Large Scale Association Analysis of Novel Genetic Loci for Coronary Artery Disease

Coronary Artery Disease Consortium*

Background—Combined analysis of 2 genome-wide association studies in cases enriched for family history recently identified 7 loci (on 1p13.3, 1q41, 2q36.3, 6q25.1, 9p21, 10q11.21, and 15q22.33) that may affect risk of coronary artery disease (CAD). Apart from the 9p21 locus, the other loci await substantive replication. Furthermore, the effect of these loci on CAD risk in a broader range of individuals remains to be determined.

Methods and Results—We undertook association analysis of single nucleotide polymorphisms at each locus with CAD risk in 11 550 cases and 11 205 controls from 9 European studies. The 9p21.3 locus showed unequivocal association (rs1333049, combined odds ratio [OR]=1.20, 95% CI [1.16 to 1.25], probability value=2.81×10^{-21}). We also confirmed association signals at 1p13.3 (rs599839, OR=1.13 [1.08 to 1.19], $P=1.44\times10^{-7}$), 1q41 (rs3008621, OR=1.10 [1.04 to 1.17], $P=1.02\times10^{-3}$), and 10q11.21 (rs501120, OR=1.11 [1.05 to 1.18], $P=4.34\times10^{-4}$). The associations with 6q25.1 (rs6922269, $P=0.020$) and 2q36.3 (rs2943634, $P=0.032$) were borderline and not statistically significant after correction for multiple testing. The 15q22.33 locus did not replicate. The 10q11.21 locus showed a possible sex interaction ($P=0.015$), with a significant effect in women (OR=1.29 [1.15 to 1.45], $P=1.86\times10^{-5}$) but not men (OR=1.03 [0.96 to 1.11], $P=0.387$). There were no other strong interactions of any of the loci with other traditional risk factors. The loci at 9p21, 1p13.3, 2q36.3, and 10q11.21 acted independently and cumulatively increased CAD risk by 15% (12% to 18%), per additional risk allele.

ConclusionsThe findings provide strong evidence for association between at least 4 genetic loci and CAD risk. Cumulatively, these novel loci have a significant impact on risk of CAD at least in European populations. **(*Arterioscler Thromb Vasc Biol*. 2009;29:774-780.)**

Key Words: coronary artery disease ■ genetics ■ risk factors

Coronary artery disease (CAD), and its main complication, myocardial infarction (MI), have a significant genetic basis. Until recently, attempts at identifying genetic variants that affect risk of these common diseases have been hampered by poor reproducibility of associations and limited coverage of the genome.[1] However, well-powered genome-wide association (GWA) studies have now identified several novel putative loci that increase risk of CAD and MI.[2–5] Specifically, combined analysis of the Wellcome Trust Case Control Consortium (WTCCC) and the German MI Family GWA studies identified 7 chromosomal loci (on 1p13.3, 1q41, 2q36.3, 6.q25.1, 9p21.3, 10q11.21, and 15q22.33), all of which showed highly significant associations with CAD.[5] The locus on chromosome 9p21.3 was also identified in 2 other GWA studies[3,4] and has since been associated with CAD, stroke, as well as abdominal aortic and intracranial aneurysms in several other cohorts.[6–8] The locus on chromosome 1p13.3 was recently shown to also be strongly associated with LDL cholesterol concentration,[9–13] reinforcing the close mechanistic association between the variability in LDL levels and CAD risk. Beyond these initial studies on the loci at 9p21.3 and 1p13.3, the reproducibility of the association with CAD risk of the other loci identified by GWA studies has not yet been studied systematically.

See accompanying article on page 615

Many of the exploratory GWA studies were carried out on patients with a high genetic burden of the disease. For example, both the WTCCC and German MI Family Study analyzed cases enriched for a positive family history of CAD.[5] Here, in one of the largest molecular-genetic experiments on CAD, we report the replication analysis of the 7 principal loci for CAD identified thus far in GWA studies,[2–5] in a broader group of CAD patients, explore their interactions with traditional risk factors, and assess their cumulative impact on CAD risk.

Materials and Methods

Study Populations

We investigated participants recruited into 9 separate studies in Europe with validated cases of CAD and appropriate controls: the

Received November 19, 2008; revision accepted January 12, 2009.

*Members of the Coronary Artery Disease Consortium are listed in the Appendix.

Correspondence to Dr Samani, the Department of Cardiovascular Sciences, University of Leicester, Clinical Sciences Wing, Glenfield Hospital, Groby Road, Leicester, LE3 9QP (E-mail njs@le.ac.uk), or Dr Schunkert, Universität zu Lübeck, Medizinische Klinik II, Ratzeburger Allee 160, 23538 Lübeck, Germany (E-mail heribert.schunkert@uk-sh.de), or Dr Deloukas, The Wellcome Trust Sanger Institute, Hinxton UK (E-mail panos@sanger.ac.uk).

Academic Medical Center Amsterdam Premature Atherosclerosis Study (AMC-PAS), the Etude Cas-Témoins sur l'infarctus du Myocarde Study (ECTIM), the European Prospective Investigation into Cancer and Nutrition Study (EPIC-Norfolk), the German MI Family Study (GerMIFS; only including subjects that did not overlap with the original GWA Study), the Cooperative Health Research in the Region of Augsburg Study (KORA/GOC), the Ludwigshafen Risk and Cardiovascular Health Study (LURIC), the MORGAM Study, which has harmonized data from prospective follow-up of population cohorts in several countries, the Population-based northern German cross-sectional study (PopGen), and the UKMI Study. Almost all of these participants were of white Northern European origin. Details of the recruitment process in each study and references for each study are provided in the supplemental materials (available online at http://atvb.ahajournals.org). For the chromosome 9p21.3 locus some of the study groups (GerMIFS, KORA/GOC, PopGen, and UKMI) overlap with a previous publication on this locus.[6] In addition, the MORGAM Study has recently reported an analysis of the association of the novel loci with disease history and risk factors at baseline, and CAD and stroke events and death during follow-up in their prospective cohorts.[14]

Definition of Phenotypes

Common criteria for CAD and MI were applied across all the studies and required validated evidence for the phenotype (see supplemental materials). Cases not meeting these criteria were excluded. Similarly, uniform criteria were defined for partitioning of participants for risk factors (see supplemental materials). Those individuals where the information was unavailable or could not be assigned according to the criteria were classified as missing for the variable.

SNP Selection and Genotyping

For the 7 loci (1p13.3, 1q41, 2q36.3, 6.q25.1, 9p21.3, 10q11.21, 15q22.33), we selected the SNP showing the strongest association with CAD in the previous GWA analysis (lead SNP).[5] In addition, we selected SNP rs10738610 in the 9p21.3 locus which had shown marginally stronger association in a fine mapping experiment using HapMap SNPs,[6] and SNP rs2972147 in the 2q36.3 locus which is a proxy for rs2943634. Finally, linkage disequilibrium analysis of the 7 loci in HapMap identified that in 3 of them, there were subsets of highly correlated SNPs ($r^2>0.8$) significantly associated with CAD/MI which were not in the haplotype block defined by the lead SNP. The most significant SNP in each of these secondary haplotype blocks was also selected for genotyping giving a total of 13 SNPs (supplemental Table I and supplemental Figure I).

Genotyping was carried out with the iPLEX assay (Sequenom) for all SNPs except rs599839 and rs2943634, which were assayed by TaqMan (Applied Biosystems) using standard protocols (assays details available by request from the authors). iPLEX genotyping was performed at the Wellcome Trust Sanger Institute in Cambridge, UK, for all studies apart from MORGAM which was genotyped at the National Public Health Institute, Helsinki, Finland (Sequenom assay) and INSERM Unit 525, Paris France (TaqMan assays).

Statistical Analyses

Analysis was performed in Stata (Stata Statistical Software Release 10, 2007, StataCorp LP). Each study was analyzed separately by unconditional logistic regression using an additive genetic model (ie, genotype codes 0, 1, 2) adjusting for center in studies involving multiple sites. Heterogeneity between the studies was tested using Cochran Q chi-squared test, and the size of the heterogeneity was measured using the I^2 statistic. Only one nonlead SNP showed strong between-study heterogeneity (see supplemental Table III). Consequently, the odds ratios (OR) for the studies were meta-analyzed under a fixed effect model. Bonferroni correction was used to adjust for the number of SNPs tested. The analysis was performed for CAD cases and then for the subset of MI cases. The assumption of an additive model was assessed in the whole dataset by comparing the fit of that model with the fit of a 2-degree of freedom pairwise comparison in a likelihood ratio test. To assess the overall strength of the association of novel loci with CAD risk, probability values from the present analysis were combined with those from the GWA studies[5] using Fisher method, both with and without adjusting the original GWAs findings for multiple comparisons using the Bonferroni method. All tests were 2-sided.

To investigate whether there was any interaction between a locus and relevant demographic characteristics or cardiovascular risk factors (age, sex, body mass index, hypertension, smoking, and diabetes) participants were divided into 2 groups on the basis of the particular covariate. The analysis for CAD was repeated on the appropriate selection of patients in the same way as for the full study, and then the results were combined into a single forest plot.

Independent effects of the associated loci were verified by including them all in a single logistic regression. Cumulative risk was assessed by forming a score based on the total number of risk alleles across 4 or 6 loci (see Results). The case/control status was then compared with the number of risk alleles in a logistic regression analysis adjusting for study and center within study. The number of risk alleles was considered both as a categorical measure and as a continuous measure in a trend test.

Power calculations were performed by simulation. Data were generated to represent studies of the same numbers of cases and controls as in the actual replication study using an additive model with a given common odds ratio and allele frequency and assuming Hardy-Weinberg equilibrium (HWE). Analysis was by logistic regression and then meta-analysis as in the main study. The number of individual studies that were significant at the 5% level was counted and whether the combined result was significant was noted. Each set of simulations was repeated 1000 times.

Table 1. Summary Characteristics of Participants Included in the Study

	Cases	Controls
No.	11 550	11 205
Males, %	79%	63%
Age (years), mean (SD)	59.5 (10.0)	55.9 (12.8)
BMI (kg/m^2), mean (SD)	27.6 (4.2)	26.6 (4.2)
Hypertension	58%	28%
Ever smoked	74%	62%
Diabetes	16%	4%
MI	59%	

The characteristics of the cases and controls in the 9 individual studies are shown in supplemental Table II. Definitions for hypertension, smoking, and diabetes are given in the supplemental materials. For cases the age given and the status for hypertension, diabetes, and smoking relate to the time of event or at time of recruitment for the prospectively ascertained studies (EPIC-Norfolk and MORGAM). Data on risk factors were unavailable for 3 of the control cohorts (see supplemental Table II).

Results

Study Participants

The characteristics of the pooled case and control participants from the 9 European studies are shown in Table 1. Additionally, the characteristics of participants in each study individually are shown in supplemental Table II. Altogether, 22 755 participants (11 550 cases and 11 205 controls) underwent genotyping from these studies. As anticipated, the prevalence of established cardiovascular risk factors was higher in CAD cases than in controls (Table 1). Of the CAD cases, 6831 (59.1%) had a confirmed history of MI and the mean age of cases at first event was 59.5 (std. dev. 10.0) years.

Table 2. Pooled Association Results for the Lead SNP* at Each Locus for CAD and Separately for MI

						CAD			MI		
SNP	Chr	Alleles a1/a2+	RA	MAF in Controls	MAF in Cases	OR (95% CI)	*P* Value	Hetero *P* Value§	OR (95% CI)	*P* Value	Hetero *P* Value§
rs599839	1p13.3	A/G	A	0.228	0.207	1.13 (1.08, 1.19)	1.44×10^{-7} (1.87×10^{-6})	0.72	1.11 (1.05, 1.18)	2.35×10^{-4} (3.06×10^{-3})	0.38
rs3008621*	1q41	G/A	G	0.136	0.122	1.10 (1.04, 1.17)	1.02×10^{-3} (1.33×10^{-2})	0.49	1.09 (1.01, 1.17)	2.11×10^{-2} (0.274)	0.24
rs2943634	2q36.3	C/A	C	0.342	0.333	1.05 (1.00, 1.09)	3.22×10^{-2} (0.419)	0.24	1.03 (0.98, 1.09)	0.218 (1.000)	0.24
rs6922269	6q25.1	G/A	A	0.263	0.272	1.05 (1.01, 1.10)	1.96×10^{-2} (0.255)	0.40	1.08 (1.03, 1.14)	3.45×10^{-3} (4.49×10^{-2})	0.66
rs1333049	9p21	G/C	C	0.458	0.505	1.20 (1.16, 1.25)	2.89×10^{-21} (3.76×10^{-20})	0.65	1.24 (1.18, 1.30)	1.28×10^{-18} (1.66×10^{-17})	0.76
rs501120	10q11.21	T/C	T	0.133	0.121	1.11 (1.05, 1.18)	4.34×10^{-4} (5.64×10^{-3})	0.04	1.15 (1.07, 1.24)	1.99×10^{-4} (2.59×10^{-3})	0.07
rs17228212	15q22.33	T/C	C	0.276	0.274	1.00 (0.95, 1.04)	0.893 (1.000)	0.11	1.02 (0.97, 1.07)	0.521 (1.000)	0.13

*Lead SNP refers to the SNP that showed the strongest association at each locus in the prior genome-wide association studies except for chromosome 1q41 where the findings relate to a related SNP (see Text). +a2 is minor allele. §*P* value for heterogeneity between studies assessed (see Statistical Methods). SNP indicates single nucleotide polymorphism; chr, chromosome, RA, risk allele; MAF, minor allele frequency; OR, odds ratio associated with the risk allele; CI, confidence intervals. The *P* values in brackets for the odds-ratios are those after Bonferroni correction for the 13 SNPs that were tested (see Methods and supplemental Table I). The CIs have not been adjusted for multiple comparisons.

Association Analysis

Genotypes in excess of 90% were obtained for all SNPs, and there was no difference in the proportion of successful genotypes between cases and controls (supplemental Table I). None of the SNPs showed significant deviation from HWE. Nominally significant association ($P<0.05$) with CAD was observed at 6 of the 7 chromosomal loci studied (Table 2). For all loci, except for chromosome 1q41, the lead SNP identified in the GWA studies[5] showed the strongest association among the genotyped SNPs (supplemental Table III). Moreover, in all instances the same allele as in the previous study showed the increased risk with CAD. Interestingly, for 1q41, no significant association was seen for the previously reported lead SNP (rs17465637; supplemental Table III); however, a SNP (rs3008621) in an adjacent haplotype block showed a significant association (Table 2). The strength of the association ranged from an OR of 1.20 (95% CI: 1.16 to 1.25) for the 9p21 locus (rs1333049, $P=1.63\times10^{-21}$) to an OR of 1.05 (1.00 to 1.09) for the locus at 2q36.3 (rs2943634, $P=0.03$) and 1.05 (1.01 to 1.10) for the locus at 6q25.1 (rs6922269 $P=0.02$). The associations for the loci at 2q36.3 and 6q25.1 were not statistically significant after Bonferroni correction for the number of SNPs tested (Table 2). We found no evidence for association with the locus at 15q22.33 (Table 2). The associations in the subset of cases with MI largely paralleled those seen for CAD (Table 2).

To examine the totality of our evidence of association for each locus, we also combined the association results from the present studies with those from the 2 original GWA studies.[5] The signals for the 6 loci that showed nominally significant association in the present study became stronger in a meta-analysis that included these prior studies, even after correction for multiple testing in the GWA studies (supplemental Table IV). There was no evidence of nonadditivity for any of the loci assessed (ie, better fit using a dominant or recessive model: supplemental Table V).

Heterogeneity and Interactions

There was no significant heterogeneity in the magnitude of the associations of the loci between the pooled studies (Table 2). However, as expected from the power calculations (see Methods and supplemental Table VI), associations were not individually significant in every study (findings for each study by locus are shown in supplemental Figure II).

We also investigated whether there was any interaction between the effect of the loci and a number of prespecified characteristics or risk factors, namely age, sex, BMI, hypertension, diabetes mellitus, and smoking on risk of CAD. The analyses are displayed in supplemental Figure III. Note that the analyses for the risk factors are limited because only demographic information (age and sex) was available from 3 of the control groups (supplemental Table II). The most notable finding was that the magnitude of the association of the locus on 10q11.21 with CAD was greater ($P=0.015$ for interaction) in women (OR=1.29 [1.15 to 1.45], $P=1.86\times10^{-5}$) compared with men (OR=1.03 [0.96 to 1.11], $P=0.387$). There was also a suggestion that the effect of the locus on 1q41 was only present in older subjects and that the effect of the chromosome 9p21 locus was stronger in women and weaker in the presence of hypertension. However, neither of these interactions was significant ($P>0.05$), and otherwise the association of the loci with CAD appeared largely independent of anthropometric characteristics and risk factors (supplemental Figure III).

Distribution of Risk Alleles and Cumulative Risk

The proportions of cases and controls carrying different number of copies of the risk alleles for the 4 most strongly associated loci (1p13.3, 1q41, 9p21, and 10q11.21) are shown in the Figure. There is a significant rightward shift in the number of risk alleles carried by cases ($P<0.0001$). Because of the high prevalence of these alleles, the majority of these European white individuals carry more than 5 out of a possible 8 alleles (Figure). To investigate the cumulative risk associated with carriage of multiple risk alleles, we estimated the ORs in individuals carrying different numbers of the risk alleles for these loci. The 4 loci act independently with a combined OR of 1.15 (1.12 to 1.18) per additional risk allele. Because of the sex-specific effect of the locus on 10q11.21, the OR per additional risk allele was higher in women (1.21 [1.15 to 1.27]) compared with men (1.12 [1.08 to 1.16]). There was no significant interaction with age ($P=0.30$).

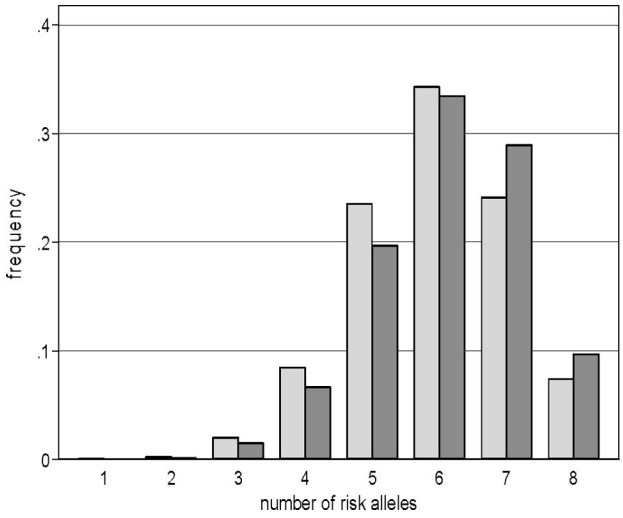


Figure. Distribution of cases (dark gray) and controls (light gray) carrying different number of risk alleles, ranging from 0 to 8, for the 4 most strongly associated loci: those on 1p13.3, 1q41, 9p21, and 10q11.21. Note the rightward shift in the distribution for the cases.

Discussion

In this study we describe a large scale evaluation of novel loci for CAD identified through previous GWA studies.[5] In addition to the 9p21 locus, which has already been robustly replicated in several other studies,[2,3,6–8] we provide compelling evidence for the association of at least 3 further loci (1p13.3, 1q41, and 10q11.21) with CAD risk. Nominal associations ($P<0.05$) were observed for 2 further loci, those at 2q36.3 and 6q25.1, but these became statistically nonsignificant after correction for the number of variants examined.

The increase in risk among the loci ranged from 5% to 20% per copy of the risk allele. These are less than those we found in the GWA studies (20% to 37%).[5] There are perhaps two main reasons for this. First, the GWA studies were carried out in relatively young cases enriched for a genetic tendency for CAD (each case had to have at least one first degree relative affected with CAD) which may have enhanced the genetic effect. Second, primary association findings by their nature tend to often be more inflated than the true degree of association. Thus, some degree of "the winner's curse" was to be expected.[15] Our present analysis was carried out in a much wider range of individuals with CAD, better reflecting the disease spectrum with regard to age of onset as well as relevant comorbidities and thus likely to provide a more reliable estimate of the association of each locus with CAD risk in general populations. Although individually the effect of carrying each risk allele is relatively modest, their importance in terms of contribution to the development of CAD and the public health needs to also take into account the prevalence of the risk alleles which range from 26% to 87% (Table 2). Thus, most European individuals carry multiple risk alleles (Figure).

Our study emphasizes the scale of endeavor required to quantify reliably the modest effect sizes which are typically being found for loci detected using GWA approaches for complex traits. Even with a combined sample size of more than 22 000 European participants, we only had sufficient (>80%) power to detect OR >1.05 across a range of allele frequencies (supplemental Table VI); hence the evidence of association for 2 of the loci (those at 2q36.3 and 6q25.1) remains inconclusive. Furthermore, simulations showed that even under the most favorable scenario, that pertaining to the locus at 9p21 with an OR of 1.20 and an allele frequency approaching 0.5, a "true" effect would not have been expected to be observed in each of the individual studies. Indeed, for the sizes of the studies involved here, the proportion of positive studies we observed for each locus was largely consistent with what might have been expected for "true" effects (supplemental Table VI). These findings are therefore remarkable in that we were able to detect a definite association with at least four of the initial 7 loci that emerged from the GWA studies in populations based in geographically and culturally different parts of Europe. This suggests that further loci with similar effect sizes await discovery in even larger analyses. Although we cannot rule out important gene–gene or gene–environment effects, our findings suggest that the loci identified affect CAD risk under a wide range of circumstances. This is also consistent with the lack of significant interactions with demographic parameters or other cardiovascular risk factors except for the locus on chromosome 10q11.21 (see below).

Our study does not identify the precise causal variant(s) at each locus. This will require resequencing of the entire recombination interval for each locus in a large set of chromosomes enriched for the risk allele to define the full spectrum of variants followed by fine mapping of the association signal. The finding at the locus on chromosome 1q41 emphasizes the importance of this work. We confirmed an association not with the GWA lead SNP but a related SNP suggesting that both markers are in linkage disequilibrium (LD) with the causal variant(s) at this locus but the strength of pair-wise LD differs.

So what are the implications of the present findings? The role in disease prediction often dominates discussion of such findings. Our analysis shows that although, cumulatively, carriage of increasing number of risk alleles imparts substantial additional risk (eg, carriage of seven risk alleles versus 4 risk alleles increases risk on average by 52%), genetic testing for the 4 most strongly replicated loci is unlikely to be sufficiently discriminatory to identify people likely to develop CAD (Figure). This lack of discriminatory capability is very similar to that seen for genetic loci underlying other complex traits such as diabetes[16] as well as with other cardiovascular risk factors (eg, plasma cholesterol level)[17] and consistent with theoretical considerations.[18] Potentially a more immediate and realistic clinical application of the findings could be to help identify people at increased coronary risk so that primary preventive measures, eg, cholesterol lowering, could be directed to those at highest genetic risk. This stratification could theoretically be carried out from a relatively young age, as DNA analysis can be done at any stage from birth. However, whether such testing is clinically beneficial and cost-effective requires much further investigation.

Perhaps, the greatest utility of these findings will come from understanding the mechanisms and pathways by which

Table 3. Genes Located Within or Adjacent to the Six Loci Associated With CAD

Chromosome	Genes
1p13.3	PSRC1, CELRS2, MYBPHL, SORT1
1q41	MIA3
2q36.3	No recognized genes
6q25.1	MTHFD1L
9p21	p16/CDKN2A, p15/CDKN2B, p14/ARF, MTAP, ANRIL
10q11.21	CXCL12

PSRC1 indicates proline/serine-rich coiled coil 1 gene; CELSR2, cadherin EGF LAG seven-pass G-type receptor 2 gene; MYBPHL, myosin binding protein H-like gene; SORT1, sortilin 1 gene; MIA3, melanoma inhibitory activity family, member 3 (MIA3) gene; MTHFD1L, methylenetetrahydrofolate dehydrogenase (NADP+dependent) 1-like gene; p16/CDKN2A, cyclin-dependent kinase inhibitor 2A gene; p15/CDKN2B, cyclin-dependent kinase inhibitor 2B gene; p14/ARF, P14 tumour suppressor gene; MTAP, methylthioadenosine phosphorylase gene; ANRIL, antisense noncoding RNA; CXCL12, chemokine (C-X-C motif) ligand 12 gene.

the loci affect CAD risk as this could provide new targets for drug development. The genes located within each locus (Table 3) have not been previously implicated in the pathogenesis of CAD. Recently, for the locus at 1p13.3, the same allele of SNP rs599839 that is associated with increased CAD risk, has been shown to be associated with higher plasma total and LDL cholesterol in multiple studies,[9–13] providing a possible explanation for the effect on CAD risk, although even for it the precise mechanism by which the locus affects LDL cholesterol and the gene(s) involved awaits elucidation.[19] The 9p21 locus shows a region of association coincident with a gene for a noncoding RNA, ANRIL.[20] Such RNAs often play a regulatory role in gene expression or translation. There is preliminary evidence that ANRIL may affect the expression of the adjacent cyclin-dependent kinase inhibitors,[20] which in turn could affect vascular remodeling processes which are important in the pathogenesis of atherosclerosis and its complications. The association signal at 1q41 lies within the melanoma inhibitory activity family, member 3 (*MIA3*) gene, which may play a similar role in cell growth or inhibition.[21] The locus at 10q11.21 lies upstream of the *CXCL12* gene which codes for stromal cell-derived factor-1 (SDF-1), a chemokine which plays a key role in stem-cell homing and tissue regeneration in ischemic cardiomyopathy[22] and in promoting angiogenesis through recruitment of endothelial progenitor cells.[23] Altogether, the findings open up the prospect of novel therapies for CAD, which may be broadly applicable, from a better understanding of the pathogenic mechanisms in the vascular wall affected by these loci.

Women are less prone to CAD than men, which could partly be attributable to differences in gene–environment interactions. Interestingly, the locus on chromosome 10q11.21 showed a stronger association in women than in men. The nature of the locus with CXCL12 as the most proximate gene (Table 3) does not suggest an immediate mechanism that could explain a gender interaction and whether this finding, which was of borderline statistical significance and would not have been significant if we had adjusted for the multiple interaction analyses carried out, represents a true sex difference in effect requires further investigation. Apart from this, we did not find any other striking interactions, although it should be noted that the lack of data on some risk factors for three control populations means that our ability to detect such interactions was constrained and further investigation in a larger sample is necessary.

In summary, through a large scale replication study we provide compelling evidence for the association of at least 4 genetic loci and risk for CAD. The findings provide a strong foundation for further investigation of these loci as risk factors for CAD and their potential value in the treatment and prevention of this common condition.

Appendix

*CAD Consortium (alphabetical order):

Philippe Amouyel, Dominique Arveiler, S. Matthijs Boekholdt, Peter Braund, Petra Bruse, Suzannah J. Bumpstead, Peter Bugert, Francois Cambien, John Danesh, Panos Deloukas, Angela Doering, Pierre Ducimetière, Ruth M. Dunn, Nour-Eddine El Mokhtari, Jeanette Erdmann, Alun Evans, Phil Ewels, Jean Ferrières, Marcus Fischer, Philippe Frossard, Stephen Garner, Christian Gieger, Mohammed J.R. Gohri, Alison H. Goodall, Anika Großhennig, Alistair Hall, Rob Hardwick, Ari Haukijärvi, Christian Hengstenberg, Thomas Illig, Juha Karvanen, John Kastelein, Frank Kee, Kay-Tee Khaw, Harald Klüter, Inke R. König, Kari Kuulasmaa, Paivi Laiho, Gérald Luc, Winfried März, Ralph McGinnis, William McLaren, Christa Meisinger, Caroline Morrison, Xiodan Ou, Willem H. Ouwehand, Michael Preuss, Carole Proust, Radhi Ravindrarajah, Wilfried Renner, Kate Rice, Jean-Bernard Ruidavets, Danish Saleheen, Veikko Salomaa, Nilesh J. Samani, Manjinder S. Sandhu, Arne S. Schäfer, Michael Scholz, Stefan Schreiber, Heribert Schunkert, Kaisa Silander, Ravi Singh, Nicole Soranzo, Klaus Stark, Birgitta Stegmayr, Jonathan Stephens, John Thompson, Laurence Tiret, Mieke D. Trip, Ellen van der Schoot, Jarmo Virtamo, Nicholas J. Wareham, H-Erich Wichmann, Per-Gunnar Wiklund, Ben Wright, Andreas Ziegler, Jaap-Jan Zwaginga

Steering Committee

H. Schunkert (Cochair), N.J. Samani (Cochair), F. Cambien, J. Danesh, P. Deloukas, J. Erdmann, A. Evans, A. Hall, C. Hengstenberg, K. Kuulasmaa, R. McGinnis, W.H. Ouwehand, D. Saleheen, M. Scholz, J. Thompson, A. Ziegler

Core Writing Group

N.J. Samani (Chair), P. Deloukas, J. Erdmann, C. Hengstenberg, K. Kuulasmaa, R. McGinnis, H. Schunkert, N. Soranzo, J. Thompson, L.Tiret, A. Ziegler

Analysis Group

R. McGinnis (Cochair), J. Thompson (Cochair), A. Ziegler (Cochair), M. Fischer, C. Gieger, A. Großhennig, I.R. König, J. Karvanen, W. McLaren, M. Preuss, M. Scholz, N. Soranzo, L. Tiret, B. Wright

DNA, Genotyping, Data QC, and Informatics

J. Erdmann (Cochair), N. Soranzo (Cochair), P. Braund, P. Bruse, S.J. Bumpstead, P. Deloukas, R.M. Dunn, P. Ewels, S. Garner, R. Hardwick, A. Haukijärvi, M.J.R. Ghori, J. Karvanen, K. Kuulasmaa, P. Laiho, R. McGinnis, W. McLaren, W. März, X. Ou, W.H.

Ouwehand, C. Proust, R. Ravindrarajah, K. Rice, D. Saleheen, M. Sandhu, A.S. Schäfer, M. Scholz, K. Silander, J. Stephens, L. Tiret, M.D. Trip

Primary Investigators of Each Participating Study

AMC-PAS: M.D. Trip, J. Kastelein
SANQUIN controls: E. van der Schoot, J.-J. Zwaginga
ECTIM: D. Arveiler, F. Cambien, A. Evans, F. Kee, G. Luc, C. Morrison, J.-B. Ruidavets
EPIC-Norfolk: M.S. Sandhu, N.J. Wareham, S.M. Boekholdt, K.-T. Khaw
GerMIFS: J. Erdmann, H. Schunkert, C. Hengstenberg
KORA/GOC: M. Fischer, C. Hengstenberg, K. Stark, C. Meisinger,
KORA S4 controls: T. Illig, A. Doering, H.-Erich Wichmann, C. Gieger
LURIC: W. März, W. Renner
German Blood Service (GerBS) controls: P. Bugert, H. Klüter
MORGAM: V. Salomaa, J. Virtamo, P. Amouyel, B. Stegmayr, A. Evans, K. Kuulasmaa, L. Tiret, J. Karvanen, K. Silander, P. Laiho, A. Haukijärvi, C. Proust, D. Arveiler, J. Ferrières, P. Ducimetière, P.-G. Wiklund
PopGen: A.S. Schäfer, N.-E. El Mokhtari, S. Schreiber
UKMI: N.J. Samani, P. Braund, R. Singh, A.H. Goodall

Affiliations

Pasteur Institute, Lille, France (P.A.); Department of Epidemiology and Public Health, Louis Pasteur University, Strasbourg, France (D.A.); Academic Medical Center, University of Amsterdam, The Netherlands (S.M.B., J. Kastelein, M.D.T.); Department of Cardiovascular Sciences, University of Leicester, UK (P. Braund, A.H.G., R.H., N.J.S., R.S.); Medizinische Klinik II, Universität zu Lübeck, Germany (P. Bruse, J.E., H.S.); The Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK (S.J.B., P. Deloukas, R.M.D., P.E., M.J.R.G., R.M., W. McLaren, X.O., R.R., K.R., N.S.); Institute of Transfusion Medicine and Immunology, University of Heidelberg, Mannheim, Germany (P. Bugert, H.K.); INSERM UMR_S 525, UPMC Univ. Paris, France (F.C., C.P., L.T.); Department of Public Health and Primary Care, University of Cambridge, UK (J.D., K.-T.K. D.S., M.S.S.); Institute of Epidemiology, Helmholtz Zentrum München—German Research Center for Environmental Health, Neuherberg, Germany (A.D., C.G., T.I., C. Meisinger, H.-E.W.); IFR69 INSERM-Paris XI University, Paris, France (P. Ducimetière); Institut für Klinische Molekularbiologie, Christian-Albrechts Universität, Kiel, Germany (N.-E.E.M., A.S.S., S.S.); Queens University Belfast, Northern Ireland (A.E., F.K.); INSERM 558, Department of Epidemiology, Paul Sabatier-Toulouse Purpan University, Toulouse, France (J.F.); Klinik und Poliklinik für Innere Medizin II, Universität Regensburg, Germany (M.F., C.H., K. Stark); Aga Khan University, Karachi, Pakistan (P.F., D.S.); Department of Hematology, University of Cambridge, UK (S.G., W.H.O., J.S.); Institut für Medizinische Biometrie und Statistik, Universität zu Lübeck, Germany (A.G., I.R.K., M.P., A.Z.); LIGHT, University of Leeds, UK (A. Hall); National Public Health Institute (KTL), Helsinki, Finland (A. Haukijärvi, J. Karvanen, K.K., P.L., V.S., K. Silander, J.V.); INSERM U508, MONICA Lille, Service d'Epidémiologie et de Santé Publique, Institut Pasteur, Lille, France (G.L.); Clinical Institute of Medical and Chemical Laboratory Diagnostics, Medical University Graz, Austria (W. März, W.R.); The MONICA Project, Glasgow Royal Infirmary, Scotland, UK (C. Morrison); INSERM U518, Faculté de Médecine, MONICA Toulouse, France (J.-B.R.); Trium Analysis Online GmbH, Munich, Germany (M.S.); FIMM, Institute for Molecular Medicine Finland, Helsinki (P.L., K. Silander); Twin Genetic Epidemiology, King's College London, UK (N.S.); Umeå University Hospital, Umeå, Sweden (B.S., P.-G.W.); Department of Health Sciences and Genetics, University of Leicester, UK (J.T., B.W.); Department of Experimental Immunohematology, Sanquin Research, Amsterdam, The Netherlands (E.v.d.S., J.J.Z.); MRC Epidemiology Unit, Cambridge, UK (N.J.W.).

Acknowledgments

We thank the participants and staff in each of the studies who contributed to the present article. We particularly thank Siv Knaappila and Minttu Jussila for technical support in MORGAM. We thank members of the MORGAM Management Group who are not coauthors: Stefan Blankenberg, Marco Ferrario, Leena Peltonen, Markus Perola, Denis Shields, Hugh Tunstall-Pedoe, and Kjell Asplund.

Sources of Funding

Data and sample collation and genotyping were funded by the EU Integrated Project *Cardiogenics* and also supported by the Wellcome Trust. The GerMIFS Study was partly funded through the German Federal Ministry of Education and Research (BMBF) in the context of the German National Genome Research Network (NGFN-2 and NGFN-plus). The MORGAM study was partly funded through the European Community's Seventh Framework Programme ENGAGE project (grant agreement HEALTH-F4-2007-201413), the Center of Excellence in Complex Disease Genetics of the Academy of Finland (CoECDG), and Finnish Foundation for Cardiovascular Research. N.J.S. holds a Chair supported by the British Heart Foundation.

Disclosures

None.

References

1. Morgan TM, Krumholz HM, Lifton RP, Spertus JA. Non-validation of reported genetic risk factors for acute coronary syndrome in a large-scale replication study. *JAMA*. 2007;297:1551–1561.
2. Wellcome Trust Case Control Consortium. Genomewide association study of 14,000 cases of seven common diseases and 3000 shared controls. *Nature*. 2007;447:661–678.
3. McPherson R, Pertsemlidis A, Kavaslar N, Stewart A, Roberts R, Cox DR, Hinds DA, Pennacchio LA, Tybjaerg-Hansen A, Folsom AR, Boerwinkle E, Hobbs HH, Cohen JC. A common allele on chromosome 9 associated with coronary heart disease. *Science*. 2007;316:1488–1491.
4. Helgadottir A, Thorleifsson G, Manolescu A, Gretarsdottir S, Blondal T, Jonasdottir A, Jonasdottir A, Sigurdsson A, Baker A, Palsson A, Masson G, Gudbjartsson DF, Magnusson KP, Andersen K, Levey AI, Backman VM, Matthiasdottir S, Jonsdottir T, Palsson S, Einarsdottir H, Gunnarsdottir S, Gylfason A, Vaccarino V, Hooper WC, Reilly MP, Granger CB, Austin H, Rader DJ, Shah SH, Quyyumi AA, Gulcher JR, Thorgeirsson G, Thorsteinsdottir U, Kong A, Stefansson K. A common variant on chromosome 9p21 affects the risk of myocardial infarction. *Science*. 2007;316:1491–1493.
5. Samani NJ, Erdmann J, Hall AS, Hengstenberg C, Mangino M, Mayer B, Dixon RJ, Meitinger T, Braund P, Wichmann HE, Barrett JH, König IR, Stevens SE, Szymczak S, Tregouet DA, Iles MM, Pahlke F, Pollard H, Lieb W, Cambien F, Fischer M, Ouwehand W, Blankenberg S, Balmforth AJ, Baessler A, Ball SG, Strom TM, Braenne I, Gieger C, Deloukas P, Tobin MD, Ziegler A, Thompson JR, Schunkert H; WTCCC and the Cardiogenics Consortium. Genomewide association analysis of coronary artery disease. *N Engl J Med*. 2007;357:443–453.
6. Schunkert H, Götz A, Braund P, McGinnis R, Tregouet D-A, Mangino M, Linsel-Nitschke P, Cambien F, Hengstenberg C, Stark K, Blankenberg S, Tiret L, Ducimetiere P, Keniry A, Ghori MJR, Schreiber S, El Mokhtari NE, Hall AS, Dixon RJ, Goodall AH, Liptau H, Pollard H, Schwarz DM, Hothorn LA, Wichmann H-E, König IR, Fischer M, M.D, Meisinger C, Ouwehand W, Cardiogenics Consortium, Deloukas P, Thompson JR, Erdmann J, Ziegler A, Samani NJ. Repeated replication and a prospective meta-analysis of the association between chromosome 9p21.3 and coronary artery disease. *Circulation*. 2008;117:1675–1684.
7. Broadbent HM, Peden JF, Lorkowski S, Goel A, Ongen H, Green F, Clarke R, Collins R, Franzosi MG, Tognoni G, Seedorf U, Rust S, Eriksson P, Hamsten A, Farrall M, Watkins H. Susceptibility to coronary artery disease and diabetes is encoded by distinct, tightly linked, SNPs in the ANRIL locus on chromosome 9p. *Hum Mol Genet*. 2008;17:806–814.
8. Helgadottir A, Thorleifsson G, Magnusson KP, Grétarsdottir S, Steinthorsdottir V, Manolescu A, Jones GT, Rinkel GJ, Blankensteijn JD, Ronkainen A, Jääskeläinen JE, Kyo Y, Lenk GM, Sakalihasan N, Kostulas K, Gottsäter A, Flex A, Stefansson H, Hansen T, Andersen G, Weinsheimer S, Borch-

Johnsen K, Jorgensen T, Shah SH, Quyyumi AA, Granger CB, Reilly MP, Austin H, Levey AI, Vaccarino V, Palsdottir E, Walters GB, Jonsdottir T, Snorradottir S, Magnusdottir D, Gudmundsson G, Ferrell RE, Sveinbjornsdottir S, Hernesniemi J, Niemelä M, Limet R, Andersen K, Sigurdsson G, Benediktsson R, Verhoeven EL, Teijink JA, Grobbee DE, Rader DJ, Collier DA, Pedersen O, Pola R, Hillert J, Lindblad B, Valdimarsson EM, Magnadottir HB, Wijmenga C, Tromp G, Baas AF, Ruigrok YM, van Rij AM, Kuivaniemi H, Powell JT, Matthiasson SE, Gulcher JR, Thorgeirsson G, Kong A, Thorsteinsdottir U, Stefansson K. The same sequence variant on 9p21 associates with myocardial infarction, abdominal aortic aneurysm and intracranial aneurysm. *Nat Genet*. 2008;40:217–224.

9. Wallace C, Newhouse SJ, Braund P, Zhang F, Tobin M, Falchi M, Ahmadi K, Dobson RJ, Marçano AC, Hajat C, Burton P, Deloukas P, Brown M, Connell JM, Dominiczak A, Lathrop GM, Webster J, Farrall M, Spector T, Samani NJ, Caulfield MJ, Munroe PB. Genome-wide association study identifies genes for biomarkers of cardiovascular disease: serum urate and dyslipidemia. *Am J Hum Genet*. 2008;82:139–149.
10. Kathiresan S, Melander O, Guiducci C, Surti A, Burtt NP, Rieder MJ, Cooper GM, Roos C, Voight BF, Havulinna AS, Wahlstrand B, Hedner T, Corella D, Tai ES, Ordovas JM, Berglund G, Vartiainen E, Jousilahti P, Hedblad B, Taskinen MR, Newton-Cheh C, Salomaa V, Peltonen L, Groop L, Altshuler DM, Orho-Melander M. Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. *Nat Genet*. 2008;40:189–197.
11. Willer CJ, Sanna S, Jackson AU, Scuteri A, Bonnycastle LL, Clarke R, Heath SC, Timpson NJ, Najjar SS, Stringham HM, Strait J, Duren WL, Maschio A, Busonero F, Mulas A, Albai G, Swift AJ, Morken MA, Narisu N, Bennett D, Parish S, Shen H, Galan P, Meneton P, Hercberg S, Zelenika D, Chen WM, Li Y, Scott LJ, Scheet PA, Sundvall J, Watanabe RM, Nagaraja R, Ebrahim S, Lawlor DA, Ben-Shlomo Y, Davey-Smith G, Shuldiner AR, Collins R, Bergman RN, Uda M, Tuomilehto J, Cao A, Collins FS, Lakatta E, Lathrop GM, Boehnke M, Schlessinger D, Mohlke KL, Abecasis GR. Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nat Genet*. 2008;40:161–169.
12. Sandhu MS, Waterworth DM, Debenham SL, Wheeler E, Papadakis K, Zhao JH, Song K, Yuan X, Johnson T, Ashford S, Inouye M, Luben R, Sims M, Hadley D, McArdle W, Barter P, Kesäniemi YA, Mahley RW, McPherson R, Grundy SM, Wellcome Trust Case Control Consortium, Bingham SA, Khaw KT, Loos RJ, Waeber G, Barroso I, Strachan DP, Deloukas P, Vollenweider P, Wareham NJ, Mooser V LDL-cholesterol concentrations: a genome-wide association study. *Lancet.* 2008;371:483–491.
13. Samani NJ, Braund PS, Erdmann J, Götz A, Tomaszewski M, Linsel-Nitschke P, Hajat C, Mangino M, Hengstenberg C, Stark K, Ziegler A, Caulfield M, Burton PR, Schunkert H, Tobin MD The novel genetic variant predisposing to coronary artery disease in the region of the *PSRC1* and *CELSR2* genes on chromosome 1 associates with serum cholesterol. *J Mol Med.* 2008;86:1233–1241.
14. Karvanen J, Silander K, Kee F, Tiret L, Salomaa V, Kuulasmaa K, Wiklund P-G, Virtamo J, Saarela O, Perret C, Perola M, Peltonen L, Cambien F, Erdmann J, Samani NJ, Schunkert H, Evans A, for the MORGAM Project. The impact of newly-identified loci on coronary heart disease, stroke and total mortality in the MORGAM prospective cohorts. *Genet Epidemiol*. 2008; Oct 31 Epub ahead of print.
15. Lohmueller KE, Pearce CI, Pike M, Lander ES, Hirshhorn JN. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nat Genet*. 2003;33:177–182.
16. Lango H, The UK Type 2 Diabetes Consortium, Palmer CN, Morris AD, Zeggini E, Hattersley AT, McCarthy MI, Frayling TM, Weedon MN. Assessing the combined impact of 18 common genetic variants of modest effect sizes on type 2 diabetes risk. *Diabetes*. 2008;57:2129–2135.
17. Kannel WB, Garcia M, McNamara PM, Pearson G. Serum lipid precursors of coronary heart disease. *Hum Pathol*. 1971;2:129–151.
18. Ortlepp JR, Lauscher J, Janssens U, Minkenberg R, Hanrath P, Hoffmann R. Analysis of several hundred genetic polymorphisms may improve assessment of the individual genetic burden for coronary artery disease. *Eur J Intern Med*. 2002;13:485–492.
19. Schadt EE, Molony C, Chudin E, Hao K, Yang X[1], Lum PY, Kasarskis A, Zhang B, Wang S, Suver C, Zhu J, Millstein J, Sieberts S, Lamb J, GuhaThakurta D, Derry J, Storey JD, Avila-Campillo I, Kruger MJ, Johnson JM, Rohl CA, van Nas A, Mehrabian M, Drake TA, Lusis AJ, Smith RC, Guengerich FP, Strom SC, Schuetz E, Rushmore TS, Ulrich R. Mapping the genetic architecture of gene expression in human liver. *PLoS Biol*. 2008;6:e107.
20. Pasmant E, Laurendeau I, Heron D, Vidaud M, Vidaud D, Bieche I. Characterization of a germ-line deletion, including the entire INK4/ARF locus, in a melanoma-neural system tumor family: identification of ANRIL, an antisense noncoding RNA whose expression coclusters with ARF. *Cancer Res*. 2007;67:3963–3969.
21. Bosserhoff AK, Buettner R. Expression, function and clinical relevance of MIA (melanoma inhibitory activity). *Histol Histopathol*. 2002;17:289–300.
22. Askari A, Unzek S, Popovic ZP, Goldman CK, Ellis SG, Thomas JD, DiCorleto PE, Topol EJ, Penn MS Effect of stromal-cell-derived factor 1 on stem-cell homing and tissue regeneration in ischaemic cardiomyopathy. *Lancet.* 2003;362:697–703.
23. Zheng H, Fu G, Dai T, Huang H. Migration of endothelial progenitor cells mediated by stromal cell-derived factor-1alpha/CXCR4 via PI3K/Akt/eNOS signal transduction pathway. *J Cardiovasc Pharmacol*. 2007;50:274–280.

Arteriosclerosis, Thrombosis, and Vascular Biology

JOURNAL OF THE AMERICAN HEART ASSOCIATION

Large Scale Association Analysis of Novel Genetic Loci for Coronary Artery Disease

Coronary Artery Disease Consortium

Arterioscler Thromb Vasc Biol. 2009;29:774-780; originally published online January 22, 2009;
doi: 10.1161/ATVBAHA.108.181388
Arteriosclerosis, Thrombosis, and Vascular Biology is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231

Print ISSN: 1079-5642. Online ISSN: 1524-4636

ATVB/2008/181388R2

LARGE SCALE ASSOCIATON ANALYSIS OF NOVEL GENETIC LOCI FOR CORONARY ARTERY DISEASE

Coronary Artery Disease Consortium

Supplementary Materials

1. Descriptions of individuals and recruitment procedures by study

AMC-PAS. Cases were recruited as part of a prospective cohort study (Academic Medical Centre Amsterdam Premature Atherosclerosis Study (AMC-PAS) with symptomatic CAD before the age of 51 years, defined as MI, coronary revascularization, or evidence of at least 70% stenosis in a major epicardial artery.[1] A collection of DNA samples from blood donors from the north-west region of the Netherlands was established as controls for this study. Participating donors were recruited at routine Sanquin Blood Bank donation sessions (Sanquin Common Controls (SANQUIN-CC). More than 95% of the controls are from the same region as the cases of the AMC-PAS cohort.

ECTIM. The ECTIM (Etude Cas-Témoin sur l'Infarctus du Myocarde) study is a case-control study of MI based on the WHO MONICA (Multinational MONItoring of trends and determinants in CArdiovascular disease) Project registers in the United-Kingdom (UK) and France[2,3]. All participants were of European descent and gave an informed consent. The recruitment was performed in two phases: in 1988-1990, men with MI, aged 25-64 yrs, were drawn from the MONICA registers of Belfast (Northern Ireland), Strasbourg, Lille and Toulouse (France). They were recruited 3 to 9 months after the event and had to satisfy the WHO MONICA criteria for definite acute MI. In each

centre, controls of similar age and sex were randomly selected in the areas covered by the MONICA registers. Controls with self-reported history of coronary heart disease were excluded from the current analysis. They were drawn from the lists of General Practitioners held by the Central Service Agency in Northern Ireland and from the electoral rolls in France. In 1997-1998, men and women from Glasgow were included in the study using the same protocol. At the same time, the sample was extended in Belfast by the inclusion of women and additional men.

EPIC-Norfolk. Participants for this study were from the Norfolk cohort of the European Prospective Investigation into Cancer and Nutrition study (EPIC-Norfolk). EPIC-Norfolk is a prospective population study of 25 663 men and women aged between 40 and 79 years, resident in Norfolk, UK, recruited from general practice registers between 1993 and 1997. It is an ethnically homogeneous Caucasian population. These participants completed a health examination. Details of recruitment, anthropometric measurements, and health examinations following standardized protocols have been published.[4] All individuals have been flagged for death certification at the UK Office of National Statistics, with vital status ascertained for the entire cohort. In addition, participants admitted to hospital were identified using their unique National Health Service number by data linkage with ENCORE (East Norfolk Health Authority database), which identifies all hospital contacts throughout England and Wales for Norfolk residents. The Norwich District Health Authority Ethics Committee approved the study. All participants gave signed informed consent.

For the purposes of this study, we used a nested case–control design based on participants who were disease-free at the baseline assessment. Details of the study design have been previously described.[5] This study represents an extended follow-up. Briefly, we excluded all individuals who reported a history of heart attack or stroke at the baseline clinic visit. Cases were individuals who developed a fatal or non-fatal CAD during an average follow-up of 11 years, until June 2006. Participants were identified as having CAD during follow-up if they had a hospital admission and/or died with CAD as the underlying cause. CAD was defined as cause of death codes ICD9 410-414 or ICD10 I20-I25, and hospital discharge codes ICD10 I20.0, I21, I22 or I23 according to the International Classification of Diseases, 9th and 10th revisions. Controls were study participants who remained free of any cardiovascular disease during follow-up (defined as ICD9 401-448 and ICD10 I10-I79). We matched one control to each case by sex, age (within 5 years), and time of enrolment (within 3 months).

GerMIFS. The German MI Family Study (GerMIFS) comprises unrelated German MI patients (age of onset < 65 years) having at least one first-degree relative with premature CAD.[6] The subjects used in this analysis are distinct from those used in the initial genome-wide association study.[7] Patients were recruited between 1997 and 2002 from hospital clinics and were studied by physical examination, blood testing, echocardiography, as well as a standardized interview. All events were validated through inspection of hospital charts. Healthy German married-in spouses from the same recruitment centre served as controls.[7]

KORA/GOC. In 1996, a total of 589 patients with history of sporadic MI prior to the age of 60 years were identified through the Augsburg, Germany, MONICA MI register and recruited. This cohort is referred to as the KORA-B cohort.[8] Diagnosis of MI was established according to the MONICA diagnostic criteria. From 2007, this cohort has been extended by additional patients attending for angiography in the Cardiology Department at the University of Regensburg, who also have a validated diagnosis of MI prior to 60 years of age Recruitment of this cohort called Go-Kard (GOC) is on-going. Controls (KORA-S4) represent a gender- and age-stratified random sample of all German residents and come from the same geographic area as the cases. They participated in the echocardiographic sub study of the third MONICA Augsburg survey 1994/1995.[9] All controls were recruited with the same protocol as the MI patients.

LURIC. The LURIC (Ludwigshafen Risk and Cardiovascular Health) study included 3,316 consecutive white patients of German ancestry hospitalized for coronary angiography between June 1997 and May 2001 in Ludwigshafen, Germany.[10] For the purpose of this analysis, only subjects with angiographically confirmed CAD (at least one coronary vessel with a stenosis > 50%) were included. The controls were from the GerBS control series that consists of healthy, unrelated blood donors. They were recruited between May-July 2004 from the southwestern area of Germany which corresponds to the geographical origin of the LURIC patients by the Institute of Transfusion Medicine and Immunology, Mannheim, Germany. According to the German guidelines for blood donation all blood donors were examined to rule out cardiovascular, malignant, and other diseases by a standard questionnaire, cardiac auscultation, blood pressure and pulse measurement.

MORGAM. MORGAM is a prospective follow-up of the respondents of representative population samples that were examined at baseline.[11] This study includes cohorts from Finland (FINRISK, ATBC), France (Lille, Strasbourg, Toulouse), Northern Sweden and UK (Belfast). Details of each cohort and the diagnostic procedures and a quality assessment of the baseline and follow-up data in each cohort have been published separately.[12,13] For its genetic component, MORGAM has a case-cohort design.[14] For the purposes of this study, cases and controls were selected from the MORGAM case-cohort set in such a way that the sets of cases and controls from each cohort would be similar in size and comparable with respect to age and sex. For each sex in each cohort, the cases and controls were selected using a stepwise procedure, starting from the youngest case. A person was eligible to be a case if: (i) he had a documented MI at baseline or had definite AMI, unstable angina, cardiac revascularization or unclassifiable death during follow-up; and (ii) had not been selected as a control for a younger case. A control was selected for each case in random from those who were at risk at the age of the onset of the event (or age at baseline if the case had a MI at baseline). The risk set, from which the controls were selected, constituted of those members of the random subcohort of the MORGAM case-cohort set who: (i) had not had documented or self-reported MI or stroke at baseline; and (ii) had not had definite or possible AMI, unstable angina or cardiac revascularization during follow-up before the age of the case; and (iii) were in the follow-up at the age of the case; and (iv) had not been selected as a case or a control at an earlier stage.

PopGen. The case sample comprised unrelated CAD patients who were recruited in Schleswig-Holstein, the northernmost region in Germany through the population-based PopGen biobank (www.popgen.de).[15] Coronary angiograms of any of the five cardiac catheterization laboratories in this geographical area were screened. Study subjects were required to demonstrate significant CAD (> 70% stenosis in at least one major epicardial coronary vessel). All subjects were recruited between 2002 and 2005. The control samples were healthy blood donors to the Blood Service of the University Hospital Schleswig-Holstein. All subjects were recruited between 2003 and 2006.

UKMI. The UK MI Study combined subjects from two previous molecular genetic studies of MI that were recruited between 1993 and 2002.[16,17] The first study[16] comprised 548 consecutive acute MI cases (age < 75 years) recruited from admissions to the coronary care units (CCUs) of the Leicester Royal Infirmary, Leicester and the Royal Hallamshire Hospital, Sheffield. Control subjects (n = 525) were recruited in each hospital from adult visitors to patients without overt CHD and were matched for age and gender with the cases. The second study[17] comprised 433 subjects. The MI cases (n = 224) were recruited retrospectively from the registries of CCUs in Leicester. All cases satisfied the WHO MONICA diagnostic criteria for acute MI. Control subjects (n = 209) were recruited from three primary care practices located within the same geographical area as the cases.

2. Definitions

Because of different recruitment strategies in the different studies and different time periods for recruitment (ranging from early 1990s to 2007), during which the definition

of MI has changed, a sub-committee of the steering committee consisting of clinicians with cardiology experience and an epidemiologist with experience on harmonization of data from different studies (CH, HS, KK, NJS) examined the definitions used in each study for both CAD phenotypes and risk factors and ensured that there was consistency across the studies. Cases that did not meet the criteria for CAD or MI were excluded.

(i) Coronary artery phenotypes

Coronary artery disease (CAD) This was defined by one of the following criteria:

- Confirmed MI (see below)
- > 50% stenosis in at least one coronary vessel at angiography with validation from hospital records
- validated history of percutaneous transluminal coronary angioplasty (PTCA) or coronary artery bypass graft surgery (CABG)
- validated angina, defined as symptoms + confirmation from at least non-invasive provocation test e.g. scintigraphy or exercise treadmill test.
- For EPIC-Norfolk and MORGAM death confirmed to be due to CAD or highly likely to be due to CAD

Myocardial infarction (MI) Before inclusion, cases with MI in each study had to fulfil **one** of the following criteria:

1) 'Definite MI' according to the 1983 WHO criteria ("MONICA")[18]

2) 'MI' according to the ESC/ACC 2000 definition of MI[19]

3) 'Definite MI' according to the 2003 AHA/WHF/ESC/CDC/NHLBI definition[20]

4) 'Clinical MI' according to the 2004 British Cardiac Society Working Group Definition[21]

5) 'Myocardial infarction' according to the 2007 ESC/ACCF/AHA/WHF definition.[22]

(ii). Risk factors

Risk factors were defined on the criteria defined below. For cases the categorisation was based on status at the time of the qualifying event, or at baseline of the prospective cohorts (EPIC-Norfolk, MORGAM):

Hypertension: Reported history of hypertension or, when such information was not available, systolic blood pressure >= 140 mmHg and/or diastolic blood pressure >= 90 mmHg.

Diabetes Mellitus. Known Diabetes Mellitus (either Type 1 or Type 2)..

Smoking Current or former smokers were categorised as smokers; non-smokers were subjects who had never smoked.

SUPPLEMENTARY REFERENCES

1. Trip MD, Smulders YM, Wegman JJ et al. Frequent mutation in the ABCC6 gene (R1141X) is associated with a strong increase in the prevalence of coronary artery disease. *Circulation* 2002 Aug 13;106(7):773-5.

2. Parra, HJ, Arveiler D, Evans AE et al. A case-control study of lipoprotein particles in two populations at contrasting risk for coronary heart disease. The ECTIM Study. *Arterioscler Thromb.* 1992;12(6):701-707.

3. Kee F, Morrison C, Evans AE et al. Polymorphisms of the P-selectin gene and risk of myocardial infarction in men and women in the ECTIM extension study. Etude cas-temoin de l'infarctus myocarde. *Heart*. 2000;84(5):548-552.

4. Day N. Oakes S, Luben R et al. EPIC-Norfolk: study design and characteristics of the cohort. European Prospective Investigation of Cancer. *Br J Cancer.*1999; 80 Suppl 1:95-103.

5. Boekholdt SM, Kuivenhoven JA, Wareham NJ et al. Plasma levels of cholesteryl ester transfer protein and the risk of future coronary artery disease in apparently healthy men and women: the prospective EPIC (European Prospective Investigation into Cancer and nutrition)-Norfolk population study. *Circulation*. 2004;110(11):1418-1423

6. Fischer M, Broeckel U, Holmer S et al., Distinct heritable patterns of angiographic coronary artery disease in families with myocardial infarction. *Circulation.* 2005; 111(7): 855-862.

7. Samani NJ, Erdmann J, Hall AS, et al. Genomewide association analysis of coronary artery disease. *N Engl J Med.* 2007;357(5):443-453.

8. Lieb W, Voss C, Ortak J et al. Impact of diabetes on QT dynamicity in patients with and without myocardial infarction: The KORA Family Heart Study. Pacing Clin Electrophysiol. 2007 Jan;30 Suppl 1:S183-7

9. Baessler A, Kwitek A, Fischer M et al. Association of the Ghrelin receptor gene region with left ventricular hypertrophy in the general population: results of the MONICA/KORA Augsburg Echocardiographic Substudy. Hypertension. 2006 May;47(5):920-927.

10. Winkelmann BR, Marz W, Boehm BO et al. Rationale and design of the LURIC study--a resource for functional genomics, pharmacogenomics and long-term prognosis of cardiovascular disease. *Pharmacogenomics.* 2001 Feb;2(1 Suppl 1):S1-S73.

11. Evans A, Salomaa V, Kulathinal S, et al. MORGAM (an international pooling of cardiovascular cohorts). *Int J Epidemiol.* 2005;34(1):21-27.

12. Kulathinal S, Niemelä M, Kuulasmaa K, contributors from Participating Centres, for the MORGAM Project. Description of MORGAM Cohorts. (2005). Available from URL:*http://www.ktl.fi/publications/morgam/cohorts/index.html*, URN:*NBN:fi-fe20051214*.

13. Niemelä M, Kulathinal S and Kuulasmaa K, editors, for the MORGAM Project. Description and quality assessment of MORGAM data. (2007). Available from URL:*http://www.ktl.fi/publications/morgam/qa/contents.htm*,URN:*NBN:fi-fe20071495*.)

14. Kulathinal S, Karvanen J, Saarela O, Kuulasmaa K, for the MORGAM Project. Case-cohort design in practice — experiences from The MORGAM Project. Epidemiol Perspect Innov 2007;4:15 doi:10.1186/1742-5573-4-15]

15. Krawczak M, Nikolaus S, von Eberstein H, El Mokhtari NE, Schreiber S. PopGen: population-based recruitment of patients and controls for the analysis of complex genotype-phenotype relationships. *Community Genet*. 2006; 9(1):55-61.

16. Steeds R, Addams M, Smith P, Channer K, Samani NJ. Distribution of Tissue Plasminogen Activator Insertion/Deletion Polymorphism in Myocardial Infarction and Control Subjects. *Thromb Haemost.* 1998; 79(5):980-84.

17. Brouilette S, Singh RK, Thompson JR, Goodall AH, Samani NJ. White Cell Telomere Length and Risk of Premature Myocardial Infarction. *Arterioscler Thromb. Vasc Biol.* 2003; 23(5):842-846.

18. Tunstall-Pedoe H, Kuulasmaa K, Amouyel P, Arveiler D, Rajakangas AM, Pajak A. Myocardial infarction and coronary deaths in the World Health Organization MONICA Project. Registration procedures, event rates and case fatality in 38 populations from 21 countries in 4 continents. *Circulation* 1994;90(1):583-612.

19. Myocardial infarction redefined--a consensus document of The Joint European Society of Cardiology/American College of Cardiology Committee for the redefinition of myocardial infarction. *Eur Heart J.* 2000; 21(18):1502-1513.

20. Luepker RV, Apple FS, Christenson RH, et al. Case Definitions for Acute Coronary Heart Disease in Epidemiology and Clinical Research Studies: A Statement From the AHA Council on Epidemiology and Prevention; AHA Statistics Committee; World Heart Federation Council on Epidemiology and Prevention. *Circulation.* 2003; 108(20):2543-2549.

21. Fox KA, Birkhead J, Wilcox R, Knight C, Barth J. British Cardiac Society Working Group on the definition of myocardial infarction. *Heart.* 2004. 90(6): 603-609.

22. Thygesen K, Alpert JS, White HD, et al. Universal definition of myocardial infarction. *Eur Heart J.* 2007; 28(20):2525-2538.

Figure legends

Figure 1. Linkage disequilibrium (LD) plots of (a) 1q41, (b) 2q36.3, (c) 6q25.1, (d) 9p21, and (e) 10q11.21 each showing the position of the lead SNP (red circle or diamond) and additional ones (blue circles) selected for genotyping.

Figure 2. Forest plots showing the association with CAD in the individual contributing studies for the loci on chromosomes 1p13.3, 1q41, 2q36.3, 6q25.1, 9p21, and 10q11.21. The non-significant results for the locus on chromosome 15q22.33 are not shown. The individual odds ratios (OR) with 95% confidence intervals are shown at the side of each study. Please refer to text for details of each study.

Figure 3. Forest plots showing the association with CAD for the loci on chromosomes 1p13.3, 1q41, 2q36.3, 6q25.1, 9p21, and 10q11.21 for pre-specified sub-groups. The individual odds ratios (OR) with 95% confidence intervals are shown at the side of each sub-group. Note the significant difference in the effect of the locus on chromosome 10q11.21 between men and women.

Supplementary Table 1 List of SNPs genotyped and relationship of additional SNPs typed to the lead SNPs

SNP	Chromosome	Type*	Controls	CAD cases	MAF in controls	MAF in CAD cases	r^2 to lead SNP (controls)	D' to lead SNP (controls)
rs599839	1p13.3	Lead	10938	11168	0.228	0.207		
rs17465637	1q41	Lead	10708	10878	0.266	0.255		
rs3008621	1q41	additional	10574	10890	0.136	0.122	0.34	0.88
rs2943634	2q36.3	Lead	10824	11040	0.342	0.333		
rs2972147	2q36.3	additional	10540	10261	0.362	0.354	0.65	0.85
rs6922269	6q25.1	Lead	10765	11115	0.263	0.272		
rs12214306	6q25.1	additional	10903	11239	0.035	0.034	0.01	0.63
rs12525353	6q25.1	additional	10852	11194	0.029	0.029	0.01	0.86
rs1333049	9p21	Lead	10870	11064	0.458	0.505		
rs10738610	9p21	additional	10368	10418	0.475	0.516	0.88	0.98
rs501120	10q11.21	Lead	10095	10478	0.133	0.121		
rs2146807	10q11.21	additional	10803	11128	0.122	0.117	0.16	0.43
rs17228212	15q22.33	Lead	10834	11212	0.276	0.274		

*Lead SNP is the SNP at each locus showing the strongest association in the combined analysis of the two genome-wide association studies (Supplementary Reference 6). The rationale for typing the additional SNPs is given in the main paper (see Methods). r^2 and D' are the two measures of linkage disequilibrium between the lead SNP and the additional SNPs. The number of controls and CAD cases with successful genotypes are shown for each SNP as are the minor allele frequencies (MAF) in controls and cases.

Supplementary Table 2 Characteristics of the cases and controls in each study

Study		Number	Male (%)	Age Mean (sd)	BMI Mean (sd)	Hypertension (%)	Smoking (%)	Diabetes (%)
AMC-PAS	Cases	744	78%	43.0(5.2)	27.1(4.1)	27%	76%	10%
	Controls	1299	67%	50.9(11.8)	-	-	-	-
ECTIM	Cases	1146	78%	55.5(8.3)	27.0(4.2)	44%	80%	13%
	Controls	1102	78%	55.7(8.5)	26.8(4.2)	29%	62%	6%
EPIC-NORFOLK	Cases	1081	69%	65.0(8.2)	27.1(3.9)	27%	72%	6%
	Controls	2175	65%	64.4(7.9)	26.4(3.5)	13%	59%	2%
GERMAN MI-FS	Cases	732	75%	58.6(8.5)	27.6(4.2)	84%	57%	19%
	Controls	739	49%	57.4(10.0)	26.8(4.0)	51%	54%	6%
KORA/GOC	Cases	809	84%	59.8(9.5)	28.3(4.0)	94%	73%	20%
	Controls	1022	77%	58.3(9.9)	28.1(4.4)	45%	62%	7%
LURIC	Cases	2038	76%	63.3(10.0)	27.4(4.0)	61%	71%	21%
	Controls	1334	50%	46.8(15.3)	-	-	-	-
MORGAM	Cases	1418	89%	60.4(7.4)	27.9(4.2)	49%	82%	13%
	Controls	1433	89%	60.2 (7.4)	27.0(4.1)	37%	73%	14%
POPGEN	Cases	2811	82%	60.9(8.4)	28.0(4.2)	74%	73%	20%
	Controls	1368	27%	49.2(17.0)	25.5(4.8)	-	-	-
UKMI	Cases	771	73%	54.4(11.1)	27.2(4.8)	31%	80%	9%
	Controls	733	69%	52.1(10.4)	26.0(3.7)	14%	59%	2%

The number of case and controls in each study and their characteristics are shown. For cases, the age given and the status for hypertension, diabetes and smoking relate to the time of event or at time of recruitment for the prospectively ascertained studies (EPIC-Norfolk and MORGAM). See text for definition of phenotype. Please note that because the controls for the AMC-PAS, LURIC, and Popgen studies are from blood donor collections there is no information on them on cardiovascular risk factors.

Supplementary Table 3 Associations in the pooled European cohorts for all SNPs.

						CAD			MI		
SNP	Chr	a1/a2+	RA	MAF in controls	MAF in cases	OR (95%CI)	P-value	Hetero P-value*	OR (95%CI)	P-value	Hetero p-value*
rs599839	1p13.3	A/G	A	0.228	0.207	1.13 (1.08,1.19)	$1.44x10^{-7}$ ($1.87x10^{-6}$)	0.72	1.11 (1.05,1.18)	$2.35x10^{-4}$ ($3.06x10^{-3}$)	0.38
rs3008621	1q41	G/A	G	0.136	0.122	1.10 (1.04,1.17)	$1.02x10^{-3}$ ($1.33x10^{-2}$)	0.49	1.09 (1.01,1.17)	$2.11x10^{-2}$ (0.274)	0.24
rs17465637	1q41	C/A	C	0.266	0.255	1.04 (0.99,1.09)	0.089 (1.000)	0.43	1.01 (0.96,1.07)	0.637 (1.000)	0.36
rs2943634	2q36.3	C/A	C	0.342	0.333	1.05 (1.00,1.09)	$3.22x10^{-2}$ (0.419)	0.24	1.03 (0.98,1.09)	0.218 (1.000)	0.24
rs2972147[#]	2q36.3	C/T	C	0.362	0.354	1.04 (1.00,1.09)	$4.29x10^{-2}$ (0.558)	<0.001	1.03 (0.98,1.08)	0.344 (1.000)	0.001
rs6922269	6q25.1	G/A	A	0.263	0.272	1.05 (1.01,1.10)	$1.96x10^{-2}$ (0.255)	0.40	1.08 (1.03,1.14)	$3.45x10^{-3}$ ($4.49x10^{-2}$)	0.66
rs12214306	6q25.1	T/C	T	0.035	0.034	1.04 (0.94,1.15)	0.474 (1.000)	0.52	1.04 (0.91,1.18)	0.559 (1.000)	0.50
rs12525353	6q25.1	C/A	C	0.029	0.029	1.03 (0.92,1.16)	0.580 (1.000)	0.65	1.08 (0.94,`.25)	0.278 (1.000)	0.74
rs1333049	9p21	G/C	C	0.458	0.505	1.20 (1.16,1.25)	$2.89x10^{-21}$ ($3.76x10^{-20}$)	0.65	1.24 (1.18,1.30)	$1.28x10^{-18}$ ($1.66x10^{-17}$)	0.76
rs10738610	9p21	A/C	C	0.475	0.516	1.19 (1.14,1.24)	$1.98x10^{-17}$ ($2.57x10^{-16}$)	0.23	1.22 (1.16,1.28)	$1.03x10^{-14}$ ($1.34x10^{-13}$)	0.17
rs501120	10q11.21	T/C	T	0.133	0.121	1.11 (1.05,1.18)	$4.34x10^{-4}$ ($5.64x10^{-3}$)	0.04	1.15 (1.07,1.24)	$1.99x10^{-4}$ ($2.59x10^{-3}$)	0.07
rs2146807	10q11.21	T/C	T	0.122	0.117	1.02 (0.96,1.09)	0.453 (1.000)	0.95	1.01 (0.94,1.08)	0.844 (1.000)	0.99
rs17228212	15q22.33	T/C	T	0.276	0.274	1.00 (0.95,1.04)	0.893 (1.000)	0.11	1.02 (0.97,1.07)	0.521 (1.000)	0.13

Legend to Supplementary Table 3. SNP, single nucleotide polymorphism; chr, chromosome; [+] a2 is minor allele; RA, risk allele; MAF, minor allele frequency; OR, odds ratio associated with the risk allele, based on a fixed effects logistic regression analysis adjusting for study and centre (for studies with multiple centres); CI, Confidence intervals. [§] p-value for heterogeneity between studies assessed (see Statistical Methods). [#]Only one SNP, rs2972147, showed very strong heterogeneity, I^2=72% in CAD, I^2=71% in MI. Analysis of this SNP using a random effects model gave an OR for CAD of 1.03 (0.95, 1.12) and for MI of 1.01 (0.92, 1.12). P-values in parentheses are after Bonferonni correction. The CIs have not been adjusted for multiple comparisons.

Supplementary Table 4 Pooled association analysis for the six loci nominally significant in the present study with those from the two previous genome-wide association studies

SNP	Previous GWA studies*	Current studies	Combined	Combined with Bonferroni correction[#]
rs599839 (1p13.3))	4.05×10^{-9}	1.44×10^{-7}	1.09×10^{-14}	1.97×10^{-9}
rs3008621 (1q41)	4.89×10^{-6}	1.02×10^{-3}	5.36×10^{-8}	4.72×10^{-3}
rs2943634 (2q36.3)	1.61×10^{-7}	3.22×10^{-2}	5.56×10^{-8}	6.27×10^{-3}
rs6922269 (6q25.1)	2.90×10^{-8}	1.96×10^{-2}	6.73×10^{-9}	8.57×10^{-4}
rs1333049 (9p21)	2.90×10^{-19}	2.89×10^{-21}	3.87×10^{-38}	9.30×10^{-34}
rs501120 (10q11.21)	9.46×10^{-8}	4.34×10^{-4}	5.40×10^{-10}	7.65×10^{-5}

*The previous GWA studies refer to the Wellcome Trust Case Control Consortium (WTCCC) study and the German MI Family Study. The pooled results from these studies have been taken from Supplementary Reference 6. [#]P values for the previous GWA studies were Bonferroni corrected for the multiple (270,000) SNPs tested in these studies.

Supplementary Table 5 Likelihood ratio test for a non-additive genetic model at each locus

SNP	P-value
rs599839 (1p13.3)	0.344
rs3008621 (1q41)	0.859
rs2943634 (2q36.3)	0.151
rs6922269 (6q25.1)	0.339
rs1333049 (9p21)	0.407
rs501120 (10q11.21)*	0.070
rs17228212 (15q22.33)	0.783

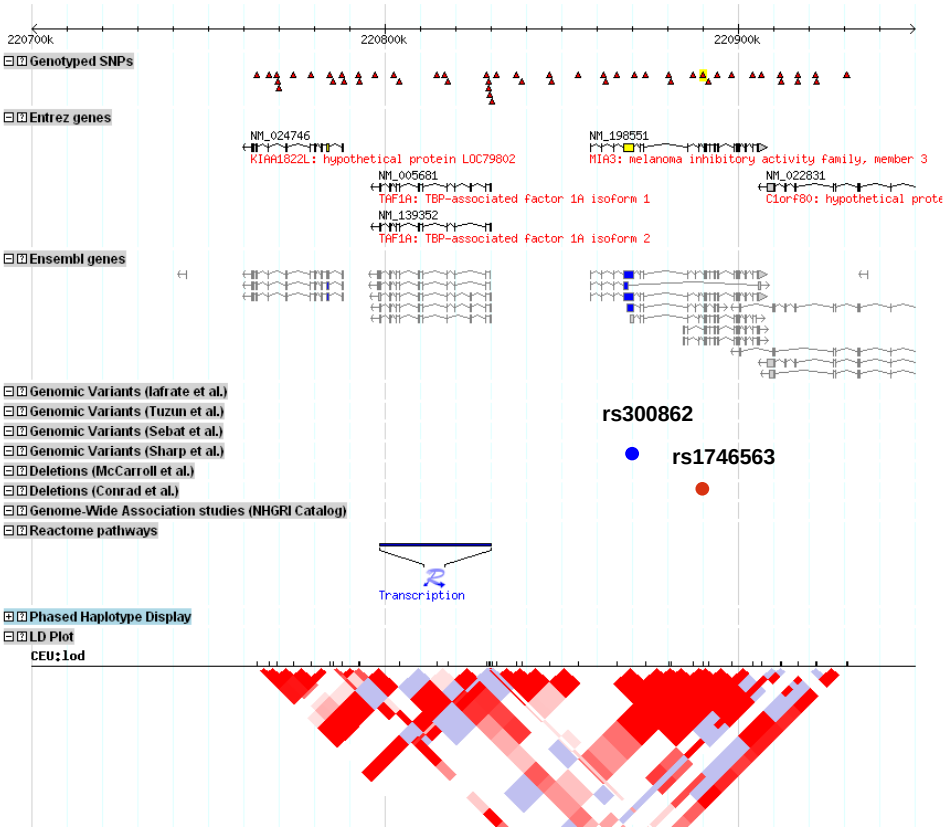
* P-values for males and females were 0.139 and 0.166 respectively

Supplementary Table 6 Probability (as percentage) that 0,1,…9 studies will be individually significant at the 5% level and the power of the pooled analysis for different combinations of per allele odds ratio (OR) and minor allele frequency (MAF)

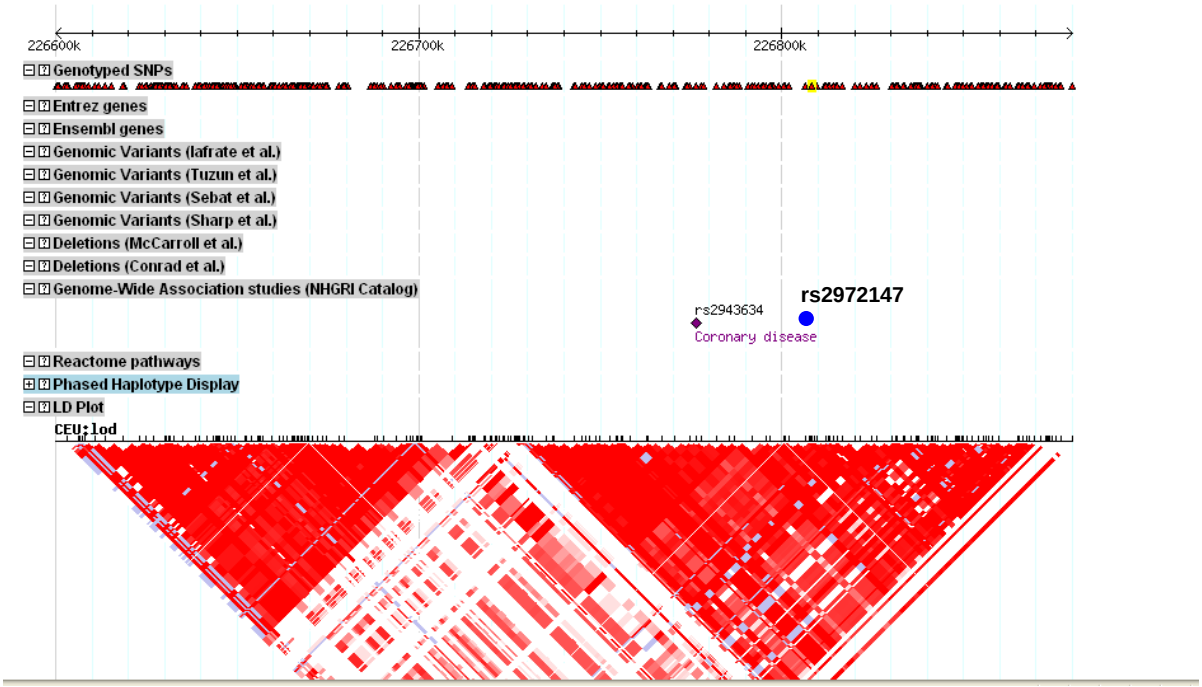
OR			**1.05**					**1.10**					**1.15**					**1.20**					**1.25**		
MAF	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**
0	50	41	34	29	27	19	5	2	1	1	3														
1	37	38	41	38	42	32	21	11	8	7	13	2				2									
2	11	16	18	25	22	31	29	22	20	20	25	8	2	2		8									
3	2	4	7	8	7	13	26	30	29	26	29	15	7	4	3	17	1				3				
4			1	1	2	5	14	20	23	23	20	27	16	15	10	25	6	2		1	11	1			
5						1	5	9	13	14	9	26	29	23	24	26	16	7	3	3	22	3			
6							1	4	5	7	2	15	26	26	30	16	27	16	11	11	29	7	3	1	1
7									1	2		6	15	23	21	5	28	31	26	28	22	25	14	7	7
8												1	5	7	9	1	16	30	40	35	11	41	37	34	35
9													1	1	2		6	14	19	23	2	23	46	59	57
Power	30	51	60	71	70	86	98	99	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100

Supplementary Figure 1

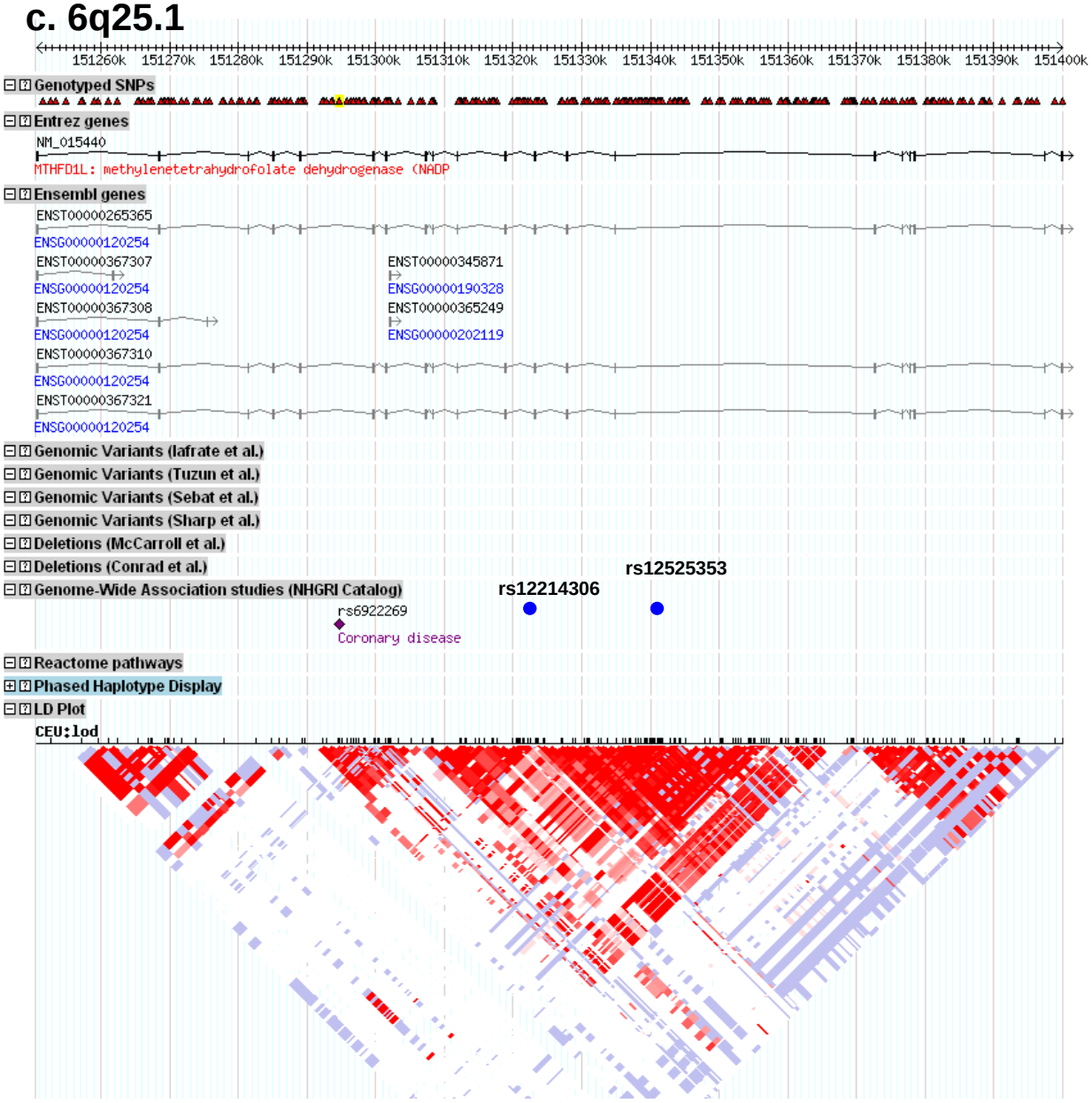
a. 1q41



b. 2q36.3



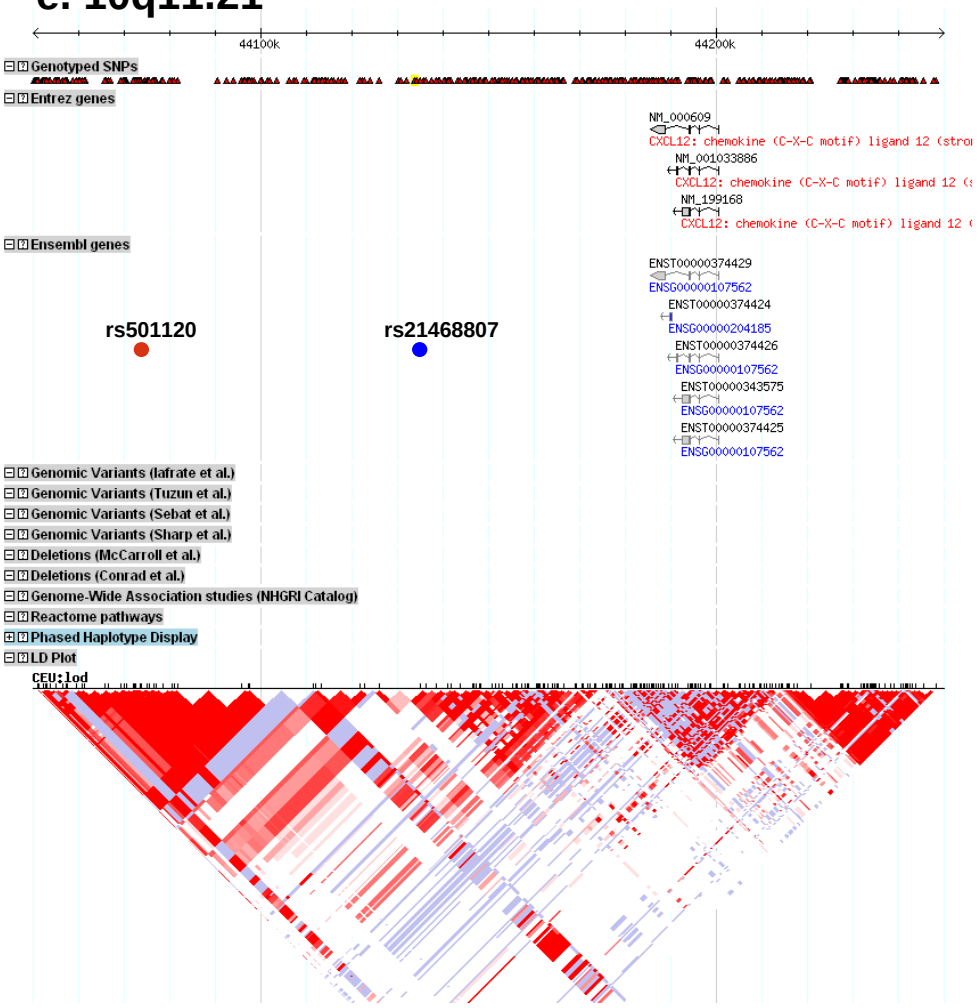
c. 6q25.1


151260k 151270k 151280k 151290k 151300k 151310k 151320k 151330k 151340k 151350k 151360k 151370k 151380k 151390k 151400k
Genotyped SNPs
Entrez genes
NM_015440
MTHFD1L: methylenetetrahydrofolate dehydrogenase (NADP
Ensembl genes
ENST00000265365
ENSG00000120254
ENST00000367307
ENSG00000120254
ENST00000345871
ENSG00000190328
ENST00000367308
ENSG00000120254
ENST00000365249
ENSG00000202119
ENST00000367310
ENSG00000120254
ENST00000367321
ENSG00000120254
Genomic Variants (Iafrate et al.)
Genomic Variants (Tuzun et al.)
Genomic Variants (Sebat et al.)
Genomic Variants (Sharp et al.)
Deletions (McCarroll et al.)
Deletions (Conrad et al.)
Genome-Wide Association studies (NHGRI Catalog)
rs12525353
rs12214306
rs6922269
Coronary disease
Reactome pathways
Phased Haplotype Display
LD Plot
CEU:lod

d. 9p21

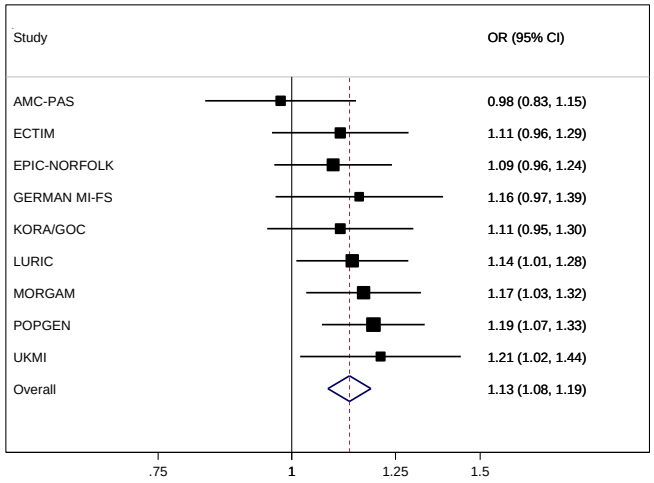

22000k 22010k 22020k 22030k 22040k 22050k 22060k 22070k 22080k 22090k 22100k 22110k 22120k 22130k 22140k 22150k 22160k
Genotyped SNPs
dbSNP SNPs
Entrez genes
Ensembl genes
ENST00000380165
ENSG00000099810
ENST00000380190
ENSG00000099810
Genomic Variants (Iafrate et al.)
Genomic Variants (Tuzun et al.)
Genomic Variants (Sebat et al.)
Genomic Variants (Sharp et al.)
Deletions (McCarroll et al.)
Deletions (Conrad et al.)
Genome-Wide Association studies (NHGRI Catalog)
rs10738610
rs10757278
Myocardial infarction
rs1333049
Coronary disease
rs10811661
Type 2 diabetes
rs10811661
Type 2 diabetes
rs10811661
Type 2 diabetes
Reactome pathways
Phased Haplotype Display
LD Plot
CEU:lod

e. 10q11.21

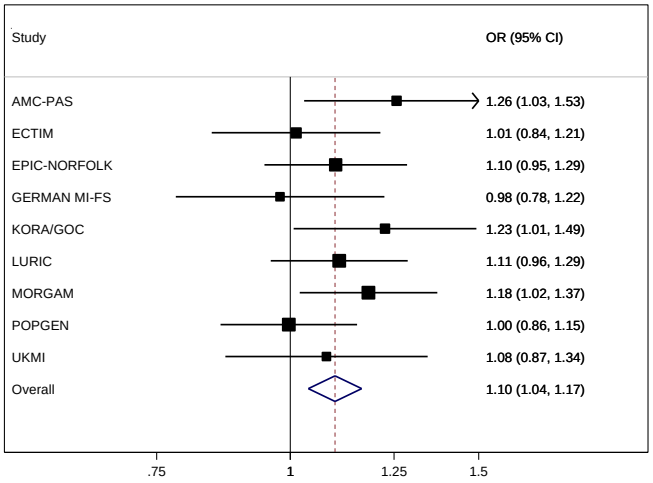

44100k
44200k
Genotyped SNPs
Entrez genes
NM_000609
CXCL12: chemokine (C-X-C motif) ligand 12 (stro
NM_001033886
CXCL12: chemokine (C-X-C motif) ligand 12 (
NM_199168
CXCL12: chemokine (C-X-C motif) ligand 12 (
Ensembl genes
ENST00000374429
ENSG00000107562
ENST00000374424
ENSG00000204185
ENST00000374426
ENSG00000107562
ENST00000343575
ENSG00000107562
ENST00000374425
ENSG00000107562
rs501120
rs21468807
Genomic Variants (Iafrate et al.)
Genomic Variants (Tuzun et al.)
Genomic Variants (Sebat et al.)
Genomic Variants (Sharp et al.)
Deletions (McCarroll et al.)
Deletions (Conrad et al.)
Genome-Wide Association studies (NHGRI Catalog)
Reactome pathways
Phased Haplotype Display
LD Plot
CEU:lod

Supplementary Figure 2

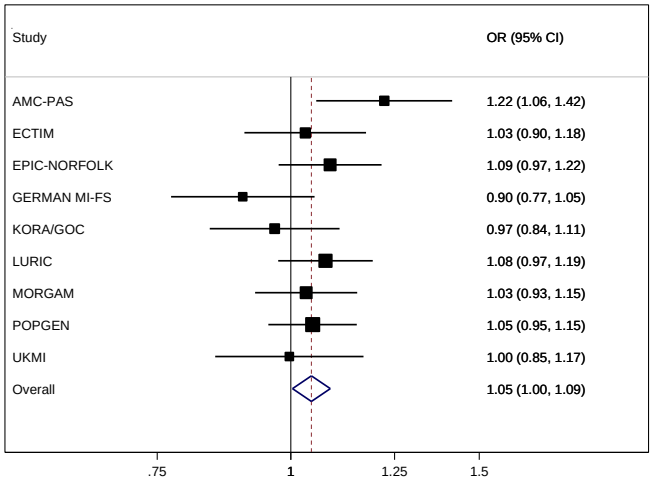
rs599839 on 1p13.3



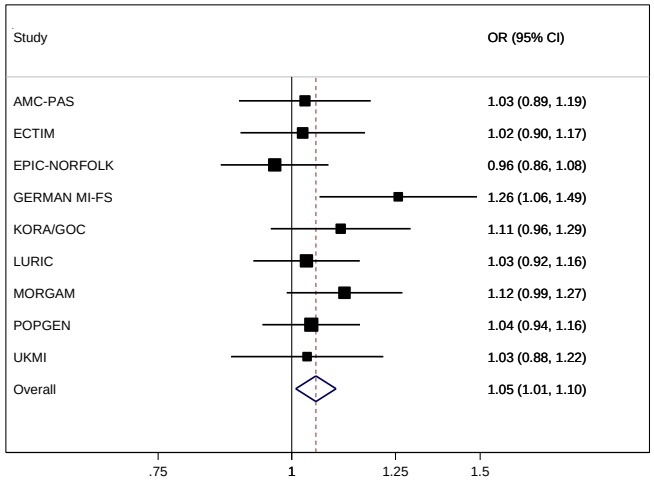
rs3008621on 1q41



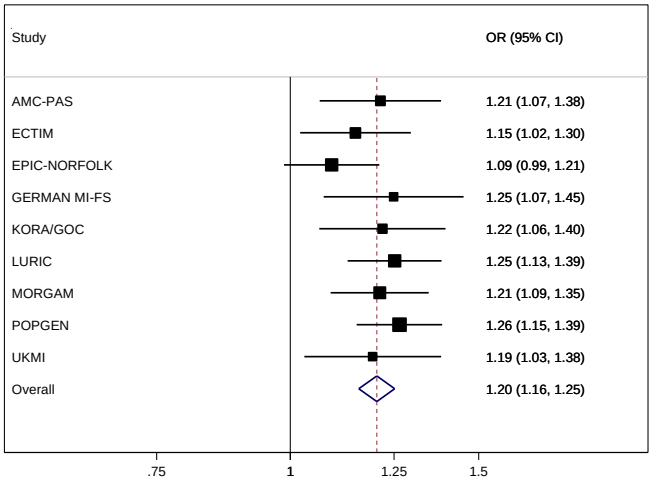
rs2943634 on 2q36.3



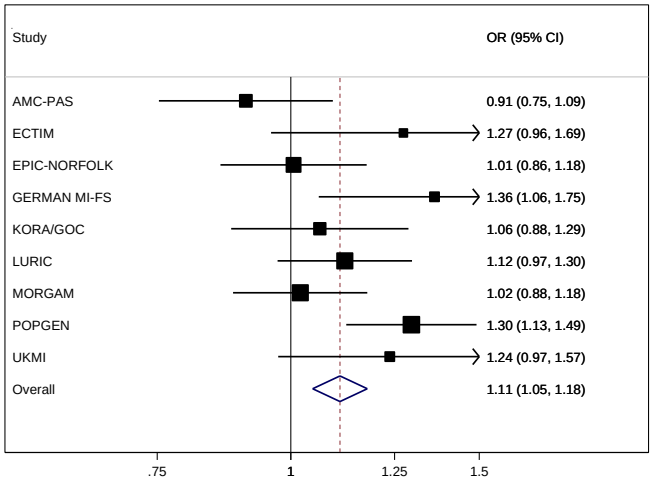
rs6922269 on 6q25.1



rs1333049 on 9p21



rs501120 on 10q11.21



Supplementary Figure 3