [32]. These data seem to support the phenotypic convergence observed between *CETP* and *F2*.

Considering the hierarchical clustering of phenotypic traits on the basis of their genetic similarity, we noticed significant clusters for: 1) CVD, T2D, HA-stroke, IGT-IR-hyperglycemia, T1D, thrombosis, hypercholesterolemia, and hyperglycemia; 2) BMI and BFR; 3) total and LDL cholesterol; and 4) HOMA-IR and FI. The significant clusters were also observed by the hierarchical clustering based on their correlations, confirming that they are not independent traits. However, previous studies have demonstrated that they have different genetic architectures [33–35]. Accordingly, we used a Bayesian network approach to verify whether the genetic similarity observed among CMTs is due to causative relationships or shared genetic architecture. In Figure 3A, we reported the best causative model among CVD, T2D, HA-stroke, IGT-IR-hyperglycemia, T1D, thrombosis, hypercholesterolemia, hyperglycemia, soybean intake, vitamin B12 intake, *IL10* rs1800896, *COL1 A1* rs1800012, *VDR* rs1544410, *ITGB3* rs5918, *APOC3* rs5128, *IL6* rs1800795, *F2* rs1799963, *F5* rs6025, and *MTR* rs180508. We observed an intricate causative network that included multiple causative relationships among

CMTs; CMTs and dietary habits (i.e., soybean and vitamin B12 intake); and CMTs and genetic variants. This causative network agrees with numerous epidemiologic and molecular studies that have demonstrated that cardiovascular and metabolic diseases are affected by shared pathogenic mechanisms (i.e., shared genetic and environmental factors) and causative relationships [36, 37]. Furthermore, *F5* rs6025 and *MTR* rs180508 have demonstrated pleiotropic effects on multiple CMTs. Previous studies reported that genetic pleiotropy (i.e., a single gene affecting multiple phenotypic traits) is common among cardiometabolic diseases, contributing to the shared pathogenesis of these traits [38,39]. Figure 3B, reports the best causative models for BMI, BFR, *COL1 A1* rs1800012, and *NOS3* rs1799983. We observed a causative relationship between BMI and BFR. Both these anthropometric measurements are related to body fat and their strong correlation is well known [40]. However, different genetic associations are present for these correlated traits. The *NOS3* rs1799983 genotype was associated with BMI, confirming its role in BMI variability [41]. The *COL1 A1* rs1800012 genotype was correlated with BFR. *COL1 A1* plays a relevant role in osteogenesis and recent studies have suggested its involvement in adipose tissue matrix remodeling [42,43]. In this second causative model, we identified two correlated traits with different genetic factors (i.e., no pleiotropic effects). In Figure 3C, we reported the best causative models among total cholesterol, LDL, *MTRR* rs1801394, *LIPC* rs1800588, *COL1 A1* rs1800012, *APOE* rs7412, *CYP1 A1* rs4646903, *ITGB3* rs5918, *APOC3* rs5128, *IL6* rs1800795, and *CETP* rs708272. The Bayesian network analysis confirmed the association between *APOC3* rs5128 and LDL and the causative relationship between LDL and total cholesterol. *APOC3* encodes a very low-density lipoprotein, and rs5128 variant is associated with lipid profile changes [44]. Our data indicated that *APOC3* rs5128 directly affects in LDL levels and indirectly other lipid levels. No pleiotropic effects are observed in this causative network.

Conclusions

This study reported on what is to our knowledge the first comprehensive analysis of dietary habits in relation to CMTs in the Turkish population. We collected about 100 characteristics for each participant. This raised two issues: reduced samples size and increased number of tests. Accordingly, our study population is not large enough to enable us to detect significant pairwise correlations after multiple testing correction. However, our aim was not to detect pairwise correlations among the characteristics tested, but to determine causative networks that integrate dietary habits, lifestyle, and genetic information. Our main results were obtained applying hierarchical clustering and Bayesian network analyses to integrate data regarding dietary habits, lifestyle, and genetic predisposition. These investigations provide novel details about the complex interactive processes at the basis of the pathogenesis of cardiometabolic diseases. A better understanding of the complex pathogenic mechanisms of cardiometabolic disorder can have relevant translational implications since it can help to develop novel therapeutic and preventive approaches. Further studies with larger sample size and genome-wide data could fully exploit our approach and shed further light on the complex causative relationships of CMTs.

Supplementary data

Supplementary data related to this article can be found at <http://dx.doi.org/10.1016/j.nut.2015.12.027>.

References

- Wang J, Ruotsalainen S, Moilanen L, Lepisto P, Laakso M, Kuusisto J. The metabolic syndrome predicts cardiovascular mortality: A 13-year follow-up study in elderly non-diabetic Finns. *Eur Heart J* 2007;28:857–64.
- Blackmore HL, Ozanne SE. Programming of cardiovascular disease across the life-course. *J Mol Cell Cardiol* 2015;83:122–30.
- Pounis G, Costanzo S, di Giuseppe R, de Lucia F, Santimone I, Sciarretta A, et al. Consumption of healthy foods at different content of antioxidant vitamins and phytochemicals and metabolic risk factors for cardiovascular disease in men and women of the Moli-sani study. *Eur J Clin Nutr* 2013;67:207–13.
- Baylin A, Deka R, Tuitele J, Viali S, Weeks DE, McGarvey ST. *INSIG2* variants, dietary patterns and metabolic risk in Samoa. *Eur J Clin Nutr* 2013;67:101–7.
- van der Klaauw AA, Farooqi IS. The hunger genes: Pathways to obesity. *Cell* 2015;161:119–32.
- Li P, Fu Y, Ru J, Huang C, Du J, Zheng C, et al. Insights from systems pharmacology into cardiovascular drug discovery and therapy. *BMC Syst Biol* 2014;8:141.
- Sharma A, Gulbahce N, Pevzner SJ, Menche J, Ladenvall C, Folkersen L, et al. Network-based analysis of genome wide association data provides novel candidate genes for lipid and lipoprotein traits. *Mol Cell Proteomics* 2013;12:3398–408.
- Keser A, Yabancı Ayhan N, Bilgic P, Tayfur M, Şimşek I. Determination of dietary status as a risk factor of cardiovascular heart disease in Turkish elderly people. *Ecol Food Nutr* 2015;54:328–41.
- White MJ, Eren F, Agirbasli D, Williams SM, Agirbasli M. *SHBG* gene polymorphism (rs1799941) associates with metabolic syndrome in children and adolescents. *PLoS One* 2015;10:E0116915.
- Alkan C, Kavak P, Somel M, Gokcumen O, Ugurlu S, Saygi C, et al. Whole genome sequencing of Turkish genomes reveals functional private alleles and impact of genetic interactions with Europe, Asia and Africa. *BMC Genomics* 2014;15:963.
- Karaca S, Cesuroglu T, Karaca M, Erge S, Polimanti R. Genetic diversity of disease-associated loci in Turkish population. *J Hum Genet* 2015;60:193–8.
- Karaca S, Bozkurt NC, Cesuroglu T, Karaca M, Bozkurt M, Eskioğlu E, et al. International warfarin genotype-guided dosing algorithms in the Turkish population and their preventive effects on major and life-threatening hemorrhagic events. *Pharmacogenomics* 2015;16:1109–18.
- Cesuroglu T, Karaca S, Erge S. A practice model for personalized healthcare with a public health genomics perspective. *Personalized Medicine* 2009;6:567–77.
- Erhardt J. *BeBis: Nutrition Data Base Software*. Istanbul; 2004.
- Karaca S, Karaca M, Cesuroglu T, Erge S, Polimanti R. *GSTM1*, *GSTP1*, and *GSTT1* genetic variability in Turkish and worldwide populations. *Am J Human Biol* 2015;27:310–6.
- Suzuki R, Shimodaira H. *Pvclust: An R package for assessing the uncertainty in hierarchical clustering*. *Bioinformatics* 2006;22:1540–2.
- Mi H, Muruganujan A, Casagrande JT, Thomas PD. Large-scale gene function analysis with the PANTHER classification system. *Nat Protoc* 2013;8:1551–66.
- Scutari M. Learning Bayesian Networks with the bnlearn R Package. *Journal of Statistical Software* 2010;35:1–22.
- Satman I, Omer B, Tutuncu Y, Kalaca S, Gedik S, Dincceg N, et al. Twelve-year trends in the prevalence and risk factors of diabetes and prediabetes in Turkish adults. *Eur J Epidemiol* 2013;28:169–80.
- Hall MA, Verma A, Brown-Gentry KD, Goodloe R, Boston J, Wilson S, et al. Detection of pleiotropy through a Phenome-wide association study (PheWAS) of epidemiologic data as part of the Environmental Architecture for Genes Linked to Environment (EAGLE) study. *PLoS Genet* 2014;10:E1004678.
- Li J, Flammer AJ, Lennon RJ, Nelson RE, Gulati R, Friedman PA, et al. Comparison of the effect of the metabolic syndrome and multiple traditional cardiovascular risk factors on vascular function. *Mayo Clin Proc* 2012;87:968–75.
- Pounis G, de Lorgeril M, Salen P, Laporte F, Krogh V, Siani A, et al. Dietary patterns and fatty acids levels of three European populations. Results from the IMMIDIET study. *Nutr Metab Cardiovasc Dis* 2014;24:883–90.
- Liu J, Sun LL, He LP, Ling WH, Liu ZM, Chen YM. Soy food consumption, cardiometabolic alterations and carotid intima-media thickness in Chinese adults. *Nutr Metab Cardiovasc Dis* 2014;24:1097–104.
- Chakrabarti S, Jahandideh F, Wu J. Food-derived bioactive peptides on inflammation and oxidative stress. *Biomed Res Int* 2014;2014:608979.
- Debrenceni B, Debrenceni L. The role of homocysteine-lowering B-vitamins in the primary prevention of cardiovascular disease. *Cardiovasc Ther* 2014;32:130–8.
- Pawlak R. Is vitamin B12 deficiency a risk factor for cardiovascular disease in vegetarians? *Am J Prev Med* 2015;48:E11–26.
- Gille D, Schmid A. Vitamin B12 in meat and dairy products. *Nutr Rev* 2015;73:106–15.
- Liu C, Battilwalla F, Li W, Lee A, Roubenoff R, Beckman E, et al. Genome-wide association scan identifies candidate polymorphisms associated with differential response to anti-TNF treatment in rheumatoid arthritis. *Mol Med* 2008;14:575–81.
- Kutikhin AG, Yuzhalin AE, Brusina EB, Ponasenko AV, Golovkin AS, Barbarash OL. Genetic predisposition to calcific aortic stenosis and mitral annular calcification. *Mol Biol Rep* 2014;41:5645–63.

- [30] Palomer X, Salvado L, Barroso E, Vazquez-Carrera M. An overview of the crosstalk between inflammatory processes and metabolic dysregulation during diabetic cardiomyopathy. *Int J Cardiol* 2013;168:3160–72.
- [31] Stankovic S, Majkic-Singh N. Genetic aspects of ischemic stroke: Coagulation, homocysteine, and lipoprotein metabolism as potential risk factors. *Crit Rev Clin Lab Sci* 2010;47:72–123.
- [32] Tripathi S, Flobak A, Chawla K, Baudot A, Bruland T, Thommesen L, et al. The gastrin and cholecystokinin receptors mediated signaling network: A scaffold for data analysis and new hypotheses on regulatory mechanisms. *BMC Syst Biol* 2015;9:40.
- [33] Shungin D, Winkler TW, Croteau-Chonka DC, Ferreira T, Locke AE, Magi R, et al. New genetic loci link adipose and insulin biology to body fat distribution. *Nature* 2015;518:187–96.
- [34] Shin SY, Fauman EB, Petersen AK, Krumsiek J, Santos R, Huang J, et al. An atlas of genetic influences on human blood metabolites. *Nat Genet* 2014;46:543–50.
- [35] Dupuis J, Langenberg C, Prokopenko I, Saxena R, Soranzo N, Jackson AU, et al. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat Genet* 2010;42:105–16.
- [36] Chan KH, Huang YT, Meng Q, Wu C, Reiner A, Sobel EM, et al. Shared molecular pathways and gene networks for cardiovascular disease and type 2 diabetes mellitus in women across diverse ethnicities. *Circ Cardiovasc Genet* 2014;7:911–9.
- [37] Holliday EG, Traylor M, Malik R, Bevan S, Falcone G, Hopewell JC, et al. Genetic overlap between diagnostic subtypes of ischemic stroke. *Stroke* 2015;46:615–9.
- [38] Gottesman O, Drill E, Lotay V, Bottinger E, Peter I. Can genetic pleiotropy replicate common clinical constellations of cardiovascular disease and risk? *PLoS One* 2012;7:E46419.
- [39] Kraja AT, Chasman DI, North KE, Reiner AP, Yanek LR, Kilpelainen TO, et al. Pleiotropic genes for metabolic syndrome and inflammation. *Mol Genet Metab* 2014;112:317–38.
- [40] Naruse R, Inukai Y, Terasawa T, Hara K, Takebayashi K, Morita M, et al. Relationship of body fat weight and body fat ratio determined by bioelectric impedance to serum adipocytokines in patients with type 2 diabetes mellitus. *Obes Res Clin Pract* 2011;5:E267–360.
- [41] Bressler J, Pankow JS, Coresh J, Boerwinkle E. Interaction between the NOS3 gene and obesity as a determinant of risk of type 2 diabetes: The Atherosclerosis Risk in Communities study. *PLoS One* 2013;8:E79466.
- [42] Lancha A, Rodriguez A, Catalan V, Becerril S, Sainz N, Ramirez B, et al. Osteopontin deletion prevents the development of obesity and hepatic steatosis via impaired adipose tissue matrix remodeling and reduced inflammation and fibrosis in adipose tissue and liver in mice. *PLoS One* 2014;9:E98398.
- [43] Gao B, Huang Q, Lin YS, Wei BY, Guo YS, Sun Z, et al. Dose-dependent effect of estrogen suppresses the osteo-adipogenic transdifferentiation of osteoblasts via canonical Wnt signaling pathway. *PLoS One* 2014;9:E99137.
- [44] Song Y, Zhu L, Richa M, Li P, Yang Y, Li S. Associations of the APOC3 rs5128 polymorphism with plasma APOC3 and lipid levels: A meta-analysis. *Lipids Health Dis* 2015;14:32.