

Genome-wide association analysis and replication of coronary artery disease in South Korea suggests a causal variant common to diverse populations

Eun Young Cho,² Yangsoo Jang,³ Eun Soon Shin,² Hye Yoon Jang,² Yeon-Kyeong Yoo,² Sook Kim,² Ji Hyun Jang,² Ji Yeon Lee,² Min Hye Yun,² Min Young Park,² Jey Sook Chae,³ Jin Woo Lim,⁴ Dong Jik Shin,⁴ Sungha Park,⁴ Jong Ho Lee,³ Bok Ghee Han,⁵ Kim Hyung Rae,⁵ Lon R Cardon,⁶ Andrew P Morris,¹ Jong Eun Lee,² Geraldine M Clarke¹

► Additional figures/tables/appendix are published online only. To view these files please visit the journal online (<http://heartasia.bmj.com>).

¹Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford, UK

²DNA Link, Seoul, Republic of Korea

³Clinical Nutrigenetics/Nutrigenomics Lab, Department of Food & Nutrition, College of Human Ecology, Yonsei University Research Institute of Science for Ageing, Yonsei University, Seoul, Republic of Korea

⁴Division of Cardiology, Cardiovascular Genome Center, Yonsei Medical Institute, Yonsei University, Seoul, Republic of Korea

⁵National Genome Research Institute, Korean National Institute of Health, Seoul, Korea

⁶GlaxoSmithKline, Philadelphia, Pennsylvania, USA

Correspondence to

Dr Geraldine M Clarke, Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, UK; gclarke@well.ox.ac.uk

EYC and YJ contributed equally.

Accepted 7 May 2010

ABSTRACT

Background Recent genome-wide association (GWA) studies have identified and replicated several genetic loci associated with the risk of development of coronary artery disease (CAD) in samples from populations of Caucasian and Asian descent. However, only chromosome 9p21 has been confirmed as a major susceptibility locus conferring risk for development of CAD across multiple ethnic groups. The authors aimed to find evidence of further similarities and differences in genetic risk of CAD between Korean and other populations.

Methods The authors performed a GWA study comprising 230 cases and 290 controls from a Korean population typed on 490 032 single nucleotide polymorphisms (SNPs). A total of 3148 SNPs were taken forward for genotyping in a subsequent replication study using an independent sample of 1172 cases and 1087 controls from the same population.

Results The association previously observed on chromosome 9p21 was independently replicated ($p=3.08e-07$). Within this region, the same risk haplotype was observed in samples from both Korea and of Western European descent, suggesting that the causal mutation carried on this background occurred on a single ancestral allele. Other than 9p21, the authors were unable to replicate any of the previously reported signals for association with CAD. Furthermore, no evidence of association was found at chromosome 1q41 for risk of myocardial infarction, previously identified as conferring risk in a Japanese population.

Conclusion A common causal variant is likely to be responsible for risk of CAD in Korean and Western European populations at chromosome 9p21.3. Further investigations are required to confirm non-replication of any other cross-race genetic risk factors.

genotyping chips containing hundreds of thousands of single nucleotide polymorphisms (SNPs), makes genome-wide association (GWA) studies both financially and practically feasible.⁵ Recent GWA and follow-up studies identified and replicated several genetic loci associated with the risk of CAD in Caucasian samples. McPherson *et al*⁶ reported two susceptibility loci at chromosome 9p21 using six independent samples from four Caucasian populations; Helgadottir *et al*⁷ described an association between MI and two SNPs located in the same region in an Icelandic population and replicated their findings using additional Icelanders and a sample of controls of European descent from the USA; the Wellcome Trust Case Control Consortium (WTCCC)⁸ identified several genetic loci, including chromosome 9p21, associated with the risk of development of CAD in a sample of white Europeans from the UK; Samani *et al*⁹ replicated WTCCC findings and identified additional loci in a white European sample from Germany; Schunkert *et al*¹⁰ subsequently conducted a meta-analysis combining results from seven case-control studies to provide unprecedented evidence for association between genetic variants at chromosome 9p21 and risk of CAD in northern European populations. More recently, GWA studies have also found evidence of association at chromosome 9p21 in Asian and Hispanic populations,^{11–16} and in addition, Hiura *et al*,¹⁶ studying a Japanese population, found evidence of association for risk of MI at chromosome 1q41.

There is, of course, a precedent for ethnic variation in the genetic aetiology of disease—for example, the genetic variants associated with risk of Crohn's disease in a Caucasian population do not even exist in an Asian population,¹⁷ but variants of the transcription factor 7-like-2 (TCF7L2) are important genetic risk factors for the development of type 2 diabetes in multiple ethnic groups.^{17 18} Here we conduct a GWA study and a follow-up replication study on independent Korean samples with the aim of finding genetic loci associated with CAD and its acute complication, MI, in a Korean population. We then compare and combine these data with data from the WTCCC and other studies to investigate the similarities and differences in the genetic risk of CAD and MI between Caucasian and Asian populations.

Coronary artery disease (CAD), including myocardial infarction (MI), is a major cause of death and disability worldwide.¹ In particular, morbidity and mortality from cardiovascular disease rapidly escalated in Korea between 1981 and 2003.² Clustering of CAD and MI in families suggests a genetic susceptibility to CAD,³ although lifestyle and environmental factors are also implicated in the development of these complex diseases.⁴ Advances in genotyping technology, creating the availability of dense

METHODS

Subjects

We conducted a two-stage genome-wide association study on Korean patients with angiographically documented CAD and healthy Korean volunteers. We used publicly available data from the WTCCC to compare and contrast associations between Korean and Caucasian populations.

Korean study

All participants in Stage 1 and Stage 2 are of Korean ancestry. Detailed descriptions of recruitment and ascertainment in both stages are given in the Supplementary Material. Informed consent was obtained from each study participant, and the study protocol was approved by the ethics committee or institutional review board in each of the participating centres. Two hundred and thirty cases were selected for Stage 1 from patients in the Cardiovascular Genome Study (CGS) at Yonsei University aged under 55 years with angiographically documented CAD. Two hundred and ninety controls were selected from healthy volunteers participating in the CGS and frequency-matched to Stage 1 cases by 5-year age category and sex. One thousand one hundred and seventy two cases were selected for Stage 2 from patients in the CGS and the Nutrition Study (NS) conducted by the Clinical Nutrigenetics and Nutrigenomics Laboratory at Yonsei University aged under 66 years with angiographically documented CAD or who had a history of MI. One thousand and eighty-seven controls were selected from healthy volunteers aged under 84 years participating in the CGS and NS, and also enrolled in the Korean Health and Genome Study (KHGS). The basic clinical characteristics of all subjects are summarised in table 1.

WTCCC study

All participants in the WTCCC study were of white European origin. Approximately 2000 cases were selected from individuals with a history of either MI or coronary revascularisation before the age of 66 years and a strong family history of CAD. The control individuals came from two sources: approximately 1500

individuals from the 1958 British Birth Cohort (58C) and 1500 individuals selected from blood donors from the UK Blood Service recruited as part of the WTCCC study. See WTCCC *et al*⁸ for details of recruitment, phenotypes and ascertainment.

SNP selection and genotyping

Korean study

All 520 Stage 1 samples were genotyped with the Affymetrix GeneChip Human Mapping 500K Array Set which comprised 490 032 genome-wide SNPs in the DNALink. Since the X chromosome must be treated differently from the autosomal chromosomes, the X chromosome SNPs were excluded from our analyses. In particular, most loci on the X chromosome are subject to X chromosome inactivation, and so it cannot be predicted which allele is active. Furthermore, because there is only one copy of X in males, sample sizes and, consequentially, power are different from that in the autosomes.

We excluded 26 samples after checks for contamination and relatedness; 494 individuals remained in Stage 1 (supplementary table 1). Four hundred and fifty two thousand six hundred and thirty six SNPs (92.4%) passed our quality control filters. Of these, 90 603 had a study wide minor allele frequency (MAF) less than 1% and were not analysed due to the small Stage 1 sample size (supplementary table 2).

SNPs for Stage 2 were selected based on the strength of association observed in Stage 1 samples. They comprised 2184 SNPs with an association test p value $<1e-03$ and, for further analysis and discovery, 964 cSNPs, linked SNPs ($r^2>0.3$) and SNPs with association test p value <0.05 that were within 50 kb of SNPs with an association test p value $<1e-04$. All 2259 Stage 2 samples were genotyped with the Affymetrix Targeted Genotyping (TG) 3K Chip array in the DNALink. We excluded six samples after checks for contamination; 2253 individuals remained in Stage 2 (supplementary table 1). Three thousand and fifty-five SNPs (97.4%) passed our quality control filters (supplementary table 3). Further details on genotyping and calling algorithms, quality control criteria (at both individual and SNP level) and validation steps are described in the supplementary appendix.

WTCCC study

All WTCCC samples were genotyped with the Affymetrix GeneChip Human Mapping 500K Array Set comprising 500 568 genome-wide SNPs. See WTCCC *et al*⁸ for detailed descriptions of sample and marker exclusions.

Confirmation and replication genotyping

To ensure genotype consistency, we regenotyped eight SNPs from Stage 1 on a different platform, GenomeLab SNPstream; a concordance rate of 99.92% was observed across platforms. To test the accuracy of the Affymetrix assays, SNPs from Stage 1 for 76 Stage 1 individuals were also genotyped on the Affymetrix TG 3K Chip. The overall concordance between genotypes generated by the Affymetrix 500K array and the Affymetrix TG chip array was 99.8% and did not differ significantly between cases and controls.

Statistical analysis

All autosomal SNPs that showed evidence of a significant association with CAD (p value $<1e-03$) in either Stage 1 or Stage 2 samples were identified. We combined data on all SNPs for which there was evidence of association in either the Stage 1 or the Stage 2 samples and identified SNPs with a significant (p value $<1e-04$) association with CAD. The p value cut-off of

Table 1 Baseline clinical characteristics of coronary artery disease cases and controls in the combined Stage 1 and Stage 2 samples

Clinical characteristic*	Controls S1=290, S2=1087, total=1377	Cases S1=230, S2=1172, total=1402	p Value†
Female (%)	23.5	14.6	$<1e-05$
Age at recruitment (years)	53.4 ± 8.58	54.5 ± 7.79	$8.3e-04$
BMI (kg/m^2)	23.5 ± 2.63	25 ± 2.93	$<1e-05$
Systolic blood pressure (mm Hg)	116.5 ± 14.51	124 ± 19.99	$<1e-05$
Diastolic blood pressure (mm Hg)	75 ± 9.39	78 ± 11.38	$<1e-05$
Triglyceride (mg/dl)	139.2 ± 88.18	161.5 ± 116.39	$<1e-05$
Total cholesterol (mg/dl)	197.1 ± 35.09	178.5 ± 40.8	$<1e-05$
Glucose (mg/dl)	88.5 ± 12.85	104.2 ± 37.55	$<1e-05$
Insulin	7.3 ± 5.54	9.7 ± 7.06	$<1e-05$
HOMA-IR‡	1.58 ± 1.30	54.5 ± 3.10	$<1e-05$
Former or current smoker (%)	60.2	72.2	$<1e-05$
Diabetes (%)	0.7	29.6	$<1e-05$
Hypertension (%)	9.9	59.8	$<1e-05$
Myocardial infarction (%)	0	53.8	$<1e-05$

*Data are shown as mean \pm SD or a percentage.

†Continuous data were tested using a two-tailed Student t test, and categorical data were tested using a χ^2 test (with 1 degree of freedom) for difference between case and control groups.

‡Homeostatic model assessment of insulin resistance. Hypertension is defined as blood pressure $>140/90$ mm Hg at the time of examination, a current history of taking blood pressure medications or a history of hypertension. Diabetes is defined as a fasting glucose level of 126 mg/dl, or patients with history of diabetes currently taking diabetic medications.

S1, Stage 1; S2, Stage 2.

1e–04 was selected to ensure that the false discovery rate (FDR), calculated according to the method of Benjamini and Hochberg,¹⁹ was less than 5%. The combined sample provides increased power to assess more accurately association as well as the potential to identify additional loci of interest. Population-attributable risk and allelic and genotype specific ORs with associated 95% CIs adjusted for age, sex and stage (1 or 2) were calculated for the identified SNPs. Finally, we combined these data with corresponding data from the Wellcome Trust Case Control Consortium (WTCCC) study (which, after quality control,⁸ comprised 1926 Caucasian case subjects with CAD and 2938 normal controls). We used a Breslow Day test to test for homogeneity of ORs across Korean and WTCCC samples and, where appropriate, calculated a pooled OR for the risk allele and 95% CI adjusted for sex and study (Korean or WTCCC). Evidence for association was assessed via trend tests (1 degree of freedom) between cases and controls. Ten SNPs from Stage 1 showed marginal significance for deviation from additivity ($p < 1e-04$). None of these SNPs were either included in Stage 2 or showed any evidence for deviation from additivity in Stage 2. Hence, an additive disease model was assumed for all single marker tests of association. Haplotype analysis was carried out at the replicated loci as described in the supplementary appendix.

RESULTS

Association between 9p21 and CAD

The most notable association with CAD was observed on chromosome 9p21.3. The strongest signal was at rs1333049, but evidence for association was seen at SNPs across approximately 50 kb (table 2). These results replicate signals first observed in

Caucasian samples and subsequently observed in a variety of Asian populations including Japanese, Korean and Chinese Han, and thus provide yet further confirmation of a major genetic risk variant at this locus common to multiple ethnic groups. SNP rs1333049 had a p value of $3.08e-07$ in the combined Korean sample, and each copy of a C allele at this locus increased the risk of CAD by 32% (95% CI 19 to 47%). Approximately 24% of study participants were homozygous for this allele, and 49% carried a single copy. The patterns of association with respect to direction and magnitude of effect across the approximately 50 kb length region were similar to that found in the WTCCC study (supplementary table 5). In the combined Korean and WTCCC sample, the p value for association with CAD at rs1333049 was $3.12e-18$. The six SNPs (rs6475606, rs4977574, rs29891168, rs1333042, rs1333048, rs1333049) showing significant association within this region were in strong linkage disequilibrium. However, three blocks could be distinguished with strong linkage disequilibrium within each block (minimum $|D'| > 0.94$) and between blocks ($|D'| = 0.80$ between blocks 1 and 2, $|D'| = 0.92$ between blocks 2 and 3) (figure 1). One SNP in block 1, three in block 2 and one in block 3 were sufficient to tag the region. Haplotype analysis showed that the association was due to two mutually exclusive haplotypes for the four SNPs (CAAG and TGGC) (supplementary appendix). In the Korean study, the OR for CAD with the TGGC haplotype (frequency 0.449 and 0.385 among case and control subjects respectively) as compared with the CAAG haplotype (frequency 0.230 and 0.284 among case and control subjects respectively) was 1.44 (95% CI 1.26 to 1.65) per copy of the haplotype ($p = 1.71e-07$) (supplementary table 6). The OR for CAD with the TGGC haplotype as compared with the CAAG haplotype in the Caucasian WTCCC

Table 2 Stage 1, Stage 2 and combined Stage 1 and Stage 2 samples association results for risk of coronary artery disease at single nucleotide polymorphisms in 9p21 with $p < 1e-04$ in the combined sample

Single nucleotide polymorphism	Allele Minor	Risk	Stage	Genotypes minor homozygous/ heterozygous/major homozygous		Minor allele frequency		OR for risk allele (95% CI)	p Value
				Control	Case	Control	Case		
rs6475606	C	T	S1	24/124/120	13/80/132	0.32	0.24	1.55 (1.16 to 2.07)	3.24e–03
			S2	123/450/474	98/458/580	0.33	0.29	1.22 (1.08 to 1.39)	2.00e–03
			S1&S2	147/574/594	111/538/712	0.33	0.28	1.27 (1.13 to 1.43)	5.75e–05
rs4977574	G	G	S1	44/133/91	63/117/45	0.41	0.54	1.71 (1.31 to 2.22)	6.74e–05
			S2	199/522/345	274/569/315	0.43	0.48	1.22 (1.08 to 1.37)	1.27e–03
			S1&S2	243/655/436	337/686/360	0.43	0.49	1.29 (1.16 to 1.44)	4.61e–06
rs2891168	G	G	S1	43/134/91	63/116/45	0.41	0.54	1.73 (1.32 to 2.25)	5.16e–05
			S2	199/524/351	274/569/316	0.43	0.48	1.23 (1.09 to 1.38)	7.59e–04
			S1&S2	242/658/442	337/685/361	0.43	0.49	1.3 (1.17 to 1.45)	2.27e–06
rs1333042	A	G	S1	27/130/106	14/90/119	0.35	0.26	1.53 (1.15 to 2.04)	3.82e–03
			S2	156/473/455	122/490/555	0.36	0.31	1.23 (1.09 to 1.39)	1.04e–03
			S1&S2	183/603/561	136/580/674	0.36	0.31	1.27 (1.13 to 1.42)	3.71e–05
rs1333048	C	C	S1	55/134/76	71/119/35	0.46	0.58	1.68 (1.29 to 2.19)	1.38e–04
			S2	232/515/326	310/574/273	0.46	0.52	1.26 (1.12 to 1.42)	1.28e–04
			S1&S2	287/649/402	381/693/308	0.46	0.53	1.32 (1.18 to 1.47)	4.95e–07
rs1333049	C	C	S1	54/137/76	72/118/33	0.46	0.59	1.75 (1.34 to 2.29)	4.51e–05
			S2	232/519/334	309/574/282	0.45	0.51	1.25 (1.12 to 1.41)	1.50e–04
			S1&S2	286/656/410	381/692/315	0.45	0.52	1.32 (1.19 to 1.47)	3.08e–07

For each SNP, the ORs (with 95% CIs) and p values represent the RR of CAD associated with each copy of the risk allele. Results are based on a trend test (1 degree of freedom) with adjustment for age, sex and, in the combined sample, stage. The minor allele frequency is defined in the controls and its frequency in that group as well as the case sample is reported. Cluster plots for each single nucleotide polymorphisms have been visually inspected and are shown in supplementary figure 4.

S1, Stage 1; S2, Stage 2; S1&S2, combined Stage 1 and Stage 2.

sample was similar, specifically 1.39 (95% CI 1.27 to 1.52), per copy of the haplotype ($p=3.25e-13$) (supplementary table 7).

In order to understand the mode of inheritance in more detail, genotypic-specific ORs were computed, and the ORs for heterozygous and homozygous carriers of the risk allele C at the most significant SNP, rs1333049, were 1.40 (95% CI 1.16 to 1.68) and 1.75 (95% CI 1.41 to 2.17) respectively, indicating a population-attributable risk of 26%.

Putative novel loci

The combined analysis of Stage 1 and Stage 2 data identified four additional loci with a modest likelihood of association with CAD ($p<1e-04$) (supplementary table 4). However, none of the four regions they represented were found to be significant in the WTCCC study (supplementary table 5). The locus on chromosome 2p16.2 involves the SPTBN1, and multiple nearby SNPs located in introns in the gene were marginally significant ($p<1e-02$) in either Stage 1 or Stage 2 samples. The SNP on chromosome 5q31 is a promoter in C5orf32. The SNP on chromosome 12p11(rs11048979 with $p=6.92e-06$ in the combined Korean sample) is an isotonic SNP in ARNTL2. The lead SNPs on chromosome 19p13 is in CACNA1A. More work is required to confirm a genetic association with CAD in these regions.

Subphenotype and subgroup analysis

Approximately 54% of subjects with confirmed evidence of CAD (CAD) in combined Stage 1 and Stage 2 samples also had a history of MI at the time of recruitment (table 1). When subjects with both CAD and MI and those with CAD only were analysed separately, the ORs for both phenotypes across all six significant loci in 9p21 and the four putative novel loci remained significant (results are shown for the six significant loci in 9p21 in supplementary table 8). Breslow–Day tests for homogeneity of ORs across the sexes were not significant, indicating that all 10 loci affected the risk of CAD to a similar extent in men and women.

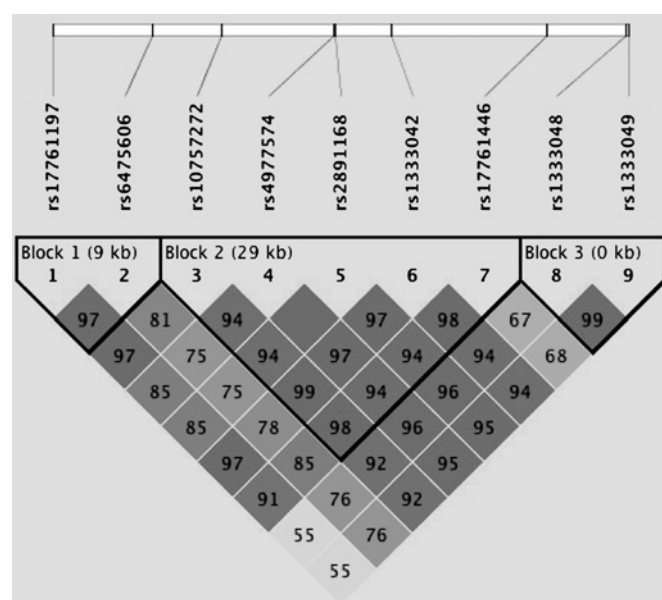


Figure 1 Analysis of linkage disequilibrium (LD). LD patterns between nine single nucleotide polymorphisms (SNPs) genotyped in the combined Stage 1 and Stage 2 Korean coronary artery disease case population in the region of interest on chromosome 9p21.3. The pairwise correlation between the SNPs was measured as $|r^2|$ and shown ($\times 100$) in each diamond.

A total of 29.6% of case subjects in combined Stage 1 and Stage 2 samples had type 2 diabetes mellitus (T2DM) at the time of recruitment (table 1). Analysis according to the presence or absence of T2DM in case subjects showed that none of the six loci on 9p21 or the four putative novel loci had an association with T2DM over and above that observed for CAD (results not shown).

Comparison with previously reported associated SNPs

Five regions previously reported to show evidence of association signals for risk of CAD or MI were not found to be significant in the Korean Stage 1 data (table 3). It can be seen that the minor allele frequencies (MAF) for each of the lead SNPs are significantly different across the Korean and Caucasian samples, and three of the five SNPs (rs599839, rs2943634 and rs6922269) have rare MAFs in the Korean sample.

DISCUSSION

Our primary objective was to identify loci with significant associations with CAD and MI in a Korean population. We also sought to compare and contrast previously reported association signals for CAD and MI with our evidence of association in a Korean population. We used a two-stage strategy to identify loci associated with CAD. The use of a genome-wide scan in Stage 1 using an enriched case sample of individuals with well-defined phenotypes maximised potential signals to enable the identification of multiple loci potentially implicated in the risk of CAD in a Korean population. The use of an independent sample of individuals drawn from the same population in Stage 2 allowed for the potential replication of significant association signals from Stage 1. A combined analysis using data from both stages provided a greater power to identify novel loci requiring further investigation to confirm their role in the aetiology of CAD. The case subjects in Stage 1 had advanced CAD with early onset possibly enhancing the power to detect an association with CAD, generating increased effect sizes and potentially biasing estimates of population attributable risk upwards suggesting that further analysis of loci in a wider range of subjects is also necessary in order to translate our findings into better prevention of treatment for CAD.

In this study, the associations between SNPs on chromosome 9p21.3 and risk of CAD were independently confirmed in a Korean sample. The magnitude and direction of the susceptibility effects in the Korean population are consistent with those reported in other studies of both Asian and Caucasian populations, providing a firm rationale for further biological experimentation. The great majority of the CAD susceptibility was seen to be encoded by two common yin–yang²⁰ haplotypes comprising four SNPs and spanning approximately 50 kb. The same risk haplotype was observed in a Caucasian population, suggesting that the causal mutation carried on this background might have occurred before the divergence of East Asians and Caucasians which took place about 15–35 000 years ago.²¹

Our findings on chromosome 9p21.3 of a replicated association in an Asian population that confers a substantial increase in risk with each additional copy of a common variant are further confirmation of a major genetic risk variant at this locus common to multiple ethnic groups. However, the newly identified putative associations on chromosomes 2p16.2, 5q3.1, 12p11 and 19p13 must be viewed with caution until they have been replicated in appropriate validation samples.

Apart from the signal at 9p21.3, we were unable to find any significant evidence of association with risk of CAD at previously identified loci in Northern European and Japanese

Table 3 Evidence of association signals in Stage 1 for risk of coronary artery disease (CAD) and myocardial infarction compared with evidence at previously reported loci

Chromosome	SNP	Population	Disease	Allele		MAF		OR for risk allele (95% CI)	p Value	Reference
				Minor	Risk	Controls	Case			
1	rs599839	K	CAD	G	A	0.07	0.04	1.95 (1.08 to 3.51)	2.72e-02	Samani <i>et al</i> ⁹
		E	CAD	G	A	0.23	0.19	1.29 (1.18 to 1.4)	4.05e-09	
1	rs17465637	K	CAD	A	C	0.45	0.39	1.24 (0.97 to 1.60)	8.60e-02	Samani <i>et al</i> ⁹
		E	CAD	A	C	0.28	0.25	1.20 (1.12 to 1.30)	1.27e-06	
		K	MI	A	C	0.45	0.38	1.31 (0.96 to 1.77)	8.81e-02	
		J	MI	A	C	0.46	0.42	1.45 (1.15 to 1.83)	2.00e-03	
2	rs2943634	K	CAD	A	C	0.07	0.07	1.01 (0.61 to 1.68)	9.66e-01	Samani <i>et al</i> ⁹
		E	CAD	A	C	0.35	0.31	1.20 (1.13 to 1.30)	1.61e-07	
6	rs6922269	K	CAD	A	G	0.02	0.02	1.09 (0.42 to 2.08)	8.61e-01	Samani <i>et al</i> ⁹
		E	CAD	A	A	0.25	0.30	1.23 (1.15 to 1.33)	2.90e-08	
10	rs501120	K	CAD	C	C	0.36	0.36	1.01 (0.78 to 1.32)	9.21e-01	Samani <i>et al</i> ⁹
		E	CAD	C	T	0.14	0.11	1.33 (1.20 to 1.48)	9.46e-08	

Lead single nucleotide polymorphisms (SNPs) represent regions found previously to be significantly associated with CAD in a given population. For each SNP, the ORs (with 95% CIs) and p values represent the RR of CAD associated with each copy of the risk allele. Results given for Stage 1 Korean data are based on a trend test (1 degree of freedom) with adjustment for age and sex. Results given for Samani *et al* are based on the Cochran–Armitage test for trend without adjustment for covariates in a combined analysis of data from the Wellcome Trust Case Control Consortium and German myocardial infarction family studies. Results given for Hiura *et al* are based on a trend test with adjustment for age, sex, diabetes, hypertension and smoking. The minor allele frequency (MAF) is defined in the controls, and its frequency in that group as well as the case sample is reported. E, European; J, Japanese; K, Korean.

populations. Nor were we able to find any significant evidence of an association with MI at SNP rs17465637 on chromosome 1q41, previously reported in a Japanese population.¹⁶ However, we note that the risk allele frequencies in the controls are similar in both populations and that the allelic ORs are of a similar magnitude and direction. Given the small sample size in Stage 1 and the consequential low power, further investigation is required to assess the significance of our negative findings.

In general, there are several possible explanations for heterogeneity of association signals across study populations. Heterogeneity of association signals between any two studies in general may simply be as a result of false positives or that one study may have a greater power than the other to detect an effect. Across studies using different populations, risks of heterogeneity in association signals are increased for a variety of reasons: variation in risk allele frequency at a causal locus may lead to a difference in power to detect an effect, the causal locus may be in LD with the observed locus, the LD may vary between the study populations, the effect may be epistatic involving one or more additional loci with risk alleles occurring at different frequencies in the study populations, or the effect may be interactive with one or more environmental variables with different levels of exposure in the study populations. Once contrary findings are replicated across ethnic groups, further work will be of interest to disentangle the potentially complex array of variables responsible for the underlying causes of ethnic differences in disease risk.

Funding Korea Health 21 R&D Project, Ministry of Health and Welfare, Republic of Korea Grant A050558; 2001 Good Health R&D Project, Ministry of Health and Welfare, Republic of Korea Grant A000385; Wellcome Trust.

Competing interests None.

Patient consent Obtained.

Ethics approval Ethics approval was provided by the Institutional Review Board of Yonsei University.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. Lopez AD, Mathers CD, Ezzati M, *et al*. Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. *Lancet* 2006;**367**:1747–57.

2. Office NS. *Annual Report on the Cause of Death Statistics*, 2003.
3. Marenberg ME, Risch N, Berkman LF, *et al*. Genetic susceptibility to death from coronary heart disease in a study of twins. *N Engl J Med* 1994;**330**:1041–6.
4. Yusuf S, Hawken S, Ounpuu S, *et al*. Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case–control study. *Lancet* 2004;**364**:937–52.
5. Christensen K, Murray JC. What genome-wide association studies can do for medicine. *N Engl J Med* 2007;**356**:1094–7.
6. McPherson R, Pertsemlidis A, Kavaslar N, *et al*. A common allele on chromosome 9 associated with coronary heart disease. *Science* 2007;**316**:1488–91.
7. Helgadottir A, Thorleifsson G, Manolescu A, *et al*. A common variant on chromosome 9p21 affects the risk of myocardial infarction. *Science* 2007;**316**:1491–3.
8. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 2007;**447**:661–78.
9. Samani NJ, Erdmann J, Hall AS, *et al*. Genomewide association analysis of coronary artery disease. *N Engl J Med* 2007;**357**:443–53.
10. Schunkert H, Gotz A, Braund P, *et al*. Repeated replication and a prospective meta-analysis of the association between chromosome 9p21.3 and coronary artery disease. *Circulation* 2008;**117**:1675–84.
11. Zhou L, Zhang X, He M, *et al*. Associations between single nucleotide polymorphisms on chromosome 9p21 and risk of coronary heart disease in Chinese Han population. *Arterioscler Thromb Vasc Biol* 2008;**28**:2085–9.
12. Shen GQ, Li L, Rao S, *et al*. Four SNPs on chromosome 9p21 in a South Korean population implicate a genetic locus that confers high cross-race risk for development of coronary artery disease. *Arterioscler Thromb Vasc Biol* 2008;**28**:360–5.
13. Shen GQ, Rao S, Martinelli N, *et al*. Association between four SNPs on chromosome 9p21 and myocardial infarction is replicated in an Italian population. *J Hum Genet* 2008;**53**:144–50.
14. Assimes TL, Knowles JW, Basu A, *et al*. Susceptibility locus for clinical and subclinical coronary artery disease at chromosome 9p21 in the multi-ethnic ADVANCE study. *Hum Mol Genet* 2008;**17**:2320–8.
15. Hinohara K, Nakajima T, Takahashi M, *et al*. Replication of the association between a chromosome 9p21 polymorphism and coronary artery disease in Japanese and Korean populations. *J Hum Genet* 2008;**53**:357–9.
16. Hiura Y, Fukushima Y, Yuno M, *et al*. Validation of the association of genetic variants on chromosome 9p21 and 1q41 with myocardial infarction in a Japanese population. *Circ J* 2008;**72**:1213–17.
17. Marsh S, McLeod HL. Crohn's disease: ethnic variation in CARD15 genotypes. *Gut* 2003;**52**:770.
18. Rees SD, Bellary S, Britten AC, *et al*. Common variants of the TCF7L2 gene are associated with increased risk of type 2 diabetes mellitus in a UK-resident South Asian population. *BMC Med Genet* 2008;**9**:8.
19. Benjamin Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing *J Royal Stat Soc Ser B* 1995;**57**:289–300.
20. Zhang J, Rowe WL, Clark AG, *et al*. Genomewide distribution of high-frequency, completely mismatching SNP haplotype pairs observed to be common across human populations. *Am J Hum Genet* 2003;**73**:1073–81.
21. Cavalli-Sforza LL, Feldman MW. The application of molecular genetic approaches to the study of human evolution. *Nat Genet* 2003;(33 Suppl):266–75.