Journal Pre-proofs

Research paper

ALDH2, *ADCY3* and *BCMO1* polymorphisms and lifestyle-induced traits are jointly associated with CAD risk in Chinese Han people

Cheng-yin Ye, Jia-rui Xin, Zheng Li, Xiao-yu Yin, Shu-li Guo, Jin-Mei Li, Tian-Yu Zhao, Li Wang, Lei Yang

PII: S0378-1119(21)00543-6
DOI: https://doi.org/10.1016/j.gene.2021.145948
Reference: GENE 145948

To appear in: *Gene Gene*

Received Date: 6 May 2021
Revised Date: 29 August 2021
Accepted Date: 30 August 2021

Please cite this article as: C-y. Ye, J-r. Xin, Z. Li, X-y. Yin, S-l. Guo, J-M. Li, T-Y. Zhao, L. Wang, L. Yang, *ALDH2*, *ADCY3* and *BCMO1* polymorphisms and lifestyle-induced traits are jointly associated with CAD risk in Chinese Han people, *Gene Gene* (2021), doi: https://doi.org/10.1016/j.gene.2021.145948

ALDH2, *ADCY3* and *BCMO1* polymorphisms and lifestyle-induced traits are jointly associated with CAD risk in Chinese Han people

Authors

Cheng-yin Ye [1, †], Jia-rui Xin[1, †], Zheng Li [2], Xiao-yu Yin[1], Shu-li Guo[1], Jin-Mei Li[1], Tian-Yu Zhao[1], Li Wang[1], Lei Yang[1, *].

1. School of Public Health, Hangzhou Normal University, Hangzhou 310000, China.

2.Wu Yun Shan Hospital, Hangzhou 31000, China.

[*] **Corresponding author**

Lei Yang, PhD, Professor

School of Public Health, Hangzhou Normal University, Yuhangtang Rd, Hangzhou, Zhejiang, China.

Email: yanglei62@hznu.edu.cn

[†] These authors contributed equally to this work.

Emails:

Cheng-yin Ye: yechengyin@hznu.edu.cn

Jia-Rui Xin: xinjiarui@stu.hznu.edu.cn

Zheng Li: lizheng@stu.hznu.edu.cn

Xiao-Yu Yin: yinxiaoyu@stu.hznu.edu.cn

Shu-li Guo: 2020011012014@stu.hznu.edu.cn

Jin-Mei Li: lijinmei@stu.hznu.edu.cn

Tian-Yu Zhao: zhaotianyu0225@foxmail.com

Li Wang: wangli0522@stu.hznu.edu.cn

Abstract

Backgrounds: To investigate associations of genetic factors and environmental factors with coronary artery disease (CAD), we collected medical reports, lifestyle details, and blood samples of 2113 individuals, and then used the polymerase chain reaction (PCR)-ligase detection reaction (LDR) to genotype the targeted 102 SNPs. **Methods:** We adopted elastic net algorithm to build an association model that considered simultaneously genetic and lifestyle/clinical factors associated with CAD in Chinese Han population. **Results:** In this study, we developed an all covariates-based model to explain the risk of CAD, which incorporated 8 lifestyle/clinical factors and a gene-score variable calculated from 3 significant SNPs (rs671, rs6751537 and rs11641677), attaining an area under the curve (AUC) value of 0.71. It is found that, in terms of genetic variants, the AA genotype of rs671 in the additive (adjusted odds ratio (OR)=2.51, p=0.008) and recessive (adjusted OR=2.12, p=0.021) models, the GG genotype of rs6751537 in the additive (adjusted OR=3.36, p=0.001) and recessive (adjusted OR=3.47, p=0.001) models, and GG genotype of rs11641677 in additive model (adjusted OR=0.39, p=0.044) was associated with increased risk of CAD. In terms of lifestyle/clinical factors, the history of hypertension (unadjusted OR=2.37, p<0.001) and dyslipidemia (unadjusted OR=1.82, p=0.007), age (unadjusted OR=1.07, p<0.001) and waist circumference (unadjusted OR=1.02, p=0.05) would significantly increase the risk of CAD, while height (unadjusted OR=0.97, p=0.006) and regular intake of chicken (unadjusted OR=0.78, p=0.008) reduced the risk of CAD. A significant interaction was found between rs671 and dyslipidemia (the relative excess risk due to interaction (RERI) = 3.36, p=0.05). **Conclusion:** In this study, we constructed an association mode and identified a set of SNPs and lifestyle/clinical risk factors of CAD in Chinese Han population. By considering both genetic and non-genetic risk factors, the built model may provide implications for CAD pathogenesis and clues for screening tool development in Chinese Han population.

Keywords: CAD, genetic, lifestyle/clinical features, elastic net

Introduction

In recent years, along with the trend of longevity and the improvement in health, cardiovascular disease, including stroke and coronary heart disease (CAD), is considered one of the leading causes of death. It is estimated that more than 7 million people die from the disease each year[1, 2]. Recent evidence from the epidemiologic literatures suggests that the mortality due to CAD is as high as 11.1 million worldwide, of which 1.3 million are in China[3, 4].

The biological and epidemiological evidence suggest that CAD is a complex disorder resulting from the interplay of genetic and non-genetic factors: (1) genetic variation; (2) physiological factors, such as hypertension, diabetes, obesity and so on; (3) smoking, drinking, diet, physical activity and other unhealthy lifestyle behaviors; (4) environmental pollution exposures; (5) interaction between various factors[5-9].

Genetic factors contribute significantly to the risk of CAD. In 2007, chromosome 9p21 was discovered and replicated as the first genetic risk variant of CAD[10, 11]. Since then, a growing number of large-scale and population-specific studies have been conducted to identify CAD-related susceptibility loci. Until recently, by using the trans-ancestry meta-analysis, studies were continually reporting results of 35 newly discovered susceptibility loci associated with CAD risk[12]. In the past two decades, more than 200 CAD susceptibility loci have been identified[13]. These risk

variants were found to affect CAD risk via various pathogeneses and mechanisms[13, 14], such as lipid metabolism (e.g., *PCSK9*[15], *LPA*[16], *APOB*[17], *APOE*[18] and *BCMO1*[19, 20]), alcohol metabolism (e.g., *ALDH2*[21, 22]), adipose tissue development (e.g., *ADCY3*[23, 24]), blood pressure (e.g., *SH2B3*[25]), etc. For instance, *PCSK9* coding protein, proprotein convertase subtilisin/kexin type 9, could affects low-density lipoprotein cholesterol (LDL-C) receptors while *PCSK9* inhibitors could be used to reduce plasma LDL-C levels and influence the incidence of CAD[15]. Aldehyde dehydrogenase 2 (coded by *ALDH2*) is an important oxidase involved in alcohol metabolism in cell mitochondria, which has a protective effect on cell damage caused by oxidative stimulation. It can also delay the process of cardiovascular diseases by reducing the occurrence of vascular endothelial inflammation[21, 22]. It is believed that, these discovered genetic risk variants can be further used to create genetic risk scores, thereby improve risk prediction beyond traditional risk factors[12].

In addition to genetic factors, many other risk factors are considered important to the incident of cardiovascular diseases, such as diabetes, obesity, physical activity, diet, smoking, and alcohol consumption[5-8]. Previous studies have shown that, the leading risk factor for cardiovascular mortality in both men and women in China and other East Asia area is hypertension[22].

Therefore, in this study, by utilizing the elastic net algorithm, we sought to build an association model of CAD that considering both genetic and clinical/lifestyle-related factors simultaneously. By doing this, we expected to identify impactful biomarkers associated with CAD risk before incidence of the disease. And then, with the identified genetic and non-genetic factors, we also plan to comprehensively explore possible interactions between these factors to assess their high-dimension impact on CAD risk, so as to provide insights into the pathogenesis of CAD.

Methods

Subjects

Cluster random sampling was used to recruit 2,323 non-CAD subjects who underwent physical examination at the community health service centers in 4 townships under the jurisdiction of a certain district of Ningbo, Zhejiang Province from April 2013 to July 2013. All subjects were unrelated and over 40 years of age. Patients diagnosed with CAD before April 2013, as well as patients with severe liver and kidney disease and malignant tumors were excluded. We collected blood samples, medical records and lifestyle details for all individuals, then used the polymerase chain reaction (PCR)-ligase detection reaction (LDR) to genotype the target single nucleotide polymorphisms (SNPs). The field epidemiological investigation mainly included basic demographic criteria and a standard questionnaire for lifestyle exposures, which was described in details in previous studies[20]. The incidence of coronary atherosclerosis was investigated in August 2016. Due to various factors such as moving house, death and so on, 10.04% of the subjects were lost to follow-up after 3 years in August 2016, leaving a total of 2113 people as our study subjects. The study design is shown in Figure 1.

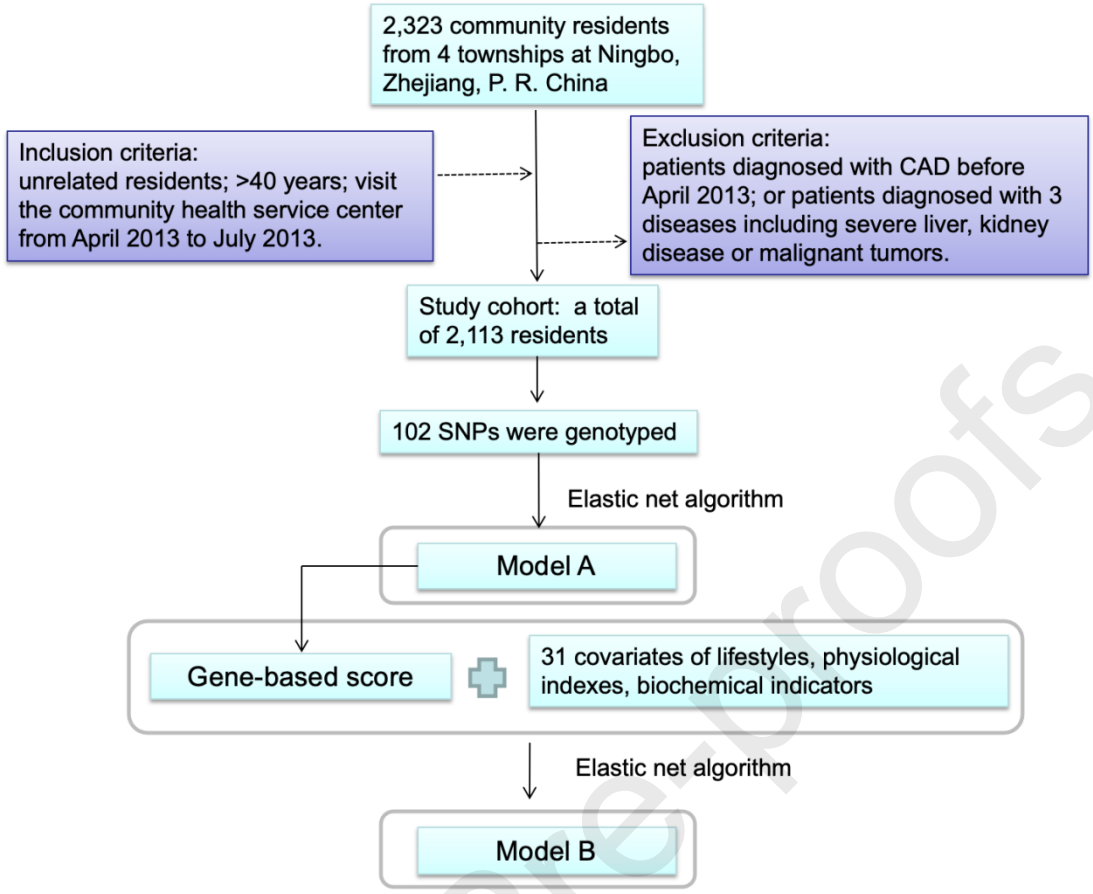

Figure 1. Study design

Ethical approval statement

This study was approved by the Medical Ethics Committee of Hangzhou Normal University (No.2013020). This study was exempted from the requirement for individual informed consent because all data were anonymized during the entire study. All methods were performed in accordance with the relevant guidelines and regulations.

Diagnostic criteria

CAD is usually diagnosed by electrocardiogram, electrocardiogram load test, dynamic electrocardiogram, ultrasonic electrocardiogram, hematologic examination, coronary computed tomography (CT), coronary angiography or intravascular imaging. In our study, the standard Judkins technique was used to performing coronary angiography, the diagnostic criteria for coronary atherosclerotic heart disease are based on the China Health and Family Planning Commission in 2010[26]. CAD was defined as more than 1 (≥) atherosclerotic plaque in a major coronary artery (≥1.5 mm lumen diameter) causing ≥50% luminal diameter stenosis by the quantitative coronary angiography (QCA) test. Asymptomatic coronary heart disease, angina pectoris or ischemic cardiomyopathy, coronary artery ischemia, myocardial infarction, all belong to coronary heart disease.

SNP selection and genotyping

The process of SNP selection is as follows: firstly, we searched literatures from large

bibliographic databases (e.g., PubMed, MEDLINE, EMBASE) to collect related polymorphisms of coronary atherosclerosis or its relevant conditions (e.g., lipid levels, hypertension, diabetes and obesity); then, we selected GeneView information, missense mutation, 3' un-translation region (UTR), 5' UTR or transcription factor binding sites of the selected SNPs from GeneCards database and NCBI database; after that, the minor allele frequency (MAF) of SNPs in Chinese population was detected from the HapMap database of the international human genome, and the SNPs with MAF value greater than 0.05 were retained. Finally, 102 SNPs located in genes such as *ALDH2*, *ABCA1*, *ALDH2*, *BCMO1*, *SLC12A3*, *PCSK9* etc., were selected for subsequent genotyping process. All genotyped SNPs were summarized in Table S1.

The isolation of genomic DNA is described in detail in previous studies[27]. All blood samples were collected by venipuncture from fasted participants, anticoagulated with EDTA and preserved at -80℃. The Tiangen Blood Genomic DNA extraction kits was used to extract DNA while the PCR-LDR reactions were adopted for SNP genotyping. Specifically, the PCR reactions were conducted in an ABI Prism 7000 Sequence Detection System with an initial melting at 94 °C for 3 min, 35 cycles of denaturation at 94 °C for 15 s, annealing at 55 °C for 15 s, extension at 72 °C for 30 s, and final extension at 72 °C for 3 min. Each reaction consisted 1 μL genomic DNA, 1.5 μL $MgCl_2$, 1.5 μL 10× PCR buffer, 0.15 μL each primer, 0.3 μL dNTPs, and 0.2 μL Taq DNA polymerase in a total volume of 15 μl. The LDR reactions were performed in 30 cycles at 94 °C for 30 s and 56 °C for 3 min. Each reaction contained 3 μL PCR product, 1 μL 10×Taq DNA ligase buffer, 5 U Taq DNA ligase, and 0.01 μL each discriminating probe in a total volume of 10 μL. Re-sequencing results for 10% of the samples showed that the concordance rates were >95% for all target SNPs.

Statistical analysis

Statistical analysis was conducted with SPSS 24.0 software (SPSS, Inc., Chicago, IL, USA) and RStudio (Version 1.1.456. RStudio: Integrated development environment for R. Boston, MA, USA; http://www.rstudio.org/). The t-test (for continuous variable) and chi-square test (for categorical variable) were used to evaluated the association between demographic characteristics or SNPs and CAD individually. Elastic net regularization, characterized by reducing over fitting and co-variate correlation, was implemented from R package *glmnet* for feature selection and multivariate CAD-association model construction[26]. When the set of features are large and may have potential high-dimensional interaction, the technique has been shown to be superior to other analysis methods[28]. To investigate the CAD risk from both genetic and clinical/lifestyle dimensions successively, we divided our modeling process into two phases. Initially, by adopting the elastic net algorithm, we derived an association model based on 102-SNP features, where the best choice of parameter was chosen according to the classification accuracy of the fitted elastic net model. Secondly, we calculated an elastic-net-driven gene score for each individual, indicating their genetic risk of CAD. Following that, by further applying elastic net algorithm, we combined the derived gene score with 31 lifestyles/clinical covariates to construct a more comprehensive CAD association model. Finally, receiver operating characteristic (ROC) curves were plotted to assess the accuracy of both genetic and full models. The impacts of identified SNPs on the risk of CAD were further evaluated by odd ratios (ORs) and 95% confidence intervals (CIs) in terms of each genetic mode (i.e., additive, recessive and dominant modes) using logistic regression analysis.

Crossover analysis[29] was adopted to explore the marginal effects and pairwise interactions among variables. In details, for a certain pair of dichotomous variables, *A* and *B*, after the 4×2 contingency table is generated (see Table S2), the number of cases or controls is counted in each cell for a particular combination of the two variables, and then ORs are calculated using logistic regression to reflect the relative risk of this combination compared to the reference group (with no exposure to both *A* and *B*). Subsequently, the interaction index, the relative excess risk due to interaction (RERI), are calculated, which represents the difference between the combined effect of the two factors and the sum of their individual effects, and is often considered as the standard measure for interaction on the additive scale with case–control studies[30]. In particular, for a certain pair of dichotomous variables, *A* and *B*, RERI is calculated from different ORs as $RERI=OR_{(A+,\ B+)}-(OR_{(A+,\ B-)}+OR_{(A-,\ B+)}-1)$, where *A*+ or *B*+ indicates an exposure to variable *A* or *B*, whereas *A*- or *B*- represents a non-exposure status. The corresponding CI and p-value of RERI are used to indicate whether this interaction effect is statistically significant compared to the potential effects in the reference group (with no exposure to both *A* and *B*) due to any variables other than the two considered factors, *A* and *B*[31]. The crossover analysis and RERI calculation were implemented by SPSS 24.0. In all analyses, p values<0.05 were considered significant statistically. The purpose of this study is to explain CAD incidence through a relatively meaningful model, such as which SNP or life behavior factors are more likely to associate to CAD incidence, not to establish a model with very good performance in predicting CAD.

Results

Clinical Characteristics

The incidence of CAD in the study population within three years was 0.47%, which is higher than that of general Chinese population (0.09% every year) reported by Chinese cardiovascular disease report in 2017[3]. This was mainly because all subjects were over 40 years old in our study cohort, which increased the risk of CAD. According to previous studies, age is one of the major risk factors for CAD[7]. The composition ratio of gender is as follows: 45.7% of the subjects were male, and 54.3% were female. The median of SBP in the case group was significantly higher than that in the control group (p<0.05). For other features, the χ^2 test found significant differences in hypertension and dyslipidemia history, current fruit and chicken intake between the case group and the control group (p<0.05) (Table 1).

Table 1. Basic characteristics

Characteristic	CAD (+)	CAD (-)	t/χ^{2a}	P
Total	100	2013		
Age, mean±SD[b], year	67.55±11.88	58.57±10.96	56.03	**<0.001**
Sex, n (%)[c]			0.46	0.500
Male	49(49.00)	917(45.50)		
Female	51(51.00)	1096(54.50)		
Height, mean±SD, year	158.08±8.27	160.45±8.30	7.67	**0.006**
Waist, mean±SD, cm	82.87±8.02	81.22±8.22	3.85	0.050
SBP, mean±SD, mmHg	140.66±16.16	133.90±19.24	11.79	**0.001**
DBP, mean±SD, mmHg	83.32±10.31	81.61±11.99	3.29	0.070
BMI, mean±SD, kg/m2	23.62±3.52	23.25±3.11	-1.06	0.291

TC, mean±SD, mmol/L	4.91±0.91	4.87±0.91	0.17	0.679
TG, mean±SD, mmol/L	1.47±0.66	1.44±0.69	0.26	0.608
HDL-C, mean±SD, mmol/L	1.26±0.29	1.30±0.31	2.13	0.145
LDL-C, mean±SD, mmol/L	3.16±0.87	3.08±0.80	1.17	0.280
Hypertension history, n (%)			17.22	**<0.001**
Yes	38(38.00)	414(20.57)		
No	62(62.00)	1599(79.43)		
Dyslipidemia history, n (%)				
Yes	70(70.00)	1131(56.18)	7.41	**0.006**
No	30(30.00)	882(43.82)		
Current fruit intake, n (%)			8.57	**0.032**
Never	38(38.00)	662(32.89)		
50-100g/d	59(59.00)	1111(55.19)		
≥100g/d	3(3.00)	240(11.92)		
Current chicken intake, n (%)			7.09	**0.008**
Never	40(40.00)	568(28.22)		
1-7 times/week	54(54.00)	1241(61.65)		
≥8times/week	6(6.00)	204(10.13)		

SD, standard deviation; SBP, systolic blood pressure; DBP, diastolic blood pressure; BMI, body mass index; TC, total cholesterol; TG, triglyceride; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol. Bold values are statistically significant with p values <0.05

[a] t/χ^2: t stands for the t-value calculated by t-test and $\chi 2$ stands for the chi-square value derived from the chi-square test;

[b] mean±SD: for continuous variables, mean and standard deviation were calculated within CAD(+) or CAD(-) subgroup.

[c] n (%): for categorical variables, n describes the number of individuals carrying this characteristic in CAD(+) or CAD(-) subgroup, and % is the constituent ratio of this characteristic within each subgroup.

The gene-based association

The variable selection of elastic net penalization is achieved by shrinking the coefficients of the variables not related to the response to zero. Thus, variables with non-zero coefficients are considered as important predictors. With the initial 102 SNP inputs, the gene-based association model (model A) finally recruited 3 SNPs with nonzero coefficients in the elastic net model. The gene scores were then calculated from the elastic net regression, where each of the 3 SNPs were weighted by their coefficients, respectively. The 3 associated SNPs were rs671 in *ALDH2*, and rs6751537 in *ADCY3*, rs11641677 in *BCMO1*. The area under the curve (AUC) value of model A was 0.59 (95% CI: 0.53-0.65) (Figure 2).

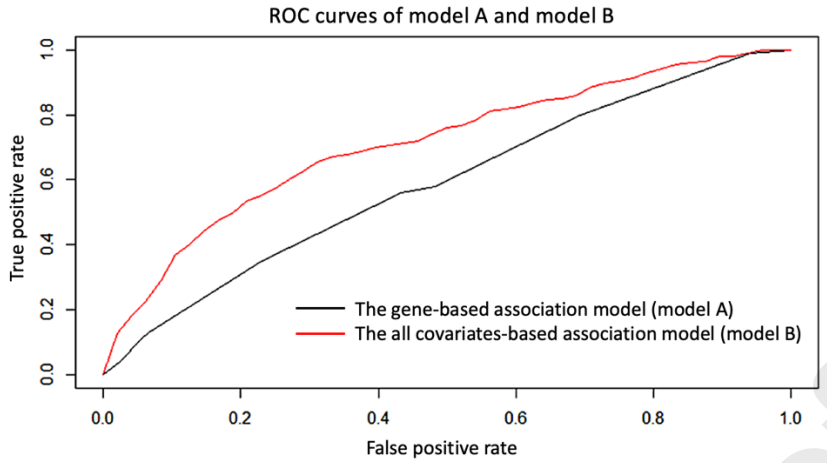

Figure 2. ROC curves of the derived gene-based association model (model A) and the subsequently constructed all covariates-based association model (model B), which attained AUCs of 0.59 and 0.71 respectively.

The association between the 3 SNPs and CAD was also examined separately under each genetic mode (Table 2). Without adjustment, the recessive models of all 3 SNPs were found to be significantly associated with CAD risk. In additive models, the AA genotype of rs671 (unadjusted OR=2.45, 95% CI=1.28-4.68, p=0.007) and the GG genotype of rs6751537 (unadjusted OR=3.63, 95% CI=1.83-7.20, p<0.001) increased the risk of CAD, while the GG genotype of rs11641677 reduced the risk of CAD (unadjusted OR=0.36, 95% CI=0.14-0.90, p =0.032). People carrying the AA genotype in terms of rs671 recessive model (unadjusted OR=2.13, 95% CI = 1.16-3.92, p=0.015), GG genotype in terms of rs6751537 recessive model (unadjusted OR=3.76, 95% CI=1.92-7.39, p<0.001) had a higher risk of CAD than each of the reference groups respectively. All of the studied SNPs in the control subjects were in Hardy–Weinberg equilibrium (p>0.05). After the adjustments of other 8 captured risk factors (i.e., age, height, SBP, waist, LDL-C, intake of chicken and history of hypertension and dyslipidemia selected by elastic net in the "**All covariates-based model**" section below), we further investigated each genetic mode for all 3 involving SNPs for their contribution to the risk of CAD. The results showed that, the recessive models of rs671 (adjusted OR=2.12, 95% CI=1.12-4.03, p=0.021) and rs6751537 (adjusted OR=3.47, 95% CI=1.70-7.09, p=0.001) were still significantly associated with CAD (Table 2). In additive models, the AA genotype of rs671 (adjusted OR=2.51, 95% CI=1.27-4.95, p=0.008) and the GG genotypes of rs6751537 (adjusted OR =3.36, 95% CI=1.62-6.94, p=0.001) still increased the risk of CAD (Table 2).

Table 2. Associations of genetic variants with CAD risk

SNP	Genotype	Unadjusted OR (95%CI)	Unadjusted *P*	Adjusted OR (95%CI)	Adjusted *P*
rs671					
additive	AG/GG	1.34(0.87-2.06)	0.186	1.41(0.91-2.20)	0.127
	AA/GG	2.45(1.28-4.68)	**0.007**	2.51(1.27-4.95)	**0.008**
dominant	AA+AG/GG	1.49(0.99-2.23)	0.055	1.56(1.03-2.38)	**0.037**
recessive	AA/AG+GG	2.13(1.16-3.92)	**0.015**	2.12(1.12-4.03)	**0.021**
rs6751537					

additive	AG/AA	0.87(0.53-1.42)	0.583	0.88(0.52-1.22)	0.623
	GG/AA	3.63(1.83-7.20)	**<0.001**	3.36(1.62-6.94)	**0.001**
dominant	GG+AG/AA	1.16(0.76-1.77)	0.500	1.16(0.75-1.79)	0.512
recessive	GG/AG+AA	3.76(1.92-7.39)	**<0.001**	3.47(1.70-7.09)	**0.001**
rs11641677					
additive	GA/AA	0.82(0.55-1.24)	0.355	0.78(0.51-1.19)	0.254
	GG/AA	0.36(0.14-0.90)	**0.032**	0.39(0.15-0.97)	**0.044**
dominant	GG+GA/AA	0.73(0.49-1.10)	0.129	0.71(0.47-1.08)	0.100
recessive	GG/GA+AA	0.40(0.16-0.99)	**0.047**	0.43(0.17-1.08)	0.072

Adjust factors: age, height, SBP, waist, LDL-C, intake of chicken, history of hypertension and dyslipidemia. p value < 0.05 was considered statistically significant.

All covariates-based model

Considering that model A only focused on the influence of genetic variants on CAD risk, we further constructed an all covariates-based model (model B), which recruit not only the gene-score derived from model A, but also comprehensively account for other 31 CAD-related physiological, biochemical and lifestyle indicators. 9 variables with nonzero coefficients were finally screened out in model B by using the same technique of elastic net. The red line represents the model B in Figure 2, and the AUC for model B was 0.71 (95% CI: 0.62-0.78). In terms of AUC value, the accuracy of model B was 20.34% higher than that of model A, showing great contribution of clinical/lifestyle factors to CAD risk. On the other hand, the AUC confidence intervals of model A and B overlapped slightly, which may be due to the relatively small sample size of the study, resulting in increased variance in AUC when measuring the models' discriminative ability. The 9 variables of model B were the gene-score, history of hypertension and dyslipidemia, SBP, age, waist, LDL-C, the intake of chicken and height. 6 of the 9 variables were found to be significantly associated with CAD individually ($p<0.05$) (Table 3). Table S3 listed the coefficients of each variable derived from the elastic net algorithm.

Table 3. Associations of gene-score and lifestyle-related factors with risk of CAD

Characters	OR	95%CI	P
Gene-score	3.65	1.36-9.79	**0.010**
Hypertension history	2.37	1.56-3.60	**<0.001**
Dyslipidemia history	1.82	1.18-2.82	**0.007**
Age	1.07	1.05-1.09	**<0.001**
Waist	1.02	1.00-1.05	0.050
SBP	1.02	1.01-1.03	0.001
LDL-C	1.14	0.89-1.46	0.280
Height	0.97	0.94-0.99	**0.006**
Current chicken intake	0.78	0.64-0.94	**0.008**

p value < 0.05 was considered statistically significant.

Interactions between gene polymorphism and other covariates

Except for a few disease cases purely associated with genetic disease or environmental factors, the vast majority of diseases are the result of interplay of genetic and environmental effects, especially for complex traits of chronic diseases. Thus, we further explored marginal effects and pairwise interactions among several variables using crossover analysis. Table 4 revealed that hypertension and dyslipidemia both have significant marginal effects to increase the risk of CAD. In particular, compared with individuals without hypertension and dyslipidemia, those with hypertension or dyslipidemia alone had a significantly increased risk of CAD by reaching an OR of 2.40 (95% CI:1.10-5.23, p=0.028) for hypertension and an OR of 1.76 (95% CI:1.03-3.03, p=0.004) for dyslipidemia, respectively, showing their significant individual effects to increase CAD risk. On the other hand, the index RERI was also adopted to investigate the interactive effect between hypertension and dyslipidemia, which was calculated as $RERI=OR_{(hypertension+,\ dyslipidemia+)}-(OR_{(hypertension+,\ dyslipidemia-)}+OR_{(hypertension-,\ dyslipidemia+)}-1)$. As a result, RERI reached 0.73 (95% CI: -1.49-2.96, p=0.051) for hypertension and dyslipidemia interaction, indicating that the estimated joint effect of hypertension and dyslipidemia on CAD was not statistically different from the sum of the estimated marginal effects of the two diseases individually, and thus no significant interaction was observed between hypertension and dyslipidemia to increase the risk of CAD. Although the interaction between hypertension and age was not significant as well according to the index RERI, compared with people without hypertension and younger than 60 years old, people without hypertension and older than 60 years old had an increased risk of CAD (OR=3.76, 95% CI=1.76-8.03, p=0.001), while the risk of CAD in people with hypertension and over 60 years old increased by 7.02 times (OR=7.02, 95% CI= 3.64-13.55, p<0.001).

Table 4. Interactions between hypertension and dyslipidemia, hypertension and age for the risk of CAD.

	Hypertension (-)		Hypertension (+)		OR (95%CI) For hypertension patients within strata of risk characters	RERI (95%CI)	p
	case/control (n)	OR (95%CI)	case/control (n)	OR (95%CI)		0.73 (-1.49-2.96)	0.051
Dyslipidemia (-)	20/730	1	10/152	**2.40 (1.10-5.23)** **p=0.028**	**2.40 (1.10-5.23)** **p=0.028**		
Dyslipidemia (+)	42/869	**1.76 (1.03-3.03）** **p=0.004**	28/262	**3.90(2.16-7.04)** **p<0.001**	**2.21 (1.34-3.64)** **p=0.002**		
OR (95%CI) for dyslipidemia within strata of risk characters		**1.76 (1.03-3.03）** **p=0.004**		1.62 (0.77-3.44) p=0.204			
Age <60 years	15/885	1	47/714	**3.88 (2.15-7.00)** **p<0.001**	**3.88 (2.15-7.00)** **p<0.001**	0.38 (-3.04-4.16)	0.843
Age≥60 years	13/204	**3.76 (1.76-8.03)** **p=0.001**	25/210	**7.02 (3.64-13.55)** **p<0.001**	1.87 (0.93-3.75) p=0.079		
OR (95%CI) for older people within strata of risk characters		**3.76 (1.76-8.03)** **p=0.001**		**1.81(1.09-3.01)** **p=0.023**			

p value < 0.05 was considered statistically significant.

We also investigated pairwise interactions between 3 SNPs (rs671, rs6751537 and rs11641677) of 3 genes (*ALDH2*, *ADCY3* and *BCMO1*) and dyslipidemia/hypertension on CAD (Table 5 and Table 6). A significant positive interaction between rs671 and dyslipidemia was observed (RERI=3.36, p=0.05) in our study, which implied that the joint effect of rs671 and dyslipidemia on the CAD risk was greater than the sum of their independent main effects (Table 5). Although the interactions between the other two SNPs and dyslipidemia didn't reach the statistical significance, the studies showed that individuals with dyslipidemia and high genetic risk were at a higher risk of developing CAD than those without dyslipidemia or having low genetic risk. For example, compared to individuals with low genetic risk of rs6751537 and without dyslipidemia, people with dyslipidemia and high genetic risk had a 7.43-fold increase in CAD risk (OR=7.43, 95% CI=3.14-17.55, p<0.001).

Although no significant interactions were identified for all of the 3 SNPs and hypertension on the CAD risk (Table 6, with p values of RERI>0.05), significant cumulative effects were still remarkable. For instance, compared to people carrying GA+GG genotype of rs11641677 and having no history of hypertension, people with AA genotype and hypertension would have a 9.13-fold increase in the risk of CAD (OR=9.13, 95% CI=2.17-38.34, p=0.003).

Table 5. Interaction between gene polymorphism and dyslipidemia for the risk of CAD

	Dyslipidemia (-)		Dyslipidemia (+)		OR (95%CI) for dyslipidemia patients within strata of genotype	RERI (95%CI)	p
	case/control(n)	OR（95%CI）	case/control(n)	OR（95%CI）			
rs671							
Non-risk allele carriers (AG+GG）	28/816		59/1065				
		1		**1.61 (1.02-2.56)**	**1.61 (1.02-2.56)**		
				p=0.041	**p=0.041**		
Risk allele carriers （AA）	2/66		11/66			**3.36 (0.12-6.84)**	**0.050**
		0.88 (0.21-3.79)		**4.86 (2.32-10.19)**	**5.5 (1.17-25.78)**		
		p=0.867		**p<0.001**	**p=0.031**		
OR (95%CI) for risk allele carriers within strata of dyslipidemia		0.88 (0.21-3.79)		**1.73 (1.23-2.45)**			
		p=0.867		**p=0.002**			
rs6751537							
Non-risk allele carriers (AG+AA）	27/852		62/1097				
		1		p=0.014	**1.78 (1.13-2.83)**		
					p=0.014		
Risk allele carriers （GG）	3/30		8/34			3.49(-3.43-10.40)	0.321
		3.16 (0.91-10.98)		**7.43 (3.14-17.55)**	2.35 (0.57-9.68)		
		p=0.071		**p<0.001**	p=0.241		
OR (95%CI) for risk allele carriers within strata of dyslipidemia		3.16 (0.91-10.98)		**4.16 (1.85-9.37)**			
		p=0.071		**p=0.001**			
rs11641677							

Non-risk allele carriers (GA+GG）	1/105		4/130				
		1		3.23 (0.36-29.34) p=0.298	3.23 (0.36-29.34) p=0.298		
Risk allele carriers （AA）	29/777		66/1001			0.77 (-3.19-4.74)	0.771
		3.92 (0.53-29.07) p=0.18		6.92 (0.95-50.39) p=0.05	**1.77 (1.13-2.76) p=0.01**		
OR (95%CI) for risk allele carriers within strata of dyslipidemia		3.92 (0.53-29.07) p=0.18		2.14(0.77-5.98) p=0.14			

p value < 0.05 was considered statistically significant.

Table 6. Interaction between gene polymorphism and hypertension for the risk of CAD

	Hypertension (-)		Hypertension (+)		OR (95%CI) for hypertension patients within strata of genotype	RERI (95%CI)	p
	case/control(n)	OR (95%CI)	case/control(n)	OR (95%CI)			
rs671							
Non-risk allele carriers (AG+GG）	52/1494		35/387				
		1		**2.60 (1.67-4.05) p<0.001**	**2.60 (1.67-4.05) p<0.001**		
Risk allele carriers （AA）	10/105		3/27			-1.14 (-5.37-3.09)	0.600
		2.74 (1.35-5.54) p=0.005		3.19 (0.94-10.86) p=0.063	1.17 (0.30-4.54) p=0.824		
OR (95%CI) for risk allele carriers within strata of hypertension		**2.74 (1.35-5.54) p=0.005**		1.23 (0.60-2.06) p=0.75			
rs6751537							
Non-risk allele carriers (AG+AA）	55/1549		34/400				
		1		**2.39 (1.54-3.72) p＜0.001**	**2.39 (1.54-3.72) p＜0.001**		
Risk allele carriers （GG）	7/50		4/14			2.7 (-6.80-12.22)	0.580
		3.94 (1.71-9.09) p=0.001		**8.05 (2.57-25.24) p<0.001**	2.04 (0.52-7.98) p=0.305		
OR (95%CI) for risk allele carriers within strata of hypertension		**3.94 (1.71-9.09) p=0.001**		**3.36 (1.05-10.78) p=0.041**			
rs11641677							
Non-risk allele carriers (GA+GG）	2/195		3/40				
		1		**7.31 (1.18-45.19)**	**7.31 (1.18-45.19)**		

			p=0.032	**p=0.032**		
Risk allele carriers （AA）	60/1404	35/373			-1.36 (-11.14-8.43)	0.791
		4.17 (1.01-17.18)	**9.13 (2.17-38.34)**	**2.19 (1.42-3.37)**		
		p=0.048	**p=0.003**	**p<0.001**		
OR (95%CI) for risk allele carriers		**4.17 (1.01-17.18)**	1.25 (0.37-4.24)			
within strata of hypertension		**p=0.048**	p=0.723			

p value < 0.05 was considered statistically significant.

Discussion

In this study, we initially developed a gene-based association model (model A) and secondly constructed an all covariates-based model (model B) to explain the risk of CAD. In particular, the all covariates-based model (model B) simultaneously considered both genetic and lifestyle/clinical factors, and finally incorporated a gene-score variable and 8 physiological, biochemical and lifestyle characteristics as significant CAD risk factors, where the gene-score variable was specifically calculated from 3 significant SNPs (rs671, rs6751537 and rs11641677) that survived in model A. With an acceptable accuracy (AUC=0.71), the constructed all covariates-based model (model B) made use of both genetic variants and easily accessible lifestyle/clinical metrics, and could serve as a useful tool to interpret risk of CAD as well as its pathogenesis.

In this study, we found that the risk of CAD in people carrying AA genotype of rs671 in *ALDH2* was 2.45 times higher than people carrying GG genotype, which was consistent with most previous genetic and expression studies[32, 33]. The distribution of aldehyde dehydrogenase 2 (coded by *ALDH2*) in the body is specific, and it is mainly concentrated in the heart, brain, liver and other organs with dense mitochondria[34]. It is an important oxidase involved in alcohol metabolism in cell mitochondria, so early exploration of the relationship between *ALDH2* and diseases had focused on areas directly related to drinking behavior, such as hepatitis and digestive system[34-36]. In addition, the mutation of rs671 gene will reduce the activity of aldehyde dehydrogenase 2, leading to blocked alcohol metabolism and accumulation of aldehydes in the liver[37], which eventually result in the dysfunction of cardiovascular system and incidence of CAD. Compared to those carrying the GG genotype, the people carrying the AA genotype are more likely to experience discomfort such as flushing, nausea, palpitations, and more severe symptoms of alcohol poisoning and alcohol allergy while drinking, and thus they were at a higher CAD risk under the same drinking habits[38]. In our study, no differences were observed in drinking behavior between the case group and the control group, which may be due to the limited sample size.

ADCY3 plays an important role in many physiological and pathophysiological processes, such as adiposity and glucose homeostasis[39]. In this study, the GG genotype of rs6751537 in *ADCY3* was significantly associated with the risk of CAD, which was confirmed in the previous study[40]. Numerous studies have shown that different genetic variants located near or in *ADCY3* are significantly associated with obesity[39]. For example, both candidate gene studies and genome-wide association approaches have demonstrated that *ADCY3* polymorphisms are associated with obesity in European and Chinese populations[40]. Epigenetic studies have indicated that increased DNA methylation levels in *ADCY3* are involved in the pathogenesis of obesity[41]. Furthermore, biological studies using animal models have implicated that adenylate cyclase 3 (coded by *ADCY3*) dysfunction will increase body weight and fat mass, while the activation of adenylate cyclase 3 can

reduce the body weight[39]. It is well known that obesity was an impactful risk factor of cardiovascular and cerebrovascular diseases, which would lead to lipid abnormalities and type 2 diabetes, and subsequently to increased risk of CAD and other cardiovascular diseases[4, 39].

There is a significant role of *BCMO1* in the central cleavage and conversion of dietary provitamin carotenoids to vitamin A (retinal)[42]. It is reported that SNPs of *BCMO1* may affect the efficiency of Beta-carotene (BC) transformation into vitamin A in vivo, and then influence blood concentrations of BC[43]. It is widely known that vitamin A is very important for vision, immune response, cell differentiation, embryonic development, and membrane and skin protection in the small intestine[42, 43]. Studies have reported[44, 45] that vitamin A has a strong association with heart health in mouse, which is mainly because vitamin A deficiency could lead to heart gene expression changes. For example, after a heart attack, the amount of vitamin A delivered from the liver to the heart will increase significantly. The strong association between rs11641677 (in *BCMO1*) and CAD found in our study may further confirm the essential role of vitamin A in heart function at a molecular level, indicating that vitamin A deficiency caused by *BCMO1* mutations may increase the risk of CAD as well as other heart diseases.

Many studies have shown that age, waist circumference, blood pressure and blood lipid can affect CAD, which was also found in our study[3, 6, 8, 46]. Besides, cardiovascular pathologies are widely regarded as strongly influenced by diet and epigenetic effects related to molecules contained in food or drinks. For instance, chicken consumption is another lifestyle protector that we found can significantly reduce the risk of CAD. It is reported that chicken consumption has a protective effect on the occurrence of many chronic diseases[45]. The main dietary guidelines for prevention of CAD developed by the National Cholesterol Education Program in the United States suggested that people should limit their total fat intake to less than 30% of their energy intake and the saturated fatty acid to 8-10%. Compared with red meat, white meat such as chicken and fish have a higher protein content and a lower fat level[47]. In addition, numerous epidemiological studies have shown that populations following the Mediterranean Diet have a reduced incidence of cardiovascular diseases, as such diet has a favorable lipid profile for cardiovascular function, with a high percentage of unsaturated fatty acids. In last decades, epigenetic interaction between the nutrients and DNA has been explored from the nutrigenomics perspective to explain some cellular and molecular phenomena underlying human pathological conditions, and reveal how food components and diet can influence the trajectory toward human health outcomes. In particular, it has been shown that, several dietary molecules can act as epigenetic modulators of DNA methylation, histone acetylation/deacetylation and small non-coding RNA action by directly adding/removing epigenetic marks or indirectly regulating relevant gene expression[48, 49]. These progress on nutrigenomic effects could help explain our findings about the diet's influence on the risk of CAD.

The RERI index is generally considered to be the most suitable parameter for investigating the importance of quantitative interaction in the additive model[30]. We studied pairwise interactions between genetic and biomedical factors, as well as the interplay of certain biomedical factors on the risk of CAD. Among these factors, the rs671-dyslipidemia interactive effect was found statistically significant. The results imply that the mutation of rs671 can lead to the obstruction of alcohol metabolism and the accumulation of acetaldehyde in the liver, thereby accelerating the deterioration of dyslipidemia, and subsequently resulting in CAD incidence. These risk factors may act in a high-dimensional interactive way to increase the risk of CAD where further molecular-level research is needed to confirm this. From a mechanistic point of view, a gene can modify its expression both

due to a polymorphism and to an epigenetic modification of its promoter. Our findings of gene mutations and their interactions with clinical conditions to increase CAD risk may indicate that, compared with the possible epigenetic modifications caused by food component or diet as we discussed above, allele polymorphisms appear to have more weight in gene expression, especially in genes involved in metabolism related to cardiovascular pathologies, which was also supported by the findings of previous studies[50].

The major limitation of this study is the relatively small sample size. As a result, risk factors with small-to-moderate effect size may not be detected as impactful CAD associators in this study, which may further lead to increased variation, reduced power, and potentially misleading findings of our results. Second, our study only focused on the Han population in southern China. Given the population's unique genetic architecture and specific environmental/lifestyle characteristics, our findings may not be directly applied to other regions and population settings. Third, our model is designed to explain correlation, not prediction, which means that only association implications can be interpreted from the results of the study, not causation. Therefore, importantly, long-term studies should be conducted in the future to identify whether these factors are the causes of CAD. Fourthly, many people may be influenced by recall bias and social desirability when filling in the lifestyle questionnaire, and several diet-related habits in our questionnaire were recorded in consumption frequency rather than quantity (such as current chicken intake), therefore, our model's external application needs to be conducted with cautions.

Conclusion

In conclusion, our study constructed an association model and identified 3 SNPs and 8 covariates that were significantly associated with the incidence of CAD in southern Chinese Han population. These results will provide theoretical insights into genetic and lifestyle/clinical risk factors of CAD and facilitate the development of early screening tools for CAD risk and illuminate the etiology of CAD.

Abbreviation list

CAD, coronary artery disease
LDL-C, low-density lipoprotein cholesterol
PCR, polymerase chain reaction
LDR, ligase detection reaction
SNP, single nucleotide polymorphisms
CT, Computed Tomography
QCA test, quantitative coronary angiography test
UTR, un-translation region
MAF, minor allele frequency
ROC, receiver operating characteristic
OR, odd ratio
CIs, confidence intervals
RERI, the relative excess risk due to interaction
SD, standard deviation
SBP, systolic blood pressure
DBP, diastolic blood pressure

BMI, body mass index
TC, total cholesterol
TG, triglyceride
HDL-C, high-density lipoprotein cholesterol
AUC, area under the curve
BC, Beta-carotene

Competing Interests
The authors declare no conflict of interest.

Authors' contributions
ZL, CY, and LY conceived and designed the study. ZL and JRX performed the research. ZL, XYY, JML and LW
analyzed and interpreted the data. ZL and CY wrote the report, which was edited by all authors.

Funding
This study was supported by the Program for Zhejiang Leading Team of Science and Technology Innovation (no. 2011R50021).

Consent for publication
Not applicable.

Acknowledgements
Not applicable.

References
1. Bibbins-Domingo K. **Aspirin Use for the Primary Prevention of Cardiovascular Disease and Colorectal Cancer: U.S. Preventive Services Task Force Recommendation Statement**. *Ann Intern Med* 2016; 164(12):836-845.
2. Go AS, Mozaffarian D, Roger VL, Benjamin EJ, Berry JD, Borden WB, et al. **Executive summary: heart disease and stroke statistics--2013 update: a report from the American Heart Association**. *Circulation* 2013; 127(1):143-152.
3. Roth GA, Nguyen G, Forouzanfar MH, Mokdad AH, Naghavi M, Murray CJL. **Estimates of global and regional premature cardiovascular mortality in 2025**. *Circulation* 2015; 132(13):1270-1282.
4. Grover SA, Dorais M, Paradis G, Fodor JG, Frohlich JJ, McPherson R, et al. **Lipid screening to prevent coronary artery disease: a quantitative evaluation of evolving guidelines**. *CMAJ* 2000; 163(10):1263-1269.
5. Kralik G, Kralik Z. **Poultry products enriched with nutricines have beneficial effects on human health**. *Med Glas (Zenica)* 2017; 14(1):1-7.
6. Nordestgaard BG, Langsted A. **Lipoprotein (a) as a cause of cardiovascular disease: insights from epidemiology, genetics, and biology**. *J Lipid Res* 2016; 57(11):1953-1975.
7. Cosselman KE, Navas-Acien A, Kaufman JD. **Environmental factors in cardiovascular disease**. *Nat Rev Cardiol* 2015; 12(11):627-642.
8. Rodondi N, Collet T-H, Nanchen D, Locatelli I, Depairon M, Aujesky D, et al. **Impact of carotid**

plaque screening on smoking cessation and other cardiovascular risk factors: a randomized controlled trial. *Arch Intern Med* 2012; 172(4):344-352.

9. Vicari RM, Bramlet D, Olivera B, Barber L, Alexander L, Parker L, et al. **Atherosclerosis and teen eating study**. *J Clin Lipidol* 2007; 1(3):194-197.

10. McPherson R, Pertsemlidis A, Kavaslar N, Stewart A, Roberts R, Cox DR, et al. **A common allele on chromosome 9 associated with coronary heart disease**. *Science* 2007; 316(5830):1488-1491.

11. Helgadottir A, Thorleifsson G, Manolescu A, Gretarsdottir S, Blondal T, Jonasdottir A, et al. **A common variant on chromosome 9p21 affects the risk of myocardial infarction**. *Science* 2007; 316(5830):1491-1493.

12. Koyama S, Ito K, Terao C, Akiyama M, Horikoshi M, Momozawa Y, et al. **Population-specific and trans-ancestry genome-wide analyses identify distinct and shared genetic risk loci for coronary artery disease**. *Nat Genet* 2020; 52(11):1169-1177.

13. Hartiala JA, Han Y, Jia Q, Hilser JR, Huang P, Gukasyan J, et al. **Genome-wide analysis identifies novel susceptibility loci for myocardial infarction**. *Eur Heart J* 2021; 42(9):919-933.

14. Roberts R. **Genetics of coronary artery disease: an update**. *Methodist Debakey Cardiovasc J* 2014; 10(1):7-12.

15. Reiss AB, Shah N, Muhieddine D, Zhen J, Yudkevich J, Kasselman LJ, et al. **PCSK9 in cholesterol metabolism: from bench to bedside**. *Clin Sci (Lond)* 2018; 132(11):1135-1153.

16. Hoover-Plow J, Huang M. **Lipoprotein(a) metabolism: potential sites for therapeutic targets**. *Metabolism* 2013; 62(4):479-491.

17. Lehner R, Lian J, Quiroga AD. **Lumenal lipid metabolism: implications for lipoprotein assembly**. *Arterioscler Thromb Vasc Biol* 2012; 32(5):1087-1093.

18. Huang Y, Mahley RW. **Apolipoprotein E: structure and function in lipid metabolism, neurobiology, and Alzheimer's diseases**. *Neurobiol Dis* 2014; 72 Pt A:3-12.

19. Lietz G, Oxley A, Boesch-Saadatmandi C, Kobayashi D. **Importance of β,β-carotene 15,15'-monooxygenase 1 (BCMO1) and β,β-carotene 9',10'-dioxygenase 2 (BCDO2) in nutrition and health**. *Mol Nutr Food Res* 2012; 56(2):241-250.

20. Kyle-Little Z, Zele AJ, Morris CP, Feigl B. **The Effect of BCMO1 Gene Variants on Macular Pigment Optical Density in Young Healthy Caucasians**. *Front Nutr* 2014; 1:22.

21. Narita M, Kitagawa K, Nagai Y, Hougaku H, Hashimoto H, Sakaguchi M, et al. **Effects of aldehyde dehydrogenase genotypes on carotid atherosclerosis**. *Ultrasound Med Biol* 2003; 29(10):1415-1419.

22. Takagi S, Iwai N, Yamauchi R, Kojima S, Yasuno S, Baba T, et al. **Aldehyde dehydrogenase 2 gene is a risk factor for myocardial infarction in Japanese men**. *Hypertens Res* 2002; 25(5):677-681.

23. Goni L, Riezu-Boj JI, Milagro FI, Corrales FJ, Ortiz L, Cuervo M, et al. **Interaction between an Genetic Variant and Two Weight-Lowering Diets Affecting Body Fatness and Body Composition Outcomes Depending on Macronutrient Distribution: A Randomized Trial**. *Nutrients* 2018; 10(6).

24. Richardson TG, Zheng J, Davey Smith G, Timpson NJ, Gaunt TR, Relton CL, et al. **Mendelian Randomization Analysis Identifies CpG Sites as Putative Mediators for Genetic Influences on Cardiovascular Disease Risk**. *Am J Hum Genet* 2017; 101(4):590-602.

25. Rudemiller NP, Lund H, Priestley JR, Endres BT, Prokop JW, Jacob HJ, et al. **Mutation of SH2B3 (LNK), a genome-wide association study candidate for hypertension, attenuates Dahl salt-sensitive hypertension via inflammatory modulation**. *Hypertension* 2015; 65(5):1111-1117.

26. Friedman J, Hastie T, Tibshirani R. **Regularization Paths for Generalized Linear Models via Coordinate Descent**. *J Stat Softw* 2010; 33(1).

27. Li Z, Ye C-Y, Zhao T-Y, Yang L. **Model of genetic and environmental factors associated with type 2 diabetes mellitus in a Chinese Han population**. *BMC Public Health* 2020; 20(1):1024.

28. Huang S, Hu C, Bell ML, Billheimer D, Guerra S, Roe D, et al. **Regularized continuous-time Markov Model via elastic net**. *Biometrics* 2018; 74(3):1045-1054.

29. Noguchi Y, Tachi T, Teramachi H. **Comparison of Signal Detection Algorithms Based on Frequency Statistical Model for Drug-Drug Interaction Using Spontaneous Reporting Systems**. *Pharm Res* 2020; 37(5):86.

30. Knol MJ, VanderWeele TJ. **Recommendations for presenting analyses of effect modification and interaction**. *Int J Epidemiol* 2012; 41(2):514-520.

31. Hosmer DW, Lemeshow S. **Confidence interval estimation of interaction**. *Epidemiology* 1992; 3(5):452-456.

32. Li R, Zhao Z, Sun M, Luo J, Xiao Y. **ALDH2 gene polymorphism in different types of cancers and its clinical significance**. *Life Sci* 2016; 147:59-66.

33. Han H, Wang H, Yin Z, Jiang H, Fang M, Han J. **Association of genetic polymorphisms in ADH and ALDH2 with risk of coronary artery disease and myocardial infarction: a meta-analysis**. *Gene* 2013; 526(2):134-141.

34. Chen IC, Kuo P-H, Yang AC, Tsai S-J, Liu T-H, Liu H-J, et al. **CUX2, BRAP and ALDH2 are associated with metabolic traits in people with excessive alcohol consumption**. *Sci Rep* 2020; 10(1):18118.

35. Hashimoto M, Watanabe M, Uematsu Y, Hattori S, Miyai N, Utsumi M, et al. **Relationships of alcohol dehydrogenase 1B (ADH1B) and aldehyde dehydrogenase 2 (ALDH2) genotypes with alcohol sensitivity, drinking behavior and problem drinking in Japanese older men**. *Environ Health Prev Med* 2016; 21(3):138-148.

36. Quillen EE, Chen X-D, Almasy L, Yang F, He H, Li X, et al. **ALDH2 is associated to alcohol dependence and is the major genetic determinant of "daily maximum drinks" in a GWAS study of an isolated rural Chinese sample**. *Am J Med Genet B Neuropsychiatr Genet* 2014; 165B(2):103-110.

37. Ma C, Yu B, Zhang W, Wang W, Zhang L, Zeng Q. **Associations between aldehyde dehydrogenase 2 (ALDH2) rs671 genetic polymorphisms, lifestyles and hypertension risk in Chinese Han people**. *Sci Rep* 2017; 7(1):11136.

38. Taylor AE, Lu F, Carslake D, Hu Z, Qian Y, Liu S, et al. **Exploring causal associations of alcohol with cardiovascular and metabolic risk factors in a Chinese population using Mendelian randomization analysis**. *Sci Rep* 2015; 5:14005.

39. Wu L, Shen C, Seed Ahmed M, Östenson CG, Gu HF. **Adenylate cyclase 3: a new target for anti-obesity drug development**. *Obes Rev* 2016; 17(9):907-914.

40. Tong T, Shen Y, Lee H-W, Yu R, Park T. **Adenylyl cyclase 3 haploinsufficiency confers susceptibility to diet-induced obesity and insulin resistance in mice**. *Sci Rep* 2016; 6:34179.

41. Rogne M, Taskén K. **Compartmentalization of cAMP signaling in adipogenesis, lipogenesis, and lipolysis**. *Horm Metab Res* 2014; 46(12):833-840.

42. Ferrucci L, Perry JRB, Matteini A, Perola M, Tanaka T, Silander K, et al. **Common variation in the beta-carotene 15,15'-monooxygenase 1 gene affects circulating levels of carotenoids: a genome-wide association study**. *Am J Hum Genet* 2009; 84(2):123-133.

43. Cai X, Lian F, Kong Y, Huang L, Xu L, Wu Y, et al. **Carotenoid metabolic (BCO1) polymorphisms and personal behaviors modify the risk of coronary atherosclerosis: a nested case-control study in**

Han Chinese with dyslipidaemia (2013-2016). *Asia Pac J Clin Nutr* 2019; 28(1):192-202.

44. Asson-Batres MA, Ryzhov S, Tikhomirov O, Duarte CW, Congdon CB, Lessard CR, et al. **Effects of vitamin A deficiency in the postnatal mouse heart: role of hepatic retinoid stores**. *Am J Physiol Heart Circ Physiol* 2016; 310(11):H1773-H1789.

45. de Paiva SAR, Zornoff LAM, Okoshi MP, Okoshi K, Matsubara LS, Matsubara BB, et al. **Ventricular remodeling induced by retinoic acid supplementation in adult rats**. *Am J Physiol Heart Circ Physiol* 2003; 284(6):H2242-H2246.

46. **Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) final report**. *Circulation* 2002; 106(25):3143-3421.

47. Akbaraly T, Würtz P, Singh-Manoux A, Shipley MJ, Haapakoski R, Lehto M, et al. **Association of circulating metabolites with healthy diet and risk of cardiovascular disease: analysis of two cohort studies**. *Sci Rep* 2018; 8(1):8620.

48. Caradonna F, Consiglio O, Luparello C, Gentile C. **Science and Healthy Meals in the World: Nutritional Epigenomics and Nutrigenetics of the Mediterranean Diet**. *Nutrients* 2020; 12(6).

49. Caradonna F, Cruciata I, Luparello C. **Nutrigenetics, nutrigenomics and phenotypic outcomes of dietary low-dose alcohol consumption in the suppression and induction of cancer development: evidence from in vitro studies**. *Crit Rev Food Sci Nutr* 2020:1-32.

50. Naselli F, Catanzaro I, Bellavia D, Perez A, Sposito L, Caradonna F. **Role and importance of polymorphisms with respect to DNA methylation for the expression of CYP2E1 enzyme**. *Gene* 2014; 536(1):29-39.

Abstract

Backgrounds: To investigate associations of genetic factors and environmental factors with coronary artery disease (CAD), We collected medical reports, lifestyle details, and blood samples of 2113 individuals, and then used the polymerase chain reaction-ligase detection reaction to genotype the targeted 102 SNPs. **Methods:** We adopted elastic net algorithm to build an association model that considered simultaneously genetic and lifestyle/clinical factors associated with CAD in Chinese Han population. **Results:** It is found that, in terms of genetic variants, the AA genotype of rs671 in the additive (adjusted OR=2.51, p=0.008) and recessive (adjusted OR=2.12, p=0.021) models, the GG genotype of rs6751537 in the additive (adjusted OR=3.36, p=0.001) and recessive (adjusted OR=3.47, p=0.001) models, and GG genotype of rs11641677 in additive model (adjusted OR=0.39, p=0.044) was associated with the increased risk of CAD. In terms of lifestyle/clinical factors, the history of hypertension (unadjusted OR=2.37, p<0.001) and dyslipidemia (unadjusted OR=1.82, p=0.007), age (unadjusted OR=1.07, p<0.001) and waist circumference (unadjusted OR=1.02, p=0.05) would significantly increase the risk of CAD, while height (unadjusted OR=0.97, p=0.006) and regular intake of chicken (unadjusted OR=0.78, p=0.008) reduced the risk of CAD. Significant interactions were found between hypertension and dyslipidemia (RERI = 4.88, p=0.014), as well as between rs671 and dyslipidemia (RERI = 3.36, p=0.05). **Conclusion:** In this study, we constructed an association mode and identified a set of SNPs and lifestyle/clinical risk factors of CAD in Chinese Han population. By considering both genetic and non-genetic risk factors, the built model may provide implications for CAD pathogenesis and clues for screening tool development in Chinese Han population.

Abbreviations list

coronary artery disease (CAD)
Aldehyde Dehydrogenase 2 (ALDH2)
Adenylate cyclase 3 (ADCY3)
Beta-carotene monooxygenase (BCMO1)
3'untranslation region (3' UTR)
5'untranslation region (5' UTR)
Minor allele frequency (MAF)
operating characteristic (ROC)
odd ratios (ORs)
confidence intervals (CIs)
relative excess risk due to interaction (RERI)

ZL, CY, and LY conceived and designed the study. ZL and JRX performed the research. SLG, XYY, JML and LW analyzed and interpreted the data. ZL and CY wrote the report, which was edited by all authors.

Highlights

- An all covariates-based association model was built to explain the risk of CAD
- The model incorporated both a gene-score and 8 lifestyle/clinical risk factors
- The gene-score was derived from *ALDH2*, *ADCY3* and *BCMO1* polymorphisms
- The findings may provide implications for CAD pathogenesis in Chinese Han population