

Shared polymorphisms and modifiable behavior factors for myocardial infarction and high cholesterol in a retrospective population study

Yulan Liang, PhD^{a,*}, Arpad Kelemen, PhD^b

Abstract

Genetic and environmental (behavior, clinical, and demographic) factors are associated with increased risks of both myocardial infarction (MI) and high cholesterol (HC). It is known that HC is major risk factor that may cause MI. However, whether there are common single nucleotide polymorphism (SNPs) associated with both MI and HC is not firmly established, and whether there are modulate and modified effects (interactions of genetic and known environmental factors) on either HC or MI, and whether these joint effects improve the predictions of MI, is understudied.

The purpose of this study is to identify novel shared SNPs and modifiable environmental factors on MI and HC. We assess whether SNPs from a metabolic pathway related to MI may relate to HC; whether there are moderate effects among SNPs, lifestyle (smoke and drinking), HC, and MI after controlling other factors [gender, body mass index (BMI), and hypertension (HTN)]; and evaluate prediction power of the joint and modulate genetic and environmental factors influencing the MI and HC.

This is a retrospective study with residents of Erie and Niagara counties in New York with a history of MI or with no history of MI. The data set includes environmental variables (demographic, clinical, lifestyle). Thirty-one tagSNPs from a metabolic pathway related to MI are genotyped. Generalized linear models (GLMs) with imputation-based analysis are conducted for examining the common effects of tagSNPs and environmental exposures and their interactions on having a history of HC or MI.

MI, BMI, and HTN are significant risk factors for HC. HC shows the strongest effect on risk of MI in addition to HTN; gender and smoking status while drinking status shows protective effect on MI. *rs16944* (gene *IL-1 β*) and *rs17222772* (gene *ALOX*) increase the risks of HC, while *rs17231896* (gene *CETP*) has protective effects on HC either with or without the clinical, behavioral, demographic factors with different effect sizes that may indicate the existence of moderate or modifiable effects. Further analysis with the inclusions of gene–gene and gene–environmental interactions shows interactions between *rs17231896* (*CETP*) and *rs17222772* (*ALOX*); *rs17231896* (*CETP*) and gender. *rs17237890* (*CETP*) and *rs2070744* (*NOS3*) are found to be significantly associated with risks of MI adjusted by both SNPs and environmental factors. After multiple testing adjustments, these effects diminished as expected. In addition, an interaction between drinking and smoking status is significant. Overall, the prediction power in successfully classifying MI status is increased to 80% with inclusions of all significant tagSNPs and environmental factors and their interactions compared with environmental factors only (72%).

Having a history of either HC or MI has significant effects on each other in both directions, in addition to HTN and gender. Genes/SNPs identified from this analysis that are associated with HC may be potentially linked to MI, which could be further examined and validated through haplotype-pairs analysis with appropriate population stratification corrections, and function/pathway regulation analysis to eliminate the limitations of the current analysis.

Abbreviations: AIC = Akaike information criterion, AUC = area under curve, BIC = Bayesian information criterion, BMI = body mass index, CHD = coronary heart disease, CI = confidence interval, DBP = diastolic blood pressure, df = degree of freedom, GLMs = generalized linear models, HC = high cholesterol, HDL = high-density lipoprotein, HTN = hypertension, LDL = low-density lipoprotein, MI = myocardial infarction, OR = odds ratio, SBP = systolic blood pressure, SD = standard deviation, SNP = single nucleotide polymorphism, TG = triglycerides.

Keywords: association analysis, gene–environment, high cholesterol, myocardial infarction, retrospective population study, single nucleotide polymorphism

Editor: Phil Phan.

YL and AK both contributed to all aspects of this research, including design, performing the study, data analysis, drafting, and revising the manuscript.

The authors declare no conflict of interests.

^aDepartment of Family and Community Health, ^bDepartment of Organizational Systems and Adult Health, University of Maryland, Baltimore, MD.

* Correspondence: Yulan Liang, University of Maryland, Baltimore, 655 W. Lombard Street, Baltimore, MD 21201-1579 (e-mail: liang@son.umaryland.edu).

Copyright © 2017 the Author(s). Published by Wolters Kluwer Health, Inc.

This is an open access article distributed under the Creative Commons Attribution-Share Alike License 4.0, which allows others to remix, tweak, and build upon the work, even for commercial purposes, as long as the author is credited and the new creations are licensed under the identical terms.

Medicine (2017) 96:37(e7683)

Received: 23 March 2017 / Received in final form: 10 July 2017 / Accepted: 12 July 2017

<http://dx.doi.org/10.1097/MD.0000000000007683>

1. Introduction

Risk prediction and assessment are central parts of common complex disease prevention, including coronary heart disease (CHD) and its sequels, such as acute myocardial infarction (MI). Refining prediction strategies remains important for targeting treatment recommendations.^[1–4] MI is a leading cause of death throughout the world. Approximately 450,000 people in the United States die from coronary disease per year.^[5] The risk of MI increases with age, while the actual incidence is dependent on predisposing risk factors for atherosclerosis.^[6] Potential risk factors of MI and atherosclerotic coronary artery disease have been reported as hyperlipidemia, diabetes mellitus, hypertension (HTN), tobacco use, male gender, age, and family history of atherosclerotic arterial disease; genetics.^[7–9] Furthermore, family history of heart attacks, lack of physical activity, alcohol consumption, obesity, and stress are also considered as potential associated factors.^[10–13] It was noted that having a history of high cholesterol (HC) was associated with a history of MI.^[14–16] The genetics on either HC^[17–21] or MI^[22,23] have been examined. There are a number of studies that reported several of the known MI/CHD loci that are also associated with lipids traits and vice versa.^[24] However, whether there are common genotypes or polymorphisms associated with both MI and HC is not confirmed or firmly established,^[25–28] which are important and ideal drug or intervention target for precision medicine.^[29–31] Furthermore, the genetic variants may play an important, but under-recognized role in modulating the effect of environmental exposures on the risk of MI and HC.^[32–35] Currently, whether the joint and modulate effects (including the interactions of both genetic and environmental factors) improve the predictions of MI is understudied. In this article, we will examine whether there are common genetic variants that contribute to the HC and MI from a selected set of genetic variants for retrospective western population studies. Furthermore, we will evaluate the prediction power of these combined effects together with the identified significant environmental factors, which target to test the central hypothesis that a combination of common single nucleotide polymorphisms (SNPs) and environmental factors contributes to the risk of HC as well as MI.

2. Methods

2.1. Data source, design, study population, and variables

The study sample was randomly selected from the general populations of Erie and Niagara counties in New York State falling into age group 35 to 79 years matched by age within 5-year difference.^[36] The study was approved by the ethical committees (institutional review board of the University at Buffalo), and all study participants signed informed consent.^[37] For the current analyses, we only used partial data of 1837 Caucasian participants that have been genotyped, of which 818 having history of MI (209 women and 609 men) and 1019 with no history of MI (608 men and 411 women). The data set includes 29 environmental variables collected through participants' clinic visits, which include health behavior or lifestyle (smoking status, alcohol consumption) in addition to demographic (gender), clinical [body mass index (BMI)], systolic blood pressure (SBP), diastolic blood pressure (DBP), HTN, family history for heart disease, triglycerides (TG), high-density lipoprotein (HDL), low-density lipoprotein (LDL), and total cholesterol. Diabetic patients were excluded from participating in the study.

2.2. Genetic marker selection

A number of genes (SNPs) in the metabolic pathways related to MI (i.e., inflammation, inflammation-mediated coagulation, and metalloproteinase activation) from the SeattleSNPs variation discovery panel database (<http://pga.gs.washington.edu/>) were pre-selected and genotyped on the basis of biological plausibility; significant statistical evidence that SNPs in the gene are consistently associated with variation in intermediate phenotype (e.g., protein levels) or with MI; using linkage disequilibrium and allele frequency ($\leq 10\%$).^[38–40] These genes and SNPs include inflammatory genes Interleukin 1 (*IL-1 β*): *rs1143634*, *rs16944*, *rs3917354*, *rs3917356*; and interleukin 6 (*IL 6*): *rs2069825*, *rs1818879*, *rs1548216*, *rs1800795*; *CETP* gene: *rs17231513*, *rs17237890*; Tissue remodeling gene, matrix metalloproteinase (*MMP3*): *rs522616*, *rs595840*, *rs602128*, *rs680753*; Coagulation gene: tissue factor (*TF*): *rs1324214*, *rs1361600*, *rs3354*, *rs3917639*, and constitutive androstane receptor (*CAR*) gene etc. Genotyping was done in duplicates (success rate 93% to 99%) using the MassARRAY System (Sequenom) in multiplexes with 5 ng of DNA.^[37,40] All SNPs are coded as dominant effects [with 1 represents the presence of at least one reference allele (*A/a*), while 0 is none].

2.3. Main outcome measures

MI cases are defined as people who used to have MI (once or multiple) before or during data collection period, and the controls are defined as people who have never had MI by the time of data collection. HC, having a history of HC (once or multiple), are used as a primary lipid outcome measure given some missing issues for other lipid measures for these data.

2.4. Data analysis

2.4.1. Imputation-based data analysis for missing data and data preprocessing. Missing data analysis was conducted to examine missing pattern and mechanism, for example, missing of BMI was complete at random ($P=.178$ for little MCAR test). Multiple imputation was performed for near significant variables ($P<.10$), missing $\leq 10\%$. These include BMI (2% missing), smoking status (0.3%), HTN (0.1%), Meds for high HTN (6.7%), HC (1.1%), family history for heart disease (8.5%), drinking status in past 1 to 2 years (0.4%), *rs3917356* (*IL-1 β*) (2.2%), *rs2069825* (*IL6*) (2.9%), *ADDU-000614* (*ADDU*) (1.1%), *rs17237890* (*CETP*) (1.7%), *rs2070744* (*NOS*) (2.3%), and *rs2077647* (*ESR1*) (1.9%). There were 3 cases with outliers in amount of ethanol ($z=23.73$, 14.33, 11.60), which were deleted, as the amount of ethanol was significantly different by MI. "Drinking status in past 1–2yrs" was recoded to make the sample size in each category less imbalanced: "Lifetime abstainers," "Irregular abstainers," and "Non-current drinker" were categorized as "non-current drinker," "current non-weekly drinker," "current weekly drinker" as "current drinker."

2.4.2. Association analysis and predictions. The unadjusted univariate association between all factors and MI was calculated and correlation between explanatory variables that may indicate collinearity were assessed pre hoc using Pearson correlation for continuous variables, Chi-square test for categorical variables, and *t* test or analysis of variance (ANOVA) for continuous and categorical variables. Generalized linear models (GLMs) were performed with the dominant genetic effect coding. Besides the genetic factors, demographic, clinical, and lifestyle factors are included as covariates and confounding factors in the multivari-

able modeling processes, and adjusted odds ratios (ORs) and 95% confidence intervals (95% CIs) were calculated. The model significance was examined using the Omnibus test and the Hosmer and Lemeshow tests. Goodness of fit of model was assessed using deviance/df, AIC, BIC, the model with smaller deviance/df, AIC and BIC values fits the data better. Nagelkerke pseudo R^2 values were reported for the model performance. The presences of outliers (absolute standardized residual values ≥ 3) that may affect the stability of the model were assessed and removed using plot of leverage value, Cooks distance, and visual inspection. All 2-way interactions between significant explanatory variables were tested and retained in the final model if they significantly improved the model with $P < .05$ based on the χ^2 test of model-fit improvement. For the prediction power of the model, random sampling was conducted to select 70% of the total sample as training data to build up the model, and the other 30% of the sample as testing data; AIC and classification/prediction accuracy are reported (see Fig. 1).

3. Results

3.1. Sample characteristics

3.1.1. The whole sample. Among all participants, the average age was 54.8 years old [standard deviation (SD)=8.9]. The majority of the sample was male (66.2%). The average BMI was 28.5 (SD=5.2), 46.3% had a history of HC, 70.5% had family history for heart disease, and 64% did not have HTN. About two-thirds of the sample drank nonweekly or weekly in the past 1 to 2 years (66.3%), and 43.5% smoked in the past, while 30.6% were current smokers.

3.1.2. MI case-control. In the MI case, 64.9% had a history of HC, while in the non-MI group, only 31.3% had HC, which indicated that HC was associated with MI. These numbers are much higher (due to the age class differences) than the numbers the National Health and Nutrition Examination Survey data reported for 2011 to 2012, which estimated 12.9% of U.S. adults aged 20 years and over (11.1% of men and 14.4% of women; 17.1% of non-Hispanic whites) had HC.^[31–35,41] There were significant differences in gender, smoking (lifetime total pack years), drinking status, and amount of ethanol, BMI, SBP, DBP, TG, total cholesterol, HDL, LDL, high blood cholesterol, whether HTN and high blood cholesterol was treated by meds,

and family history on heart disease between MI cases and controls (all $P < .05$). Within the MI case group, two-thirds of participants had HC (65.9%), while one-third of participants in the control group had HC (31.6%), which is significantly different ($\chi^2=210.895$, $P < .001$). These descriptive statistics confirm that HC is associated with MI for this Caucasian population. Summary of environmental characteristics for the total sample, the MI cases, and controls is summarized in Table 1.

3.2. Univariate analysis and multivariable for MI

Table 2 provides a summary of genetic characteristics in MI cases and control, as well as the full list of pre-selected and genotyped gene/SNPs. The unadjusted univariate association between tagSNPs and MI calculated using χ^2 tests are included in Table 2. Risk of MI was associated with SNPs ($P < .05$): *rs3917356* (*IL-1 β* gene), *rs17231513* (*CETP_1*), *rs17237890* (*CETP_3*), and *rs2070744* (*NOS_2*), while *s2069825* (*IL6*), *ADDU-000614* (*ADDU*), and *rs9340799* (*ESR1*) were near significantly different by MI case and control groups ($P < .10$). None of the rest SNPs was significantly or nearly significantly associated with a risk of MI, which were excluded from multivariable analysis. Finally, 14 genetic and environmental variables ($P < .10$), including BMI, smoking status, HBP, Meds for HTN, HC, family history for heart disease, drinking status in past 1 to 2 years, *rs3917356* (*IL-1 β*), *rs2069825* (*il6*), *ADDU-000614* (*ADDU*), *rs17237890* (*CETP*), *rs2070744* (*NOS*), and *rs2077647* (*ESR1*), were included in the follow-up multivariable analysis after multiple imputation.

3.3. Multivariable analysis for MI

3.3.1. Without SNPs/genes. Age, BMI, gender, HTN, HC, family history about heart disease, high TG, and high alcohol consumption in past 30 days explained 28.5% of the variance (Nagelkerke R squared) in the risk of MI with 72% predictive power for MI status. The strongest factor for MI was the HC [OR: 112.51, 95% CI: 23.10–548.04]. In addition, interactions between HC and age are significant (OR: 0.941, 95% CI: 0.92–0.97).

3.3.2. With both SNPs/genes and the environmental factors with main effects.

GLM showed that 8 variables, including gender, smoking status, HTN, HC, family history on heart disease, drinking status in past 1 to 2 years, *rs17231513* (*CETP*), and *rs2070744* (*NOS*) had significant effects in predicting a risk of MI (see Model 1 in Table 3). The model was significant ($P < .001$ in Omnibus Tests, $P = .671$ in Hosmer and Lemeshow Test) and explained 40.6% of total variance in risk of MI (Nagelkerke $R^2 = .406$), which is higher than without genes/SNPs (28.5%). The prediction power in successfully classifying MI is increased to 78.8% when including both SNPs and environmental factors [area under curve (AUC)=0.788 (0.751–0.826)] compared with environmental factors (72%).

HC shows the strongest effect on risk of MI (OR = 5.397, 95% CI 4.069–7.160) compared with other clinical, behavioral, and demographic factors. Smoking status (OR=3.399, 95% CI 2.376–4.862) for current smoker compared with nonsmoker, HTN (OR = 2.420, 95% CI 1.819–3.219), and gender being male (OR = 4.252, 95% CI 3.087–5.857) are found to be significantly associated with risk of MI. Drinking status in past 1 to 2 years shows protective effect (OR = .672, 95% CI 0.502–0.901). *rs17231513* (*CETP*) (OR = 1.431, 95% CI 1.047–1.956) and *rs2070744* (*NOS*) (OR = 1.413, 95% CI 1.063–1.877) are found

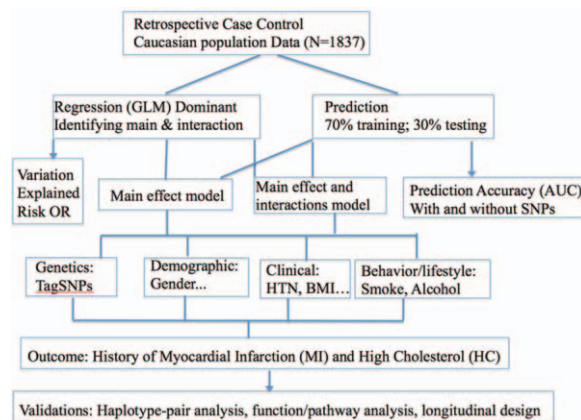


Figure 1. Association and prediction analysis of shared polymorphisms and modifiable environmental (behavior) factors for myocardial infarction and high cholesterol.

Table 1**Summary of clinical, behavior, and demographic characteristics and their unadjusted univariate association with MI cases/control status (N = 1837).**

Continuous variables	Myocardial infarction (MI)				t	P
	Control (n = 1019)	Case (n = 818)	Total	Missing		
	Mean (SD)	Mean (SD)	Mean (SD)	n (%)		
Age, year	54.9 (8.9)	54.7 (8.9)	54.8 (8.9)	0	0.641	.522
Age difference	0.3 (.8)	0.2 (.7)	0.3 (.7)	0	1.9	.058
BMI, kg/m ²	28.3 (5.5)	28.8 (4.7)	28.5 (5.2)	38 (2.0)	-2.19	.029
SBP, mm Hg	120.9 (14.9)	114.2 (14.2)	117.9 (14.9)	21 (1.1)	9.73	<.001
DBP, mm Hg	74.2 (9.8)	70.5 (8.8)	72.6 (9.6)	20 (1.0)	8.385	<.001
Smoke, Lifetime total pack-years, y	15.7 (18.6)	25.0 (24.1)	19.8 (21.7)	2 (0.1)	-9.127	<.001
Triglycerides, mg/dL	151.0 (103.0)	180.4 (116.0)	163.2 (109.5)	114 (6.2)	-5.527	<.001
Total cholesterol, mg/dL	224.2 (41.0)	192.7 (41.5)	210.3 (44.0)	816 (44.4)	12.094	<.001
HDL, mg/dL	53.0 (15.4)	41.8 (12.0)	48.2 (15.1)	862 (46.9)	12.842	<.001
LDL, mg/dL	144.1 (37.3)	118.4 (36.3)	133.1 (39.1)	862 (46.9)	10.787	<.001
Ethanol, ounce/past 30 d	11.8 (28.3)	7.4 (18.4)	9.9 (24.5)	10 (0.5)	4.008	<.001
Alcohol, ounce/past 1–2 y	158.1 (324.1)	175.8 (399.6)	166.0 (359.8)	7 (0.4)	-1.021	.308
Categorical variables	n (%)	n (%)	n (%)	n (%)	χ ²	P
Gender				0	43.693	<.001
Female	411 (66.3)	209 (33.7)	620 (33.8)			
Male	608 (50.0)	609 (50.0)	1217 (66.2)			
HTN				2 (0.1)	77.627	<.001
No	743 (63.2)	432 (36.8)	1175 (64.0)			
Yes	276 (41.8)	384 (58.2)	660 (35.9)			
Meds for high BP				44 (6.7)	3.990	.026
No	41 (15.9)	36 (10.9)	77 (4.2)			
Yes	217 (84.1)	319 (89.1)	536 (29.2)			
High blood cholesterol				20 (1.1)	210.895	<.001
No	692 (71.6)	275 (28.4)	967 (52.6)			
Yes	319 (37.5)	531 (62.5)	850 (46.3)			
Meds for high BC				359 (42.2)	74.031	<.001
No	100 (59.9)	69 (14.1)	170 (9.3)			
Yes	65 (38.9)	255 (78.7)	320 (17.4)			
Family history for heart disease				157 (8.5)	12.928	<.001
No	251 (65.4)	133 (34.6)	384 (20.9)			
Yes	711 (54.9)	585 (45.1)	1296 (70.5)			
Smoking status				5 (0.3)	176.548	<.001
Nonsmoker	267 (56.8)	203 (43.2)	470 (25.6)			
Former smokers	561 (70.2)	238 (29.8)	799 (43.5)			
Current smoker	191 (33.9)	372 (66.1)	563 (30.6)			
Drinking status in past 1–2 y				7 (0.4)	11.921	.018
Lifetime abstainers	15 (46.9)	17 (53.1)	32 (1.7)			
Irregular abstainers	39 (54.9)	32 (45.1)	71 (3.9)			
Noncurrent drinker	252 (49.4)	258 (50.6)	510 (27.8)			
Current nonweekly drinker	246 (58.0)	178 (42.0)	424 (23.1)			
Current weekly drinker	461 (58.1)	332 (41.9)	793 (43.2)			

Note: The unadjusted univariate association between environmental variables and MI calculated using Chi-squared tests for categorical variables and *t* tests for continuous variables. Fisher exact test was used if no Chi-squared statistic value was provided. The continuity correction was applied for factors involving 2 × 2 categories. BC = blood cholesterol; BMI = body mass index; BP = blood pressure; DBP = diastolic blood pressure; HDL = high-density lipoprotein; HTN = hypertension; LDL = low-density lipoprotein; MI = myocardial infarction; SBP = systolic blood pressure; SD = standard deviation.

to be significantly associated with risk of MI, which are also the most significant SNPs in the unadjusted association analysis. Two other significant SNPs from unadjusted analysis diminish due to the inclusion of conditional or modifiable effects.

3.3.3. With both main effects and interactions of gene and environmental factors. After entering all 2-way interaction effects, the interaction between drinking and smoking status was the only significant effect in the model, which improved the total percent of variance explained by 0.4%. With the interaction effect added, 41% of total variance in risk of MI was explained (Nagelkerke $R^2 = 0.410$) (see Table 3 Model 2). Similarly, the model fits the data better (Deviance/df = 1.130, AIC/BIC =

691.788/763.237 = 0.906). Overall, the prediction power in successfully classifying MI status is increased to 80% with inclusions of all significant SNPs and environmental factors and their interactions (AUC = 0.799, 0.763–0.836) compared with-out interactions (78.8%).

HC shows the strongest effect on risk of MI (OR = 5.397, 95% CI 4.069–7.160) compared with other clinical, behavioral, and demographic factors. Smoking status (OR = 3.399, 95% CI 2.376–4.862) for current smokers compared with nonsmokers, HTN (OR = 2.420, 95% CI 1.819–3.219), and being male (OR = 4.252, 95% CI 3.087–5.857) are found to be strongly associated with risk of MI ($P < .001$). Drinking status in the past 1 to 2 years shows protective effect (OR = 0.672, 95% CI 0.502–0.901).

Table 2

Summary of genetic characteristics in MI cases and control subjects (N = 1837), the unadjusted univariate association between SNPs and MI calculated using Chi-squared tests.

Genes/SNPs	Control n (%)	Case n (%)	Total n (%)	Missing n (%)	χ^2	P
IL1B (rs1143634)	423 (42.6)	352 (43.9)	775 (42.2)	43 (2.3)	0.275	.600
IL1B (rs16944)	533 (53.6)	455 (56.7)	988 (53.8)	40 (2.2)	1.672	.196
IL1B rs3917356	224 (22.6)	144 (17.9)	368 (20.0)	40 (2.2)	5.724	.017
IL6 (IL6-07592RF)	655 (65.5)	542 (67.3)	1197 (65.2)	32 (1.7)	0.589	.443
IL6 (rs2069825)	468 (47.6)	416 (52.0)	884 (48.1)	53 (2.9)	3.303	.069
TF (F3-0012457)	408 (41.1)	348 (43.2)	756 (41.2)	38 (2.1)	0.713	.398
TF (rs1361600)	669 (71.9)	535 (73.5)	1204 (65.5)	179 (9.7)	0.421	.517
TF (rs3354)	394 (42.3)	331 (43.7)	725 (39.5)	148 (8.1)	0.257	.612
ADDU (ADDU-000614)	312 (31.1)	267 (32.9)	579 (31.5)	21 (1.1)	0.594	.441
ALOX (rs17216522)	86 (8.6)	69 (8.6)	155 (8.4)	35 (1.9)	0.000	1.000
ALOX (rs17222772)	545 (54.4)	420 (51.7)	965 (52.5)	23 (1.3)	1.177	.278
ALOX (rs17222814)	195 (19.9)	161 (20.2)	356 (19.4)	63 (3.4)	0.008	.928
ALOX (rs17222828)	740 (74.8)	571 (72.4)	1311 (71.4)	59 (3.2)	1.240	.266
CETP (rs17231513)	628 (63.6)	552 (68.8)	1180 (64.2)	47 (2.6)	5.231	.022
CETP (rs17231896)	552 (55.3)	416 (51.9)	968 (52.7)	37 (2.0)	1.840	.175
CETP (rs17237890)	727 (72.7)	631 (78.3)	1358 (73.9)	31 (1.7)	7.174	.007
NOS (rs1799983)	567 (57.1)	450 (56.1)	1017 (55.4)	42 (2.3)	0.139	.709
NOS (rs2070744)	619 (62.3)	542 (67.7)	1161 (63.2)	43 (2.3)	5.283	.022
ESR1 (rs2077647)	730 (73.3)	566 (70.4)	1296 (70.5)	37 (2.0)	1.709	.191
ESR1 (rs9340799)	583 (58.5)	436 (54.2)	1019 (55.5)	35 (1.9)	3.200	.074
MMP3 (rs522616)	302 (38.0)	256 (42.2)	558 (30.4)	436 (23.7)	2.425	.119
MMP3 (rs595840)	736 (74.7)	599 (75.3)	1335 (72.7)	56 (3.0)	0.041	.840
MMP3 (rs602128)	748 (74.7)	605 (75.0)	1353 (73.7)	29 (1.6)	0.004	.949
MMP3 (rs680753)	119 (14.1)	105 (15.0)	224 (12.2)	293 (15.9)	0.221	.638

MI = myocardial infarction; SNP = single nucleotide polymorphism.

Interaction between drinking and smoking status is the only significant protective effect (OR=0.459, 95% CI 0.220–0.956) for current smokers who drank in the past 1 to 2 years. Two polymorphisms, *rs17237890* (gene *CETP*) (OR=1.42, 95% CI 1.04–1.95) and *rs2070744* (gene *NOS3*) (OR=1.41, 95% CI 1.06–1.87), are found to be significantly associated with a risk of MI adjusted by other SNPs and environmental factors.

3.3.4. Univariate and Multivariable analysis for HC. Table 4 provides the univariate analysis and the unadjusted risk using χ^2 tests, ORs, and 95% CIs for HC for each selected individual genes/SNPs. When adjusted by other genes (SNPs) (see Table 5),

with dominant genetic model, the presence of at least one reference allele of *rs16944* (*IL-1 β*) increases the risk of HC by 27% (OR=1.27, 95% CI 1.05–1.55), and *rs17222772* (*ALOX*) increases the odds of the likelihood of HC (OR=1.42, 95% CI 1.03–1.97), respectively. The presence of at least one reference allele, of SNP *rs17231896* (*CETP*), reduces the risk by 21% (OR=0.79, 95% CI: 0.65–0.96). While adjusting for the clinical, behavioral, and demographic factors (see Table 6), the risk of HC for *rs17231896* (*CETP*) is 33% lower (OR=0.67, 95% CI: 0.51–0.88). The increase of effect size (from 21% to 33%) without or with the inclusion of the environmental factors may indicate modifiable effects exist. Table 7 provides the adjusted

Table 3

Significant clinical, behavior, demographic, and genetic factors associated with the risk of MI (N = 1834).

Variables	Model 1 (main effects)				Model 2 (main effects and interaction)			
	B	Odds ratio	95% CI		B	Odds ratio	95% CI	
			Lower	Upper			Lower	Upper
Gender (male vs female)	1.447	4.252 [‡]	3.087	5.857	1.451	4.269 [‡]	3.092	5.896
Hypertension (yes/no)	0.884	2.420 [‡]	1.819	3.219	0.887	2.427 [‡]	1.822	3.231
High cholesterol (yes vs no)	1.686	5.397 [‡]	4.069	7.160	1.682	5.378 [‡]	4.053	7.138
Family history of heart disease (no vs yes)	0.459	1.582 [†]	1.133	2.208	0.451	1.570 [†]	1.123	2.194
Drinking status in the past 1–2 y (current drinker vs noncurrent drinker)	–0.397	0.672 [†]	0.502	0.901	–0.405	0.667 [†]	0.495	0.899
Smoking status (current smoker vs nonsmoker)	1.223	3.399 [‡]	2.376	4.862	1.714	5.551 [‡]	3.068	10.041
Smoking status \times Drinking status					–0.779	0.459 [*]	0.220	0.956
<i>CETP</i> (<i>rs17237890</i>)	0.358	1.431 [*]	1.047	1.956	0.353	1.423 [*]	1.040	1.947
<i>NOS_2</i> (<i>rs2070744</i>)	0.346	1.413 [*]	1.063	1.877	0.343	1.409 [*]	1.060	1.873

Area under curve (AUC) and classification accuracy are similar for with or without interaction effects, both with or without environmental factors and genetics are different (72% vs 80%).

CI = confidence interval; MI = myocardial infarction.

* $P < .05$.

† $P < .01$.

‡ $P < .001$.

Table 4**Unadjusted odds ratios and 95% confidence intervals for HC by genes/SNPs (N = 1837).**

Genes/SNPs	χ^2	Odds ratio	95% confidence interval	
			Lower	Upper
<i>IL1B</i> (<i>rs16944</i>)	5.306*	1.252	1.038	1.511
<i>IL1B</i> (<i>rs3917356</i>)	1.968‡	1.148	0.953	1.383
<i>IL6</i> (<i>rs2069825</i>)	2.873†	1.190	0.979	1.448
<i>IL6</i> (<i>07592RF</i>)	3.347†	1.195	0.992	1.439
<i>IL6</i> (<i>rs2069825</i>)	3.009†	1.185	0.983	1.429
<i>ALOX</i> (<i>rs17222828</i>)	1.662‡	1.157	0.935	1.433
<i>ALOX</i> (<i>rs17222772</i>)	3.615†	1.359	1.002	1.842
<i>ALOX</i> (<i>rs17222828</i>)	2.265‡	1.160	0.962	1.399
<i>CETP</i> (<i>rs17231896</i>)	4.797*	0.807	0.668	0.974
<i>NOS3</i> (<i>rs1799983</i>)	1.683‡	1.138	0.943	1.374
<i>ESR1</i> (<i>rs9340799</i>)	2.277‡	0.862	0.714	1.039
<i>ESR1</i> (<i>rs2077647</i>)	6.074*	0.746	0.594	0.937

HC = high cholesterol; SNP = single nucleotide polymorphism.

* $P < .05$.† $P < .10$.‡ $P < .2$.

ORs and 95% CIs for HC predicted by significant genetic, demographic, behavioral, and clinical factors, with both main and interactions effects included. Both BMI and HTN are significant risk factors for HC. With the inclusions of gene–gene and gene–environmental interactions, the multiplicative effects between *rs17231896* (*CETP*) and *rs17222772* (*Alox*); *rs17231896* (*CETP*) and gender; MI and age; *rs9340799* (*ESR1*) and family history of heart diseases were found significantly associated with HC. As expected, after multiple

testing (e.g., Bonferroni or false discover rate) adjustments, the significant SNPs for MI or HC diminished to be statistically nonsignificant.

4. Discussion

Having a history of either HC or MI has a significant effect on each other in both directions. Common environmental (e.g., smoke, HTN, gender) and genetic factors (gene *CETP*) for both HC and MI are identified from separate analysis with or without adjustment of other significant factors. The association between *CETP* polymorphism (*rs17231896*, *rs17237890*, and *rs17231513*) has protective effects on HC, which is consistent and replicated with the meta-analysis study from ARIC data using genome-wide association study for the association between *CETP* and HDL, in which 3 SNPs from *CETP*, *rs708272*, ($OR = 0.95$), *rs5882* ($OR = 0.94$), and *rs1800775* ($OR = 0.95$) have been found to be significantly associated with HDL adjusted by both gene and environmental interactions.^[32] From a functional content point of view, *CETP* is known to regulate the process of transporting cholesterol from the peripheral arteries to the liver, which helps reduce the risk of CHD. Other studies have also found that *CETP* expression is regulated by multiple functional SNPs, affecting splicing and transcription, with increased or decreased *CETP* function, although they focus on selected different variants (*rs247616*, *rs173539*).^[42–44]

The modulations and joint effects among genes, clinical, behavioral, and other factors on HC and MI are complex. A study from 1995 demonstrated how the *CETP* gene regulating above process is influenced (positively) by alcohol, but subsequent studies have not been able to fully replicate the result.^[45] Our analysis confirms that alcohol and smoking habits have

Table 5**Adjusted odds ratios and 95% confidence intervals for HC predicted by SNPs/genes only adjusted by other gene/SNPs (N = 1837).**

Gene/SNPs	B (SE)	Wald	Odds ratio	95% confidence interval	
				Lower	Upper
<i>IL1B</i> (<i>rs16944</i>)	0.241 (0.100)	5.749*	1.272	1.045	1.549
<i>ALOX</i> (<i>rs17222772</i>)	0.353 (0.167)	4.484*	1.424	1.027	1.974
<i>CETP</i> (<i>rs17231896</i>)	−0.238 (0.101)	5.589*	0.788	0.647	0.960
<i>ESR1</i> (<i>rs9340799</i>)	−0.188 (0.101)	3.508	0.828	0.680	1.009

HC = high cholesterol; SE = standard error; SNP = single nucleotide polymorphism.

* $P < .05$.**Table 6****Adjusted odds ratios and 95% confidence intervals for HC predicted by genetic, demographic, behavior, and clinical factors with main effects (N = 1837).**

Variables	B (SE)	Wald	Odds ratio	95% confidence interval	
				Lower	Upper
History of MI	1.47 (0.14)	107.27*	4.35	3.295	5.749
HTN	0.63 (0.15)	17.25*	1.88	1.395	2.527
BMI	0.03 (0.01)	5.62‡	1.03	1.006	1.061
Age	−0.00 (0.01)	0.08	0.998	0.982	1.013
Gender/Male	−0.41 (0.15)	7.41†	0.66	0.492	0.891
<i>CETP</i> (<i>rs17231896</i>)	−0.39 (0.14)	8.27†	0.67	0.511	0.881
<i>ESR1</i> (<i>rs9340799</i>)	−0.25 (0.14)	3.25	0.78	0.595	1.022

Adjusted for history of MI, history of high BP, BMI, age, gender/male, *CETP* (*rs17231896*), *ESR1* (*rs9340799*).

BMI = body mass index; HC = high cholesterol; HTN = hypertension; MI = myocardial infarction; SE = standard error.

* $P < .001$.† $P < .01$.‡ $P < .05$.

Table 7

Adjusted odds ratios and 95% confidence intervals for HC predicted by genetic, demographic, behavior, and clinical factors, with main and interactions effects included (N = 1837).

Variables	B (SE)	Wald	Odds ratio	95% confidence interval	
				Lower	Upper
No history of MI	-4.555 (1.614)	7.966 [†]	0.011	0.000	0.249
No MI × age	0.064 (0.018)	12.132 [*]	1.066	1.028	1.105
Male × <i>CETP</i> (<i>rs17231896</i>)	0.749 (0.345)	4.716 [‡]	2.114	1.076	4.155
Fam_heart × <i>ESR1</i> (<i>rs9340799</i>)	-0.886 (0.383)	5.346 [‡]	0.412	0.194	0.874
<i>ALOX</i> (<i>rs17222772</i>) × <i>CET P</i> (<i>rs17231896</i>)	-1.178 (0.555)	4.506 [‡]	0.308	0.104	0.914

MI adjusted for age, gender, history of MI, BMI, family history of heart disease, hypertension, current drinker in past 1–2 years, *IL1B* (*rs16944*), *ALOX* (*rs17222772*), *CETP* (*rs17231896*), *ESR1* (*rs9340799*). BMI=body mass index; HC=high cholesterol; MI=myocardial infarction; SE=standard error.

^{*} *P* < .001.

[†] *P* < .01.

[‡] *P* < .05.

effects on MI when both genetic and other factors are adjusted. With inclusions of genes and their interactions, these joint and modulating genetic and lifestyle, clinical factors improve the prediction power for MI and HC. Moreover, both gene–gene and gene–environmental interaction effects were found to be significantly associated with MI and HC, which could be further examined through pathway/regulations and function analysis.^[46]

It is known that appropriate multiplicity adjustment (either false discovery rate or Bonferroni procedure) is crucial to guarantee the replicability and reproducibility of findings, which should be conducted in large-scale genome-wide association analysis to avoid the potential false positives for multi-trait association analysis.^[47] Despite the protective effect of *CETP* on HC, *rs17237890* (different SNP of *CETP*) has shown the increased risk on MI when other significant lifestyle factors were included such as smoking and alcohol assumptions. Future haplotype-pair analysis (to reduce the number of tests and increase the expected effect using the haplotypes instead of single genotypes) could be conducted for further validating plausibly gene (i.e., *CETP*) without using all the gene set (for better power) to see whether consistently associated with all the studied phenotypes.^[40,48] In addition, within populations of European Caucasian decent, population stratification issue will be assessed and weighted using genomic control with the inflation factor to avoid the spurious association findings.^[49,50]

The common alleles/genes associated with both MI and HC discovered include *rs16944* (*IL-1β*) with an independent effect without other genetic or environmental factors. However, when adding others, some effects either diminished to statistically nonsignificant or effect size modified, which may indicate some under-recognized role of modulate and modifiable effect of environmental exposures for the genes on the risk of both conditions. The diminished effects (OR with or close to zero) or other potentially missing heritable components of disease etiology may be better further fully explored either through epigenetics analysis^[46] with longitudinal data or more complex network and pathway-based analysis, for example, using Bayesian networks for analyzing the direct and indirect probabilistic causal associations to dissect the complex relationships among the significant factors (age, gender, BMI, HTN, smoking, alcohol drinking status, *CETP* gene, CHD/MI).

One limitation of this study is the notable missing data for lipid measures, such as total cholesterol, HDL, LDL, and TG (missing ranged from 44.4% to 46.9%), which is why having a history of HC (once or multiple) is chosen as the primary lipid outcome measure in this paper. Many published works regarding lipid–SNP associations use those measures as primary outcome.

Furthermore, for HC treated by meds (statin treatment or other cholesterol-lowering therapies) variable, only 26.6% of participants reported presence or absence such medication, which does not make it feasible to include it in the estimation model. Cases of familial dyslipidemia are not collected for this study and these related lipid variables may change estimated effects and should be addressed and taken into account in future studies.

References

- Wray NR, Goddard ME, Visscher PM. Prediction of individual genetic risk to disease from genome-wide association studies. *Genome Res* 2007;17:1520–8.
- Smith JA, Ware EB, Middha P, et al. Current applications of genetic risk scores to cardiovascular outcomes and subclinical phenotypes. *Curr Epidemiol Rep* 2015;2:180–90.
- Paynter NP, Chasman DI, Paré G, et al. Association between a literature-based genetic risk score and cardiovascular events in women. *JAMA* 2010; 303:631–637.
- Holm H, Thorleifsson G, Stefansson K. Genetic risk score and cardiovascular events in women. *JAMA* 2010;303:2032–3.
- Mozaffarian D, Benjamin EJ, Go AS, et al. Heart disease and stroke statistics–2015 update: a report from the American Heart Association. *Circulation* 2015;131:e29–322.
- Canto JG, Rogers WJ, Goldberg RJ, et al. NIMI Investigators Association of age and sex with myocardial infarction symptom presentation and in-hospital mortality. *J Am Med Assoc* 2012;307:813–22.
- Yusuf S, Hawken S, Ounpuu S, et al. INTERHEART Study Investigators. Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. *Lancet* 2004;364:937–52.
- Merry AH, Boer JM, Schouten LJ, et al. Smoking, alcohol consumption, physical activity, and family history and the risks of acute myocardial infarction and unstable angina pectoris: a prospective cohort study. *BMC Cardiovasc Disord* 2011;11:13.
- Kessler T, Erdmann J, Schunkert H. Genetics of coronary artery disease and myocardial infarction. *Curr Cardiol Rep* 2013;15:368.
- Heidemann C, Hoffmann K, Klipstein-Grobusch K, et al. Potentially modifiable classic risk factors and their impact on incident myocardial infarction: results from the EPIC-potsdam study. *Eur J Cardiovasc Prev Rehab* 2007;14:65–71.
- Kathiresan S, Srivastava D. Genetics of human cardiovascular disease. *Cell* 2012;148:1242–57.
- Abdulla J, Kober L, Abildstrom SZ, et al. Impact of obesity as a mortality predictor in high-risk patients with myocardial infarction or chronic heart failure: a pooled analysis of five registries. *Eur Heart J* 2008; 29:594–601.
- Labounty TM, Gomez MJ, Achenbach S, et al. Body mass index and the prevalence, severity, and risk of coronary artery disease: an international multicentre study of 13 874 patients. *Eur Heart J Cardiovasc Imaging* 2013;14:456–63.
- Lindquist P, Bengtsson C, Lissner L, et al. Cholesterol and triglyceride concentration as risk factors for myocardial infarction and death in women, with special reference to influence of age. *J Intern Med* 2002; 251:484–9.

- [15] Heart Protection Study Collaborative Group. MRC/BHF Heart Protection Study of cholesterol lowering with simvastatin in 20,536 high-risk individuals: a randomised placebo-controlled trial. *Lancet* 2002;360:7–22.
- [16] Li JX, Cao J, Lu XF, et al. The effect of total cholesterol on myocardial infarction in Chinese male hypertension population. *Biomed Environ Sci* 2010;23:37–41.
- [17] Kim DS, Burt AA, Ranchalis JE, et al. Novel gene-by-environment interactions: APOB and NPC1L1 variants affect the relationship between dietary and total plasma cholesterol. *J Lipid Res* 2013;54:1512–20.
- [18] Kathiresan S, Melander O, Guiducci C, et al. Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. *Nat Genet* 2008;40:189–97.
- [19] Kathiresan S, Melander O, Anevski D, et al. Polymorphisms associated with cholesterol and risk of cardiovascular events. *N Engl J Med* 2008;358:1240–9.
- [20] Kathiresan S, Willer CJ, Peloso GM, et al. Common variants at 30 loci contribute to polygenic dyslipidemia. *Nat Genet* 2009;41:56–65.
- [21] Sandhu MS, Waterworth DM, Debenham SL, et al. LDL-cholesterol concentrations: a genome-wide association study. *Lancet* 2008;371:483–549.
- [22] Matsuoka R, Abe S, Tokoro F, et al. Association of six genetic variants with myocardial infarction. *Int J Mol Med* 2015;35:1451–9.
- [23] Wallace C, Newhouse SJ, Braund P, et al. Genome-wide association study identifies genes for biomarkers of cardiovascular disease: serum urate and dyslipidemia. *Am J Hum Genet* 2008;82:139–49.
- [24] Keenan TE, Rader DJ. Genetics of lipid traits and relationship to coronary artery disease. *Curr Cardiol Rep* 2013;15:396.
- [25] Keavney B, Palmer A, Parish S, et al. International Studies of Infarct Survival (ISIS) Collaborators. Lipid-related genes and myocardial infarction in 4685 cases and 3460 controls: discrepancies between genotype, blood lipid concentrations, and coronary disease risk. *Int J Epidemiol* 2004;33:1002–13.
- [26] Morgan TM, Krumholz HM, Lifton RP, et al. Nonvalidation of reported genetic risk factors for acute coronary syndrome in a large-scale replication study. *JAMA* 2007;297:1551–61.
- [27] Cohen JC, Boerwinkle E, Mosley TH, et al. Sequence variations in PCSK9, low LDL, and protection against coronary heart disease. *N Engl J Med* 2006;354:1264–72.
- [28] Willer CJ, Sanna S, Jackson AU, et al. Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nat Genet* 2008;40:161–9. (2007).
- [29] Holmes M, Keating B, Asselbergs F, et al. Mendelian randomization of blood lipids for coronary heart disease. *Eur Heart J* 2015;36:539–50.
- [30] Brown G, Albers JJ, Fisher LD, et al. Regression of coronary artery disease as a result of intensive lipid-lowering therapy in men with high levels of apolipoprotein B. *N Engl J Med* 1990;323:1289–98.
- [31] Ha V, Sievenpiper J, de Souza RJ, et al. Effect of dietary pulse intake on established therapeutic lipid targets for cardiovascular risk reduction: a systematic review and meta-analysis of randomized controlled trials. *CMAJ* 2014;186:E252–62.
- [32] Brautbar A, Ballantyne CM, Lawson K, et al. Impact of adding a single allele in the 9p21 locus to traditional risk factors on reclassification of coronary heart disease risk and implications for lipid-modifying therapy in the Atherosclerosis Risk in Communities (ARIC) study. *Circ Cardiovasc Genet* 2009;2:279–85.
- [33] Teslovich TM, Musunuru K, Smith AV, et al. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* 2010;466:707–13.
- [34] Mäkelä KM, Seppälä I, Hernesniemi JA, et al. Genome-wide association study pinpoints a new functional apolipoprotein B variant influencing oxidized low-density lipoprotein levels but not cardiovascular events: AtheroRemo Consortium. *Circ Cardiovasc Genet* 2013;6:73–81.
- [35] Elbers CC, Guo Y, Tragante V, et al. Gene-centric meta-analysis of lipid traits in African, East Asian and Hispanic populations. *PLoS One* 2012;7:e50198.
- [36] Clarke A, Donahue RP, Dorn J, et al. Franco OH Cross-cultural comparison of correlates of quality of life and health status: the Whitehall II Study (UK) and the Western New York Health Study (US). *Eur J Epidemiol* 2012;27:255–65.
- [37] McCann SE, Sempos C, Freudenheim JL, et al. Alcoholic beverage preference and characteristics of drinkers and nondrinkers in western New York (United States). *Nutr Metab Cardiovasc Dis* 2003;13:2–11.
- [38] Bis JC, Heckbert SR, Smith NL, et al. Variation in inflammation-related genes and risk of incident non-fatal myocardial infarction or ischemic stroke. *Atherosclerosis* 2008;198:166–73.
- [39] Gigante B, Bennet AM, Leander K, et al. The interaction between coagulation factor 2 receptor and interleukin 6 haplotypes increases the risk of myocardial infarction in men. *PLoS One* 2010;5:e11300.
- [40] Gaetano M, Quacquaruccio G, Di Castelnuovo A, et al. Haplotypes and haplotype-pairs of IL-1 beta and IL-6 genes and risk of non fatal myocardial infarction in the Western New York Acute MI Study. *Thromb Haemost* 2011;106:1231–3.
- [41] Roger VL, Go AS, Lloyd-Jones DM, et al. American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Heart disease and stroke statistics—2012 update: a report from the American heart association. *Circulation* 2012;125:e2–20.
- [42] Suhy A, Hartmann K, Papp A, et al. Regulation of cholesteryl ester transfer protein expression by upstream polymorphisms: reduced expression associated with rs247616. *Pharmacogenet Genomics* 2015;25:394–401.
- [43] Suhy A, Hartmann K, Newman L, et al. Genetic variants affecting alternative splicing of human cholesteryl ester transfer protein. *Biochem Biophys Res Commun* 2014;443:1270–4.
- [44] Papp AC, Pinsonneault JK, Wang D, et al. Transfer protein (CETP) polymorphisms affect mRNA splicing, HDL levels, and sex-dependent cardiovascular risk. *PLoS One* 2012;7:e31930.
- [45] Tolstrup JS, Gronbaek M, Nordestgaard BG. Alcohol intake, myocardial infarction, biochemical risk factors, and alcohol dehydrogenase genotypes. *Circ Cardiovasc Genet* 2009;2:507–14.
- [46] Zanon P, Khetarpal SA, Larach DB, et al. CHD Exome+ Consortium; CARDIoGRAM Exome Consortium; Global Lipids Genetics Consortium. Rare variant in scavenger receptor BI raises HDL cholesterol and increases risk of coronary heart disease. *Science* 2016;351:1166–71.
- [47] Peterson C, Bogomolov M, Benjamini Y, et al. Many phenotypes without many false discoveries: error controlling strategies for multitrait association studies. *Genet Epidemiol* 2016;40:45–56.
- [48] De Gaetano M, Donati MB, Trevisan M, et al. Tissue factor gene polymorphisms and haplotypes and the risk of ischemic vascular events: four studies and a meta-analysis. *J Thromb Haemost* 2009;7:1465–71.
- [49] Campbell CD, Ogburn EL, Lunetta KL, et al. Demonstrating stratification in a European American population. *Nat Genet* 2005;37:868–72.
- [50] Price AL, Butler J, Patterson N, et al. Discerning the ancestry of European Americans in genetic association studies. *PLoS Genet* 2008;4: e236. <https://doi.org/10.1371/journal.pgen.0030236>.