ORIGINAL ARTICLE

A genome-wide association study of a coronary artery disease risk variant

Ji-Young Lee^{1,16}, Bok-Soo Lee^{2,16}, Dong-Jik Shin^{3,16}, Kyung Woo Park^{4,16}, Young-Ah Shin¹, Kwang Joong Kim¹, Lyong Heo¹, Ji Young Lee¹, Yun Kyoung Kim¹, Young Jin Kim¹, Chang Bum Hong¹, Sang-Hak Lee³, Dankyu Yoon⁵, Hyo Jung Ku², Il-Young Oh⁴, Bong-Jo Kim¹, Juyoung Lee¹, Seon-Joo Park¹, Jimin Kim¹, Hye-kyung Kawk¹, Jong-Eun Lee⁶, Hye-kyung Park¹, Jae-Eun Lee¹, Hye-young Nam¹, Hyun-young Park⁷, Chol Shin⁸, Mitsuhiro Yokota⁹, Hiroyuki Asano¹⁰, Masahiro Nakatochi¹¹, Tatsuaki Matsubara¹², Hidetoshi Kitajima¹³, Ken Yamamoto¹³, Hyung-Lae Kim¹⁴, Bok-Ghee Han¹, Myeong-Chan Cho¹⁵, Yangsoo Jang^{3,17}, Hyo-Soo Kim^{4,17}, Jeong Euy Park^{2,17} and Jong-Young Lee^{1,17}

Although over 30 common genetic susceptibility loci have been identified to be independently associated with coronary artery disease (CAD) risk through genome-wide association studies (GWAS), genetic risk variants reported to date explain only a small fraction of heritability. To identify novel susceptibility variants for CAD and confirm those previously identified in European population, GWAS and a replication study were performed in the Koreans and Japanese. In the discovery stage, we genotyped 2123 cases and 3591 controls with 521 786 SNPs using the Affymetrix SNP Array 6.0 chips in Korean. In the replication, direct genotyping was performed using 3052 cases and 4976 controls from the KItaNagoya Genome study of Japan with 14 selected SNPs. To maximize the coverage of the genome, imputation was performed based on 1000 Genome JPT + CHB and 5.1 million SNPs were retained. CAD association was replicated for three GWAS-identified loci (1p13.3/*SORT1* (rs599839), 9p21.3/*CDKN2A/2B* (rs4977574), and 11q22.3/*PDGFD* (rs974819)) in Koreans. From GWAS and a replication, SNP rs3782889 showed a strong association (combined $P = 3.95 \times 10^{-14}$), although the association of SNP rs3782889 doesn't remain statistically significant after adjusting for SNP rs11066015 (proxy SNP with *BRAP* ($r^2 = 1$)). But new possible CAD-associated variant was observed for rs9508025 (*FLT1*), even though its statistical significance did marginally reach at the genome-wide a significance level (combined $P = 6.07 \times 10^{-7}$). This study shows that three CAD susceptibility loci, which were previously identified in European can be directly replicated in Koreans and also provides additional evidences implicating suggestive loci as risk variants for CAD in East Asian.

Journal of Human Genetics (2013) 58, 120–126; doi:10.1038/jhg.2012.124; published online 31 January 2013

Keywords: coronary artery disease; genome-wide association study; polymorphism

INTRODUCTION

Coronary artery disease (CAD) being the leading causes of disability and mortality world-wide,¹ is a complex polygenic disease in which genetic factors have a significant role in disease etiology.² It has been estimated that heritable factors account for 30–60% of the inter-individual variation in the risk of CAD.³ Until now, genome-wide

association studies (GWAS) have identified over 30 common variants that are associated with the risk of coronary artery disease, as reported in NHGRI catalog (<http://www.genome.gov/gwasstudies>) and literature review.⁴ A number of loci associated with CAD have been found but the most studies have recently been reported to identify susceptibility loci in population of European descent.^{5–11} GWASs in

¹Center for Genome Science, National Institute of Health, Chungcheongbuk-do, Korea; ²Division of Cardiology, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Korea; ³Cardiology Division, Department of Internal Medicine, Cardiovascular Genome Center, Yonsei University College of Medicine, Seoul, Korea; ⁴Department of Internal Medicine, Cardiovascular Center, Seoul National University Hospital, Seoul, Korea; ⁵Interdisciplinary Program in Bioinformatics, Seoul National University, Seoul, Korea; ⁶DNA link, Seoul, Korea; ⁷Division of Cardiovascular Diseases, Center for Biomedical Sciences, National Institute of Health, Chungcheongbuk-do, Korea; ⁸Division of Pulmonary and Critical Care Medicine, Korea University Ansan Hospital, Ansan, Korea; ⁹Department of Genome Science, Aichi-Gakuin University, School of Dentistry, Nagoya, Japan; ¹⁰Department of Cardiology, Graduate School of Medicine, Nagoya University, Nagoya, Japan; ¹¹Department of Biotechnology, Nagoya University School of Engineering, Nagoya, Japan; ¹²Department of Internal Medicine, Aichi-Gakuin University, School of Dentistry, Nagoya, Japan; ¹³Research Center for Genetic Information, Medical Institute of Bioregulation, Kyushu University, Fukuoka, Japan; ¹⁴Department of Biochemistry, School of Medicine, Ewha Womans University, Seoul, Korea and ¹⁵National Institute of Health, Osong Health Technology Administration complex, Chungcheongbuk-do, Korea

¹⁶These authors contributed equally to this work.

¹⁷These authors jointly directed this work.

Correspondence: Dr J-Y Lee, Center for Genome Science, National Institute of Health, Chungcheongbuk-do 363–951, Republic of Korea.

E-mail: leeyj63@nih.go.kr

Received 23 July 2012; revised 2 October 2012; accepted 13 October 2012; published online 31 January 2013

Asians have led to the discovery of genetic variants in *BRAP* at 12q24.12 (Ozaki *et al.*¹²) and *C5orf105* at 6p24.1 that are associated with CAD.¹³ However, association of newly discovered loci previously identified in population of European descents were much smaller than those associated with CAD risk in population of Asian descents or could not be confirmed in Asians given the difference in linkage disequilibrium (LD) patterns between populations or overlapped between two populations.¹⁴ In addition, reproducible evidence of disease association has been acquired at a few candidate loci previously identified by GWAS studies.^{15,16}

In this study, we conducted GWAS and a replication study to identify common CAD susceptibility loci and validate the previously reported loci using Affymetrix Genome-Wide Human SNP array 6.0 with 2293 CAD patients from Genomics Research in Cardiovascular Disease (GenRIC) and 4302 healthy controls from a large urban cohort, the Korea Genome Epidemiology Study (KoGES) as Stage I. By analyzing data from GWAS scan, replication was performed using 3052 cases and 4976 controls from the KItaNagoya Genome (King) study of Japan. From GWAS and a replication analysis, we found evidence for genetic variants that may be associated with CAD risk.

MATERIALS AND METHODS

Study population

The study protocol was approved by the institutional review boards at Korea National Institute of Health and at each collaborating institute. Informed consent was obtained from all participants. To identify common susceptibility loci, a GWA scan (GWAS) was conducted with 2293 CAD patients from GenRIC working groups consisting with three teaching hospitals (Samsung Medical Center, Seoul National University Hospital and Yonsei University College of Medicine) in Korea. CAD was confirmed by standard coronary angiography. Subjects with myocardial infarction (31%), stable angina (41%) and unstable angina (28%) were classified as CAD subjects. The diagnosis of myocardial infarction was based on typical chest pain with a duration >30 min, on characteristic electrocardiographic patterns of acute myocardial infarction, and on elevated creatine kinase-MB and troponin I levels. Subjects with familial hypercholesterolemia, known vasculitides, end-stage renal disease and congenital heart disease were excluded.

KoGES is an ongoing cohort study in Korea that started in 2001 and is aimed at understanding the causes of disease and identifying disease risk factors. The participants of the large urban cohort visited one of the several centers located in four north-central (Seoul) and south-eastern (Busan) regions of South Korea to undergo a health examination that included a clinical test and physical measurements. All procedures were standardized across the centers through the development of a standard protocol, as well as the training of research coordinators and research assistants. For replication, samples were used from the KItaNagoya Genome (KING) study, which is a community-based prospective observational study of individuals who underwent community-based annual health checkups in Kitagoya City, Japan, between May 2005 and December 2007, that is aimed at identifying factors contributing to the genetic basis of cardiovascular disease and its risk factors (Supplementary Table S3). The replication cohort has been described in a previous study.¹⁷

Genotyping and quality control of the GWAS and replication

Genetic variants in samples from 2293 cases and 4302 controls were genotyped using the Affymetrix Genome-Wide Human SNP array 6.0. The BirdSeed (<http://www.broadinstitute.org/mpg/birdsuite/birdseed.html>) genotyping algorithm was used for genotype calling. The PLINK program (Ver. 1.06) and R statistics (<http://www.r-project.org/>) (Ver. 2.10.1) were used for quality control procedures. Samples with any of the following were removed: (i) gender inconsistencies ($n = 23$ in cases, $n = 12$ in controls), (ii) call rate <95% ($n = 94$ in cases, $n = 443$ in controls), (iii) outliers in a heterozygosity plot ($n = 41$ in cases, $n = 25$ in controls), (iv) cryptic first-degree relative >0.8 ($n = 9$ in cases, $n = 33$ in controls), or (v) outliers in a multidimensional scaling plot ($n = 3$ in

cases, $n = 26$ in controls). Samples with a history of cancer and CAD were also excluded (in controls, $n = 172$).

Samples in which the genotype-deduced gender differed from the clinical record were excluded. The calculation of heterozygosity and identity-by-state (IBS) were performed on the basis of the method reported by the WTCCC.⁷ To eliminate the genetic influence of sample contamination, duplications and cryptic first-degree relative sibling pairs, genome-wide average IBS values were calculated for each pair of individuals in the present GWA case-control data set using pruned SNPs (51 195 SNPs in 2123 cases and 74 965 SNPs in 3703 controls), which tend to represent weak LD blocks. From the IBS analysis, individuals who shared a too high degree of IBS were excluded. To evaluate differences in population structure, an multidimensional scaling calculation was performed using pruned SNPs.

SNPs were filtered if: (1) the call rate was <95% ($n = 92\,682$ in cases, $n = 170\,631$ in controls), (2) the minor allele frequency (MAF) was <1% ($n = 83\,902$ in cases, $n = 112\,231$ in controls), (3) the difference between case and control was missing ($P < 5 \times 10^{-5}$) (41 142 SNPs in cases and controls), (4) differential genotype calling rate between the cases and controls (case missing rate >1% or control missing rate >1%, and missing P -value < 5×10^{-6}), and (5) significant deviation from Hardy-Weinberg equilibrium (HWE $P \leq 1 \times 10^{-6}$) were also filtered. After quality control, 521 786 autosomal SNPs in 2123 cases and in 3591 controls remained for association analysis.

For replication, fourteen significant candidate loci in GWAS scan were examined for replication using 3052 cases and 4976 controls using TaqMan SNP genotyping assay (Supplementary Figure S1). The minimum genotyping success rate in replication study was 0.993 and Hardy-Weinberg equilibrium test P -value showed >0.05 in control group, suggesting no fault in genotyping procedure.

SNP selection for validation study using GWAS

We selected the SNPs if they were implicated in previous GWAS studies and reported from the National Human Genome Research Institute catalog (<http://www.genome.gov/gwastudies>). We included SNPs only if they have assigned reference allele, defined MAF, and estimated OR or β -coefficient and 95% confidence interval (CI). All SNPs successfully imputed were used (imputation QC $R^2 < 0.3$).

Association analysis of the GWAS

Association analysis was performed for 521 786 common SNPs that passed QC criteria for cases and controls under an additive model (one degree of freedom) of logistic regression with adjustment for age and gender in 2123 cases and 3591 controls. The genomic inflation factor λ was calculated as 1.08 (Supplementary Figure S2), and for examining a population stratification, population substructures using multidimensional scaling and principal component analyses were performed. It is on the basis of pairwise IBS which showed that, apart for some outliers, all subjects in the present study cluster closely with HapMap Asians (Supplementary Figure S3). Detailed method was described in previous study.¹⁸ The statistical power of detecting the ORs reported in previous GWASs was calculated by using Quanto version 1.2.4. (<http://hydra.usc.edu/gxe>) (Supplementary Table S6).

Association of loci with established cardiovascular risk factors was examined in control samples from the GWAS. For quantitative traits (the levels of high-density lipoprotein, low-density lipoprotein, total cholesterol, triglycerides, blood pressure and body mass index), linear regressions were used, whereas, for the binary trait (hypertension), a logistic regression model was applied (Supplementary Table S5).

Single-nucleotide polymorphism imputation

To infer the genotype of SNPs that were not observed in the Affymetrix Genome-wide Human SNP array 6.0 platform, SNP imputation was performed by using a Markov chain Monte Carlo models (MCMC), as implemented in IMPUTE ver2 (http://mathgen.stats.ox.ac.uk/impute/impute_v2.html). On the basis of NCBI build 37, the phased JPT+CHB data from 1000 Genomes Phase 1 (interim) were used as a reference panel, which consisted of over 7.5 million SNPs. After excluding imputed SNPs with an

Table 1 CAD associations of previously identified loci in European GWA studies and evaluated loci in the current GWA study

Chr	Reported genes/region	SNP	Loci reported in published GWAS				Current GWAS		
			RAF	OR (95%CI)	P-value	Ref ^a	RAF	OR (95%CI)	P-value ^c
1	PCSK9	rs11206510	0.82 (T)	1.08 (1.05–1.11)	9.10×10^{-8}	1	0.94	1.09 (0.98–1.23)	0.12
1	PPAP2B	rs17114036	0.91 (A)	1.17 (1.13–1.22)	3.81×10^{-19}	1	0.98	1.47 (1.09–1.99)	0.01
1	SORT1	rs599839	0.78 (A)	1.11 (1.08–1.15)	2.89×10^{-10}	1	0.94	1.34 (1.12–1.61)	1.3×10^{-3}
1	MIA3	rs17465637	0.74 (C)	1.14 (1.09–1.20)	1.36×10^{-8}	2	0.56	1.04 (0.93–1.16)	0.48
2	ABCG8 ^b	rs4299376	0.29 (C)	1.07 (1.04–1.11)	1.36×10^{-8}	3	—	—	—
2	WDR12	rs6725887	0.15 (C)	1.14 (1.09–1.19)	1.12×10^{-9}	1	0.01	1.38 (0.99–1.89)	0.05
3	MRAS	rs2306374	0.18 (T)	1.12 (1.07–1.16)	3.34×10^{-8}	1	0.02	0.97 (1.05–1.11)	0.84
5	IL5	rs2706399	0.51 (G)	1.07 (1.03–1.10)	2.10×10^{-6}	3	0.14	0.97 (1.05–1.11)	0.57
6	PHACTR1	rs12526453	0.67 (C)	1.10 (1.06–1.13)	1.15×10^{-9}	1	0.99	0.94 (1.05–1.11)	0.77
6	ANKS1A	rs17609940	0.75 (G)	1.07 (1.05–1.10)	1.36×10^{-8}	1	0.97	1.04 (0.81–1.33)	0.76
6	TCF21 ^d	rs12190287	0.62 (C)	1.08 (1.06–1.10)	1.07×10^{-12}	1	0.61	1.04 (0.94–1.16)	0.45
6	LPA ^d	rs3798220	0.02 (C)	1.51 (1.33–1.70)	3.00×10^{-11}	1	0.07	0.93 (1.05–1.11)	0.43
7	7q22	rs10953541	0.80 (C)	1.08 (1.05–1.11)	3.12×10^{-8}	4	0.87	0.95 (1.05–1.11)	0.40
7	ZC3HC1	rs11556924	0.62 (C)	1.09 (1.07–1.12)	9.18×10^{-18}	1	0.96	1.11 (0.90–1.36)	0.34
8	TRIB1	rs10808546	0.52 (A)	1.06 (1.03–1.10)	6.50×10^{-7}	3	0.83	1.00 (0.90–1.11)	0.93
9	CDKN2A/2B	rs4977574	0.46 (G)	1.29 (1.23–1.36)	1.35×10^{-22}	1	0.46	1.26 (1.16–1.36)	1.36×10^{-8}
9	CDKN2A/2B	rs1333049	0.55 (C)	1.37 (1.26–1.48)	1.80×10^{-14}	5	0.49	1.26 (1.17–1.34)	6.05×10^{-9}
9	ABO	rs579459	0.21 (C)	1.10 (1.07–1.13)	4.08×10^{-14}	1	0.27	1.08 (0.99–1.18)	0.07
10	KIAA1462 ^b	rs2505083	0.38 (C)	1.07 (1.04–1.09)	3.87×10^{-8}	4	—	—	—
10	CXCL12	rs1746048	0.87 (C)	1.09 (1.07–1.13)	2.93×10^{-10}	1	0.67	0.91 (1.05–1.11)	0.02
10	LIPA ^d	rs1412444	0.42 (A)	1.09 (1.07–1.12)	2.76×10^{-13}	3	0.70	0.87 (0.78–0.97)	0.01
10	CYP17A1-NT5C2	rs12413409	0.89 (G)	1.12 (1.08–1.16)	1.03×10^{-9}	1	0.76	1.09 (0.98–1.23)	0.12
11	PDGFD	rs974819	0.32 (T)	1.07 (1.04–1.09)	2.41×10^{-9}	4	0.63	1.15 (1.06–1.24)	7.9×10^{-4}
11	APOA1-C3-A4-A5	rs964184	0.13 (G)	1.13 (1.10–1.16)	1.02×10^{-17}	1	0.22	1.13 (1.03–1.24)	0.01
12	SH2B3 ^b	rs3184504	0.44 (T)	1.07 (1.04–1.10)	6.35×10^{-6}	1	—	—	—
13	COL4A1-A2	rs4773144	0.44 (G)	1.07 (1.05–1.09)	3.84×10^{-9}	1	0.39	1.12 (1.04–1.21)	5.05×10^{-3}
14	HHLPL1	rs2895811	0.43 (C)	1.07 (1.05–1.10)	1.14×10^{-10}	1	—	—	—
15	ADAMTS7	rs3825807	0.57 (A)	1.08 (1.06–1.10)	1.07×10^{-12}	1	0.88	0.94 (1.05–1.11)	0.10
17	SMG6-SRR	rs216172	0.37 (C)	1.07 (1.05–1.09)	1.15×10^{-9}	1	0.28	1.11 (1.01–1.21)	0.02
17	PEMT	rs12936587	0.56 (G)	1.07 (1.05–1.09)	4.45×10^{-10}	1	0.90	1.06 (0.93–1.22)	0.36
17	GIP-ATP	rs46522	0.53 (T)	1.06 (1.04–1.08)	1.81×10^{-8}	1	0.73	0.94 (1.05–1.11)	0.16
19	LDLR	rs1122608	0.77 (G)	1.14 (1.09–1.18)	9.73×10^{-10}	1	0.90	0.94 (1.05–1.11)	0.35
21	MRPS6 ^b	rs9982601	0.15 (T)	1.18 (1.12–1.24)	4.22×10^{-10}	1	—	—	—

Abbreviations: CAD, coronary artery disease; GWAS, genome-wide association studies; OR, odds ratio; RAF, risk allele frequency; SNP, single-nucleotide polymorphism; 95% CI, 95% confidence interval.

^aCARDIoGRAM¹, Erdmann *et al.*², IBC 50 K CAD consortium³, C4D⁴ Samani *et al.*⁵

^bMonomorphic in the Asian reference (JPT + CHB).

^cAll alleles are calculated using imputation data (Reference based on 1,000 Genomes Phase I (interim) JPT + CHB).

^dSNPs were not on the current GWAS. SNPs (rs12193973 for TCF21, rs9457925 for LPA and rs2246833 for LIPA ($r^2 > 0.2$)) were used to calculate.

imputation quality score below a set threshold ($R^2 < 0.3$), call rate of < 0.9 in either cases or controls, MAF of < 0.01 in either cases or controls, Hardy–Weinberg equilibrium P of $< 1 \times 10^{-3}$ in controls, we retained a total of 5125961 genotyped and imputed autosomal SNPs. Plots were drawn using the LocusZoom standalone version (<http://genome.sph.umich.edu/wiki/LocusZoom>) on the basis of 1000 Genome JPT + CHB for all SNPs in Figure 1.

Replication for SNP selection and combined analysis

Following the GWAS scan, SNPs for replication were selected on the basis of the following criteria among the 521,786 SNPs directly genotyped that had passed QC procedure: SNPs (a) with an MAF of $> 5\%$ (b) with very clear genotyping clusters (c) with P -value of $< 5 \times 10^{-4}$, and (d) not in strong LD ($R^2 < 0.5$) with any of the GWAS-identified risk variants. In addition, when multiple SNPs showed LD within 100kb ($R^2 > 0.2$), the SNP with the lowest P -value was selected for replication. SNPs with MAF of between 1% and 5% is $\sim 50,000$ and, among them, fifty SNPs (0.1%) have the P -value of $< 5 \times 10^{-4}$. And mostly, these SNPs didn't meet other replication criteria. For replication study, we did performed logistic regression by adjusting age and gender. The

combined analysis was performed by fixed effects and weighted numbers of samples, and combined P -values and ORs were calculated using the Mantel–Haenszel test in RMETA in the R package. In addition, as we found CAD association of variant previously identified on 12q24 in the Japanese, we examined independent test at chromosomal position 12q24.11 using SNP (rs11066011) as a proxy of *BRAP*.

RESULTS

Validation of previously identified coronary artery disease susceptibility loci in Koreans

Table 1 presents the test of associations between the previously reported variants in European population and CAD risk in 2123 cases and 3591 controls. Among the CAD-associated loci previously reported in Europeans, most strong association was replicated at the *CDKN2A/B* locus in the Koreans. Among 33 SNPs reported in published GWASs, four SNPs located in three loci showed associations with CAD risk (after considering multiple testing, $P < 0.05/$

28 = 0.0018). Suggestive findings of replication were observed for six loci—*PPAP2B*, *CXCL12*, *LIPA*, *APOA1-C3-A4-A5*, *COL4A1-A2*, *SMG6-SRR*—for CAD association ($p < 0.05$). Most significant findings were observed in 9p21.3 (rs1333049 at *CDKN2A/2B*) (OR = 1.26, 95% CI 1.17–1.34 and $P = 6.05 \times 10^{-9}$), following the same locus (rs4977574 at *CDKN2A/2B*) (OR = 1.26, 95% CI 1.16–1.36 and $P = 1.36 \times 10^{-8}$) almost similar to the magnitude and direction of previous reports. The *SORT1* (rs599839) and *PDGFD* (rs974819) were also confirmed to be associated with CAD risk although the magnitude of the two SNPs was smaller than the previously reported in Europeans. All the remaining 11 SNPs were not replicated in our study ($P > 0.05$), and the effect sizes were calculated < 1.11 , which are lower than those of estimated in European descents. Regional plots and association by chromosomal position for three loci are shown in Figure 1.

GWA analysis of CAD

Multiple genomic locations were shown to be potentially associated with the risk of CAD (Figure 2). We observed a strong association signals in two chromosomal regions 9p21.3 and 12q24: six SNPs in two genomic loci, *ACAD10*, *C12orf51*, *CDKN2A/B* and *RPL6-PTPN11*, reached the genome-wide significant P -value of 5.0×10^{-8} . In Supplementary Table S1, a summary of the two genomic loci and their SNPs is shown. The top SNP (rs11066015 at 12p24) with a P -value of $P = 4.51 \times 10^{-11}$ is located within the *ACAD10* gene, which is adjacent to *ALDH2*. We also found two other SNPs, rs2074356 and rs11066280 (located -477 kb and -650 kb from rs11066015, respectively) to be significantly associated with CAD risk (OR = 1.42, 95% CI 1.28–1.58 and $P = 6.73 \times 10^{-11}$; OR = 1.33, 95% CI 1.21–1.47 and $P = 1.44 \times 10^{-8}$). Therefore, we performed conditional association analyses using the case–control sample (Supplementary Table S2). The

Figure 1 Regional plots of the $-\log P$ -values for three replicated loci. (a–c) Results ($-\log P$) are shown for the association of directly genotyped and imputed SNPs for a 1-Mb region centered on SNP reported in previous genome-wide association studies (GWAS) (diamond). Genotyped SNPs are represented as square and imputed SNPs are marked by circle.

Figure 2 Manhattan plots for genome-wide association studies (GWAS) of coronary artery disease (CAD) in the Koreans. The horizontal solid red line indicates the preset threshold of $P = 5 \times 10^{-8}$.

results from these analyses showed that the associations for rs11066015, rs2074356 and rs11066280 were not significant at the genome-wide level after adjustment for each other's effect, suggesting that they are unlikely to be independent signals. Our GWAS signals reaching genome-wide significance levels were included another five SNPs mapping to chromosome 9p21.3, which were previously discovered in association with CAD.

Replication and meta-analysis

To search additional independent genetic risk variants in Koreans, we selected the most suggestive 14 SNPs from GWAS scan and genotyped in an independent set of 3052 cases and 4976 controls from the KING study of Japanese individuals. Three loci were found to have statistically significant associations with the risk of CAD ($P < 0.05$). The other 10 SNPs showed no significant association in replication set ($P > 0.05$). Most significant finding were observed in 12q24.11 (rs3782889 at *MYL2*) (OR = 1.26, 95% CI 1.18–1.35 and $P = 1.65 \times 10^{-10}$) (Table 2). The combined analysis also reached the genome-wide significant levels ($P = 3.95 \times 10^{-14}$). Test for heterogeneity suggested no difference in genetic effects across between GWAS and replication study for rs3782889 (P for heterogeneity = 0.8577). Suggestive findings of replication analysis were observed for rs9508025 in *FLT1* (OR = 1.11, 95% CI 1.04–1.18 and $P = 0.0023$). Although the *MYL2* was in completely different LD block from that of the *BRAP*, and the association of SNP rs3782889 with CAD doesn't remained statistically significant after adjusting for the effect of the SNP rs11066015 (Supplementary Table S4).

Imputation analysis of CAD

After SNP quality control and mapping of genomic positions to build37, a total of 5 125 961 SNPs for 5714 samples were retained as input genotype data for imputation. Among previously identified CAD loci, imputed SNPs for 9p21.3 and 12q24 loci were satisfied the P -value of GWAS significance (5×10^{-8}) (Supplementary Table S7).

DISCUSSION

We conducted a GWAS study on the Koreans population with a larger number of subjects in the discovery screen than a previous study.¹⁴ We evaluated whether 33 SNPs in GWAS-identified loci from European GWAS are also relevant in Korean populations and four SNPs at three loci were successfully replicated in Koreans. The main locus is the chromosome 9 loci, which contains *CDKN2ABS* gene, which was previously identified in CAD and myocardial infarction.^{7,11,14,15,19} SNPs for rs1333049 and rs4977574 at 9p21.3 was also identified in Koreans and the effect size was similar to that of European population. Other important loci included the *SORT1* locus on chromosome 1 and *PDGFD* locus on chromosome 11. Suggestive loci were observed for six loci—*PPAP2B*, *CXCL12*, *LIPA*, *APOA1-C3-A4-A5*, *COL4A1-A2*, *SMG6-SRR*—for CAD association ($P < 0.05$). The effect sizes of confirmed variants were similar to or rather smaller than those of initially identified. This phenomenon was frequently observed in validation studies using different ethnic population from the population used for initial finding.⁵ However, the associations of other SNPs were not significant in the study possibly owing to the limited power of current study. For the established risk locus within *TCF21* (rs12190287), *LPA* (rs3798220) and *LIPA* (rs1412444), we could not evaluate the same SNPs as reported because they were not genotyped or successfully imputed (imputation QC $R^2 < 0.3$). Therefore, we selected rs12193973 as a proxy SNP, capturing *TCF21* for rs12190287 ($R^2 = 0.24$ in CHB + JTP; $R^2 = 0.35$ in CEU) and rs9457925 in replacement for *LPA* rs3798220 ($R^2 = 0.31$ in CHB + JTP; $R^2 = \text{none}$), and rs2246833 in replacement for *LIPA* rs1412444 ($R^2 = 0.86$ in CHB + JTP; $R^2 = 1.00$ in CEU) on the basis of 1000 genomes pilot1 data.

The 12q24 locus appeared to overlap with the three reported in Europeans and Asians^{9,20,21} although this locus may have pleiotropic effects with several phenotypes including our study.¹⁸ Especially, the 12q24 haplotype confers risk alleles for CAD, much more significantly in the Japanese than in the Europeans.^{9,20} And previous Japanese study reported significant association with CAD risk at SNPs (rs3782886 and rs11066001) in *BRAP* (*BRCA1-associated protein*) and SNP (rs671) in *ALDH2*,¹⁴ located on 12q24. SNP (rs3782886) in *BRAP* identified in the Japanese GWA study is polymorphic in East Asians; conversely, a lead SNP (rs3184504 in *SH2B3*) in the European GWA study is polymorphic only in Europeans. In our study, the strong evidence of CAD association was identified via GWAS scan ($P = 4.51 \times 10^{-11}$ at rs11066015 in *ACAD10*, $P = 6.73 \times 10^{-11}$ at rs2074356 in *C12orf51* and $P = 1.44 \times 10^{-8}$ rs11066280 in *RPL6-PTPN11*). The SNP (rs11066015 in *ACAD10*) had a complete LD with SNP (rs11066001 in *BRAP* ($R^2 = 1$)), and the SNPs (rs2074356 in *C12orf51*, rs11066280 in *RPL6-PTPN11*) had a strong LD with SNP (rs671 in *ALDH2* ($r^2 = 0.8$)). Previous studies have also shown significant evidence supporting signatures of natural selection and pleiotropic effects for this region (for example, CAD,^{9,14,20} blood pressure^{21–23} and the regulation of plasma lipid levels^{21,24,25}). All variants at 12q24 that are associated with CAD in Europeans are not polymorphic in the Koreans, whereas all CAD-associated variants at 12q24 in the Koreans are monomorphic in Europeans. In the present study, rs11066280 were confirmed to be associated with CAD and blood pressure supporting previous report.²⁶

Even though the association of rs3782889 (in *MYL2*) was identified with Korean GWAS and directly replicated in the Japanese, it couldn't escape the possibility that the potential presence of LD block around *MYL2* is not independent of the one around *BRAP* previously identified in Japanese.

Table 2 Results of a meta-analysis in the GWAS and the replication sample

GWAS										Replication				Meta-analysis			
SNP	Chr	Gene	Risk/Non-risk	RAF	N	OR (95% CI)	P	RAF	N	OR (95% CI)	P	OR (95% CI)	P	Het(P)			
rs41391154	3p26.1	GRM7	A/T	0.93/0.91	2114/3580	1.37 (1.19-1.59)	1.50E-05	0.11/0.11	3050/4976	1.08 (0.97-1.19)	0.1418	1.20 (1.09-1.30)	3.72E-05	0.0073			
rs2068230	3q23	ATP1B3	T/A	0.64/0.60	2117/3561	1.18 (1.09-1.28)	5.56E-05	0.56/0.56	3048/4973	1.02 (0.95-1.09)	0.6087	1.08 (1.03-1.14)	1.84E-03	0.0058			
rs2122149	4q13.1		A/T	0.19/0.16	2067/3501	1.27 (1.14-1.41)	9.63E-06	0.22/0.21	3047/4973	1.07 (0.99-1.15)	0.1103	1.15 (1.08-1.22)	1.70E-05	0.0104			
rs9368386	6p22.3	FLJ22536	A/T	0.92/0.90	2111/3587	1.32 (1.15-1.52)	7.63E-05	0.89/0.89	3050/4974	1.04 (0.94-1.15)	0.4804	1.15 (1.06-1.25)	1.05E-03	0.0064			
rs219822	7q22.1	TRRAP	A/G	0.45/0.41	2121/3582	1.20 (1.11-1.29)	7.06E-06	0.41/0.40	3049/4976	1.01 (0.94-1.07)	0.8398	1.08 (1.03-1.14)	1.63E-03	0.0011			
rs12705702	7q31.1		C/T	0.56/0.51	2117/3582	1.20 (1.11-1.30)	3.49E-06	0.43/0.43	3050/4975	1.01 (0.94-1.07)	0.8755	1.08 (1.03-1.14)	1.24E-03	0.0008			
rs12114277	8q22.3	UBR5	G/A ^a	0.58/0.54	2093/3484	1.21 (1.11-1.31)	4.24E-06	0.45/0.44	3046/4975	1.04 (0.98-1.11)	0.1940	1.11 (1.05-1.16)	7.52E-05	0.0044			
rs1111782	9p21.2	TEK	G/A	0.57/0.53	2105/3515	1.20 (1.11-1.30)	4.77E-06	0.59/0.58	3050/4976	1.06 (0.99-1.13)	0.0710	1.12 (1.06-1.17)	1.42E-05	0.0152			
rs1163072	10q24.33	C10orf26	T/C ^a	0.51/0.47	2121/3574	1.20 (1.11-1.29)	6.70E-06	0.46/0.45	3046/4975	1.04 (0.97-1.11)	0.2425	1.11 (1.05-1.16)	1.11E-04	0.0065			
rs17101534	10q26.12		T/C	0.72/0.69	2114/3583	1.16 (1.06-1.27)	7.45E-05	0.74/0.73	3050/4973	1.02 (0.95-1.10)	0.5085	1.07 (1.02-1.14)	9.16E-03	0.0245			
rs886126	12q24.11	CUX2	T/C	0.66/0.62	2112/3558	1.22 (1.12-1.32)	2.58E-06	0.70/0.68	3047/4975	1.09 (1.01-1.17)	0.0183	1.14 (1.08-1.20)	1.03E-06	0.0442			
rs3782889	12q24.11	MYL2	C/T	0.21/0.17	2108/3572	1.25 (1.13-1.38)	1.24E-05	0.29/0.24	3048/4976	1.26 (1.18-1.36)	1.65E-10	1.26 (1.19-1.34)	3.95E-14	0.8577			
rs9508025	13q12.3	FLT1	C/G	0.50/0.46	2122/3587	1.17 (1.08-1.26)	7.42E-05	0.43/0.41	3048/4974	1.11 (1.04-1.18)	0.0023	1.14 (1.08-1.20)	6.07E-07	0.3127			
rs9944810	18q21.31	ALPK2	A/C	0.59/0.55	2120/3583	1.17 (1.08-1.27)	7.73E-05	0.57/0.57	3049/4976	1.02 (0.95-1.08)	0.6495	1.08 (1.03-1.14)	2.24E-03	0.0071			

Abbreviations: Chr Chromosomal band; GWAS, genome-wide association studies; Het (P), heterogeneity; P, OR, odds ratio; RAF, risk allele frequency; 95% CI, 95% confidence interval. Combined P-value and OR were obtained using the Cochran-Mantel-Haenszel test statistics under a fixed effects model.

ORs and P-values of GWAS were calculated using logistic regression adjusting age and sex.

^aRisk alleles of rs12114277 and rs1163072 in replication stage of Japan were not consistent with those of GWAS cohort in Korea.

As an approach to examining clinical relevance of replicated SNPs, we assessed whether they are also associated with established epidemiological risk factors. SNP for rs599839, which is replicated in GWAS showed significant evidence of association with the levels of low-density lipoprotein cholesterol and total cholesterol, which are similar to those of previous report.²⁷ In case of rs974819, we identified this as a new possible CAD-associated variant even though its statistical significance did reach at the marginal genome-wide a significance level ($P = 6.07 \times 10^{-7}$). In addition, this variant was also associated with diastolic blood pressure. These lipid and blood pressure risk loci may also have multiple phenotypes that might be tracking with CAD. As a result, interpretation of these findings is complex. From a qualitative perspective, these findings may suggest that some, but not all, biological mechanisms involved in low-density lipoprotein cholesterol regulation and blood pressure may be implicated in the etiology of CAD, which consist of previous CAD associations.^{14,21,28}

The present study has all the potential biases for the failure of replication of the previously identified loci in European descents. First, owing to the difference in underlying genomic structure between ethnicities, bias could lead to cover tagging SNPs, while the reported SNPs could work effectively in European descents. Another potential reasons are that, variants previously examined may not be strongly associated with CAD risk in Asian such as in the null association of rs17114036 (*PPAP2B*) for which our study have 90% of power to detect an OR as small as 1.5. The same null association was also found in the study conducted in Japanese.¹⁴ In addition, the risk profiles of genetic variants may be overestimated by different ethnic populations, suggesting that the relative contribution of risk variants to the pathogenesis of CAD is varied between ethnicity. Especially, sampling bias (population stratification) between cases and controls might be influencing the observed results. Last, we could not exclude the effect size of previously identified reports be exaggerated caused by 'Winner's curse'.²⁹ When we interpret this genetic association study, these limitations should be considered. The multi-ethnic studies will be required to contribute a better understanding of genetic architecture of CAD and pathogenesis.

CONCLUSIONS

In summary, we have confirmed four SNPs in three loci of CAD that were initially identified in Europeans, provided additional evidence suggesting association in the risk of CAD susceptibility variants. Future studies including fine mapping study, functional assay and replication study with large sample sizes from diverse ethnic populations need to validate our results.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank all participants and investigators of the GenRIC study and the Korea Genome Epidemiology Study (KoGES). This work was supported by grants from the Korea Centers for Disease Control and Prevention (4845-301), an intramural grant from the Korea National Institute of Health (2011-N73007-00, 2012-N73003-00).

1 Lopez, A. D., Mathers, C. D., Ezzati, M., Jamison, D. T. & Murray, C. J. Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. *Lancet* **367**, 1747-1757 (2006).

- 2 Wang, Q. Molecular genetics of coronary artery disease. *Curr. Opin. Cardiol.* **20**, 182–188 (2005).
- 3 Marenberg, M. E., Risch, N., Berkman, L. F., Floderus, B. & de Faire, U. Genetic susceptibility to death from coronary heart disease in a study of twins. *N. Engl. J. Med.* **330**, 1041–1046 (1994).
- 4 Peden, J. F. & Farrall, M. Thirty-five common variants for coronary artery disease: the fruits of much collaborative labour. *Hum. Mol. Genet.* **20**, R198–R205 (2011).
- 5 Schunkert, H., König, I. R., Thompson, J. R., Erdmann, J., Absher, D., Assimes, T. L. *et al.* Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nat. Genet.* **43**, 333–338 (2011).
- 6 Preuss, M., König, J. R., Thompson, J. R., Erdmann, J., Absher, D., Assimes, T. L. *et al.* Design of the Coronary ARtery Disease Genome-Wide Replication And Meta-Analysis (CARDIoGRAM) Study: a genome-wide association meta-analysis involving more than 22 000 cases and 60 000 controls. *Circ. Cardiovasc. Genet.* **3**, 475–483 (2010).
- 7 Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447**, 661–678 (2007).
- 8 Erdmann, J., Grosshennig, A., Braund, P. S., König, I. R., Hengstenberg, C., Hall, A. S. *et al.* New susceptibility locus for coronary artery disease on chromosome 3q22.3. *Nat. Genet.* **41**, 280–282 (2009).
- 9 Gudbjartsson, D. F., Bjornsdottir, U. S., Halapi, E., Helgadóttir, A., Sulem, P., Jonsdóttir, G. M. *et al.* Sequence variants affecting eosinophil numbers associate with asthma and myocardial infarction. *Nat. Genet.* **41**, 342–347 (2009).
- 10 Tregouet, D. A., König, I. R., Erdmann, J., Munteanu, A., Braund, P. S., Hall, A. S. *et al.* Genome-wide haplotype association study identifies the SLC22A3-LPAL2-LPA gene cluster as a risk locus for coronary artery disease. *Nat. Genet.* **41**, 283–285 (2009).
- 11 Coronary Artery Disease (C4D) Genetics Consortium. A genome-wide association study in Europeans and South Asians identifies five new loci for coronary artery disease. *Nat. Genet.* **43**, 339–344 (2011).
- 12 Ozaki, K., Sato, H., Inoue, K., Tsunoda, T., Sakata, Y., Mizuno, H. *et al.* SNPs in BRAP associated with risk of myocardial infarction in Asian populations. *Nat. Genet.* **41**, 329–333 (2009).
- 13 Wang, F., Xu, C. Q., He, Q., Cai, J. P., Li, X. C., Wang, D. *et al.* Genome-wide association identifies a susceptibility locus for coronary artery disease in the Chinese Han population. *Nat. Genet.* **43**, 345–349 (2011).
- 14 Takeuchi, F., Yokota, M., Yamamoto, K., Nakashima, E., Katsuya, T., Asano, H. *et al.* Genome-wide association study of coronary artery disease in the Japanese. *Eur. J. Hum. Genet.* **20**, 333–340 (2012).
- 15 Schunkert, H., Erdmann, J. & Samani, N. J. Genetics of myocardial infarction: a progress report. *Eur. Heart. J.* **31**, 918–925 (2010).
- 16 Saleheen, D., Alexander, M., Rasheed, A., Wormser, D., Soranzo, N., Hammond, N. *et al.* Association of the 9p21.3 locus with risk of first-ever myocardial infarction in Pakistanis: case-control study in South Asia and updated meta-analysis of Europeans. *Arterioscler. Thromb. Vasc. Biol.* **30**, 1467–1473 (2010).
- 17 Asano, H., Izawa, H., Nagata, K., Nakatochi, M., Kobayashi, M., Hirashiki, A. *et al.* Plasma resistin concentration determined by common variants in the resistin gene and associated with metabolic traits in an aged Japanese population. *Diabetologia* **53**, 234–246 (2010).
- 18 Cho, Y. S., Go, M. J., Kim, Y. J., Heo, J. Y., Oh, J. H., Ban, H. J. *et al.* A large-scale genome-wide association study of Asian populations uncovers genetic factors influencing eight quantitative traits. *Nat. Genet.* **41**, 527–534 (2009).
- 19 Kathiresan, S., Voight, B. F., Purcell, S., Musunuru, K., Ardissino, D., Mannucci, P. M. *et al.* Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. *Nat. Genet.* **41**, 334–341 (2009).
- 20 Soranzo, N., Spector, T. D., Mangino, M., Kuhnel, B., Rendon, A., Teumer, A. *et al.* A genome-wide meta-analysis identifies 22 loci associated with eight hematological parameters in the HaemGen consortium. *Nat. Genet.* **41**, 1182–1190 (2009).
- 21 Kato, N., Takeuchi, F., Tabara, Y., Kelly, T. N., Go, M. J., Sim, X. *et al.* Meta-analysis of genome-wide association studies identifies common variants associated with blood pressure variation in east Asians. *Nat. Genet.* **43**, 531–538 (2011).
- 22 Newton-Cheh, C., Johnson, T., Gateva, V., Tobin, M. D., Bochud, M., Coin, L. *et al.* Genome-wide association study identifies eight loci associated with blood pressure. *Nat. Genet.* **41**, 666–676 (2009).
- 23 Levy, D., Ehret, G. B., Rice, K., Verwoert, G. C., Launer, L. J., Dehghan, A. *et al.* Genome-wide association study of blood pressure and hypertension. *Nat. Genet.* **41**, 677–687 (2009).
- 24 Kim, Y. J., Go, M. J., Hu, C., Hong, C. B., Kim, Y. K., Lee, J. Y. *et al.* Large-scale genome-wide association studies in East Asians identify new genetic loci influencing metabolic traits. *Nat. Genet.* **43**, 990–995 (2011).
- 25 Teslovich, T. M., Musunuru, K., Smith, A. V., Edmondson, A. C., Stylianou, I. M., Koseki, M. *et al.* Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* **466**, 707–713 (2010).
- 26 Lu, X., Wang, L., Chen, S., He, L., Yang, X., Shi, Y. *et al.* Genome-wide association study in Han Chinese identifies four new susceptibility loci for coronary artery disease. *Nat. Genet.* **44**, 890–894 (2012).
- 27 Wang, A. Z., Li, L., Zhang, B., Shen, G. Q. & Wang, Q. K. Association of SNP rs17465637 on chromosome 1q41 and rs599839 on 1p13.3 with myocardial infarction in an American caucasian population. *Ann. Hum. Genet.* **75**, 475–482 (2011).
- 28 Kannel, W. B., Dawber, T. R., Kagan, A., Revotskie, N. & Stokes, 3rd J. Factors of risk in the development of coronary heart disease—six year follow-up experience. The Framingham Study. *Ann. Intern. Med.* **55**, 33–50 (1961).
- 29 Zhong, H. & Prentice, R. L. Correcting "winner's curse" in odds ratios from genomewide association findings for major complex human diseases. *Genet. Epidemiol.* **34**, 78–91 (2010).

Supplementary Information accompanies the paper on Journal of Human Genetics website (<http://www.nature.com/jhg>)