nature genetics

Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease

We performed a meta-analysis of 14 genome-wide association studies of coronary artery disease (CAD) comprising 22,233 individuals with CAD (cases) and 64,762 controls of European descent followed by genotyping of top association signals in 56,682 additional individuals. This analysis identified 13 loci newly associated with CAD at $P < 5 \times 10^{-8}$ and confirmed the association of 10 of 12 previously reported CAD loci. The 13 new loci showed risk allele frequencies ranging from 0.13 to 0.91 and were associated with a 6% to 17% increase in the risk of CAD per allele. Notably, only three of the new loci showed significant association with traditional CAD risk factors and the majority lie in gene regions not previously implicated in the pathogenesis of CAD. Finally, five of the new CAD risk loci appear to have pleiotropic effects, showing strong association with various other human diseases or traits.

It has been estimated that heritable factors account for 30%–60% of the inter-individual variation in the risk of coronary artery disease (CAD)[1]. Recently, genome-wide association studies (GWAS) have identified several common variants that associate with risk of CAD[2]. However, in aggregate, these variants explain only a small fraction of the heritability of CAD, probably partly due to the limited power of previous studies to discover effects of modest size. Recognizing the need for larger studies, we formed the transatlantic Coronary ARtery DIsease Genome wide Replication and Meta-analysis (CARDIoGRAM) consortium[3]. We perfomed a meta-analysis of 14 GWAS of CAD comprising 22,233 cases and 64,762 controls, all of European ancestry (**Supplementary Table 1a–c** and **Supplementary Fig. 1**). We then genotyped the lead SNPs within the most promising previously unidentified loci as well as a subset of previously reported CAD loci in up to 56,682 additional subjects (approximately half cases and half controls) (**Supplementary Table 2a,b**). Lastly, we explored potential mechanisms and intermediate pathways by which previously unidentified loci may mediate risk.

Nine of the twelve loci previously associated with CAD through individual GWAS achieved genome-wide significance ($P < 5 \times 10^{-8}$) in our initial meta-analysis (**Table 1** and **Supplementary Table 3**). We were, however, unable to test the previously reported association with a haplotype and a rare SNP in *LPA* in our GWAS data[4,5], but we observed robust association with the rare *LPA* variant in our replication samples through direct genotyping (**Table 1**). Thus, 10 of the 12 loci previously associated with CAD at a genome-wide significance level surpassed the same threshold of significance in CARDIoGRAM.

We selected 23 new loci with a significance level of $P < 5 \times 10^{-6}$ in the meta-analysis for follow up (Online Methods and **Supplementary Note**). Taking the number of loci into consideration, our replication study had >90% power to detect effect sizes observed in the GWAS meta-analysis. Of the 23 loci, 13 replicated using our *a priori* definition of a validated locus, that is, showing independent replication after Bonferroni correction and also achieving $P < 5 \times 10^{-8}$ in the combined discovery and replication data (**Table 2**, **Fig. 1** and **Supplementary Figs. 2** and **3**). Results for all loci from the replication phase are shown in **Supplementary Tables 4** and **5**.

The 13 new loci had risk allele frequencies ranging from 0.13 to 0.91 and were associated with a 6% to 17% increase in the risk of CAD per allele (**Table 2**). Out of the 13 new loci, the additive model appeared most appropriate for 6 whereas the recessive model performed best at 5 and the dominant model at 2 loci (**Supplementary Table 6**).

In sub-group analyses, 20 out of 22 loci with $P < 5 \times 10^{-8}$ (known and new loci combined; for one locus, age subgroups were not available) had higher odds ratios for early onset than for late onset CAD ($P = 1.2 \times 10^{-4}$ for observed versus expected; **Supplementary Table 7**). The CAD loci showed consistent associations irrespective of case definition, although the odds ratios for most individual SNPs tended to be slightly greater for cases with angiographically proven CAD than for cases with unknown angiographic status ($P = 0.019$ for observed versus expected) (**Supplementary Table 8**). In contrast, sub-group analyses in males and females revealed no sex-specific effects for any risk alleles (**Supplementary Table 7**) or for their observed versus expected pattern of association ($P = 0.4$).

Among 7,637 CAD cases and 7,523 controls for whom we had individual level genotype data, the minimum and maximum number of risk alleles observed per individual was 15 and 37, respectively, when considering 23 CAD susceptibility loci. The mean weighted risk score was significantly higher for cases than for controls ($P < 10^{-20}$). Furthermore, being in the top tenth percentile or lowest tenth percentile of the weighted score was associated with an odds ratio for CAD of 1.88 (95% CI 1.67–2.11) and 0.55 (95% CI 0.48–0.64), respectively, compared to the fiftieth percentile. The change in odds ratio for CAD across a broader spectrum of categories of the weighted score is shown in **Supplementary Figure 4**.

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

Received 10 August 2010; accepted 10 February 2011; published online 6 March 2011; doi:10.1038/ng.784

Table 1 Association evidence in CARDIoGRAM for previously published loci for coronary disease (previously reported with genome-wide significance, $P < 5 \times 10^{-8}$)

Band	SNP	Gene(s) in region	*n*	Risk allele frequency (risk allele)	CARDIoGRAM		Reference
					OR (95% CI)	*P*	OR (95% CI)
1p32.3	rs11206510[a]	*PCSK9*	102,352	0.82 (T)	1.08 (1.05–1.11)	9.10×10^{-8}	1.15 (1.10–1.21)[26]
1p13.3	rs599839[b]	*SORT1*	83,873	0.78 (A)	1.11 (1.08–1.15)	2.89×10^{-10}	1.29 (1.18–1.40)[21]
1q41	rs17465637[c]	*MIA3*	25,197	0.74 (C)	1.14 (1.09–1.20)	1.36×10^{-8}	1.20 (1.12–1.30)[21]
2q33.1	rs6725887[b]	*WDR12*	77,954	0.15 (C)	1.14 (1.09–1.19)	1.12×10^{-9}	1.16 (1.10–1.22)[26]
3q22.3	rs2306374[b]	*MRAS*	77,843	0.18 (C)	1.12 (1.07–1.16)	3.34×10^{-8}	1.15 (1.11–1.19)[23]
6p24.1	rs12526453[b]	*PHACTR1*	83,050	0.67 (C)	1.10 (1.06–1.13)	1.15×10^{-9}	1.13 (1.09–1.17)[26]
6q25.3	rs3798220[d]	*LPA*	32,584	0.02 (C)	1.51 (1.33–1.70)	3.00×10^{-11}	1.92 (1.48–2.49)[5]
9p21.3	rs4977574[b]	*CDKN2A, CDKN2B*	84,256	0.46 (G)	1.29 (1.23–1.36)	1.35×10^{-22}	1.25 (1.18–1.31) 1.37 (1.26–1.48)[20,21,27,28]
10q11.21	rs1746048[a]	*CXCL12*	136,416	0.87 (C)	1.09 (1.07–1.13)	2.93×10^{-10}	1.33 (1.20–1.48)[21]
12q24.12	rs3184504[b]	*SH2B3*	67,746	0.44 (T)	1.07 (1.04–1.10)	6.35×10^{-6}	1.13 (1.08–1.18)[38]
19p13.2	rs1122608[b]	*LDLR*	49,693	0.77 (G)	1.14 (1.09–1.18)	9.73×10^{-10}	1.14 (1.09–1.19)[26]
21q22.11	rs9982601[b]	*MRPS6*	46,230	0.15 (T)	1.18 (1.12–1.24)	4.22×10^{-10}	1.19 (1.13–1.27)[26]

Data taken from [a]the combined analysis, [b]the meta-analysis, [c]only genotyped data from a subset of studies and [d]the replication.

Three of the new risk alleles were associated with differences in traditional CAD risk factors (**Table 3** and **Supplementary Tables 9**,**10**). The risk allele on chromosome 11q23.3 (rs964184, *ZNF259*, *APOA5-APOA4-APOC3-APOA1* gene region) was associated with increased low-density lipoprotein (LDL) cholesterol and decreased high-density lipoprotein (HDL) cholesterol (and previously, with increased triglycerides)[6]. The risk allele on chromosome 9q34.2 (rs579459, *ABO*) was associated with increased LDL and total cholesterol in a direction consistent with the association of these SNPs with CAD risk (**Table 3**). The variant rs12413409 on chromosome 10q24.32 representing the *CYP17A1-CNNM2-NT5C2* gene region was associated with hypertension.

In silico interrogation revealed that the lead SNPs at 4 of the 13 new loci were either non-synonymous coding variants or were in high linkage disequilibrium (LD) with such SNPs. Specifically, the lead SNPs at 7q32.2 (rs11556924) and 15q25.1 (rs3825807) encoded changes in *ZC3HC1* (p.Arg363His) and *ADAMTS7* (p.Ser214Pro), respectively, whereas the lead SNP at 14q32.2 (rs2895811) is in strong LD ($r^2 = 0.82$) with the p.Val691Ala variant in *HHIPL1*. Lastly, the lead SNP at 17q21.32 (rs46522) is in strong LD ($r^2 = 0.94$) with two potential functional variants in *GIP*: p.Ser103Gly (rs2291725) and a variant influencing the splice site of intron 3 (rs2291726) leading to a truncated transcript (**Supplementary Table 11**)[7].

We next analyzed data from three genome-wide studies that also assessed gene expression in multiple tissues to assess potential effects of new loci on the expression of regional genes (**Supplementary Note**)[8,9]. Three of the new CAD risk variants showed convincing association with regional gene expression (*cis* effect) by either representing the most significant expressed SNP in the region or by being in high LD ($r^2 \geq 0.85$) with the strongest expressed SNP in the region: rs12190287 at 6q23.2 (*TCF21*), rs12936587 at 17p11.2 (*RASD1, SMCR3* and *PEMT*) and rs46522 at 17q21.32 (*UBE2Z*) (**Supplementary Table 12**). Subsequent interrogation of our new loci in a genome-wide map of allelic expression imbalance provided further support for the expression quantitative trait locus findings at the 17q21.32 locus[10]. This analysis also provided strong evidence of *cis* effects for the 17p13.3 locus lead SNP (rs216172) on the expression of *SMG6* (**Supplementary Note** and **Supplementary Table 13**)[10].

We identified five new loci (9q34, 10q24, 11q23, 15q25 and 17p13) at which the CAD risk variant was fully or strongly correlated ($r^2 > 0.8$)

Table 2 New loci for coronary disease

Band	SNP	Gene(s) in region	Risk allele frequency (risk allele)	Meta-analysis		Replication		Combined analysis	
				P	*n*	*P*	*n*	OR (95% CI)	*P*
1p32.2	rs17114036	*PPAP2B*	0.91 (A)	1.43×10^{-8}	80,870	3.18×10^{-12}	52,356	1.17 (1.13–1.22)	3.81×10^{-19}
6p21.31	rs17609940	*ANKS1A*	0.75 (G)	2.21×10^{-6}	83,997	1.18×10^{-3}	53,415	1.07 (1.05–1.10)	1.36×10^{-8}
6q23.2	rs12190287	*TCF21*	0.62 (C)	4.64×10^{-11}	78,290	3.25×10^{-4}	52,598	1.08 (1.06–1.10)	1.07×10^{-12}
7q32.2	rs11556924	*ZC3HC1*	0.62 (C)	2.22×10^{-9}	80,011	7.37×10^{-10}	54,189	1.09 (1.07–1.12)	9.18×10^{-18}
9q34.2	rs579459	*ABO*	0.21 (C)	1.16×10^{-7}	77,138	7.02×10^{-8}	46,840	1.10 (1.07–1.13)	4.08×10^{-14}
10q24.32	rs12413409	*CYP17A1, CNNM2, NT5C2*	0.89 (G)	1.47×10^{-6}	80,940	1.38×10^{-4}	48,801	1.12 (1.08–1.16)	1.03×10^{-9}
11q23.3	rs964184	*ZNF259, APOA5-A4-C3-A1*	0.13 (G)	8.02×10^{-10}	82,562	2.20×10^{-9}	52,930	1.13 (1.10–1.16)	1.02×10^{-17}
13q34	rs4773144	*COL4A1, COL4A2*	0.44 (G)	4.15×10^{-7}	77,113	1.31×10^{-3}	37,618	1.07 (1.05–1.09)	3.84×10^{-9}
14q32.2	rs2895811	*HHIPL1*	0.43 (C)	2.67×10^{-7}	63,184	4.59×10^{-5}	51,054	1.07 (1.05–1.10)	1.14×10^{-10}
15q25.1	rs3825807	*ADAMTS7*	0.57 (A)	9.63×10^{-6}	80,849	1.39×10^{-8}	48,803	1.08 (1.06–1.10)	1.07×10^{-12}
17p13.3	rs216172	*SMG6, SRR*	0.37 (C)	6.22×10^{-7}	57,235	2.11×10^{-4}	54,303	1.07 (1.05–1.09)	1.15×10^{-9}
17p11.2	rs12936587	*RASD1, SMCR3, PEMT*	0.56 (G)	4.89×10^{-7}	76,952	1.35×10^{-4}	52,648	1.07 (1.05–1.09)	4.45×10^{-10}
17q21.32	rs46522	*UBE2Z, GIP, ATP5G1, SNF8*	0.53 (T)	3.57×10^{-6}	83,867	8.88×10^{-4}	53,766	1.06 (1.04–1.08)	1.81×10^{-8}

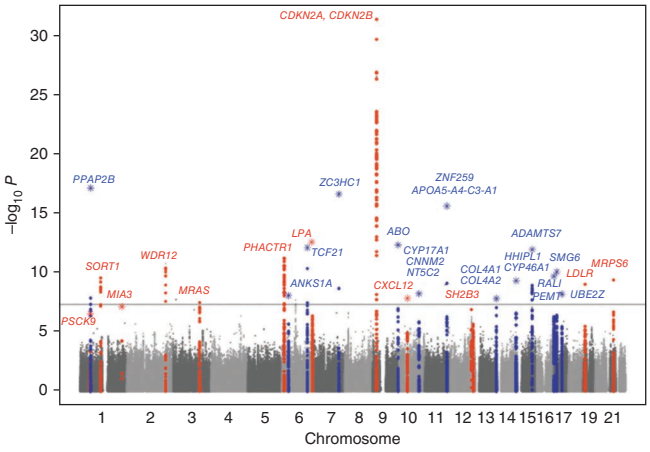

Figure 1 Graphical summary (Manhattan plot) of genome-wide association results. The *x* axis represents the genome in physical order; the *y* axis shows $-\log_{10} P$ for all SNPs. Data from the discovery phase are shown in circles, and data from the combined discovery and replication phases are shown in stars. Genes at the significant loci are listed above the signals. Known loci are shown in red and newly discovered loci are shown in blue.

with variants that have previously been associated with other traits or diseases[11]. These traits include cerebral and abdominal aneurysm, aortic root size, celiac disease, lung adenocarcinoma, type 1 diabetes, venous thrombosis, LDL cholesterol, HDL cholesterol, triglycerides, smoking, blood pressure, soluble levels of adhesion molecules, phytosterols (sitosterol and campesterol), angiotensin-converting enzyme activity, coagulation factor VIII and von Willebrand factor (at $P < 5 \times 10^{-8}$; for references see **Supplementary Table 14**). Thus, a substantial subset of the new CAD risk loci appear to have pleiotropic effects. We illustrate a particularly striking example at the *ABO* locus in **Figure 2**.

The present genomic analysis of more than 135,000 individuals revealed three major findings. First, we more than doubled the number of loci with firm association to CAD. Specifically, our study yielded 13 previously unidentified loci and confirmed 10 previously reported loci. Second, we found that only a minority of the established and new loci appear to act through traditional risk factors but that the majority reside in gene regions that were not previously suspected in the pathogenesis of CAD. Third, a substantial proportion of the CAD risk variants were also strongly associated with various other human disease traits in GWAS.

We anticipated that some of the genetic risk loci for CAD would act through established CAD risk factors, such as LDL cholesterol or blood pressure, which themselves have a significant genetic determination. Indeed, three of the new risk loci (11q23.3, 9q34.2 and 10q24.32) showed such associations. An association with higher LDL cholesterol or lipoprotein (a) concentration has also been found for four previously discovered risk variants, including the *PCSK9* locus, that missed the genome-wide significance level by a small margin in the present study (**Table 1**)[4,12,13]. On the other hand, 17 out of the 23 confirmed loci appear to act through mechanisms that are independent of traditional risk factors. Elucidation of these mechanisms is critical for a more complete understanding of CAD and identification of further therapeutic targets.

We explored several molecular mechanisms by which the new loci could affect CAD risk. We show that some lead SNPs—or linked variants—affect the primary structure of the protein product in which the variant is located, whereas in other instances, the risk variant is associated with expression of a specific gene or genes in one or more tissues. A more detailed discussion of the genes in each locus is provided in the **Supplementary Note**. Although these data help to prioritize genes for follow-up functional studies, it should be emphasized that substantial work is still necessary to define the mechanisms involved for each of the new loci, as exemplified recently for the chromosome 1p13 locus[14,15].

We also observed that 8 of 23 CAD loci (5 of the 13 new loci, 9q34, 10q24, 11q23, 15q25 and 17p13; and 3 of the 10 established loci, 1p13, 9p21.3 and 12q24) not only affect the risk of CAD but also associate with multiple other diseases and traits (**Supplementary Table 14**). Each of these findings requires further analysis to determine whether co-localization of SNPs for CAD and other traits points to intermediate phenotypes and, thus, mechanistic links in a joint etiology, results from pleiotropic effects of a single allele affecting multiple phenotypes, or identifies chromosomal regions harboring multiple genes and alleles that participate in the regulation of multiple independent traits through diverse mechanisms.

By design, our study focused on common risk variants. Assuming a heritability of 40% for CAD[1], the lead SNPs of previously established loci combined with the loci discovered in this study explain approximately 10% of the additive genetic variance of CAD. Our inability to explain a greater fraction of CAD heritability even after a large meta-analysis and replication effort is in line with the results for most other complex traits examined by current GWAS methods[16]. These results suggest that many other common susceptibility variants of similar or lower effects and/or rare variants contribute to risk of CAD.

The clinical utility of CAD risk alleles for the prediction of risk may be best determined in samples that are independent from this discovery study. In order to provide a framework for future research, we explored a weighted score based on the 23 CAD risk variants validated in this investigation. We observed a greater than threefold difference in CAD risk between the top and bottom 10% of the risk scores, although this may be a slight overestimation, as we extracted the risk scores from a subset of the discovery sample (**Supplementary Fig. 4**). Nonetheless, this increase in risk is at least comparable to that of several other traditional risk factors for CAD including hypertension, diabetes and smoking[13]. Whether risk allele information may improve the performance of current risk profiling strategies for CAD prediction[17,18] and whether such an approach is cost effective requires

Table 3 Effects of new CAD loci on traditional risk factors in combined analysis of ARIC and KORA F3 and F4 (*n* = 13,171)

SNP	Band	Gene(s) in region	Phenotype	β (95% CI)[a]	*P*
rs579459	9q34.2	*ABO*	Total cholesterol	1.720 (0.554–2.885)	0.0038
			LDL cholesterol	1.538 (0.468–2.608)	0.0049
rs12413409	10q24.32	*CYP17A1, CNNM2, NT5C2*	Hypertension	0.141 (0.044–0.238)	0.0043
rs964184[b]	11q23.3	*ZNF259, APOA5-A4-C3-A1*	HDL cholesterol	−1.926 (−2.441 to −1.411)	2.28×10^{-13}
			Total cholesterol	4.578 (3.191–5.964)	9.84×10^{-11}
			LDL cholesterol	1.699 (0.417–2.980)	0.0094

Results from fixed-effects meta-analysis based on β coefficients and standard errors from linear (for total, LDL and HDL cholesterol) and logistic (for hypertension) regression analysis of the single studies for which meta-analytic $P < 0.01$. LDL, low-density lipoprotein; HDL, high-density lipoprotein.
[a]Estimated pooled regression coefficients with 95% confidence intervals. Cholesterol levels are in mg/dl. [b]Previous genome-wide studies have demonstrated strong association of rs964184 with triglycerides[39].

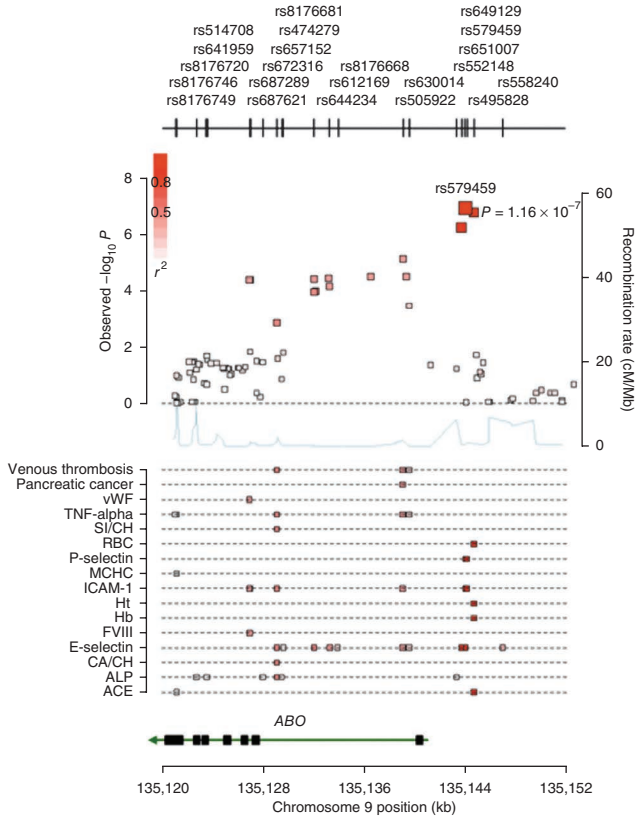

Figure 2 Example of overlapping association signals for multiple traits at the *ABO* gene region on chromosome 9q34. In the upper panel, the association signal for coronary disease at the *ABO* gene region in CARDIoGRAM and the positions and rs numbers of SNPs in this region are shown. The size of the boxes illustrates the number of individuals available for this respective SNP. In the lower panel, all SNPs with *P* values at the genome-wide significance level of $P < 5 \times 10^{-8}$ based on the National Human Genome Research Institute GWAS catalog (accessed on 28 June 2010) for all diseases and traits are shown. The degree of linkage disequilibrium (r^2) between the lead SNPs for coronary disease and the other traits is reflected by the color of the squares (upper panel) and the small bars (lower panel), from dark red (high LD) to faint red (low LD). SI/CH, sitosterol normalized to cholesterol; CA/CH, campesterol normalized to cholesterol; ALP, alkaline phosphatase; ACE, angiotensin converting enzyme; FVIII, coagulation factor VIII; vWF, von Willebrand factor.

further evaluation in prospective studies. Our findings provide a firm framework for such research.

In summary, our large-scale GWAS meta-analysis uncovered the association with CAD of 13 new chromosomal loci. We observed only limited association between these CAD SNPs and traditional risk factors, suggesting that most SNPs act through previously unidentified pathways. Elucidation of the mechanisms by which these loci affect CAD risk carries the potential for better prevention and treatment of this common 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 thank the participants and staff in each of the studies who contributed to the present article. The sources of funding are listed in the supplementary materials.

AUTHOR CONTRIBUTIONS

Manuscript writing: H.S., I.R.K., S.K., M.P.R., T.L.A., H.H., A.F.R.S., P. Deloukas, R.R., R.M., J.E., N.J.S.

GWAS meta-analysis samples, genotyping and analysis: H.S., I.R.K., S.K., M.P.R., T.L.A., H.H., M.P., A.F.R.S., M.B., C.G., D. Absher, D. Ardissino, K.A., S.G.B., A.J.B., J.C.B., E.B., P.S.B., M.S.B., L.C., A. Deghan, S.D., P. Diemert, J.D., A. Doering, N.E.E.M., R.E., S.E., M.F., A.R.F., S.G., J.R.G., E.H., A.H., T.I., C.I., M.A.K., J.W.K., A.K., R.L., M.L., W.L., C.L., C.M., T. Meitinger, O.M., V.M., K.M., T. Morgan, J.N., C.P.N., A.P., L.Q., D.J.R., V.S., A. Schäfer, A. Schillert, S.S., J.S., S.M.S., D.S.S., K.S., G. Thorgeirsson, G. Thorleifsson, M.T., A.G.U., B.F.V., G.A.W., H.E.W., C.W., P.S.W., J.C.M.W., B.J.W., T.Z., A.Z., F.C., L.A.C., T.Q., W.M., C.H., S.B., A.S.H., P.D., U.T., R.R., J.R.T., C.J.O., R.M., J.E., N.J.S.

Replication phase samples, genotyping and analysis: H.S., I.R.K., S.K., M.P.R., H.H., M.P., H.A., S.A., K.A., T.L.A., J.L.A., D. Ardissino, D. Absher, T.A.B., L.C.B, D.M.B., K.B., S.M.B., M.J.B., I.B., J.F.C., R.W.D., G.D., R.D., S.G.E., J.C.E., U.d.F., B.G., D.G., V.G., N.H., S.L.H., B.D.H., C.I., G.T.J., J.W.J., L.M.K., J.W.K., J.J.P.K., K.-T.K., G.K., D.L., K.L., P.L.-N., A.J.L., P.M.M., N.M., P.P.M., P.A.M., T. Morgan, T. Meitinger, T.W.M., J.B.M., S.C., M.M.N., O.O., F.P., R.S.P., C.C.P., A.A.Q., L.S.R., F.R.R., D.R., M.L.S., M.S.S., S. Sivapalaratnam, T.B.S., J.D.S., N.S., J.A.S., T.Q., K. Stark, K. Stirrups, M. Stoll, W.H.W.T., A.M.v.R., N.J.W., S.Y., P.D., U.T., R.R., R.M., J.E., N.J.S.

Analysis group: I.R.K., M.P., D. Absher, L.C., E.H., M.L., K.M., A. Schillert, G. Thorleifsson, B.F.V., G.A.W., L.A.C., J.R.T.

Biological analyses: H.S., T.L.A., H.H., M.B., C.G., Z.A., P.S.B., V.C., J.F., S.G., P.L.-N., G.L., S.M., C.R., E.S., M.T., F.C., A.H.G., T.Q., C.H., W.H.O., P.D., U.T., J.E., N.J.S.

CARDIoGRAM consortium executive group: H.S., S.K., M.P.R., J.E., N.J.S.

CARDIoGRAM consortium steering group: H.S., I.R.K., S.K., M.P.R., T.L.A., E.B., R.L., A.Z., C.H., A.S.H., U.T., J.R.T., R.M., J.E., N.J.S.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details accompany the full-text HTML version of the paper at http://www.nature.com/naturegenetics/.

Published online at http://www.nature.com/naturegenetics/.
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/.

1. 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).
2. Schunkert, H., Erdmann, J. & Samani, N.J. Genetics of myocardial infarction: a progress report. *Eur. Heart J.* **31**, 918–925 (2010).
3. Preuss, M. *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).
4. Clarke, R. *et al.* Genetic variants associated with Lp(a) lipoprotein level and coronary disease. *N. Engl. J. Med.* **361**, 2518–2528 (2009).
5. Trégouët, D.A. *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).
6. Psaty, B.M. *et al.* Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium: design of prospective meta-analyses of genome-wide association studies from 5 cohorts. *Circ. Cardiovasc. Genet.* **2**, 73–80 (2009).
7. Nitz, I. *et al.* Association analyses of GIP and GIPR polymorphisms with traits of the metabolic syndrome. *Mol. Nutr. Food Res.* **51**, 1046–1052 (2007).
8. Emilsson, V. *et al.* Genetics of gene expression and its effect on disease. *Nature* **452**, 423–428 (2008).
9. Zhong, H. *et al.* Liver and adipose expression associated SNPs are enriched for association to type 2 diabetes. *PLoS Genet.* **6**, e1000932 (2010).
10. Ge, B. *et al.* Global patterns of *cis* variation in human cells revealed by high-density allelic expression analysis. *Nat. Genet.* **41**, 1216–1222 (2009).
11. Hindorff, L.A. *et al.* Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc. Natl. Acad. Sci. USA* **106**, 9362–9367 (2009).
12. Brown, M.S. & Goldstein, J.L. Expression of the familial hypercholesterolemia gene in heterozygotes: mechanism for a dominant disorder in man. *Science* **185**, 61–63 (1974).
13. Linsel-Nitschke, P. *et al.* Lifelong reduction of LDL-cholesterol related to a common variant in the LDL-receptor gene decreases the risk of coronary artery disease—a Mendelian Randomisation study. *PLoS ONE* **3**, e2986 (2008).
14. Musunuru, K. *et al.* From noncoding variant to phenotype via SORT1 at the 1p13 cholesterol locus. *Nature* **466**, 714–719 (2010).

15. Linsel-Nitschke, P., Samani, N.J. & Schunkert, H. Sorting out cholesterol and coronary artery disease. *N. Engl. J. Med.* **363**, 2462–2463 (2010).
16. Manolio, T.A. *et al.* Finding the missing heritability of complex diseases. *Nature* **461**, 747–753 (2009).
17. Wilson, P.W. *et al.* Prediction of coronary heart disease using risk factor categories. *Circulation* **97**, 1837–1847 (1998).
18. Ripatti, S. *et al.* A multilocus genetic risk score for coronary heart disease: case-control and prospective cohort analyses. *Lancet* **376**, 1393–1400 (2010).

Heribert Schunkert[1,119*], Inke R König[2,119], Sekar Kathiresan[3–5,119], Muredach P Reilly[6,119], Themistocles L Assimes[7,119], Hilma Holm[8,119], Michael Preuss[1,2], Alexandre F R Stewart[9], Maja Barbalic[10], Christian Gieger[11], Devin Absher[12], Zouhair Aherrahrou[1], Hooman Allayee[13], David Altshuler[5,14], Sonia S Anand[15], Karl Andersen[16,17], Jeffrey L Anderson[18], Diego Ardissino[19], Stephen G Ball[20,21], Anthony J Balmforth[22], Timothy A Barnes[23], Diane M Becker[24], Lewis C Becker[24], Klaus Berger[25], Joshua C Bis[26], S Matthijs Boekholdt[27,28], Eric Boerwinkle[10], Peter S Braund[23], Morris J Brown[29], Mary Susan Burnett[30], Ian Buysschaert[31,32], Cardiogenics, John F Carlquist[18], Li Chen[33], Sven Cichon[34–36], Veryan Codd[23], Robert W Davies[37], George Dedoussis[38], Abbas Dehghan[39,40], Serkalem Demissie[41,42], Joseph M Devaney[30], Patrick Diemert[1], Ron Do[43], Angela Doering[11], Sandra Eifert[44], Nour Eddine El Mokhtari[45], Stephen G Ellis[46], Roberto Elosua[47], James C Engert[43,48], Stephen E Epstein[30], Ulf de Faire[49,50], Marcus Fischer[51], Aaron R Folsom[52], Jennifer Freyer[1], Bruna Gigante[49,50], Domenico Girelli[53], Solveig Gretarsdottir[8], Vilmundur Gudnason[17,54], Jeffrey R Gulcher[8], Eran Halperin[55–57], Naomi Hammond[58], Stanley L Hazen[59], Albert Hofman[39], Benjamin D Horne[18], Thomas Illig[11], Carlos Iribarren[60], Gregory T Jones[61], J Wouter Jukema[62,63], Michael A Kaiser[23], Lee M Kaplan[64], John J P Kastelein[65], Kay-Tee Khaw[66], Joshua W Knowles[7], Genovefa Kolovou[67], Augustine Kong[8], Reijo Laaksonen[68], Diether Lambrechts[32], Karin Leander[49], Guillaume Lettre[69,70], Mingyao Li[71], Wolfgang Lieb[1], Christina Loley[1,2], Andrew J Lotery[72,73], Pier M Mannucci[74], Seraya Maouche[1], Nicola Martinelli[53], Pascal P McKeown[75], Christa Meisinger[11], Thomas Meitinger[76,77], Olle Melander[78], Pier Angelica Merlini[79], Vincent Mooser[80], Thomas Morgan[81], Thomas W Mühleisen[34,35], Joseph B Muhlestein[18], Thomas Münzel[82], Kiran Musunuru[3–5], Janja Nahrstaedt[1,2], Christopher P Nelson[22], Markus M Nöthen[34,35], Oliviero Olivieri[53], Riyaz S Patel[83,84], Chris C Patterson[75], Annette Peters[11], Flora Peyvandi[85], Liming Qu[71], Arshed A Quyyumi[83], Daniel J Rader[6,86], Loukianos S Rallidis[87], Catherine Rice[58], Frits R Rosendaal[88–90], Diana Rubin[91], Veikko Salomaa[92], M Lourdes Sampietro[93], Manj S Sandhu[94,95], Eric Schadt[96,97], Arne Schäfer[98], Arne Schillert[2], Stefan Schreiber[98], Jürgen Schrezenmeir[99,100], Stephen M Schwartz[26], David S Siscovick[26], Mohan Sivananthan[101], Suthesh Sivapalaratnam[27], Albert Smith[17,54], Tamara B Smith[102], Jaapjan D Snoep[88], Nicole Soranzo[58], John A Spertus[103], Klaus Stark[51], Kathy Stirrups[58], Monika Stoll[104], W H Wilson Tang[46], Stephanie Tennstedt[1], Gudmundur Thorgeirsson[16,17], Gudmar Thorleifsson[8], Maciej Tomaszewski[23,105], Andre G Uitterlinden[39,40,106], Andre M van Rij[61], Benjamin F Voight[4,5,107], Nick J Wareham[108], George A Wells[37], H-Erich Wichmann[11,44,109], Philipp S Wild[82], Christina Willenborg[1,2], Jaqueline C M Witteman[39,40], Benjamin J Wright[110], Shu Ye[111], Tanja Zeller[82], Andreas Ziegler[2], Francois Cambien[112], Alison H Goodall[23,105], L Adrienne Cupples[41,42], Thomas Quertermous[7], Winfried März[113–115], Christian Hengstenberg[51], Stefan Blankenberg[82], Willem H Ouwehand[58,116], Alistair S Hall[21], Panos Deloukas[58], John R Thompson[117], Kari Stefansson[8,17], Robert Roberts[9], Unnur Thorsteinsdottir[8,17], Christopher J O'Donnell[42,119], Ruth McPherson[9,118,119], Jeanette Erdmann[1,119] & Nilesh J Samani[23,105,119] for the CARDIoGRAM Consortium[120]

[1]Universität zu Lübeck, Medizinische Klinik II, Lübeck, Germany. [2]Institut für Medizinische Biometrie und Statistik, Universität zu Lübeck, Lübeck, Germany. [3]Cardiovascular Research Center and Cardiology Division, Massachusetts General Hospital, Boston, Massachusetts, USA. [4]Center for Human Genetic Research, Massachusetts General Hospital, Boston, Massachusetts, USA. [5]Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA. [6]The Cardiovascular Institute, University of Pennsylvania, Philadelphia, Pennsylvania, USA. [7]Department of Medicine, Stanford University School of Medicine, Stanford, California, USA. [8]deCODE genetics, Reykjavik, Iceland. [9]The John and Jennifer Ruddy Canadian Cardiovascular Genetics Centre, University of Ottawa Heart Institute, Ottawa, Ontario, Canada. [10]University of Texas Health Science Center, Human Genetics Center, Houston, Texas, USA. [11]Institute of Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany. [12]Hudson Alpha Institute, Huntsville, Alabama, USA. [13]Department of Preventive Medicine, University of Southern California, Los Angeles, California, USA. [14]Department of Molecular Biology and Center for Human Genetic Research, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts, USA. [15]Population Health Research Institute, Hamilton Health Sciences and McMaster University, Hamilton, Ontario, Canada. [16]Department of Medicine, Landspitali University Hospital, Reykjavik, Iceland. [17]University of Iceland, Faculty of Medicine, Reykjavik, Iceland. [18]Cardiovascular Department, Intermountain Medical Center, Cardiology Division, University of Utah, Salt Lake City, Utah, USA. [19]Division of Cardiology, Azienda Ospedaliero-Universitaria di Parma, Parma, Italy. [20]LIGHT Research Institute, Faculty of Medicine and Health, University of Leeds, Leeds, UK. [21]Division of Cardiovascular and Neuronal Remodelling, Multidisciplinary Cardiovascular Research Centre, Leeds Institute of Genetics, Health and Therapeutics, University of Leeds, Leeds, UK. [22]Division of Cardiovascular and Diabetes Research, Multidisciplinary Cardiovascular Research Centre, Leeds Institute of Genetics, Health and Therapeutics, University of Leeds, Leeds, UK. [23]Department of Cardiovascular Sciences, University of Leicester, Clinical Sciences Wing, Glenfield Hospital, Leicester, UK. [24]The Johns Hopkins University School of Medicine, Division of General Internal Medicine, Baltimore,

Maryland, USA. [25]Institute of Epidemiology and Social Medicine, University of Münster, Münster, Germany. [26]Cardiovascular Health Resarch Unit and Department of Medicine, University of Washington, Seattle, Washington, USA. [27]Department of Vascular Medicine, Academic Medical Center, Amsterdam, The Netherlands. [28]Department of Cardiology, Academic Medical Center, Amsterdam, The Netherlands. [29]Clinical Pharmacology Unit, University of Cambridge, Cambridge, UK. [30]Cardiovascular Research Institute, Medstar Health Research Institute, Washington Hospital Center, Washington, DC, USA. [31]Department of Cardiology, University Hospital Gasthuisberg, Leuven, Belgium. [32]Vesalius Research Center, VIB-K.U.Leuven, Leuven, Belgium. [33]Cardiovascular Research Methods Centre, University of Ottawa Heart Institute, Ottawa, Ontario, Canada. [34]Institute of Human Genetics, University of Bonn, Bonn, Germany. [35]Department of Genomics, Life and Brain Center, University of Bonn, Bonn, Germany. [36]Institute of Neuroscience and Medicine (INM-1), Research Center Juelich, Juelich, Germany. [37]The Cardiovascular Research Methods, University of Ottawa Heart Institute, Ottawa, Ontario, Canada. [38]Department of Dietetics-Nutrition, Harokopio University, Athens, Greece. [39]Department of Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands. [40]Member of Netherlands Consortium for Healthy Aging (NCHA) sponsored by Netherlands Genomics Initiative (NGI), Leiden, The Netherlands. [41]Department of Biostatistics, Boston University School of Public Health, Boston, Massachusetts, USA. [42]National Heart, Lung, and Blood Institute's Framingham Heart Study, Framingham, Massachusetts, USA. [43]Department of Human Genetics, McGill University, Montreal, Quebec, Canada. [44]Klinikum Grosshadern, Munich, Germany. [45]Klinik für Innere Medizin, Kreiskrankenhaus Rendsburg, Rendsburg, Germany. [46]Department of Cardiovascular Medicine, Cleveland Clinic, Cleveland, Ohio, USA. [47]Cardiovascular Epidemiology and Genetics Group, Institut Municipal d'Investigació Mèdica, Ciber Epidemiología y Salud Pública (CIBERSP), Barcelona, Spain. [48]Department of Medicine, McGill University, Montreal, Canada. [49]Division of Cardiovascular Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden. [50]Department of Cardiology, Karolinska University Hospital, Stockholm, Sweden. [51]Klinik und Poliklinik für Innere Medizin II, Regensburg, Germany. [52]University of Minnesota School of Public Health, Division of Epidemiology and Community Health, School of Public Health (A.R.F.), Minneapolis, Minnesota, USA. [53]Department of Medicine, University of Verona, Verona, Italy. [54]Icelandic Heart Association, Kopavogur, Iceland. [55]The Blavatnik School of Computer Science, Tel-Aviv University, Tel-Aviv, Israel. [56]Department of Molecular Microbiology and Biotechnology, Tel-Aviv University, Tel-Aviv, Israel. [57]International Computer Science Institute, Berkeley, California, USA. [58]Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, UK. [59]Lerner Research Institute, Cleveland Clinic, Cleveland, Ohio, USA. [60]Division of Research, Kaiser Permanente of Northern California, Oakland, California, USA. [61]Surgery Department, Dunedin School of Medicine, University of Otago, Dunedin, New Zealand. [62]Department of Cardiology C5-P, Leiden University Medical Center, Leiden, The Netherlands. [63]Durrer Center for Cardiogenetic Research, Amsterdam, The Netherlands. [64]Massachusetts General Hospital, Boston, Massachusetts, USA. [65]Department of Vascular Medicine, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands. [66]Department of Public Health and Primary Care, Strangeways Research Laboratory, University of Cambridge, Cambridge, UK. [67]1st Cardiology Department, Onassis Cardiac Surgery Center, Athens, Greece. [68]Science Center, Tampere University Hospital, Tampere, Finland. [69]Montreal Heart Institute, Montréal, Québec, Canada. [70]Département de Médecine, Université de Montréal, succursale Centre-ville, Montréal, Québec, Canada. [71]Biostatistics and Epidemiology, University of Pennsylvania, Philadelphia, Pennsylvania, USA. [72]Clinical Neurosciences Division, School of Medicine, University of Southampton, Southampton, UK. [73]Southampton Eye Unit, Southampton General Hospital, Southampton, UK. [74]Scientific Direction, Istituto di Ricovero e Cura a Carattere Scientifico (IRCCS) Fondazione Cà Granda, Ospedale Maggiore Policlinico, Milano, Italy. [75]Centre for Public Health, Queen's University Belfast, Institute of Clinical Science, Belfast, Ireland, UK. [76]Institute of Human Genetics, Helmholtz Zentrum München, Deutsches Forschungszentrum für Umwelt und Gesundheit, Neuherberg, Germany. [77]Institute of Human Genetics, Technische Universität München, Klinikum rechts der Isar, Munich, Germany. [78]Department of Clinical Sciences, Hypertension and Cardiovascular Diseases, Scania University Hospital, Lund University, Malmö, Sweden. [79]Division of Cardiology, Azienda Ospedaliera Niguarda Ca'Granda, Milan, Italy. [80]Genetics Division and Drug Discovery, GlaxoSmithKline, King of Prussia, Pennsylvania, USA. [81]Department of Pediatrics, Vanderbilt University School of Medicine, Nashville, Tennessee, USA. [82]Medizinische Klinik und Poliklinik, Universitätsmedizin Mainz, Johannes-Gutenberg Universität Mainz, Germany. [83]Emory University School of Medicine, Atlanta, Georgia, USA. [84]Cardiff University, Cardiff, Wales, UK. [85]A. Bianchi Bonomi Hemophilia and Thrombosis Center, Department of Medicine and Medical Specialties, Fondazione IRCCS Cà Granda, Ospedale Maggiore Policlinico, Università degli Studi di Milano and Luigi Villa Foundation, Milan, Italy. [86]The Institute for Translational Medicine and Therapeutics, School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA. [87]Second Department of Cardiology, Attikon Hospital, School of Medicine, University of Athens, Athens, Greece. [88]Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, The Netherlands. [89]Department of Thrombosis and Haemostasis, Leiden University Medical Center, Leiden, The Netherlands. [90]Einthoven Laboratory for Experimental Vascular Medicine, Leiden University Medical Center, Leiden, The Netherlands. [91]Medizinische Klinik I, Universitätsklinikum Schleswig-Holstein, Campus Kiel, Kiel, Germany. [92]Chronic Disease Epidemiology and Prevention Unit, Department of Chronic Disease Prevention, National Institute for Health and Welfare, Helsinki, Finland. [93]Department of Human Genetics and Cardiology, Leiden University Medical Center, Leiden, The Netherlands. [94]Manjinder S. Sandhu, Genetic Epidemiology Group, Wellcome Trust Sanger Institute, Cambridge, UK. [95]Department of Public Health and Primary Care, Strangeways Research Laboratory, University of Cambridge, Cambridge, UK. [96]Pacific Biosciences, Menlo Park, California, USA. [97]Sage Bionetworks, Palo Alto, California, USA. [98]Institut für Klinische Molekularbiologie, Christian-Albrechts Universität, Kiel, Germany. [99]Institute of Physiology and Biochemistry of Nutrition, Max Rubner-Institute, Kiel, Germany. [100]Clinical Research Center Kiel, Kiel Innovation and Technology Center, Kiel, Germany. [101]Cardiology Division, Leeds Teaching Hospitals National Health Service Trust, Leeds, UK. [102]Laboratory of Epidemiology, Demography and Biometry, Intramural Research Program, National Institute on Aging, National Institutes of Health, Bethesda, Maryland, USA. [103]Mid America Heart Institute and University of Missouri-Kansas City, Kansas City, Missouri, USA. [104]Leibniz-Institute for Arteriosclerosis Research, University of Münster, Münster, Germany. [105]Leicester National Institute for Health Research Biomedical Research Unit in Cardiovascular Disease, Glenfield Hospital, Leicester, UK. [106]Department of Internal Medicine, Erasmus Medical Center, Rotterdam, The Netherlands. [107]Department of Medicine, Harvard Medical School, Boston, Massachusetts, USA. [108]Medical Research Council Epidemiology Unit, Institute of Metabolic Science, Addenbrooke's Hospital, Cambridge, UK. [109]Institute of Medical Information Science, Biometry and Epidemiology, Ludwig-Maximilians-Universität München, München, Germany. [110]Department of Cardiovascular Surgery, University of Leicester, Leicester, UK. [111]William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK. [112]INSERM UMRS 937, Pierre and Marie Curie University, Université Pierre et Marie Curie (UPMC)-Paris 6, Faculté de Médecine Pierre et Marie Curie, Paris, France. [113]Synlab Center of Laboratory Diagnostics Heidelberg, Heidelberg, Germany. [114]Clinical Institute of Medical and Chemical Laboratory Diagnostics, Medical University of Graz, Graz, Austria. [115]Institute of Public Health, Social and Preventive Medicine, Medical Faculty Manneim, University of Heidelberg, Heidelberg, Germany. [116]Department of Haematology, University of Cambridge and NHS Blood and Transplant, Cambridge, UK. [117]Department of Health Sciences, University of Leicester, Leicester, UK. [118]Atherogenomics Laboratory, University of Ottawa Heart Institute, Ottawa, Ontario, Canada. [119]These authors contributed equally to this work. [120]A full list of members is provided in the **Supplementary Note**. Correspondence should be addressed to H.S. (heribert.schunkert@uk-sh.de) or N.J.S. (njs@le.ac.uk).

ONLINE METHODS

Study design. Details of the design of CARDIoGRAM have been published previously[3]. Briefly, the CARDIoGRAM consortium combined GWAS data from: the Atherosclerotic Disease VAscular functioN and genetiC Epidemiology study (ADVANCE)[19]; the CADomics, Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE)[6] Consortium; the deCODE CAD study[20]; the German Myocardial Infarction Family Studies (GerMIFS) I, II and III (KORA)[21–23]; the Ludwigshafen Risk and Cardiovascular Health Study (LURIC)/AtheroRemo 1 and 2 (ref. 24); MedStar[25]; the Myocardial Infarction Genetics Consortium (MIGen)[26]; the Ottawa Heart Genomics Study (OHGS)[27]; PennCATH[25]; and the Wellcome Trust Case Control Consortium (WTCCC)[21,28] (**Supplementary Table 1a**). Only subjects reporting European ancestry were included. The ethics boards of all participating institutions approved the study protocols, and all participants gave written informed consent. Funding came largely through public resources but included two companies (deCODE and GlaxoSmithKline; **Supplementary Note**). CARDIoGRAM agreed *a priori* not to seek patent protection for any of the findings.

The 14 studies provided a discovery sample size of 22,233 cases and 64,762 controls. Following the discovery phase, we attempted to replicate our most promising findings in more than 50,000 additional subjects (approximately half cases and half controls) derived from multiple studies (**Supplementary Table 1b**).

We studied the association between the newly discovered susceptibility SNPs and traditional risk factors in two population-based samples: the Atherosclerosis Risk in Communities (ARIC) study (9,713 individuals)[29] and the Augsburg MONICA/KORA (MONItoring trends and determinants in CArdiovascular disease/KOoperative Gesundheitsforschung in der Region Augsburg) study (3,458 individuals)[30] (**Supplementary Table 1c**). We also studied the association between coronary disease risk loci and gene expression in four studies with both genome-wide SNP and genome-wide transcript expression data[8,9,31].

Genotyping. Genotyping in individual discovery GWAS was carried out on Affymetrix or Illumina platforms (**Supplementary Table 1b**). Approximately 2.3 million imputed genotypes were generated using the MACH[32], IMPUTE[33] or BIMBAM[34] imputation algorithms. Quality control was performed at individual sites and centrally to assure standardized data formats (**Supplementary Table 5**)[3]. In the replication phase, SNPs were genotyped using Sequenom, Taqman or Centaurus platforms or extracted *in silico* from existing independent GWAS data (**Supplementary Table 2b**).

Statistical methods. The primary analysis comprised two stages: a meta-analysis of the GWAS for discovery and a replication stage in additional cohorts for confirmation of the most interesting new SNP associations from the discovery analysis. For both stages, we performed analyses in each study separately according to an *a priori* standard operating procedure. For this, log-additive model frequentist tests adjusting for age (onset of the first event for cases or time of recruitment for controls) and gender and taking into account the uncertainty of possibly imputed genotypes were performed using using logistic regression. Quality control of these data was then performed centrally according to previously agreed criteria including check of consistency of the given alleles across all studies, quality of the imputation, deviation from Hardy-Weinberg equilibrium in the controls, minor allele frequency and call rate. Furthermore, at least half of the studies (>7) had to contribute genotyping data to result in a high-quality SNP. In every study, the variance inflation factor, λ, was estimated from genotyped SNPs and used for adjustment (**Supplementary Table 5**)[35].

We then performed a meta-analysis separately for every SNP from each study that passed the quality criteria. Our default meta-analysis used a fixed-effect model with inverse variance weighting and a calculation of two homogeneity statistics: Cochran's Q and I^2. When there was no indication for heterogeneity for a SNP (P for Q >0.01), the fixed effect model was maintained. When heterogeneity was present (P for Q <0.01), we performed an outlier test to compare the results of each study with the average of the results from all other studies. If an outlier was detected with $P < 0.01$ divided by the number of studies providing data for the SNP, the study with the most extreme result was excluded and the meta-analysis was repeated. If no outliers were detected but heterogeneity was apparent, we adopted and reported a random-effects model (DerSimonian-Laird) for that SNP. In addition to the study-wise adjustment by the variance inflation factor λ, we also estimated λ from the meta-analysis to be 1.1. Power estimates for this approach are provided in reference 3.

A locus was defined *a priori* as a new validated locus if it (i) harbored a SNP with evidence of association in the replication sample (one sided $P < 0.05$ adjusted for the number of loci taken forward); (ii) a significance level of association of $P < 5 \times 10^{-8}$ (Bonferroni correction for 1,000,000 independent SNPs) in the combined analysis of the discovery and replication samples; and (iii) not previously reported at genome-wide significance in a prior publication.

We also performed the following subgroup analyses at all loci with genome-wide significance: comparison of

- female cases with female controls
- male cases with male controls
- cases with later age of onset (>50 years) with all controls
- cases with earlier age of onset (< = 50 years) with all controls
- cases with myocardial infarction with controls
- cases with angiographically defined CAD with controls

We tested for differences in the odds ratios between subgroups by counting the number of SNPs that had an odds ratio that was higher in one subgroup compared to another. This number was formally compared to that expected by chance using a binomial distribution probability test.

In addition to the primary analysis that was based on an additive genetic model, we analyzed association using dominant and recessive genetic models for the newly identified regions. Moreover, to simultaneously estimate the genetic effect and the genetic model, we applied a meta-analytic method[36,37]. Here, we model the two log odds ratios corresponding to the heterozygous (θ_1) and homozygous effects (θ_2) together in a bivariate response. To identify the most likely mode of inheritance, these were statistically compared in the following way: if $\theta_1 = 0$ and $\theta_2 > 0$, we used a recessive model; if $\theta_1 > 0$ and $\theta_2 > 0$ but $\theta_1 = \theta_2$, we used a dominant model; and if $\theta_1 > 0$ and $\theta_2 > 0$ but $\theta_1 < \theta_2$, we used an additive model.

In a subset of studies with 7,637 cases and 7,534 controls (ADVANCE, GERMIFS-I, II, III (KORA), LURIC/AtheroRemo 1, MedStar, OHGS 1 and PennCATH), we counted the number of risk alleles at known (n = 10) and new (n = 13) loci carried by each individual in the study. We computed a weighted score for each individual by multiplying each genotype with the respective log odds ratio from the discovery meta-analysis. Cases and controls were compared in a linear regression predicting the weighted score from case-control status and adjusting for study. Finally, we partitioned the subjects into five categories according to their weighted score based on the following thresholds: <1.70, 1.70–2.05, 2.05–2.40, 2.40–2.75, 2.75–3.10, 3.10–3.45 and >3.45. These thresholds were chosen to illustrate the effect across the entire range while still including a reasonable number of subjects per category. The odds ratios for coronary disease were estimated to compare individuals in the upper and lower categories with those in the middle category. The results obtained through this will be biased upwards, as data from probands from the meta-analysis discovery stage was used.

The association of new loci for established cardiovascular risk factors including low-density lipoprotein (LDL) cholesterol, high-density lipoprotein (HDL) cholesterol, total cholesterol, hypertension, body mass index (BMI), diabetes mellitus type II and smoking was examined in ARIC, KORA F3 and KORA F4. We modeled the relationship between quantitative traits (LDL, HDL, total cholesterol and BMI) and these SNPs using linear regression and binary traits (hypertension, diabetes and smoking) and these SNPs using logistic regression. We then combined the respective regression estimates from each study in a meta-analysis using inverse variance weighting.

We examined the association between CAD associated SNPs and expression of genes (*cis* expression quantitative trait locus effects) in three genome-wide experiments of gene expression: the deCODE study of blood and adipose tissue among Icelandic families[8], the Massachusetts General Hospital study of liver, omental and subcutaneous adipose tissue among subjects undergoing Roux-en-Y gastric bypass surgery[9], and the Cardiogenics study of monocytes and macrophages among healthy subjects and individuals with CAD

(**Supplementary Note**). We then looked for supporting evidence of *cis* effects on gene expression in a smaller experiment that produced the first crude genome-wide map of allelic expression imbalance using 53 lymphoblastoid cell lines derived from donors of European descent (**Supplementary Note**)[10].

We extracted all SNPs (n = 112) in high LD ($r^2 \geq 0.8$ in the HapMap European CEU population (HapMap3_r2)) with our lead SNPs to identify any which might affect protein structure. Lastly, we searched the National Human Genome Research Institute catalog of published GWAS[11] for SNPs showing genome-wide significance ($P < 5 \times 10^{-8}$) with other diseases and traits located at the 13 new coronary disease loci within a window of 1 Mb centered on the lead SNP.

19. Assimes, T.L. *et al.* Susceptibility locus for clinical and subclinical coronary artery disease at chromosome 9p21 in the multi-ethnic ADVANCE study. *Hum. Mol. Genet.* **17**, 2320–2328 (2008).
20. Helgadottir, A. *et al.* A common variant on chromosome 9p21 affects the risk of myocardial infarction. *Science* **316**, 1491–1493 (2007).
21. Samani, N.J. *et al.* Genomewide association analysis of coronary artery disease. *N. Engl. J. Med.* **357**, 443–453 (2007).
22. Erdmann, J. *et al.* Genome-wide association study identifies a new locus for coronary artery disease on chromosome 10p11.23. *Eur. Heart J.* **32**, 158–168 (2011).
23. Erdmann, J. *et al.* New susceptibility locus for coronary artery disease on chromosome 3q22.3. *Nat. Genet.* **41**, 280–282 (2009).
24. Winkelmann, B.R. *et al.* Rationale and design of the LURIC study—a resource for functional genomics, pharmacogenomics and long-term prognosis of cardiovascular disease. *Pharmacogenomics* **2**, S1–S73 (2001).
25. Lehrke, M. *et al.* CXCL16 is a marker of inflammation, atherosclerosis and acute coronary syndromes in humans. *J. Am. Coll. Cardiol.* **49**, 442–449 (2007).
26. 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).
27. McPherson, R. *et al.* A common allele on chromosome 9 associated with coronary heart disease. *Science* **316**, 1488–1491 (2007).
28. WTCCC. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447**, 661–678 (2007).
29. Anonymous. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. *Am. J. Epidemiol.* **129**, 687–702 (1989).
30. Wichmann, H.E., Gieger, C. & Illig, T. KORA-gen–resource for population genetics, controls and a broad spectrum of disease phenotypes. *Gesundheitswesen* **67** Suppl 1, S26–S30 (2005).
31. Pastinen, T. & Hudson, T.J. *Cis*-acting regulatory variation in the human genome. *Science* **306**, 647–650 (2004).
32. Li, Y., Willer, C.J., Ding, J., Scheet, P. & Abecasis, G.R. MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet. Epidemiol.* **34**, 816–834 (2010).
33. Marchini, J., Howie, B., Myers, S., McVean, G. & Donnelly, P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat. Genet.* **39**, 906–913 (2007).
34. Servin, B. & Stephens, M. Imputation-based analysis of association studies: candidate regions and quantitative traits. *PLoS Genet.* **3**, e114 (2007).
35. Devlin, B. & Roeder, K. Genomic control for association studies. *Biometrics* **55**, 997–1004 (1999).
36. Bagos, P.G. A unification of multivariate methods for meta-analysis of genetic association studies. *Stat. Appl. Genet. Mol. Biol.* **7**, Article31 (2008).
37. Bagos, P.G. & Nikolopoulos, G.K. A method for meta-analysis of case-control genetic association studies using logistic regression. *Stat. Appl. Genet. Mol. Biol.* **6**, Article17 (2007).
38. Gudbjartsson, D.F. *et al.* Sequence variants affecting eosinophil numbers associate with asthma and myocardial infarction. *Nat. Genet.* **41**, 342–347 (2009).
39. Kathiresan, S. *et al.* Common variants at 30 loci contribute to polygenic dyslipidemia. *Nat. Genet.* **41**, 56–65 (2009).

doi:10.1038/ng.784