Genetics

Genetic Risk Score and Risk of Myocardial Infarction in Hispanics

Lu Qi, MD, PhD; Jiantao Ma, MD; Qibin Qi, PhD; Jaana Hartiala, MS; Hooman Allayee, PhD; Hannia Campos, PhD

Background—Genome-wide association studies have identified loci associated with coronary heart disease in whites of European ancestry. This study evaluated whether genetic markers previously identified in whites are associated with nonfatal acute myocardial infarction (MI) in Hispanics.

Methods and Results—Cases (n=11005) with a first nonfatal acute MI and population-based controls (n=2096) living in Costa Rica were studied. Fourteen single-nucleotide polymorphisms were genotyped. Seven single-nucleotide polymorphisms at 3 independent loci showed significant associations with MI. The odds ratios for the loci with the strongest associations were 1.16 (95% confidence interval [CI], 1.05 to 1.27) for rs4977574 (CDKN2A/2B), 1.15 (95% CI, 1.03 to 1.29) for rs646776 (CELSR2-PSRC1-SORT1), and 1.22 (95% CI, 1.08 to 1.38) for rs501120 (CXCL12); the corresponding PARs were 6.8%, 10.5%, and 15.2%; respectively. We developed a genetic risk score by summing the number of the top 3 associated risk alleles. The OR for MI per genetic risk score unit was 1.18 (95% CI, 1.11 to 1.25; P=4.83×10⁻⁸). Discrimination of MI was significantly improved (P=0.02) when the genetic risk score was added to a model including clinical predictors. However, the increase in the area under the receiver-operating characteristic curve after the genetic risk score was added was moderate, from 0.67 (95% CI, 0.65 to 0.69) to 0.68 (95% CI, 0.66 to 0.70).

Conclusions—These results indicate both the consistency and disparity of genetic effects on risk of MI between Hispanic and white populations. The improvement in the identified genetic markers on discrimination of MI in Hispanics was modest. (Circulation. 2011;123:374-380.)

Key Words: epidemiology ■ ethnicity ■ genetic association ■ genetics ■ myocardial infarction ■ Hispanic

Clinical Perspective on p 380

In the past few years, genome-wide association studies (GWAS) have identified several novel susceptibility loci for CHD.⁷–¹¹ Although most of the identified genetic variants showed reliably consistent associations with CHD in whites of European ancestry, little is known about their effects in other populations such as Hispanics. To date, most variants identified have shown modest effects on cardiovascular risk. Nevertheless, when multiple genetic markers are considered together, they might be useful to improve the identification of individuals at high risk of disease.¹² However, few studies have investigated the joint genetic effects in the discrimination of CHD status.¹³

The purpose of the present study was to examine the associations between the common single-nucleotide polymorphisms (SNPs) reported by GWAS in whites and the risk of nonfatal acute myocardial infarction (MI) in 1898 cases with MI and 2096 population-based controls from a Hispanic Costa Rican population, which has a genetic structure that is different from that of whites.¹⁴,¹⁵ We also examined the joint genetic effects of these SNPs and their roles in the discrimination of MI.

Methods

Study Population

The Hispanic study population used in the present study has been described previously.¹⁶,¹⁷ Briefly, the participants, from 34 counties...
in the Central Valley of Costa Rica, covered a full range of socioeconomic levels, as well as urban, peri-urban, and rural lifestyles. Eligible case subjects were adult residents who were diagnosed as survivors of a first acute MI by 2 independent cardiologists at any of the 6 recruiting hospitals in the catchment area between 1994 and 2004. All cases met the World Health Organization criteria for MI, which require typical symptoms plus either elevations in cardiac enzyme levels or diagnostic changes in the ECG. One free-living control subject for each case, matched for age (±5 years), sex, and area of residence (county), was randomly selected with the use of information available at the National Census and Statistics Bureau of Costa Rica. Because of the comprehensive social services provided in Costa Rica, all persons had access to medical care without regard to income. Therefore, control subjects came from the source population that gave rise to the cases and were not likely to have had CHD that was not diagnosed because of poor access to medical care. In total, 1989 MI cases and 2096 controls with genotyping data were included in the present study. All subjects gave informed consent on documents approved by the Human Subjects Committee of the Harvard School of Public Health and the University of Costa Rica.

All study participants were visited in their homes for collection of data. Information on sociodemographic characteristics, smoking, physical activity, and medical history was collected through an interview using closed-ended questionnaires. Alcohol consumption was assessed by a validated food frequency questionnaire. All anthropometric measurements were taken on subjects wearing light clothing and no shoes, collected in duplicate, and averaged out for analyses. Nonstretching fiberglass or metal tapes were used to measure the waist (smallest horizontal trunk–circumference) and hip (largest horizontal circumference around the hip and buttocks) girths. Biological samples were always collected in the morning after an overnight fast. Blood samples (20 mL) were drawn in 0.1% EDTA-containing tubes after a 12- to 14-hour fast and immediately stored at 4°C. Within 36 hours, the samples were centrifuged at 2500 rpm for 20 minutes at 4°C to isolate and place in aliquots plasma and white blood cells. The samples were then sealed and stored under N₂ at −80°C until analysis. Data and sample collection took place, on average, 3 weeks after hospital discharge.

SNP Selection and Genotype Determination
Genomic DNA was extracted from the buffy coat fraction of centrifuged blood with the QIAmp Blood Kit (Qiagen, Chatsworth, CA). For the present study, we selected SNPs previously associated with coronary artery disease and/or MI from GWAS.7–11 To increase the a priori likelihood of detecting associations in Costa Ricans, we selected SNPs that were associated with coronary artery disease and/or MI in at least 2 studies and had values of \( r^2 > 0.8 \) in the white, African, and Asian populations from the HAPMAP project for each respective locus. As a result, 14 SNPs were genotyped in the Costa Ricans: rs4977574, rs10757274, rs2383206, rs1333049 (CDKN2A/2B), rs646776, rs599839 (CELSR2-PSRC1-SORT1), rs501120, rs1746048 (CXCL12), rs2259816 (HNF1A, C12orf43), rs9818870 (MRAS), rs2048327 (SLC22A3), rs3127599 (LPA2), rs7767084, and rs1075578 (LPA). Genotyping was performed with the TaqMan Allelic Discrimination System from Applied Biosystems, Inc (Foster City, CA).13,19 using custom genotyping assays from the “assays by design” service of Applied Biosystems, Inc. Replicate quality control samples yielded >99% concordance, and the overall call rate was >95%.

Statistical Analyses
A \( \chi^2 \) test was used to assess whether the SNPs were in Hardy-Weinberg equilibrium and to determine differences in genotype frequencies between MI cases and controls. Because some of the matched case-control pairs were broken as a result of missing genotyping, an unconditional logistic regression was used to calculate odds ratios (ORs), adjusting for age, sex, area of residence, waist-to-hip ratio, smoking, alcohol consumption, physical activity, total calories, and family history of CHD. Secondary analyses used conditional logistic regression with 1864 case-control pairs for whom complete genotypes were available generated similar results. General linear models were used to compare mean values of quantitative traits across groups. Population-attributable risk (PAR%) was estimated for SNPs as follows: \( \text{PAR} \% = 100 \times \frac{p(1-OR)-1}{p(OR)-1} \), where \( p \) is the frequency of the allele associated with MI among the control subjects. Haplotype analysis was conducted with the THESIAS program, which is based on the stochastic EM algorithm.20

A genetic risk score (GRS) was calculated with the 3 SNPs showing the strongest association with MI. For this purpose, we assumed that each SNP was independently associated with risk according to an additive genetic model, which performs well even when the true genetic model may not be known or may be incorrectly specified.21 We assumed that each SNP contributed equally to the risk of MI and calculated the GRS by summing the number of risk alleles at each locus. In sensitivity analyses, we also calculated a weighted GRS by multiplying the number of risk alleles at each locus (0, 1, or 2) for the corresponding \( \beta \) coefficient from additive multivariate logistic regression model and then summing the products. We used receiver-operating characteristic curve analysis (plots were made with the receiver-operating characteristic curve function in Proc Logistic, SAS version 9.2, SAS Institute, Inc, Cary, NC) and calculated the area under the curve (also known as the C statistic) to evaluate discrimination. We tested the null hypothesis of no difference between the AUCs from models incorporating conventional risk factors (age, sex, area of residence, waist-to-hip ratio, family history of MI, smoking, alcohol intake, total calories, and physical activity) with and without the GRS. The SAS statistical package was used for all analyses. Two-sided values of \( P < 0.05 \) were considered significant.

Results
Table 1 shows the characteristics of the participants by MI status. The MI cases engaged in less physical activity, consumed more alcohol and total calories, had higher waist-to-hip ratio, and were more likely to be current smokers compared with the controls. All 14 SNPs tested were commonly distributed in the study samples, with minor allele frequency ranging from 0.09 to 0.50, and fit Hardy-Weinberg equilibrium. However, the allele frequencies of some SNPs, including rs646776, rs599839, rs9818870, rs2048327,
rs3127599, rs10755578, and rs2383206, did show significant differences ($P < 0.05$) between our Hispanic population and the white population used in the HAPMAP project (Centre d’Etude du polymorphisme Humain Trios of European Descent; Table 2). The 2 SNPs at chromosome 1 (rs646776 and rs599839) and the 2 SNPs at chromosome 10 (rs501120 and rs1746048) are in strong linkage disequilibrium (LD; $r^2 > 0.8$), whereas the 4 SNPs at chromosome 6 (rs2048327, rs3127599, rs7767084, and rs10755578) are in weak LD ($r^2 < 0.4$). At chromosome 9, SNPs rs10757274 and rs4977574 are in nearly perfect LD ($r^2 = 0.99$), and the pairwise $r^2$ among other SNPs are all < 0.7.

We first examined the associations between individual genetic variants and MI risk. Seven SNPs at 3 loci, including CELSR2-PSRC1-SORT1 (rs646776), CXCL12 (rs501120, rs1746048), and CDKN2A/2B (rs10757274, rs4977574, rs2383206, and rs1333049), showed significant associations with MI risk, with ORs (95% confidence interval [CI]) ranging from 1.12 (95% CI, 1.02 to 1.24) to 1.22 (95% CI, 1.08 to 1.38; Table 2). Importantly, the direction of these associations was consistent with previous GWAS, and adjustment for matching variables and conventional risk factors such as age, sex, area of residence, waist-to-hip ratio, smoking, alcohol consumption, physical activity, total calories, and family history of CHD did not change the results.

Previous studies had reported that SNPs at the SLC22A3-LPAL2-LPA gene cluster on chromosome 6q26 were associated with CHD risk as part of 2 risk haplotypes. Therefore, we inferred the same haplotypes based on rs2048327, rs3127599, rs7767084, and rs10755578 (Table 3). The most common haplotype in the Costa Ricans was TCTC, similar to white populations. Compared with TCTC, the ORs for the 2 previously reported risk haplotypes CCTC and CTTG were 1.05 (95% CI, 0.92 to 1.20) and 0.98 (95% CI, 0.81 to 1.17), respectively.
respectively. In addition, we observed that a rarer haplotype, TTTC, was also associated with an increased risk of MI (OR, 1.94; 95% CI, 1.15 to 3.28) compared with the most common haplotype, TCTC. However, the overall test for these haplotype associations did not achieve significance ($P = 0.15; \chi^2 = 10.8$; with 7 df).

We next evaluated the joint effects of the best-associated SNPs at the 3 loci showing evidence of association with MI (rs4977574 at CDKN2A/2B; rs646776 at CELSR2-PSRC1-SORT1, and rs501120 at CXCL12). We calculated a GRS representing the sum of the risk alleles and observed that increasing GRS was significantly associated with higher risk of MI, with an OR of 1.18 (95% CI, 1.11 to 1.25), corresponding to 1 risk allele ($P = 4.83 \times 10^{-8}$; Table 4). Compared with subjects with GRS of 1 (1.3% of the control subjects), there was a stepwise increase in the risk of MI with increasing GRS, and those with GRS of 5 (25.6% of the control subjects) and 6 (7.4% of the control subjects) had a nearly 2-fold higher risk of MI (Table 4). Adjustment for covariates did not appreciably change these results (Figure 1 and Table 4). We did not detect interactions between GRS and waist-to-hip ratio (low versus high by median value), smoking (current versus never and past smoker), alcohol consumption (low versus high by median value), and physical activity (low versus high by median value).

![Figure 1](http://circ.ahajournals.org/)

**Figure 1.** Combined effect of risk alleles on MI. For each individual, the number of best-associated risk alleles of 3 loci CDKN2A/2B, CELSR2-PSRC1-SORT1, and CXCL12 was summed to represent an individual’s GRS (range, 1 to 6). Individuals in each risk allele category are shown along the x axis, and the odds ratios are plotted on the y axis (left), with the error bars representing the 95% CIs. The histogram (y axis on right) represents the frequency of each GRS category.

![Figure 2](http://circ.ahajournals.org/)

**Figure 2.** Receiver-operating characteristic curves for MI. The curves are based on logistic regression models incorporating conventional risk factors (age, sex, area of residence, waist-to-hip ratio, family history of MI, smoking, alcohol intake, total calories, and physical activity) with and without the GRS. AUC indicates area under the curve.

We further examined whether addition of the genetic markers improved the discrimination of MI status. Figure 2 presents the receiver-operating characteristic curves for the logistic regression model incorporating conventional clinical risk factors (age, waist-to-hip ratio, smoking, alcohol consumption, physical activity, total calories, and family history of CHD) with and without the inclusion of the GRS. The area under the receiver-operating characteristic curve, which measures how well the model discriminates between case subjects and control subjects, was 0.67 (95% CI, 0.65 to 0.69) for the conventional model and slightly but significantly increased to 0.68 (95% CI, 0.66 to 0.70) when the GRS was added ($P = 0.02$).

We finally assessed whether the genetic markers could account for part of the variance of CHD family history and explain its predisposing effect on MI. A family history of CHD was significantly associated with an increased risk of MI (OR, 1.63; 95% CI, 1.30 to 2.05) with adjustment for other covariates. However, we did not observe a significant

<table>
<thead>
<tr>
<th>GRS</th>
<th>Cases (n=1911)</th>
<th>Controls (n=2024)</th>
<th>OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>16 (0.8)</td>
<td>26 (1.3)</td>
<td>1.0</td>
</tr>
<tr>
<td>2</td>
<td>83 (4.3)</td>
<td>138 (6.8)</td>
<td>0.98 (0.49–1.93)</td>
</tr>
<tr>
<td>3</td>
<td>344 (18.0)</td>
<td>423 (20.9)</td>
<td>1.32 (0.70–2.50)</td>
</tr>
<tr>
<td>4</td>
<td>684 (35.8)</td>
<td>768 (37.9)</td>
<td>1.45 (0.77–2.72)</td>
</tr>
<tr>
<td>5</td>
<td>618 (32.3)</td>
<td>519 (25.6)</td>
<td>1.93 (1.03–3.65)</td>
</tr>
<tr>
<td>6</td>
<td>166 (8.7)</td>
<td>150 (7.4)</td>
<td>1.80 (0.93–3.48)</td>
</tr>
<tr>
<td>Per allele</td>
<td>1.18 (1.11–1.25)</td>
<td>1.19 (1.12–1.27)</td>
<td>1.18 (1.11–1.25)</td>
</tr>
</tbody>
</table>

$P$ for trend: $4.83 \times 10^{-8}$, $1.49 \times 10^{-7}$

In GRS analyses, individuals with missing genotyping were excluded.

*Adjusted for age, sex, area of residence, waist-to-hip ratio, smoking, alcohol consumption, physical activity, and family history of CHD.
association of family history with the individual genetic markers and GRS.

Discussion
In this case-control genetic study, we report that common variants of 3 loci identified from GWAS in whites, CELSR2-PSRC1-SORT1, CXCL12, and CDKN2A/2B, were also significantly associated with MI risk in a Hispanic population. In addition, we found that the improvement in MI case-control discrimination by the addition of a GRS combining 3 genetic markers significantly associated with MI was modest although statistically significant.

The directions of the associations for the 3 significantly associated loci were consistent with previous reports in Europeans.9,10 Although the effect sizes of CDKN2A/2B variants appear weaker than those observed in Europeans, those of CELSR2-PSRC1-SORT1 and CXCL12 are comparable.9,10 Thus, our results indicate that the genetic effects for these susceptibility loci for CHD may persist across different ethnic groups.

SNPs in some other loci such as HNF1A and MRAS were not significantly related to CHD risk in Hispanics. In addition, we did not confirm the associations between haplotypes at the SLC22A3-LPA-LPA locus11 with MI risk. One possibility for these discrepancies may be due in part to ethnic differences in genetic structure. Such population differences in genetic associations have been exemplified by some studies.22 It is likely that most identified SNPs in white populations themselves are not the causal variants but the markers in LD with the effective variants. Our study samples are derived from the admixture of a relatively small number of founders of Southern European, Amerindian, and West African origin. This admixture is likely to result in unique underlying linkage structure. Interestingly, we did not find regional differences in the admixture proportions or between cases and controls. The mean individual ancestry proportions in cases and controls were 57.5% versus 57.8% for the European, 38.4% versus 38.3% for the Amerindian, and 4.1% versus 3.8% for the West African ancestry.15 Thus, it is possible that the markers we tested are correlated with the causal variants, if any, in white but not Hispanic populations. These observations highlight the importance of comprehensively fine-mapping susceptibility loci across different ethnicities.

Another potential explanation for the heterogeneity in the genetic associations across ethnicities could be modification by environmental factors such as dietary and behavioral habits.23,24 Future studies are warranted to investigate such potential gene-environment interactions and to improve our understanding of the influence of environment on the penetrance of genetic risk factors. Finally, it is possible that the definitions of disease outcome were not entirely identical in previous GWAS7–11 and in our study. This may also partly explain some between-study variation in the associations.

In the analyses of the joint effects of these variants on MI risk, we found that a GRS based on the 3 MI-associated loci was associated with a significant 18% increased MI risk per risk allele. Individuals harboring 5 or 6 risk alleles, which constituted 25.6% and 7.4% of control subjects, had up to a 2-fold higher risk of MI compared with those with 1 risk allele (1.3% of controls). These findings suggest that a significant fraction of this Hispanic population carries multiple risk alleles and may help to identify individuals with a high genetic predisposition to MI. However, our analyses indicate that the addition of genetic risk factors did not materially improve the discrimination of MI status, although the test for such an improvement was statistically significant. These results are in line with observations that currently identified genetic variants might have low discriminatory ability and contribute modestly to disease prediction.12,25

The genetic makers tested in our study did not explain the association between family history of CHD and MI risk. A similar pattern has been observed for other common disorders such as type 2 diabetes mellitus,12 suggesting the existence of other, as-yet unidentified genetic determinants for the “missing heritability.”26

Several limitations of the present study need to be considered. The genetic effects on complex diseases such as CHD are in general modest. Although the present study had a reasonable number of MI cases, we cannot exclude the possibility that some SNPs associated with MI have small effect size and therefore are not detectable in this sample. In addition, because we designed our study with specific prior hypotheses based on SNPs reported in European whites, we interpreted our findings primarily on the basis of consistency between our results and those in previous GWAS.7–11 However, even after consideration of a conservative Bonferroni adjustment for the number of statistical tests carried out (0.05/14=0.004), 2 loci on chromosomes 10q11 and 9p21.3 were still significantly associated with MI, demonstrating the robust nature of these results. Lastly, because of potential differences in LD structure between Hispanics and Europeans, it is possible that the SNPs we examined are not correlated with the causal genetic variants in our study samples compared with European whites.14 In addition, we acknowledge that some reported CHD-associated genetic variants were not included in our study, which warrants future studies to fine-map these previously reported regions by comprehensive genotyping and to test more additional relevant loci.

Conclusions
Our findings demonstrate that certain susceptibility loci identified in Europeans may also increase MI risk in Hispanics. The improvement in disease discrimination by incorporating associated genetic markers with conventional risk factors was modest. Our data suggest that more association analyses in various ethnicities are required before generalizable conclusions can be made regarding genetic effects on CHD across multiple populations.

Acknowledgments
We thank all the participants of the study.
Sources of Funding

This study was supported by grant HL071981 from the National Institutes of Health, an American Heart Association Scientist Development Award (0730094N), and grant DK46200 from the Boston Obesity Nutrition Research Center. A portion of this work was conducted in a facility constructed with support from the National Institutes of Health Research Facilities Improvement Program (RR10600-01, CA62528-01, RR14514-01) of the National Center for Research Resources.

Disclosures

None.

References


Coronary heart disease disproportionately affects certain ethnic groups. Hispanics are affected by excessive rates of cardiovascular risk factors such as diabetes mellitus, overweight and obesity, dyslipidemia, and hypertension. This study evaluated whether genetic markers identified from genome-wide association studies in whites were associated with myocardial infarction (MI) in Hispanics. We determined 14 variations in 1989 cases with a first nonfatal acute MI and 2096 population-based controls. Our results indicated that 7 single-nucleotide polymorphisms at 3 independent loci were associated with MI risk. A genetic risk score summing the number of the associated risk alleles was also associated with MI, per unit related to 18% (11% to 25%) increased risk. The improvement in genetic markers on discrimination of MI, represented by the area under the receiver-operating characteristic curve, was modest but significant. Our findings highlight the importance of examining ethnic differences in genetic susceptibility to MI and indicate that consideration of multiple susceptibility signals may identify individuals with a markedly increased risk of MI.
Genetic Risk Score and Risk of Myocardial Infarction in Hispanics
Lu Qi, Jiantao Ma, Qibin Qi, Jaana Hartiala, Hooman Allayee and Hannia Campos

Circulation. published online January 17, 2011;
Circulation is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2011 American Heart Association, Inc. All rights reserved.
Print ISSN: 0009-7322. Online ISSN: 1524-4539

The online version of this article, along with updated information and services, is located on the
World Wide Web at:
http://circ.ahajournals.org/content/early/2011/01/17/CIRCULATIONAHA.110.976613

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation is online at:
http://circ.ahajournals.org/subscriptions/