

A genome-wide association study in Europeans and South Asians identifies five new loci for coronary artery disease

The Coronary Artery Disease (C4D) Genetics Consortium*

Genome-wide association studies have identified 11 common variants convincingly associated with coronary artery disease (CAD)^{1–7}, a modest number considering the apparent heritability of CAD⁸. All of these variants have been discovered in European populations. We report a meta-analysis of four large genome-wide association studies of CAD, with ~575,000 genotyped SNPs in a discovery dataset comprising 15,420 individuals with CAD (cases) (8,424 Europeans and 6,996 South Asians) and 15,062 controls. There was little evidence for ancestry-specific associations, supporting the use of combined analyses. Replication in an independent sample of 21,408 cases and 19,185 controls identified five loci newly associated with CAD ($P < 5 \times 10^{-8}$ in the combined discovery and replication analysis): *LIPA* on 10q23, *PDGFD* on 11q22, *ADAMTS7-MORF4L1* on 15q25, a gene rich locus on 7q22 and *KIAA1462* on 10p11. The CAD-associated SNP in the *PDGFD* locus showed tissue-specific *cis* expression quantitative trait locus effects. These findings implicate new pathways for CAD susceptibility.

Genome-wide association studies (GWAS) in CAD have detected common variants with odds ratios of 1.1–1.3 which in aggregate explain only a small proportion of the predicted genetic risk. We hypothesized that the discovery of new susceptibility loci of smaller effect sizes (and, hence, identification of new CAD-related pathways) would be aided by conducting much larger studies in addition to an emphasis on early onset CAD and clearly defined clinical endpoints.

For the discovery stage, 8,424 cases of European ancestry were recruited by the Precocious Coronary Artery Disease (PROCARDIS) study and Heart Protection Study (HPS), and 6,996 cases of South Asian ancestry (chiefly from Pakistan and India) were recruited by the Pakistan Risk of Myocardial Infarction Study (PROMIS) and London Life Sciences Prospective Population (LOLIPOP) study. All studies recruited controls, or supplemented their data with genotypes from common controls, from within the same self-reported ethnic or

linguistic groups from which cases were recruited (Fig. 1). Overall, 81% of the cases had a prior history of myocardial infarction and the remainder had confirmed diagnoses of symptomatic CAD (angina or coronary artery revascularization), with an average age at first event under 60 years (Supplementary Table 1).

All individuals were typed using whole-genome Illumina BeadChips, allowing for a meta-analysis of actual genotypes rather than imputed data. This enabled analysis of low frequency variants (1–5%), which have typically been excluded from GWAS either due to sample size or because imputation has been required to combine data from different genotyping platforms.

As there is population substructure within India and Pakistan⁹, principal component analysis¹⁰ was used in the PROMIS and LOLIPOP studies to identify ancestry informative principal components (Online Methods and Supplementary Fig. 1), which were then used to adjust for population substructure in regression analyses.

Genotypes were tested for association with CAD in the four discovery studies (Online Methods and Supplementary Table 2). Association tests for 574,919 SNPs were entered into pre-specified fixed-effects meta-analyses with study-level correction for genomic control; the meta-analysis groups were: (i) all four studies combined; (ii) the two European studies; and (iii) the two South Asian studies. The genomic control parameters

Discovery			Replication		
	Cases	Controls		Cases	Controls
<i>European studies</i>					
PROCARDIS	5,720 ^a	4,381 ^b		HPS (non-GWAS)	9,248 7,692 ^d
HPS (GWAS)	2,704	2,887 ^c		COROGENE	2,172 1,579
<i>South Asian studies</i>					
PROMIS	4,255	4,098		ISIS	1,971 1,434
LOLIPOP	2,741	3,696		SHEEP/SCARF	1,533 1,893
TOTAL	15,420	15,062	59 SNPs selected for replication →	PROCARDIS TDTs	1,143 1,143
				GISSI-P	979 452
				AMC-PAS	597 1,285
				THISEAS	406 879
				<i>South Asian studies</i>	
				PROMIS	2,092 1,728
				INTERHEART	1,267 1,100
				TOTAL	21,408 19,185

Figure 1 Studies contributing to the discovery and replication meta-analyses. ^aIncludes 2,133 cases who are either full or half siblings of another case. ^bIncludes 2,697 controls from the National Blood Service. ^cIncludes 2,887 controls from the 1958 British Birth cohort. ^dIncludes 5,157 controls from the UK Twins study and 2,535 additional independent PROCARDIS controls not used in the discovery analysis.

*A full list of authors and affiliations appears at the end of the paper.

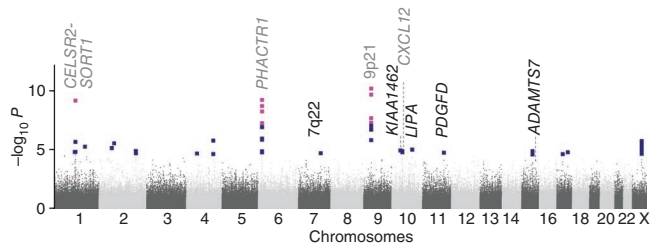


Figure 2 Genome-wide Manhattan plot of P values for all studies (European and South Asian). The $-\log_{10} P$ for 574,919 SNPs from the meta-analysis of the PROCARDIS, HPS, PROMIS and LOLIPOP studies. The y axis is truncated at $-\log_{10} P$ of 12; rs9349379 at the *PHACTR1* locus ($P = 5.8 \times 10^{-19}$) and 15 SNPs at the 9p21 locus ($7.9 \times 10^{-13} > P > 1.3 \times 10^{-25}$) exceed the truncation. SNP associations with CAD that exceeded the genome-wide significance threshold ($P < 5.0 \times 10^{-8}$) are shown in magenta; P values between $P = 4.5 \times 10^{-5}$ and $P = 5.0 \times 10^{-8}$ are shown in blue. The locations of the new replicated loci are annotated in black and previously reported CAD loci (Table 1) with $P < 4.0 \times 10^{-5}$ in the meta-analysis of all studies together are annotated in gray.

(λ_{GC}) after the study-level genomic control correction were 1.03 for the four studies combined, 1.01 for the European studies only and 1.01 for the South Asian studies only; a further meta-analysis-level genomic control correction was applied to each meta-analysis. Heterogeneity was tested between the four discovery GWAS and between the European and South Asian meta-analyses. Manhattan plots are shown in Figure 2 for the meta-analysis of all studies combined and in Supplementary Figure 2 for the ancestry-specific meta-analyses, and the quantile-quantile plots are shown in Supplementary Figure 3.

We confirmed the power and representative nature of our discovery-stage studies with data supporting the relevance of 11 known CAD susceptibility loci with comparable effect sizes to those reported previously (Table 1). We saw directionally consistent effects in the European and South Asian populations for all 11 loci.

We selected 59 SNPs from 50 loci that showed potential new associations from the meta-analysis of the European and South Asian studies (41 SNPs; $P < 1.0 \times 10^{-4}$), the European only meta-analysis (8 SNPs; $P < 3.0 \times 10^{-5}$), the South Asian only meta-analysis (6 SNPs; $P < 3.0 \times 10^{-5}$) and three loci with strong biological plausibility but only suggestive P values (4 SNPs). These SNPs were tested in ten replication studies involving a total of 21,408 CAD cases and 19,185 controls largely by *de novo* genotyping (Fig. 1). The meta-analysis

of the replication association results is shown in Supplementary Table 3. Five SNPs in the newly associated loci achieved the pre-specified threshold for replication ($P < 8.5 \times 10^{-4}$; which is $P < 0.05$ after Bonferroni correction for 59 independent tests), and each also achieved conventional genome-wide significance ($P < 5.0 \times 10^{-8}$), with P values ranging from 2.8×10^{-13} to 3.9×10^{-8} for the combined discovery and replication meta-analysis (Fig. 3). Apparent heterogeneity between the European and South Asian effect for rs4380028 in the *ADAMTS7-MORF4L1* locus in the discovery meta-analysis was not supported by the independent replication. We observed no evidence of ancestry-specific heterogeneity for any of the other previously unidentified loci in either the discovery or replication meta-analyses.

In addition to the five newly associated loci, rs9349379, located in an intron of *PHACTR1*, was significantly associated in the replication alone ($P = 9.9 \times 10^{-10}$) and in the combined discovery and replication ($P = 8.7 \times 10^{-26}$) meta-analyses, and rs17114046, in an intron of *PPAP2B*, showed consistent support in the discovery and replication meta-analyses but fell outside the pre-determined significance level in both the replication alone ($P = 1.1 \times 10^{-3}$) and in the combined discovery and replication ($P = 2.5 \times 10^{-7}$) meta-analyses. It is plausible that other SNPs in the replication study that had suggestive associations have a real effect on CAD; a quantile-quantile plot (Supplementary Fig. 3) of the replication study results shows an overdispersion of the test statistic.

The regional association plots for each confirmed locus are shown in Figure 4. We performed analyses conditioning on the replicated SNP at each new locus in all four discovery studies, and the meta-analysis of these results revealed no evidence of additional independent associations with CAD.

We investigated associations between the expression levels of all genes within 200 kb of each of the confirmed risk SNPs in tissue samples of aortic media and adventitia, mammary artery, carotid plaque, liver, adipose tissue, transformed lymphoblastoid cell lines and skin. The Bonferroni-adjusted significance threshold was $P < 3.1 \times 10^{-4}$ ($P < 0.05$ over 163 tests). Estimates of genotype effect on gene expression are reported in Supplementary Table 4 for the most significantly associated gene in each search window and, when different from the risk SNP, the most significant expression quantitative trait locus (eQTL) SNP.

None of these loci has any previously reported associations with established CAD risk factors (lipids, blood pressure, glucometabolic traits or body mass index). rs4380028 in the *ADAMTS7-MORF4L1*

Table 1 Evidence in the discovery studies for 11 previously reported common variants associated with CAD in GWAS

Gene or locus	SNP	Risk Allele	C4D odds ratio (95% CI)				Odds ratio (95% CI)	P
			European studies		South Asian studies			
			PROCARDIS	HPS	PROMIS	LOLIPOP		
9p21	rs4977574	G	1.27 (1.19–1.35)	1.22 (1.13–1.32)	1.16 (1.09–1.24)	1.16 (1.08–1.25)	1.20 (1.16–1.25)	1.62×10^{-25}
<i>CELSR2-PSRC1-SORT1</i>	rs646776	T	1.32 (1.22–1.42)	1.08 (0.98–1.18)	1.09 (1.01–1.17)	1.07 (0.99–1.17)	1.14 (1.09–1.19)	6.05×10^{-10}
<i>PHACTR1</i>	rs1332844 ^a	T	1.17 (1.09–1.25)	0.99 (0.92–1.07)	1.15 (1.07–1.23)	1.09 (1.01–1.18)	1.11 (1.07–1.15)	5.82×10^{-8}
<i>WDR12</i>	rs6725887	C	1.12 (1.02–1.23)	1.14 (1.03–1.28)	1.13 (0.96–1.33)	0.99 (0.81–1.20)	1.11 (1.05–1.19)	6.19×10^{-4}
<i>SLC5A3-MRPS6-KCNE2</i>	rs7278204 ^b	G	1.21 (1.11–1.33)	1.04 (0.93–1.16)	0.97 (0.86–1.09)	1.07 (0.94–1.23)	1.09 (1.03–1.15)	3.13×10^{-3}
<i>MRAS</i>	rs1199338 ^c	C	1.13 (1.03–1.23)	1.07 (0.96–1.18)	1.07 (0.97–1.18)	1.01 (0.90–1.13)	1.08 (1.02–1.13)	4.34×10^{-3}
<i>LDLR</i>	rs2228671 ^d	C	1.16 (1.05–1.28)	0.99 (0.89–1.12)	1.06 (0.94–1.20)	1.13 (0.98–1.30)	1.09 (1.02–1.15)	5.88×10^{-3}
<i>CXCL12</i>	rs1746048	C	1.10 (1.00–1.21)	1.06 (0.94–1.19)	1.05 (0.99–1.13)	1.03 (0.96–1.12)	1.06 (1.01–1.10)	8.52×10^{-3}
<i>MIA3</i>	rs17011666 ^e	A	1.15 (1.02–1.29)	NA	NA	1.07 (0.99–1.15)	1.09 (1.02–1.16)	1.13×10^{-2}
<i>SH2B3</i>	rs3184504	T	NA	1.05 (0.97–1.13)	1.06 (0.96–1.16)	1.04 (0.93–1.16)	1.05 (1.00–1.11)	6.61×10^{-2}
<i>PCSK9</i>	rs11206510	T	1.13 (1.04–1.23)	0.97 (0.88–1.07)	1.03 (0.90–1.17)	1.02 (0.88–1.18)	1.05 (1.00–1.11)	7.00×10^{-2}

Odds ratios per risk allele and 95% CIs for 11 SNPs previously associated with CAD^{1,3–7,19} are reported in the discovery studies separately and overall.

^aTagging published SNP ($r^2 = 0.90$) rs12526453 (risk allele C). ^bTagging published SNP ($r^2 = 0.85$) rs9982601 (risk allele T). ^cTagging published SNP ($r^2 = 1$) rs9818870 (risk allele T).

^dTagging published SNP ($r^2 = 0.73$) rs1122608 (risk allele G). ^eTagging published SNP ($r^2 = 0.57$) rs17465637 (risk allele C). NA, not available.

SNP Gene or locus	Risk allele	Ethnic group	Discovery				Replication				Discovery + replication			
			Allele freq	Cases/ controls	Odds ratio (95% CI)	<i>P</i>	Ethnic <i>P</i> _{het}	Cases/ controls	Odds ratio (95% CI)	<i>P</i>	Ethnic <i>P</i> _{het}	Odds ratio (95% CI)	<i>P</i>	
rs1412444 <i>LIPA</i>	T	European	0.34	8,412/7,250	1.08 (1.05–1.12)	1.0 × 10 ⁻⁵	0.32	17,224/15,120	1.10 (1.07–1.14)	4.52 × 10 ⁻⁹	0.26	1.09 (1.07–1.12)	2.76 × 10 ⁻¹³	
		S Asian	0.51	6,993/7,788										1,878/1,555
		All	0.42	15,405/15,038										19,102/16,675
rs974819 <i>PDGFD</i>	T	European	0.29	8,413/7,253	1.09 (1.05–1.13)	1.5 × 10 ⁻⁵	0.88	16,651/14,845	1.07 (1.04–1.11)	4.72 × 10 ⁻⁵	0.80	1.07 (1.04–1.09)	2.41 × 10 ⁻⁹	
		S Asian	0.35	6,996/7,790										1,571/1,362
		All	0.32	15,409/15,043										18,222/16,207
rs4380028 <i>ADAMTS7- MORF4L1</i>	C	European	0.60	8,417/7,253	1.07 (1.03–1.11)	3.32 × 10 ⁻⁴ *	3 × 10 ⁻³	17,968/16,314	1.08 (1.04–1.11)	2.81 × 10 ⁻⁶	0.40	1.07 (1.05–1.10)	3.71 × 10 ⁻⁹	
		S Asian	0.70	6,989/7,784										3,192/2,678
		All	0.65	15,406/15,037										21,160/18,992
rs10953541 7q22	C	European	0.75	8,413/7,253	1.10 (1.05–1.15)	1.33 × 10 ⁻⁵	0.59	16,858/14,837	1.07 (1.03–1.11)	3.65 × 10 ⁻⁴	0.66	1.08 (1.05–1.11)	3.12 × 10 ⁻⁸	
		S Asian	0.84	6,994/7,792										2,057/1,702
		All	0.80	15,407/15,045										18,915/16,539
rs2505083 <i>KIAA1462</i>	C	European	0.42	8,416/7,254	1.08 (1.05–1.12)	8.78 × 10 ⁻⁶	0.99	17,800/16,076	1.05 (1.02–1.09)	6.11 × 10 ⁻⁴	0.84	1.07 (1.04–1.09)	3.87 × 10 ⁻⁸	
		S Asian	0.34	6,987/7,789										2,073/1,707
		All	0.38	15,403/15,043										19,873/17,783

Figure 3 Newly identified loci and variants associated with CAD in European, South Asian and all studies. Odds ratios per copy of the risk allele are indicated by squares (size inversely proportional to the variance), with horizontal lines indicating 95% CIs. Odds ratios and 95% CIs for all participants are indicated by diamonds. Allele frequencies (allele freq) are given for the risk allele. *P* values for heterogeneity between European and South Asian (S Asian) results are reported (ethnic *P*_{het}). *The South Asian only discovery *P* value for rs4380028 was *P* = 3.6 × 10⁻⁶.

locus is ~200 kb downstream of a robust QTL for cigarettes smoked per day, but a conditional analysis for cigarettes smoked per day showed no attenuation of the CAD risk (*P* = 0.38).

Inspection of the regional association plots of each locus (Fig. 4), reinforced by our own or previously published eQTL data, suggests that specific genes can be implicated at two of the new loci (rs1412444 and rs974819).

rs1412444 is in an intron of the *LIPA*, the lysosomal acid lipase gene. The risk allele of this SNP has been strongly linked with increased

expression level of *LIPA* mRNA in circulating monocytes¹¹, was the lead eQTL SNP for *LIPA* in our data and had suggestive association with increased expression in liver (*P* < 1.6 × 10⁻³; Supplementary Table 4). Loss-of-function alleles in *LIPA* cause a Mendelian cholesterol ester storage disorder with hypercholesterolaemia and increased atherosclerosis, but rs1412444 is, at most, very weakly associated with LDL cholesterol levels¹¹. This suggests that risk is mediated not by interfering with hepatic cholesterol ester hydrolysis (and thereby LDL receptor downregulation) but rather by some other mechanism.

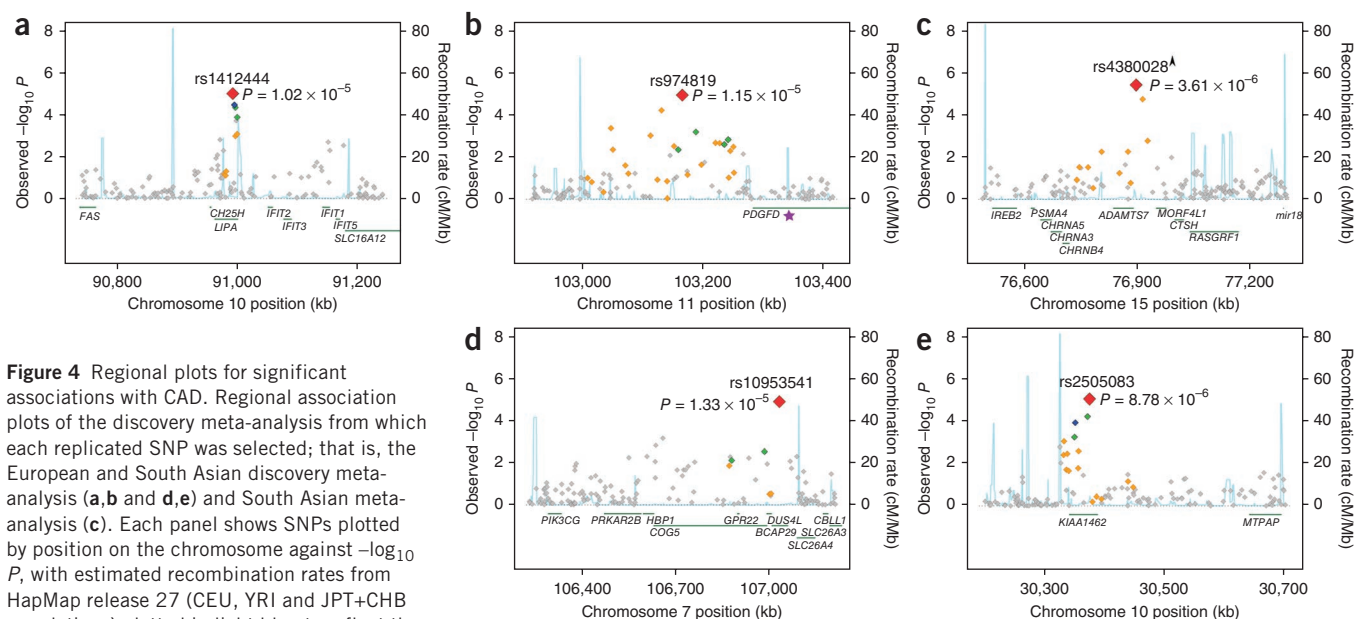


Figure 4 Regional plots for significant associations with CAD. Regional association plots of the discovery meta-analysis from which each replicated SNP was selected; that is, the European and South Asian discovery meta-analysis (a, b and d, e) and South Asian meta-analysis (c). Each panel shows SNPs plotted by position on the chromosome against $-\log_{10} P$, with estimated recombination rates from HapMap release 27 (CEU, YRI and JPT+CHB populations) plotted in light blue to reflect the local LD structure on a secondary y axis. The most significant lead SNP (red diamond) at each locus is annotated with its discovery *P* value, and flanking SNPs are color-coded to represent the pairwise r^2 measure of LD with the lead SNP: blue, $r^2 \geq 0.8$; green, $0.5 \leq r^2 < 0.8$; orange, $0.2 \leq r^2 \leq 0.5$; gray, $r^2 < 0.2$. a, b and d, e report r^2 values calculated from HapMap2 CEU reference samples, and c reports r^2 values calculated from HapMap2 GIH reference samples. Green bars represent RefSeq genes in the region. All positions are on NCBI build 36. The purple star in b represents significant eQTL association in aortic media. The black arrowhead in c represents the South Asian only meta-analysis *P* value.

The nearest gene to rs974819 is *PDGFD*, 117 kb downstream in an adjacent block of linkage disequilibrium (LD) (Fig. 4). We found a significant *PDGFD* eQTL for rs974819 in aortic media ($P < 2.3 \times 10^{-7}$), with suggestive associations in aortic adventitia ($P < 7.7 \times 10^{-4}$) and mammary artery ($P < 7.2 \times 10^{-4}$) (Supplementary Table 4). In all three tissues, the risk allele was associated with increased expression, and *PDGFD* was one of the top 10% most highly expressed transcripts. Platelet-derived growth factor D, encoded by *PDGFD*, is expressed in several cell types in atherosclerotic plaques and is predicted to stimulate atherosclerosis by influencing matrix metalloproteinase activity and monocyte migration¹² and by inhibiting smooth muscle cell gene expression¹³.

rs4380028 is 7.6 kb upstream of *ADAMTS7*. *ADAMTS7* is a metalloproteinase that accumulates in carotid artery neointima after injury and is upregulated by platelet-derived growth factor¹⁴. The CAD association at the *ADAMTS7* locus appears to be independent of the adjacent QTL for cigarette smoking, which has been attributed to the *CHRNA5-CHRNA3-CHRNA4* nicotinic acid receptor cluster¹⁵. The signal at rs4380028 therefore appears new but cannot be attributed to a specific gene.

At the 7q22 locus, the lead SNP, rs10953541, is within an intron of *BCAP29*, which encodes B-cell receptor-associated protein 29, but this SNP is also in strong LD with five other protein coding genes (*PRKAR2B*, *HBP1*, *COG5*, *GPR22* and *DUS4L*). We observed no unique eQTL effect. One promising biological candidate is *GPR22*, which encodes a G-protein-coupled receptor expressed in coronary arteries and heart¹⁶.

rs2505083 is in an intron of *KIAA1462*, a widely expressed and evolutionarily conserved gene. The function of the 1,359-amino-acid protein encoded by *KIAA1462* is largely unknown, with no recognizable functional domains and little homology to other protein families.

The genetic evidence for rs17114046 in *PPAP2B* is further supported by the association of the risk allele, with a suggestive ($P < 3.9 \times 10^{-3}$) 30% increase in *PPAP2B* expression in atherosclerotic plaque (Supplementary Table 4). *PPAP2B* encodes phosphatidic acid phosphatase type 2B, a membrane glycoprotein that hydrolyzes bioactive lipids involved in signaling¹⁷. It is expressed in the adherens junctions of endothelial cells and is believed to influence endothelial cell adhesion and migration and vasculogenesis¹⁸.

In summary, five new loci passed both the pre-specified significance threshold for the stand-alone replication and the conventional threshold for genome-wide significance in the combined discovery and replication data, yielding a substantial increase in the number of confirmed susceptibility loci for CAD. We did not find any susceptibility variants with material differences in effect size or allele frequency between South Asians and Europeans. We note, however, the potential limitation that current genome-wide arrays may not capture all important variants in South Asians. Nevertheless, all of the known and new variants were significantly associated with CAD risk in both the European and South Asian populations in the current study, indicating the importance of the genes associated with CAD beyond the European ancestry groups in which they were first defined.

As is seen in the present study, the effect sizes of previously unidentified CAD-associated genes discovered by GWAS have become progressively smaller, suggesting that there may not be large-effect common variants remaining to be discovered, but rather that a large number of common variants of small effect may contribute to CAD risk. Reliable detection of differences of only 5–10% in the per-allele risk of CAD has previously been difficult. However, the availability of large-scale GWAS carried out in populations of different ancestry, and our demonstration that results from such populations

can be informatively combined in genetic discovery, suggests that even broader collaborations would identify additional variants that influence CAD risk. Greater understanding of the genetic variants underlying CAD, and particularly the pathways involved, may lead to development of new therapeutic approaches to help address the world's leading cause of death.

URLs. TRANSMIT software, <http://www-gene.cimr.cam.ac.uk/clayton/software/transmit.txt>; METAL software, <http://www.sph.umich.edu/csg/abecasis/Metal/>.

Note added in proof: Since this manuscript was submitted *ADAMTS7* (ref. 20) and *KIAA1462* (ref. 21) have been independently reported as loci for coronary artery disease.

METHODS

Methods and any associated references are available in the online version of the paper at <http://www.nature.com/naturegenetics/>.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We are grateful to all of the study participants in the studies contributing to these meta-analyses and to all laboratory staff and former colleagues who have contributed to these studies over many years. This study makes use of data generated by the Wellcome Trust Case-Control Consortium (WTCCC); a full list of the investigators who contributed to the generation of the data is available from www.wtccc.org.uk. This study also makes use of data generated by the UK Twins study, The Twin Research Unit, King's College London, UK.

The PROCARDIS study was supported by the European Community Sixth Framework Program (LSHM-CT-2007-037273), AstraZeneca, the British Heart Foundation, the Oxford British Heart Foundation Centre of Research Excellence, the Wellcome Trust (075491/Z/04), the Swedish Research Council, the Knut and Alice Wallenberg Foundation, the Swedish Heart-Lung Foundation, the Torsten and Ragnar Söderberg Foundation, the Strategic Cardiovascular Program of Karolinska Institutet and Stockholm County Council, the Foundation for Strategic Research and the Stockholm County Council (560283).

The Heart Protection Study (ISRCTN48489393) was funded by the UK Medical Research Council, the British Heart Foundation, Merck & Co and Roche Vitamins Ltd. Genotyping and analysis was supported by a grant to Oxford University and the Centre National de Génotypage (CNG) from Merck & Co and the Oxford BHF Centre of Research Excellence.

The PROMIS study was funded by unrestricted grants to investigators at the University of Cambridge, UK and at the Centre for Non-Communicable Diseases, Pakistan. Genotyping was funded by the Wellcome Trust.

The LOLIPOP study is supported by the National Institute for Health Research Comprehensive Biomedical Research Centre Imperial College Healthcare NHS Trust, Ealing Hospital NHS Trust, the British Heart Foundation (SP/04/002), the Medical Research Council (G0700931, G0601966), the Wellcome Trust (084723/Z/08/Z) and the National Institute for Health Research (RP-PG-0407-10371). P.E. is a National Institute for Health Research Senior Investigator. This work was facilitated by Barts and The London National Institute for Health Biomedical Research Unit. We thank the participants and research staff who made the study possible.

The COROGENE-FINRISK study was supported in part by the Aarno Koskela Foundation and the Finnish Foundation for Cardiovascular Research.

The Biobank of Karolinska Carotid Endarterectomies (BiKE) and Advanced Study of Aortic Pathology (ASAP) eQTL study was supported by the Swedish Heart-Lung Foundation, the Swedish Research Council, the European Commission (FAD, Health-F22008-200647), (AtheroRemo HEALTH-2007-A-201668), DASTI (Danish Agency for Science, Technology and Innovation) and a donation from F. Lundberg.

The Multiple Tissue Human Expression Resource (MuTHER) study was supported by the Wellcome Trust (081917/Z/07/Z).

N. Soranzo is supported by the Wellcome Trust (Core Grant Number 091746/Z/10/Z).

AUTHOR CONTRIBUTIONS

Steering and writing committee: J.F.P., J.C.H., D.S., J.C.C., J.H., N. Soranzo, R. Collins, J.D., P. Elliott, M.F., K.S., W.Z., A. Hamsten, S. Parish, M.L., H.W. (Chair), R. Clarke, P. Deloukas, J.S.K.

Corresponding authors: H.W., D.S., R. Collins, J.S.K.

Analysis committee: J.C.H., W.Z., N. Soranzo, J.F.P., D.S., J.C.C., S. Parish, M.F. (Chair).

Statistical genetics and bioinformatics: HPS: J.C.H., S. Parish; LOLIPOP: W.Z.; PROCARDIS: A. Goel, H.O., R.J.S., S.H., A.M., A. Helgadottir, J.O., M.F., J.F.P.; PROMIS: D.S., K.S., M.M., S. Potter, S.E.H., P. Deloukas.

Genotyping: CNG: J.H., M.D., M.L.; Karolinska: R.J.S.; Oxford: S.J., H.O.; Uppsala: T.A., A.C.S.; WTSI: R.G., S. Bumpsted, E.G., S.E., P.D.

Expression QTL analyses: L.F., T.K., A.F.C., A. Gabrielsen, U.S., the MuTHER consortium, P. Eriksson.

Discovery cohorts: HPS: J.C.H., S. Parish, A.O., R. Clarke, L.B., P.S., J.A., R.P., R. Collins; LOLIPOP: J.C.C., G. Abecasis, N.A., M.C., P. Donnelly, P. Elliott, P.F., A.S.K., M.I.C., N.J.S., J. Scott, J. Sehmi, W.Z., J.S.K.; PROCARDIS: Sweden: A. Silveira, M.L.H., F.M.v.H., G.O., A. Hamsten; Germany: S. Rust, G. Assmann, U.S.; Italy: S. Barlera, G.T., M.G.F.; UK: R. Clarke, P.L., J.C.H., R. Collins, J.F.P., F.R.G., M.F., H.W.; PROMIS: D.S., A.R., M.Z., N. Shah, M.S., N.H.M., M.A., K.S.Z., A. Samad, M. Ishaq, A.R.G., F.M., N.J.S., P.M.F., P.D., J.D.

Replication cohorts: UK Twins: N. Soranzo, T.S.; COROGENE-FINRISK: L.P., M.S.N., J. Sinisalo, V.S., S. Ripatti; ISIS: J.C.H., D.B., S. Parish; SHEEP/SCARF: K.L., B.G., U.d.F.; GISSI-P: S. Pietri, F.G., R.M.; AMC-PAS: S.S., J.J.P.K., M.D.T.; THISEAS: E.V.T., G.V.D.; INTERHEART: J.C.E., S.Y., S.S.A.

For further details on author contributions, see the **Supplementary Note**.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at <http://www.nature.com/naturegenetics/>.

Reprints and permissions information is available online at <http://npg.nature.com/reprintsandpermissions/>.

1. Erdmann, J. *et al.* New susceptibility locus for coronary artery disease on chromosome 3q22.3. *Nat. Genet.* **41**, 280–282 (2009).
2. Helgadottir, A. *et al.* A common variant on chromosome 9p21 affects the risk of myocardial infarction. *Science* **316**, 1491–1493 (2007).
3. Kathiresan, S. *et al.* Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. *Nat. Genet.* **41**, 334–341 (2009).
4. McPherson, R. *et al.* A common allele on chromosome 9 associated with coronary heart disease. *Science* **316**, 1488–1491 (2007).
5. Samani, N.J. *et al.* Genome-wide association analysis of coronary artery disease. *N. Engl. J. Med.* **357**, 443–453 (2007).

Steering and Writing committee: John F Peden^{1,2,54}, Jemma C Hopewell^{3,54}, Danish Saleheen^{4,5,54}, John C Chambers^{6,7,54}, Jorg Hager^{8,54}, Nicole Soranzo^{9,54}, Rory Collins^{3,54}, John Danesh^{5,54}, Paul Elliott^{6,10,54}, Martin Farrall^{1,2,54}, Kathy Stirrups^{9,54}, Weihua Zhang^{6,7,54}, Anders Hamsten^{11,12,54}, Sarah Parish^{3,54}, Mark Lathrop^{8,54}, Hugh Watkins (Chair)^{1,2,54}, Robert Clarke^{3,54}, Panos Deloukas^{9,54} & Jaspal S Kooner^{7,13,54}

Statistical genetics and bioinformatics: Anuj Goel^{1,2}, Halit Ongen^{1,2}, Rona J Strawbridge^{11,12}, Simon Heath⁸, Anders Mälarstig^{11,12}, Anna Helgadottir^{1,2}, John Öhrvik^{11,12}, Muhammed Murtaza⁹, Simon Potter¹⁴ & Sarah E Hunt¹⁴

Genotyping: Marc Delepine⁸, Shapour Jalilzadeh^{1,2}, Tomas Axelsson¹⁵, Ann-Christine Syvanen¹⁵, Rhian Gwilliam¹⁴, Suzannah Bumpstead¹⁴, Emma Gray¹⁴ & Sarah Edkins¹⁴

Expression QTL analyses: Lasse Folkersen^{11,12}, Theodosios Kyriakou^{1,2}, Anders Franco-Cereceda¹⁶, Anders Gabrielsen¹⁷, Udo Seedorf¹⁸, the MuTHER consortium & Per Eriksson^{11,12}

Discovery cohorts: Alison Offer³, Louise Bowman³, Peter Sleight², Jane Armitage³, Richard Peto³, Goncalo Abecasis¹⁹, Nabeel Ahmed²⁰, Mark Caulfield²¹, Peter Donnelly^{22,23}, Philippe Froguel²⁴, Angad S Kooner²⁰, Mark I McCarthy^{22,25,26}, Nilesh J Samani^{27,28}, James Scott¹³, Joban Sehmi¹³, Angela Silveira^{11,12}, Mai-Lis Hellénus²⁹, Ferdinand M van 't Hooft^{11,12}, Gunnar Olsson^{30,31}, Stephan Rust¹⁸, Gerd Assmann¹⁸, Simona Barlera³², Gianni Tognoni^{32,33}, Maria Grazia Franzosi³², Pamela Linksted³, Fiona R Green³⁴, Asif Rasheed⁴, Moazzam Zaidi⁴, Nabi Shah⁴, Maria Samuel⁴, Nadeem H Mallick³⁵, Muhammad Azhar³⁵, Khan S Zaman³⁶, Abdus Samad³⁷, Mohammad Ishaq³⁷, Ali R Gardezi³⁸, Fazal-ur-Rehman Memon³⁹ & Philippe M Frossard⁴

Replication cohorts: Tim Spector⁴⁰, Leena Peltonen^{9,41,45}, Markku S Nieminen⁴², Juha Sinisalo⁴³, Veikko Salomaa⁴⁴, Samuli Ripatti^{44,45}, Derrick Bennett³, Karin Leander⁴⁶, Bruna Gigante⁴⁶, Ulf de Faire⁴⁶,

Silvia Pietri³², Francesca Gori³², Roberto Marchioli⁴⁷, Suthesh Sivapalaratnam⁴⁸, John J P Kastelein⁴⁸, Mieke D Trip⁴⁸, Eirini V Theodoraki⁴⁹, George V Dedoussis⁴⁹, Jamie C Engert^{50,51}, Salim Yusuf⁵² & Sonia S Anand^{52,53}

¹Department of Cardiovascular Medicine, The Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK. ²Department of Cardiovascular Medicine, University of Oxford, John Radcliffe Hospital, Headington, Oxford, UK. ³Clinical Trial Service Unit, University of Oxford, Oxford, UK. ⁴Center for Non-Communicable Diseases Pakistan, Karachi, Pakistan. ⁵Department of Public Health and Primary Care, University of Cambridge, Strangeways Research Laboratory, Cambridge, UK. ⁶Epidemiology and Biostatistics, Imperial College London, Norfolk Place, London, UK. ⁷Cardiology, Ealing Hospital National Health Service (NHS) Trust, Middlesex, UK. ⁸Commissariat à l'Energie Atomique (CEA) Genomics Institute-Centre National de Génotypage, Evry Cedex, France. ⁹Human Genetics, Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, UK. ¹⁰Medical Research Council-Health Protection Agency (MRC-HPA) Centre for Environment and Health, Imperial College London, London, UK. ¹¹Atherosclerosis Research Unit, Department of Medicine, Karolinska Institutet, Karolinska University Hospital, Stockholm, Sweden. ¹²Center for Molecular Medicine, Karolinska University Hospital, Stockholm, Sweden. ¹³National Heart and Lung Institute, Imperial College London, London, UK. ¹⁴Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, UK. ¹⁵Department of Medical Sciences, Molecular Medicine, Uppsala University, Uppsala, Sweden. ¹⁶Thoracic Surgery Unit, Department of Molecular Medicine and Surgery, Karolinska Institutet, Karolinska University Hospital, Stockholm, Sweden. ¹⁷Cardiology Unit, Department of Medicine, Karolinska Institutet, Karolinska University Hospital, Stockholm, Sweden. ¹⁸Gesellschaft für Arterioskleroseforschung e.V., Leibniz-Institut für Arterioskleroseforschung an der Universität Münster (LIFA), Münster, Germany. ¹⁹Department of Biostatistics, University of Michigan, Ann Arbor, Michigan, USA. ²⁰Cardiology, Ealing Hospital NHS Trust, Middlesex, UK. ²¹William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK. ²²Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK. ²³Department of Statistics, Oxford University, Oxford, UK. ²⁴Genomic Medicine, Imperial College London, London, UK. ²⁵Oxford Centre for Diabetes, Endocrinology and Metabolism, Oxford University, Oxford, UK. ²⁶Oxford National Institute of Health Research (NIHR) Biomedical Research Centre, Churchill Hospital, Headington, UK. ²⁷Department of Cardiovascular Sciences, University of Leicester, Glenfield Hospital, Leicester, UK. ²⁸Leicester NIHR Biomedical Research Unit in Cardiovascular Disease, Glenfield Hospital, Leicester, UK. ²⁹Cardiology Unit, Department of Medicine, Karolinska Institutet, Karolinska University Hospital, Stockholm, Sweden. ³⁰Cardiovascular Drug Research at the Department of Medicine, Solna, Sweden. ³¹Cardiovascular and Gastrointestinal Innovative Medicines, Global Research and Development, AstraZeneca, Sweden. ³²Department of Cardiovascular Research, Istituto Mario Negri, Milano, Italy. ³³Consorzio Mario Negri Sud, Santa Maria Imbaro (Chieti), Italy. ³⁴Biochemical Sciences Division, Faculty of Health and Medical Sciences, University of Surrey, Guilford, UK. ³⁵Department of Cardiology, Punjab Institute of Cardiology, Jail Road, Lahore, Pakistan. ³⁶Department of Cardiology, National Institute of Cardiovascular Diseases, Karachi, Pakistan. ³⁷Department of Cardiology, Karachi Institute of Heart Diseases, Federal B. Area, Karachi, Pakistan. ³⁸Department of Cardiology, Ch. Pervaiz Elahi Institute Of Cardiology, Multan, Pakistan. ³⁹Department of Cardiology, Red Crescent Institute of Cardiology, Latifabad, Hyderabad, Pakistan. ⁴⁰Department of Twin Research and Genetic Epidemiology, Kings College London, London, UK. ⁴¹Department of Medical Genetics, University of Helsinki and the Helsinki University Central Hospital, Helsinki, Finland. ⁴²Division of Cardiology, Helsinki University Central Hospital, Helsinki, Finland. ⁴³University Central Hospital, Cardiovascular Laboratory, Helsinki, Finland. ⁴⁴National Institute for Health and Welfare, Helsinki, Finland. ⁴⁵Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, Finland. ⁴⁶Cardiovascular Epidemiology Unit, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden. ⁴⁷Laboratory of Clinical Epidemiology of Cardiovascular Disease, Consorzio Mario Negri Sud, Santa Maria Imbaro, Italy. ⁴⁸Vascular Medicine, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands. ⁴⁹Department of Dietetics-Nutrition, Harokopio University, Athens, Greece. ⁵⁰McGill University Department of Medicine, Montreal, Quebec, Canada. ⁵¹McGill University Department of Human Genetics, Montreal, Quebec, Canada. ⁵²Population Health Research Institute, Hamilton Health Sciences, McMaster University, Hamilton, Canada. ⁵³Department of Medicine, McMaster University, Hamilton, Canada. ⁵⁴These authors contributed equally to this work. Correspondence should be addressed to H.W. (hugh.watkins@cardiov.ox.ac.uk), D.S. (DS436@medschl.cam.ac.uk), R.C. (rory.collins@ctu.ox.ac.uk) or J.S.K. (j.kooner@imperial.ac.uk).

ONLINE METHODS

Discovery. Cases and controls in the discovery studies. The PROCARDIS study comprised individuals recruited from the UK, Italy, Sweden and Germany²². All cases had a diagnosis of CAD before age 66 and 80% had a sibling in whom CAD had been diagnosed before age 66. CAD was defined as clinically documented evidence of myocardial infarction (80%), coronary artery bypass graft (10%), acute coronary syndrome (6%), coronary angioplasty (1%) or stable angina (hospitalization for angina or documented obstructive coronary disease) (3%) before age 66. The mean age of onset of CAD was 53.2 years (standard deviation (s.d.), 7.2 years).

The Heart Protection Study (HPS) was a large UK-based cholesterol-lowering trial involving participants with a history of myocardial infarction, unstable or stable angina, coronary artery bypass graft or coronary angioplasty (as well as individuals with prior history of stroke or hypertension)²³. Among these, 2,704 CAD cases were genotyped and compared with 2,887 controls from the UK 1958 British Birth Cohort. Of those individuals genotyped, the mean age of CAD onset was 58.8 years (s.d., 8.4 years), and 92% of CAD cases had a history of myocardial infarction, revascularization or hospitalization for angina.

The PROMIS study was a case-control study of myocardial infarction carried out among South Asians living in urban Pakistan²⁴. The 4,253 cases had myocardial infarction, and the average age of disease onset was 53.8 years (s.d., 10.6 years). The 4,130 controls were matched to cases by sex and age and recruited in the same hospitals as the index cases. The major ethnic groups of the participants in the PROMIS study were as follows: Urdu (42%), Punjabi (28%), Pathan (8%) and Sindhi (8%).

The LOLIPOP study was a case-control study of CAD carried out among South Asians, a population category defined by having all four grandparents born on the Indian subcontinent, living in the UK²⁵. The 2,741 cases had a history of myocardial infarction, coronary artery revascularization (coronary artery bypass grafting or percutaneous coronary intervention) or angiographically confirmed coronary artery stenosis greater than 50%. The cases were compared with 3,696 controls, free from CAD, recruited from the same ethnic or linguistic groups. The mean age of entry into the study was 59.3 years (s.d., 9.7 years). For further details of the discovery cohorts, see the **Supplementary Note**.

Genotyping in the discovery studies. The genotyping platforms, centers and quality control parameters are summarized in **Supplementary Table 2**. Additional quality control was done by testing for different allelic frequencies between centers, bead arrays, batches or BeadChips, and a further 3,883 SNPs were removed.

To identify ancestry outliers, principal component analysis using EIGENSOFT (3.0)¹⁰ was used to compare all samples with reference samples from the HapMap YRI, CHB, JPT and CEU panels. Samples with eigenvalues inconsistent with self-reported ancestry were removed. To detect and correct for population stratification, a panel of independent SNPs were used to generate ancestry-informative principal components. South Asian samples were plotted by their self-reported ethnic or linguistic group and case or control status (**Supplementary Fig. 1**).

Ethics. All studies were collected under the approval of the appropriate ethics committee and all participants gave informed consent for each study.

Statistical analyses. Discovery genome-wide analyses. Association analyses of CAD were carried out after exclusion of SNPs that failed quality control. Analyses were performed using logistic models to estimate regression coefficients and their standard errors. Additive genetic effects were modeled by defining continuous variables with levels 0, 1 and 2 corresponding to genotypes AA, AB and BB. The test statistics for each study were inspected for over-dispersion, and the genomic control parameter (λ_{GC}), an estimate of the variance inflation, was between 1.03 to 1.08 (**Supplementary Table 2**).

Clustering and covariate adjustment. The multi-center PROCARDIS study included country of origin as a categorical main effect to model differences in SNP allele frequencies indirectly across the populations. The familial relatedness in the PROCARDIS study was taken into account using a robust (sandwich) estimator of the variance. The PROMIS and LOLIPOP studies included ancestry-informative principal components to absorb population stratification.

Discovery meta-analysis. A fixed-effects inverse variance-weighted meta-analysis as implemented by METAL (see URLs) was used to combine the individual studies concurrently as European studies only, South Asian studies only and all studies. The individual studies were subjected to study-level correction for genomic control (using the sample-specific λ_{GC} described above). The test statistics for each meta-analysis were inspected for over-dispersion (**Supplementary Fig. 3**). The λ_{GC} of the three meta-analyses were calculated ($\lambda_{GC} = 1.01$ Europeans only, $\lambda_{GC} = 1.01$ South Asians only and $\lambda_{GC} = 1.03$ for all studies), and an additional genomic control correction was applied using these λ_{GC} factors (double genomic control correction). Heterogeneity between the European and the South Asian studies and between all studies was assessed using the Cochran's Q statistic.

Replication. Cases and controls in the replication studies. Independent replication was sought in a total of 21,408 CAD cases and 19,185 ancestry-matched controls obtained from eight European and two South Asian studies. A total of 18,049 CAD cases and 16,357 controls of European ancestry were included from HPS²³, COROGENE-FINRISK²⁶, International Study of Infarct Survival (ISIS)²⁷, Stockholm Heart Epidemiology Programme and Stockholm Coronary Artery Risk Factor study SHEEP/SCARF^{28,29}, Precocious Coronary Artery Disease Transmission Disequilibrium Test (PROCARDIS TDT) cohort³⁰, Gruppo Italiano per lo Studio della Sopravvivenza nell'Infarto Miocardico Prevenzion (GISSI-P)³¹, Academic Medical Center Amsterdam Premature Atherosclerosis Study (AMC-PAS)³², The Hellenic study of Interactions between SNPs & Eating in Atherosclerosis Susceptibility (THISEAS)³³, and 3,359 cases and 2,828 controls of South Asian ancestry from the PROMIS²⁴ and INTERHEART³⁴ studies.

Selection of replication SNPs. SNPs were ranked by the association P value in the discovery meta-analysis. The genotype clusters for all SNPs with an association with CAD of $P < 1.0 \times 10^{-4}$ were manually checked, and 13 SNPs were removed. With the exception of rs9349379 at the *PHACTR1* locus, SNPs that were located within a locus that had been previously reported to be associated with CAD (**Table 1**), or SNPs in strong LD ($r^2 > 0.5$) with the most significant SNP at each locus, were removed. All remaining SNPs with $P < 3.0 \times 10^{-5}$ in either the European studies only or South Asian studies or with $P < 1.0 \times 10^{-4}$ in the all-studies meta-analysis were considered for replication. When a SNP could not be multiplexed, alternative tagging SNPs ($r^2 > 0.8$) were considered.

Genotyping in the replication studies. Genotyping of the replication samples was performed by primer extension and MALDI-TOF mass spectrometry using Sequenom iPLEX technology as two multiplexes containing 59 replication SNPs and three gender-specific polymorphisms. At the Centre National de Génotypage (CNG), the assay for rs9349379 was replaced by a Kaspar SNP assay supplied by KBioscience Ltd. The PROMIS replication genotypes included data from 1,005 Illumina Human660 arrays which were not available during the discovery GWAS. The COROGENE replication genotypes were obtained *in silico* from Illumina Human670 GWAS data.

The allelic intensities of each SNP assay were plotted and, where necessary, genotypes were manually called. SNPs with genotype calls that deviated from Hardy-Weinberg equilibrium with $P < 8.5 \times 10^{-4}$ or had minor allele frequency <1% were excluded. The call rate cutoff for SNPs was empirically established in each study. SNP genotypes that passed quality control were used to determine a study-specific sample rate cutoff. Details of replication genotyping and related quality control steps are summarized in **Supplementary Table 2**. The PROCARDIS TDT study was a family based collection, and an additional quality control step was applied whereby families with greater than three Mendelian misinheritances were excluded.

Statistical analyses. Replication analyses. Logistic regression models were used to perform association analyses of CAD and to estimate the per-allele effect and standard error. The family based transmission disequilibrium test was used to estimate the association in the PROCARDIS TDT samples. This test is robust to the presence of population structure and was performed using TRANSMIT (see URLs) written by David Clayton.

Replication meta-analysis. A fixed-effects inverse-variance-weighted meta-analysis was used to combine the results for each SNP across all replication studies with available data. Heterogeneity between all studies and between the European and the South Asian studies was assessed using Cochran's Q statistic.

Conditional analysis. Conditional analyses were performed in each of the discovery studies. The position of the recombination hotspots flanking each of the SNPs selected for replication was determined, and those SNPs that were located between the two flanking hotspots were tested for association with CAD with the equivalent model used in the discovery analysis but conditioning on the genotype of the SNP chosen for replication.

eQTL analyses. Tissue samples for gene expression biobanks were obtained as described previously³⁵. Briefly, tissue biopsies were taken from patients undergoing carotid endarterectomy (plaque $n = 117$) or valve surgery (liver $n = 152$, aorta media $n = 117$, aorta adventitia $n = 103$ and mammary artery $n = 88$). Extracted RNA was hybridized to Affymetrix HG-U133 plus 2.0 microarrays (plaque) or Affymetrix ST 1.0 Exon arrays (liver, aorta and mammary artery), and obtained scans were robust multichip average (RMA) normalized and log₂ transformed. DNA was extracted from circulating blood cells and hybridized to Illumina Human610w-Quad Beadarrays. In the MuTHER study, RNA levels were measured in lymphoblastoid cell lines ($n = 826$), skin ($n = 705$) and fat biopsies ($n = 825$) from 850 well-phenotyped female twins (1/3 MZ and 2/3 DZ) from the TwinsUK resource using Illumina's whole-genome expression array HumanHT-12 version 3 as previously described³⁶. Genotyping was performed in parallel using Illumina's whole-genome arrays. Associations between genotype and expression of genes were assessed using an additive linear model.

22. Broadbent, H.M. *et al.* Susceptibility to coronary artery disease and diabetes is encoded by distinct, tightly linked SNPs in the *ANRIL* locus on chromosome 9p. *Hum. Mol. Genet.* **17**, 806–814 (2008).
23. Heart Protection Study Collaborative Group. MRC/BHF Heart Protection Study of cholesterol lowering with simvastatin in 20,536 high-risk individuals: a randomised placebo-controlled trial. *Lancet* **360**, 7–22 (2002).
24. Saleheen, D. *et al.* The Pakistan Risk of Myocardial Infarction Study: a resource for the study of genetic, lifestyle and other determinants of myocardial infarction in South Asia. *Eur. J. Epidemiol.* **24**, 329–338 (2009).
25. Chambers, J.C. *et al.* Common genetic variation near *MC4R* is associated with waist circumference and insulin resistance. *Nat. Genet.* **40**, 716–718 (2008).
26. Ripatti, S. *et al.* A multilocus genetic risk score for coronary heart disease: case-control and prospective cohort analyses. *Lancet* **376**, 1393–1400 (2010).
27. Clarke, R. *et al.* Lymphotoxin-alpha gene and risk of myocardial infarction in 6,928 cases and 2,712 controls in the ISIS case-control study. *PLoS Genet.* **2**, e107 (2006).
28. Reuterwall, C. *et al.* Higher relative, but lower absolute risks of myocardial infarction in women than in men: analysis of some major risk factors in the SHEEP study. The SHEEP Study Group. *J. Intern. Med.* **246**, 161–174 (1999).
29. Samnegard, A. *et al.* Serum matrix metalloproteinase-3 concentration is influenced by MMP-3 -1612 5A/6A promoter genotype and associated with myocardial infarction. *J. Intern. Med.* **258**, 411–419 (2005).
30. PROCARDIS Consortium. A trio family study showing association of the lymphotoxin-alpha N26 (804A) allele with coronary artery disease. *Eur. J. Hum. Genet.* **12**, 770–774 (2004).
31. GISSI-Prevenzione Investigators. Dietary supplementation with n-3 polyunsaturated fatty acids and vitamin E after myocardial infarction: results of the GISSI-Prevenzione trial. Gruppo Italiano per lo Studio della Sopravvivenza nell'Infarto miocardico. *Lancet* **354**, 447–455 (1999).
32. Trip, M.D. *et al.* Frequent mutation in the *ABCC6* gene (R1141X) is associated with a strong increase in the prevalence of coronary artery disease. *Circulation* **106**, 773–775 (2002).
33. Theodoraki, E.V. *et al.* Fibrinogen beta variants confer protection against coronary artery disease in a Greek case-control study. *BMC Med. Genet.* **11**, 28 (2010).
34. Mente, A. *et al.* Metabolic syndrome and risk of acute myocardial infarction a case-control study of 26,903 subjects from 52 countries. *J. Am. Coll. Cardiol.* **55**, 2390–2398 (2010).
35. Folkersen, L. *et al.* Association of genetic risk variants with expression of proximal genes identifies novel susceptibility genes for cardiovascular disease. *Circ. Cardiovasc. Genet.* **3**, 365–373 (2010).
36. Stranger, B.E. *et al.* Population genomics of human gene expression. *Nat. Genet.* **39**, 1217–1224 (2007).