

Published in final edited form as:

Circulation. 2008 April 1; 117(13): 1675–1684. doi:10.1161/CIRCULATIONAHA.107.730614.

Repeated Replication and a Prospective Meta-Analysis of the Association Between Chromosome 9p21.3 and Coronary Artery Disease

Heribert Schunkert, MD, Anika Götz, MSc, Peter Braund, MSc, Ralph McGinnis, PhD, David-Alexandre Tregouet, PhD, Massimo Mangino, PhD, Patrick Linsel-Nitschke, MD, Francois Cambien, MD, Christian Hengstenberg, MD, Klaus Stark, PhD, Stefan Blankenberg, MD, Laurence Tiret, PhD, Pierre Ducimetiere, MD, Andrew Keniry, PhD, Mohammed J.R. Ghorri, PhD, Stefan Schreiber, MD, Nour Eddine El Mokhtari, MD, Alistair S. Hall, FRCP, Richard J. Dixon, PhD, Alison H. Goodall, PhD, Henrike Liptau, MD, Helen Pollard, MSc, Daniel F. Schwarz, MSc, Ludwig A. Hothorn, PhD, H.-Erich Wichmann, MD, Inke R. König, PhD, Marcus Fischer, MD, Christa Meisinger, MD, Willem Ouwehand, FRCPATH, Panos Deloukas, PhD, John R. Thompson, PhD, Jeanette Erdmann, PhD, Andreas Ziegler, PhD, and Nilesh J. Samani, FMedSci on behalf of the Cardiogenics Consortium

From Medizinische Klinik II (H.S., A.G., P.L.-N., H.L., J.E.) and Institut für Medizinische Biometrie und Statistik (A.G., D.F.S., I.R.K., A.Z.), Universität zu Lübeck, Lübeck, Germany; Department of Cardiovascular Sciences (P.B., M.M., R.J.D., A.H.G., H.P., N.J.S.) and Department of Health Sciences and Genetics (J.R.T.), Glenfield Hospital, University of Leicester, Leicester, UK; Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton (R.M., A.K., M.J.R.G., P.D.), and Department of Haematology, University of Cambridge and National Health Service Blood and Transplant, Cambridge Centre (W.O.), Cambridge, UK; INSERM and Université Pierre et Marie Curie-Paris 6, Paris, France (D.-A.T., F.C., L.T.); Klinik und Poliklinik für Innere Medizin II, Universität Regensburg, Regensburg, Germany (C.H., K.S., M.F.); II. Medizinische Klinik, Johannes Gutenberg University, Mainz, Germany (S.B.); Unité INSERM 780–Université Paris Sud XI, Villejuif, France (P.D.); Institut für Klinische Molekularbiologie, UK-SH, Campus Kiel, Kiel, Germany (S.S., N.E.E.M.); C-NET Group, Leeds Institute for Genetics and Therapeutics, Faculty of Medicine and Health, University of Leeds, Leeds, UK (A.S.H.); Institut für Biostatistik, Leibniz Universität Hannover, Hannover, Germany (L.A.H.); Institut für Epidemiologie, Helmholtz Zentrum München, Deutsches Forschungszentrum für Umwelt und Gesundheit, Neuherberg, Germany (H.-E.W., C.M.); Institute of Medical Informatics, Biometry, and Epidemiology, Ludwig Maximilians University, Munich, Germany (H.-E.W.).

Abstract

Background—Recently, genome-wide association studies identified variants on chromosome 9p21.3 as affecting the risk of coronary artery disease (CAD). We investigated the association of this locus with CAD in 7 case-control studies and undertook a meta-analysis.

Methods and Results—A single-nucleotide polymorphism (SNP), rs1333049, representing the 9p21.3 locus, was genotyped in 7 case-control studies involving a total of 4645 patients with

© 2008 American Heart Association, Inc.

Reprint requests to Dr Heribert Schunkert, Medizinische Klinik II, Universität zu Lübeck, Lübeck, Germany (e-mail heribert.schunkert@innere2.uniluebeck.de); or Dr Nilesh Samani, Department of Cardiovascular Sciences, University of Leicester, Clinical Sciences Wing, Glenfield Hospital, Groby Rd, Leicester, LE3 9QP, UK (e-mail njs@le.ac.uk).

Disclosures
None.

myocardial infarction or CAD and 5177 controls. The mode of inheritance was determined. In addition, in 5 of the 7 studies, we genotyped 3 additional SNPs to assess a risk-associated haplotype (ACAC). Finally, a meta-analysis of the present data and previously published samples was conducted. A limited fine mapping of the locus was performed. The risk allele (C) of the lead SNP, rs1333049, was uniformly associated with CAD in each study ($P < 0.05$). In a pooled analysis, the odds ratio per copy of the risk allele was 1.29 (95% confidence interval, 1.22 to 1.37; $P = 0.0001$). Haplotype analysis further suggested that this effect was not homogeneous across the haplotypic background (test for interaction, $P = 0.0079$). An autosomal-additive mode of inheritance best explained the underlying association. The meta-analysis of the rs1333049 SNP in 12 004 cases and 28 949 controls increased the overall level of evidence for association with CAD to $P = 6.04 \times 10^{-10}$ (odds ratio, 1.24; 95% confidence interval, 1.20 to 1.29). Genotyping of 31 additional SNPs in the region identified several with a highly significant association with CAD, but none had predictive information beyond that of the rs1333049 SNP.

Conclusion—This broad replication provides unprecedented evidence for association between genetic variants at chromosome 9p21.3 and risk of CAD.

Keywords

chromosomes; coronary disease; genetics; meta-analysis; myocardial infarction; risk factors

Coronary artery disease (CAD) and its main complication, myocardial infarction (MI), have a strong genetic basis.¹ Hitherto, despite almost 2 decades of intensive research, the molecular basis underlying the inherited component of CAD and MI remained largely unexplained. Numerous candidate genes have been implicated, but few, if any, displayed a reproducible association between risk alleles and CAD or MI in replication studies.² These rather disappointing results have called into question the credibility of candidate gene analysis for clinical risk prediction of CAD and MI.

The development of high-density genotyping arrays now provides improved resolution for an unbiased genome-wide assessment of common variants associated with common diseases. Using this technology, several recent studies revealed a strong association between a chromosomal locus on chromosome 9p21.3 and CAD and MI.³⁻⁶ To further elucidate the role of this chromosomal locus in the molecular genetics of CAD, we investigated variants at this locus in 7 additional European studies and undertook a meta-analysis of all studies published to date.

Methods

Study Populations

We studied subjects collected in 7 different studies across Europe (the German Myocardial Infarction Family Study [GerMIFS] II,^{7,8} the United Kingdom Myocardial Infarction [UK MI] study,^{9,10} AtheroGene,¹¹ the Left Main Disease study, the MONICA/KORA study,^{12,13} PopGen,¹⁴ and the Prospective Epidemiological Study of Myocardial Infarction [PRIME]¹⁵). All subjects were of northern European origin. The large majority of cases had MI as defined by World Health Organization criteria ($n = 3544$). The remaining cases had evidence of CAD based on either a revascularization procedure or anginal symptoms with a positive stress test ($n = 1101$). Details of each study are given in Table 1 and the online-only Data Supplement.

Genotyping

In the GerMIFS II, the UK MI study, AtheroGene, the MONICA/KORA study, PopGen, and PRIME samples, genotyping was performed with TaqMan technology (Applied

Biosystems, Darmstadt, Germany). All single-nucleotide polymorphisms (SNPs) constituting the ACAC haplotype (rs7044859, rs1292136, rs7865618, and rs1333049) were genotyped in these studies except the AtheroGene and PRIME samples, in which only the lead SNP, rs1333049, was assessed. TaqMan genotyping assays with probes labeled with the fluorophores FAM and VIC were purchased from Applied Biosystems. Genotyping was performed on 384-well plates prepared with pipetting robots. The Universal PCR Master Mix from Applied Biosystems was used in a 5- μ L total reaction volume with 10 ng DNA per reaction. Allelic discrimination was measured automatically on the ABI Prism 7900HT (Applied Biosystems) with Sequence Detection Systems 2.1 software (auto caller confidence level, 95%).

The genotypes of the Left Main Disease study samples were extracted from the GeneChip 5.0 Mapping Array from Affymetrix (Santa Clara, Calif) after extensive quality-control procedures (see Samani et al³ for details). The concordance between genotypes generated with the TaqMan and the data from the 500K Array was >98% as assessed in 1000 individuals studied with both techniques.

Single-Study Analysis

For each study group, we investigated deviation from Hardy-Weinberg equilibrium by a permutation-based goodness-of-fit test with 20 000 replicates.¹⁶ Differences in genotype frequencies of rs1333049 were compared between cases and controls for every study by use of the 2-sided asymptotic Cochran-Armitage trend test. Heterozygous and homozygous odds ratios (ORs) for the risk allele and corresponding 95% confidence intervals (CIs) were estimated.

Prospectively Planned Pooled Analysis

The present study was planned as a type IV meta-analysis according to Blettner et al.¹⁷ Three different approaches were used for the prospectively planned pooled analysis of the 7 case-control studies; they are described in detail in the online-only Data Supplement.

First, we investigated the additive effect of rs1333049 on CAD/MI by random-effect (RE) logistic regression models with adjustments for study. The logistic regression framework also was used to identify the underlying genetic model.¹⁸ Fixed-effect (FE) logistic regression models were used as the sensitivity analysis. *P* values reported here are derived from the more conservative RE models unless indicated otherwise (eg, Table 2). Second, we followed the procedure of Minelli et al¹⁹ and estimated the ratio (λ) of the log OR for the heterozygous individuals compared with the homozygous individuals to determine the most likely mode of inheritance. Third, we used a modification of the MAX test approach of Freidlin et al.²⁰ This novel method allows derivation of global and multiplicity-adjusted *P* values and a fairly accurate model selection when the mode of inheritance is unknown.

Subgroup Analyses

RE logistic regression models with adjustments for study were evaluated for CAD cases without MI and for CAD cases with MI. Furthermore, for all 7 studies, we applied RE logistic regression models with adjustments for study of available cardiovascular risk factors and tested for an interaction between rs1333049 and each risk factor separately. Values of interaction tests of *P*<0.05 were considered significant. We examined gender, age (>55 years), obesity (body mass index >30 kg/m²), smoking (ever), hypertension (systolic blood pressure >140 mm Hg, diastolic blood pressure >90 mm Hg, or receiving treatment for these conditions), and hyperlipidemia (total cholesterol >200 mg/dL, low-density lipoprotein cholesterol >130 mg/dL, or receiving lipid-lowering treatment). ORs and corresponding 95% CIs and *P* values are shown. In addition, we performed in the control group linear

regression models with FE for rs1333049 and RE for study of some additional continuous risk factors: systolic blood pressure, diastolic blood pressure, total cholesterol, high-density lipoprotein cholesterol, and low-density lipoprotein cholesterol.

Combination of Cardiovascular Risk Factors

For all studies, a multiple RE logistic regression was estimated for rs1333049 with adjustments for the dichotomous risk factors given above. In addition, we estimated the population-attributable fractions according to rs1333049.²¹ Adjustments of population-attributable fractions were made on these cardiovascular risk factors.

Haplotype Analysis

Haplotype association analysis was performed with the THESIAS software implementing the Stochastic-EM algorithm.^{22,23} Using this approach, we were able to simultaneously estimate haplotype frequencies and haplotype effects on CAD under the assumption of additive haplotype effects of haplotypes on CAD. A global test of association between CAD and the haplotypes derived from rs7044859, rs1292136, rs7865618, and rs1333049 was first performed using the likelihood ratio test. The likelihood ratio test statistic also was used for testing the frequency difference of a specific haplotype between cases and controls and for investigating the heterogeneity of the effect of a single SNP according to the different haplotype backgrounds on which it can be carried out.

Meta-Analysis

The same approach used for the prospectively planned pooled analysis was used for a type III meta-analysis according to Blettner et al.¹⁷ Therefore, all previously reported replication studies (as of November 15, 2007) in individuals of European descent, including the 7 novel case-control studies studied here and the replication studies used in previously reported genome-wide association studies (GerMIFS I, ie, replication study used in Samani et al³; OHS-2, ARIC, CCHS, DHS, and OHS-3, ie, replication studies used in McPherson et al⁵; and Iceland B, Atlanta, Philadelphia, and Durham, ie, replication studies used in Helgadottir et al⁶), were analyzed. In some of the 12 004 cases and 28 949 controls, the genotypes of rs1333049 were not available. Here, we used reported genotype frequencies for the marker with highest linkage disequilibrium (LD; $r^2=0.87$ for rs238206 [used in studies from McPherson et al⁵]; $r^2=1$ for rs10757278 [used in studies from Helgadottir et al⁶]; Figure 1).

Fine Mapping

We examined HapMap SNPs at 9p21.3 (between 21 934 317 and 22 174 018 bp), the region delimited by significant pairwise LD ($D'>0$, $LOD \geq 2$) with SNPs showing CAD association ($P<10^{-3}$) in the Wellcome Trust Case Control Consortium (WTCCC) scan. We selected a total of 31 SNPs (online-only Data Supplement Table I) for fine-map genotyping if a) they had higher r^2 values with regional SNPs showing CAD association ($P<10^{-3}$) than nonassociation and b) they were not covered at $r^2>0.9$ by another SNP already genotyped or slated for fine-map genotyping. SNPs were genotyped by Sequenom assay in 1137 WTCCC CAD cases and 1143 controls.⁴ SNPs with P values stronger than that of our lead SNP in this initial analysis were further evaluated in the GerMIFS II and UK MI.

The authors had full access to and take full responsibility for the integrity of the data. All authors have read and agree to the manuscript as written.

Results

The recruitment strategy, main phenotypes, and source of cases and controls in each study, as well as the main demographical characteristics for all case-control samples, are shown in Table 1.

The chromosomal structure at the chromosome 9p21.3 locus as outlined by Samani et al³ on the basis of HapMap data²⁴ is shown in Figure 1. The lead SNP, rs1333049, identified by Samani et al³ and further evaluated in this study, is located in a haplotype block with high LD ($D' > 0.87$ and $r^2 > 0.55$). The same block harbors multiple other SNPs associated with CAD,^{3,4} including those identified by McPherson et al⁵ (rs10757274 and rs2383206) and Helgadottir et al⁶ (rs10116277, rs1333040, rs2383207, and rs10757278). In moderate LD (average $|D'| \approx 0.60$) with this main block is a neighboring LD block that harbors further SNPs associated with CAD.³ The SNPs in this second block generate 5 haplotypes with a frequency $> 2\%$ that explain $\approx 97\%$ of its haplotypic diversity.³ Three SNPs (rs7044859, rs1292136, and rs7865618) are sufficient to characterize the main haplotypes of this second block.³ Together with the lead SNP from the first block (rs1333049), these 3 SNPs constitute the ACAC haplotype (Figure 1). As a consequence of this chromosomal structure, we performed association analysis with CAD in relation to the lead SNP and the risk haplotype using the 4 tagging SNPs.

Association With the Lead SNP, rs1333049

In total, we studied 4645 subjects with CAD and 5177 controls. Of these individuals, 1101 (23.7%) had CAD without a history of MI, and 3544 (76.3%) had validated MI. No deviation from Hardy-Weinberg equilibrium was detected in any of the study groups.

The lead SNP, rs1333049, was significantly associated with CAD in each of the case-control samples (Table 2). Allele frequencies in the control groups were homogeneous ($P=0.20$). In the pooled analysis, the results for FE and RE logistic regression models with adjustments for study showed similar effects of rs1333049 on CAD. The OR per 1 copy of the risk allele for rs1333049 for RE logistic regression models was 1.29 (95% CI, 1.22 to 1.37; $P=0.0001$); the OR was virtually identical for the FE logistic regression model (OR, 1.29; 95% CI, 1.22 to 1.37; $P=1.17 \times 10^{-17}$).

Cardiovascular Risk Factors

The lead SNP (rs1333049) also was similarly associated with CAD when subjects with a history of MI were excluded from the analysis (OR, 1.38; 95% CI, 1.24 to 1.54; $P=0.0095$; Table 2). The same was true when only subjects with a history of MI were considered in the association analysis (OR, 1.27; 95% CI, 1.19 to 1.35; $P=0.0003$; Table 2).

The large sample size allowed investigation of further subgroups based on the underlying risk profile. Although the study had enough power ($> 90\%$) in most instances (online-only Data Supplement Table II), we observed no interaction on the risk of CAD between the lead SNP, rs1333049, and available traditional risk factors, including gender ($P=0.67$), age > 55 years ($P=0.07$), body mass index > 30 kg/m² ($P=0.29$), smoking ($P=0.75$), hypertension ($P=0.55$), and hyperlipidemia ($P=0.77$; Table 3). Indeed, in every subgroup analyzed, ORs were rather similar and statistically significant ($P < 0.05$; Table 3). Furthermore, no interaction was found between the number of traditional cardiovascular risk factors and rs1333049 ($P=0.54$).

We also analyzed the association of rs1333049 with several quantitative cardiovascular risk factors (systolic and diastolic blood pressures, total cholesterol, and low- and high-density lipoprotein cholesterol). This analysis was restricted to controls to minimize the impact of

treatment or disease. rs1333049 was not associated with any of these risk factors (online-only Data Supplement Table III).

After adjustment for these covariates, the OR for the association of rs1333049 with disease remained unchanged (OR per 1 risk allele, 1.32; 95% CI, 1.24 to 1.42; $P=0.0004$). As estimated from this adjusted analysis, the population-attributable fraction of the lead SNP, rs1333049, for CAD was 22% (95% CI, 16 to 28).

Genetic Model

The logistic regression framework to identify the underlying genetic model rejected an autosomal-dominant and an autosomal-recessive mode of inheritance (each $P<0.0001$). The most likely model is an additive genetic model because it is the only genetic model not rejected at the 5% test level ($P=0.38$). Whereas the results using the Minelli et al19 approach supported the additive and the dominant genetic model at the 5% test level, the approach of Hothorn strongly supports the additive genetic model (online-only Data Supplement Table IV).

Haplotype Association Analysis

Eight haplotypes with a frequency $>2\%$ could be inferred from the 4 SNPs studied (supplemental Table V). Their haplotype frequency distribution was highly significantly different between cases and controls ($P=3.38\times 10^{-16}$). In particular, the ACAC haplotype composed of the rs7044859-A, rs1292136-C, rs7865618-A, and rs1333049-C alleles was more frequent in cases than in controls and was associated in the pooled sample with an increased risk of 1.23 (95% CI, 1.15 to 1.33; $P<10^{-7}$) compared with the second most frequent TTGG haplotype (Figure 2). In a further analysis, we tested the effect of the rs1333049-C allele on the 3 TTA-, TTG-, and ACA- backgrounds separately (Table 4). For a given haplotypic background, the effect of the C allele was very homogeneous across the 5 replication samples. Whereas the C allele was not associated with CAD when it was carried by either the TTA- (OR, 0.99; 95% CI, 0.76 to 1.30; $P=0.94$) or TTG- (OR, 1.12; 95% CI, 0.96 to 1.31; $P=0.14$) background, it was strongly associated with an increased risk of CAD (OR, 1.44; 95% CI, 1.27 to 1.62; $P=4.08\times 10^{-9}$) when it was carried by the ACA- haplotype. Because the test for interaction between these 3 ORs was significant ($P=0.0079$), we have an indication for a non homogeneous effect of the rs1333049 SNP according to its haplotypic backgrounds, suggesting that this SNP alone may not be sufficient to fully explain the association with CAD.

Meta-Analysis

In total, the meta-analysis comprised the aforementioned 7 case-control studies and the replication samples of previously published reports (Figure 3; Samani et al3, McPherson et al5, and Helgadottir et al6). We analyzed rs1333049, rs2383206, and rs107572786 were 12 004 cases and 28 949 controls. The overall level of evidence for association of the respective lead SNP at the chromosome 9p21.3 locus with CAD increased to $P=6.04\times 10^{-10}$ with an OR per 1 risk allele of 1.24 (95% CI, 1.20 to 1.29) in the RE model and $P=6.36\times 10^{-39}$ with an OR per 1 risk allele of 1.24 (95% CI, 1.20 to 1.29) in the FE model. Including additional association data from the initial genome-wide association studies3-6 increased the P value to $P=1.62\times 10^{-12}$ (OR, 1.27; 95% CI, 1.23 to 1.31) and $P=9.82\times 10^{-60}$ (OR, 1.27; 95% CI, 1.24 to 1.31) in the RE and FE models, respectively.

Fine Mapping

Genotyping of 31 additional SNPs in a sample of 1137 WTCCC cases and 1143 controls revealed that 2 SNPs, rs1537378 and rs10738610, had P values for CAD that were

marginally stronger than that of the lead SNP (rs1333049, $P=1.02\times 10^{-7}$; rs1537378, $P=8.47\cdot 10^{-9}$; rs10738610, $P=1.66\cdot 10^{-8}$; online-only Data Supplement Table I). We further studied these 2 SNPs in 1535 cases and 1983 controls (GerMIFS and UK MI samples); in these replication samples, however, neither SNP provided statistically significant prediction of CAD beyond that of the lead SNP, rs1333049 (GerMIFS: rs1537378, $P=0.17$; rs10738610, $P=0.23$; UK MI samples: rs1537378, $P=0.62$; rs10738610, $P=0.57$).

Discussion

This analysis of several distinct case-control studies provides evidence for consistent associations between a locus on chromosome 9p21.3 and CAD and MI. The magnitude and direction of the effect in each of the present study samples are consistent with the 4 previous genome-wide analyses that uniformly identified this locus as the strongest genetic signal for CAD.³⁻⁶ Together, the comprehensive replication across multiple samples evaluated in the present studies provides unequivocal evidence that variants at this locus increase the risk of CAD and MI in individuals of European ancestry.

We observed associations of the chromosome 9p21.2 locus with CAD in both cross-sectional case-control studies and subjects who developed the disease in the prospective PRIME study. Likewise, we found similar associations in men and women and in various subgroups defined by age or other cardiovascular risk factors. We also confirmed that the associations are independent of these risk factors.

The striking consistency of the association across this broad range of subjects, coupled with the high frequency of the risk allele and the substantial increase in the probability to develop the disease by each allele, suggests that genotyping for this locus could have important clinical utility in risk prediction in the future if the findings are supported in further large population-based studies. However, it also is evident from the large number of SNPs in this region with highly significant association with CAD that the causally responsible variant and the related mechanism remain to be identified. The high LD within the region is the most likely explanation of this finding. The region on chromosome 9p21.3 associated with CAD is located in 2 blocks of strong intrablock and interblock LD.^{3,4} We have previously shown that the association signal was stronger for a haplotype (ACAC) derived from 4 SNPs (the last one being the rs1333049 SNP) across the 2 blocks. Here, we confirm this finding by showing that the effect of the lead SNP, rs1333049, was restricted mainly to the ACA-haplotypic background. This result suggests that the lead SNP, rs1333049, alone does not fully explain the association with CAD. Deep resequencing of cases carrying this haplotype is required to obtain the full spectrum of variation in this region and to identify the causal variant(s) that affect risk.

If the effect of the locus is due largely to a single variant, we also show evidence suggesting that the most likely mode of inheritance is additive. This information may help to define the role of chromosome 9p21.3 variants in algorithms for cardiovascular risk prediction.

Beyond the 7 case-control samples, we undertook a meta-analysis of all previously published replication data. The findings from this analyses based on 12 004 cases and 28 949 controls confirm the strength of the association of the locus with the risk of CAD and provide reliable and robust estimates of the per-allele risk associated with the locus. These findings provide a firm rationale for further interrogation of the locus.

The validity of our findings is further strengthened by virtually identical estimates of the ORs and the 95% CIs from the FE and RE models (Table 2). The FE model provides more impressive P values because of the crucial assumption that the ORs of all the different studies are identical. In contrast to FE models, RE models allow for between-study

variabilities resulting, for example, from differences in ascertainment schemes or inclusion and exclusion criteria. This statistic is more conservative, and *P* values from RE models are therefore substantially larger than *P* values from FE models (Table 2) or the originally combined *P* values reported by Samani et al.³

Although this broad replication offers unprecedented information for a better understanding of the molecular genetic architecture of CAD, the underlying mechanism is as yet elusive. The region is defined by 2 flanking recombination hotspots and contains the coding sequences of genes for 2 cyclin-dependent kinase inhibitors, *CDKN2A* (encoding p16INK4a) and *CDKN2B* (encoding p15INK4b). These genes are known to play an important role in the regulation of the cell cycle and belong to a family of genes that have been implicated in the pathogenesis of atherosclerosis through their role in transforming growth factor- β -induced growth inhibition. However, the most strongly associated SNPs lie considerably upstream of these genes, and the nearest signal is 10 kb upstream of *CDKN2B*. Although an effect through regulation of one or both of the cyclin genes is possible, other explanations need to be considered, including the involvement of the *MTAP* gene or expression sequence tags located in the region.

Our studies, like the previous studies, have analyzed primarily subjects of north European origin.²³ In fact, data from the HapMap consortium suggest that the frequency of the rs1333049-C allele differs from 0.17 in Yoruba to 0.48 in Han Chinese to 0.51 in Japanese to 0.49 in Europeans.²⁵ Thus, the risk of CAD and MI related to the chromosome 9p21.3 locus may vary among different ethnic groups. Indeed, the role of this locus in other ethnic groups remains to be investigated.

The present study is focused on large case-control samples with CAD and MI. Although these related pheno-types display similar patterns of association, further studies are needed in a wider scope of atherosclerotic complications to fully understand the pathophysiological role of the chromosome 9p21.3 locus. Moreover, larger prospective studies are needed to precisely define the incremental information obtained when risk variants at chromosome 9p21.3 are included in algorithms predicting CAD events such as the Framingham or the Prospective Cardiovascular Münster (PROCAM) study risk scores.

Conclusions

Human chromosome 9p21.3 harbors a locus with a common variant that significantly increases the risk of CAD. This finding opens up the possibility of better gene-based prediction and potentially a more targeted preventive treatment for this common condition.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Petra Bruse, Kristin Blankenberg, Viviane Nicaud, and Maylis de Suremain for assistance. We thank the investigators and participants in each of the studies for allowing us to carry out this project.

Sources of Funding

This study was supported by the Cardiogenics Integrated Project (LSH-2006-037593) of the European Union. Dr Samani holds a chair supported by the British Heart Foundation.

References

1. Marenberg ME, Risch N, Berkman LF, Floderus B, de Faire U. Genetic susceptibility to death from coronary heart disease in a study of twins. *N Engl J Med*. 1994; 330:1041–1046. [PubMed: 8127331]
2. Morgan TM, Krumholz HM, Lifton RP, Spertus JA. Nonvalidation of reported genetic risk factors for acute coronary syndrome in a large-scale replication study. *JAMA*. 2007; 297:1551–1561. [PubMed: 17426274]
3. Samani NJ, Erdmann J, Hall AS, Hengstenberg C, Mangino M, Mayer B, Dixon RJ, Metinger T, Braund P, Wichmann H-E, Barrett J, König IR, Stevens S, Szymczak S, Tregouet D-A, Iles MM, Pahlke F, Pollard H, Lieb M, Cambien F, Fischer M, Ouwehand W, Blankenberg S, Balmforth A, Baessler A, Ball S, Strom TM, Braenne I, Gieger C, Deloukas P, Tobin M, Ziegler A, Thompson J, Schunkert H, the WTCCC and Cardiogenics Consortium. Genomewide Association Analysis of Coronary Artery Disease. Genomewide association analysis of coronary artery disease. *N Engl J Med*. 2007; 357:497–499. [PubMed: 17634453]
4. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*. 2007; 447:661–678. [PubMed: 17554300]
5. 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. [PubMed: 17478681]
6. Helgadóttir A, Thorleifsson G, Manolescu A, Gretarsdóttir S, Blondal TA. Common variant on chromosome 9p21 affects the risk of myocardial infarction. *Science*. 2007; 316:1491–1493. [PubMed: 17478679]
7. Broeckel U, Hengstenberg C, Mayer B, Holmer S, Martin LJ, Comuzzie AG, Blangero J, Nürnberg P, Reis A, Riegger GA, Jacob HJ, Schunkert H. A comprehensive linkage analysis for myocardial infarction and its related risk factors. *Nat Genet*. 2002; 30:210–214. [PubMed: 11818963]
8. Fischer M, Broeckel U, Holmer S, Baessler A, Hengstenberg C, Mayer B, Erdmann J, Klein G, Riegger G, Jacob HJ, Schunkert H. Distinct heritable patterns of angiographic coronary artery disease in families with myocardial infarction. *Circulation*. 2005; 111:855–862. [PubMed: 15710764]
9. 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; 79:980–984. [PubMed: 9609233]
10. Brouillette S, Singh RK, Thompson JR, Goodall AH, Samani NJ. White cell telomere length and risk of premature myocardial infarction. *Arterioscler Thromb Vasc Biol*. 2003; 23:842–846. [PubMed: 12649083]
11. Blankenberg S, Godefroy T, Poirier O, Rupprecht HJ, Barbaux S, Bickel C, Nicaud V, Schnabel R, Kee F, Morrison C, Evans A, Lackner KJ, Cambien F, Münzel T, Tiret L, the AtheroGene Investigators. Haplotypes of the caspase-1 gene, plasma caspase-1 levels, and cardiovascular risk. *Circ Res*. 2006; 99:102–108. [PubMed: 16778130]
12. Lowel H, Meisinger C, Heier M, Hormann A. The population-based acute myocardial infarction (ami) registry of the MONICA/KORA Study region of Augsburg. *Gesundheitswesen*. 2005; 67:S31–S37. [PubMed: 16032515]
13. Holmer SR, Hengstenberg C, Kraft HG, Mayer B, Pöll M, Kürzinger S, Fischer M, Löwel H, Klein G, Riegger GA, Schunkert H. Association of polymorphisms of the apolipoprotein(a) gene with lipoprotein(a) levels and myocardial infarction. *Circulation*. 2003; 107:696–701. [PubMed: 12578871]
14. Krawczak M, Nikolaus S, von Eberstein H, Croucher P, El Mokhtari N, Schreiber S. PopGen: population-based recruitment of patients and controls for the analysis of complex genotype-phenotype relationships. *Community Genet*. 2006; 9:55–66. [PubMed: 16490960]
15. Ducimetière P, Ruidavets JB, Montaye M, Haas B, Yarnell J, the PRIME Study Group. Five-year incidence of angina pectoris and other forms of coronary heart disease in healthy men aged 50–59 in France and Northern Ireland: the Prospective Epidemiological Study of Myocardial Infarction (PRIME) Study. *Int J Epidemiol*. 2001; 30:1057–1062. [PubMed: 11689522]

16. Ziegler, A.; König, IR. A Statistical Approach to Genetic Epidemiology: Concepts and Applications. Wiley-VCH; Weinheim, Germany: 2006.
17. Blettner M, Sauerbrei W, Schlehofer B, Scheuchpflug T, Friedenreich C. Traditional reviews, meta-analyses and pooled analyses in epidemiology. *Int J Epidemiol.* 1999; 28:1–9. [PubMed: 10195657]
18. Bagos PG, Nikolopoulos GK. A method for meta-analysis of case-control genetic association studies using logistic regression. *Stat Appl Genet Mol Biol.* 2007; 6 Article 17.
19. Minelli C, Thompson JR, Abrams KR, Thakkinstian A, Attia J. The choice of a genetic model in the meta-analysis of molecular association studies. *Int J Epidemiol.* 2005; 34:1319–1328. [PubMed: 16115824]
20. Freidlin B, Zheng G, Li ZH, Gieger C. Trend tests for case-control studies of genetic markers: power, sample size and robustness. *Hum Hered.* 2002; 53:146–152. [PubMed: 12145550]
21. Brady AR. Adjusted population attributable fractions from logistic regression. *Stat Tech Bull.* 1998; 42:8–12.
22. Tregouet DA, Garelle V. A new JAVA interface implementation of THESIAS: testing haplotype effects in association studies. *Bioinformatics.* 2007; 23:1038–1039. [PubMed: 17308338]
23. Tregouet DA, Barbaux S, Escolano S, Tahri N, Golmard JL, Tiret L, Cambien F. Specific haplotypes of the P-selection gene are associated with myocardial infarction. *Hum Mol Genet.* 2002; 11:2015–2023. [PubMed: 12165563]
24. International HapMap Consortium. A haplotype map of the human genome. *Nature.* 2005; 437:1299–1320. [PubMed: 16255080]
25. NCI-NHGRI Working Group on Replication in Association Studies. Chanock SJ, Manolio T, Boehnke M, Boerwinkle E, Hunter DJ, Thomas G, Hirschhorn JN, Abecasis G, Altshuler D, Bailey-Wilson JE, Brooks LD, Cardon LR, Daly M, Donnelly P, Fraumeni JF, Freimer NB, Gerhard DS, Gunter C, Guttmacher AE, Guyer MS, Harris EL, Hoh J, Hoover R, Kong CA, Merikangas KR, Morton CC, Palmer LJ, Phimister EG, Rice JP, Roberts J, Rotimi C, Tucker MA, Vogan KJ, Wacholder S, Wijsman EM, Winn DM, Collins FS. Replicating genotype-phenotype associations. *Nature.* 2007; 447:655–660. [PubMed: 17554299]

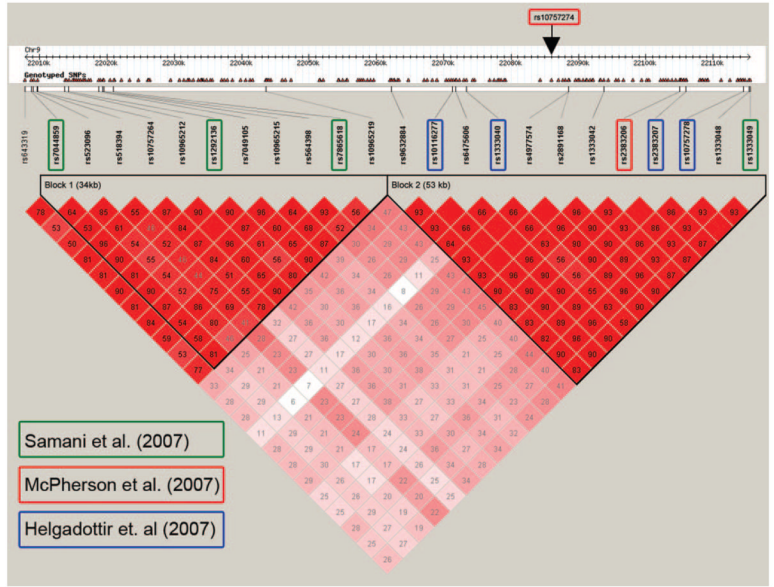


Figure 1. HapMap data (r^2)₂₂ of all SNPs of the chromosome 9 region (listed in online-only Data Supplement Table I) of Samani et al.³ and the lead SNPs of McPherson et al.⁵ and Helgadottir et al.⁶ The lead SNPs of former studies and the present study are marked.

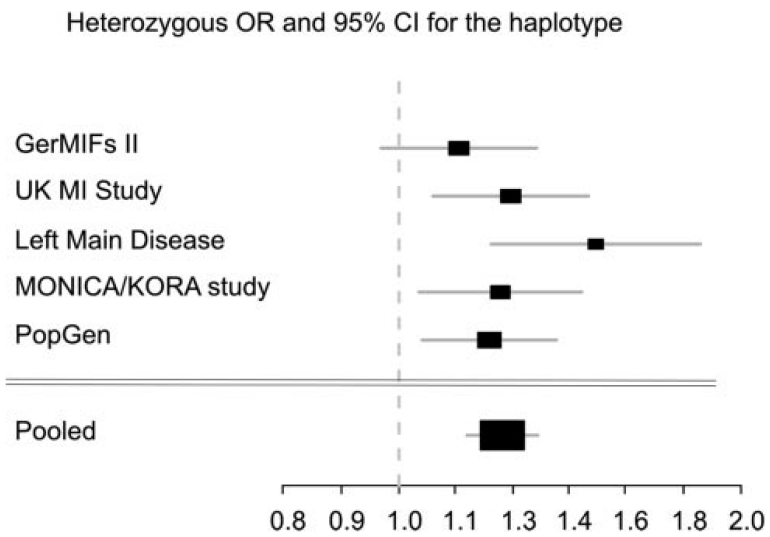


Figure 2. Haplotypic ORs and 95% CIs associated with the ACAC haplotypes under the assumption of additive haplotype effects. Boxes indicate the relative sizes of the samples.

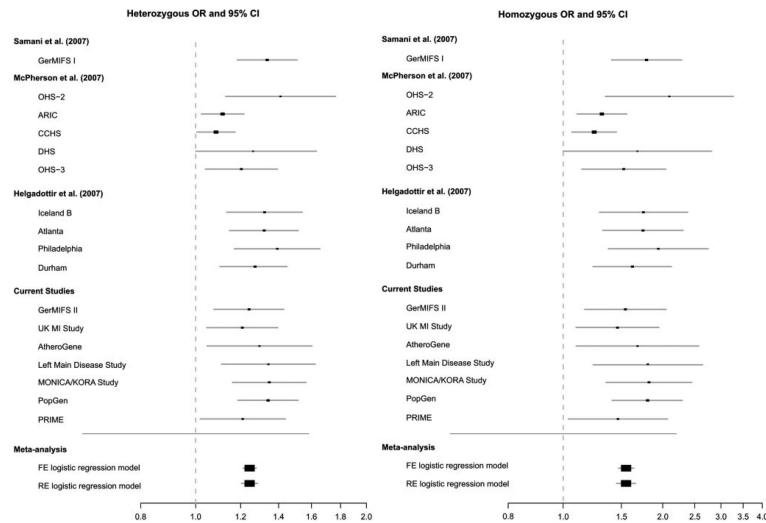


Figure 3. Heterozygous and homozygous ORs and 95% CIs under the assumption of an additive model. Boxes indicate the relative sizes of the samples. Effects were estimated from frequency tables presented on replication samples of the previously published studies (GerMIFS I: rs13330493; OHS-2, ARIC, CCHS, DHS, OHS-3: rs23832065; Iceland B, Atlanta, Philadelphia, Durham: rs107572786).

Table 1

Baseline Characteristics of the Study Populations Stratified by Affection Status

	GenMFS II		UK MI Study		AtheroGene		Left Main Disease Study		MONICA/KORA Study		PopGen		PRIME	
	Cases	Controls	Cases	Controls	Cases	Controls	Cases	Controls	Cases	Controls	Cases	Controls	Cases	Controls
Subject characteristic														
Individuals, n (%)	707	911	772	734	392	355	483	480	589	1056	1085	1021	617	620
Men	548 (77.5)	374 (41.1)	565 (73.2)	504 (68.7)	301 (76.8)	264 (74.4)	406 (84.1)	244 (50.8)	514 (87.3)	488 (46.2)	898 (82.8)	1021 (100)	617 (100)	620 (100)
Women	159 (22.5)	537 (58.9)	207 (26.8)	230 (31.3)	91 (23.2)	91 (25.6)	77 (15.9)	236 (49.2)	75 (12.7)	568 (53.8)	187 (17.2)	0	0	0
Mean (SD) age at onset, y	51.5 (7.6)	...	NA	...	NA	...	60.9 (11.6)	...	51.0 (6.6)	...	47.6 (6.3)	...	NA	...
Mean (SD) age at exam (all), y	58.9 (8.6)	57.7 (9.9)	57.6 (10.8)	55.7 (10.9)	58.2 (10.4)	58.5 (6.0)	68.6 (9.1)	61.7 (7.3)	56.3 (7.3)	51.1 (13.5)	54.2 (6.0)	56.7 (6.3)	55.3 (2.9)	55.2 (2.8)
Study information														
Recruitment strategy	Age of onset <65 y	Healthy, married-in spouses of MI cases	Age of onset <65 y	Age and gender matched	Age at exam <75 y	Subjects from routine checkup	Left main disease	Randomly identified through official population registries	Age of onset <60 y	Randomly identified through official population registries	Age of onset <55 y	Randomly identified through official population registries	10-y Follow-up study of men 50–59 y of age	Time and age matched
Recruitment period	1997–2002	1997–2002	1993–2000	1993–2000	1996–2004	1996–2004	2001–2005	2002–2005	1996	1994–1995	2002–2005	2002–2005	1991–1994	1991–1994
Main phenotype	MI	...	MI	...	CAD	...	Left main disease	...	MI	...	CAD	...	CAD	...
Ancestry	Germany	Germany	Leicester, Great Britain	Leicester, Great Britain	Mainz, Germany	...	Schleswig-Holstein, Germany	...	Augsburg, Germany	...	Schleswig-Holstein, Germany	...	Lille, Strasbourg, Toulouse (France); Belfast (Northern Ireland)	Lille, Strasbourg, Toulouse (France); Belfast (Northern Ireland)
Phenotype, n (%)														
MI	707 (100)	...	751 (100)	...	282 (71.9)	...	57 (11.8)	...	589 (100)	...	727 (67.0)	...	284 (46.0)	...
Family history of CAD	707 (100)	...	278 (36.0)	...	141 (36.0)	...	Sporadic	...	Sporadic	...	221 (20.4)	...	NA	...
CV risk factor, n (%)	357 (50.5)	463 (50.8)	468 (60.6)	266 (36.2)	268 (68.4)	86 (24.2)	NA	NA	456 (77.4)	600 (56.8)	868 (80.0)	838 (82.1)	476 (77.1)	426 (68.7)
Diabetes mellitus	129 (18.2)	53 (5.8)	70 (9.1)	14 (1.9)	57 (14.5)	10 (2.8)	152 (31.5)	6 (1.3)	92 (15.6)	29 (2.7)	198 (18.2)	38 (3.7)	49 (7.9)	36 (5.8)

	GerMIES II		UK MI Study		AtheroGene		Left Main Disease Study		MONICA/KORA Study		PopGen		PRIME	
	Cases	Controls	Cases	Controls	Cases	Controls	Cases	Controls	Cases	Controls	Cases	Controls	Cases	Controls
Hypertension	565 (79.9)	455 (49.9)	229 (30.6)	106 (14.5)	244 (62.2)	89 (25.1)	459 (95.0)	140 (29.2)	522 (88.6)	456 (43.2)	787 (72.5)	529 (51.8)	183 (29.7)	117 (18.9)
Hyperlipidemia	479 (67.8)	79 (8.7)	558 (89.1)	511 (72.2)	212 (54.1)	109 (30.7)	314 (65.0)	297 (61.9)	489 (83.0)	820 (77.7)	967 (89.1)	809 (79.2)	240 (38.9)	188 (30.3)
Obesity	168 (23.8)	166 (18.2)	224 (29.0)	85 (11.6)	83 (21.2)	79 (22.3)	116 (24.0)	79 (16.5)	182 (31.0)	220 (20.8)	350 (32.3)	183 (17.9)	123 (19.9)	81 (13.1)

Values are given as mean (SD) when appropriate. Affection status of the family members was based on reported history. Diabetes mellitus is defined as glucose >130 g/mL, known diabetes, or receiving treatment. Hypertension was defined as systolic blood pressure >140 mm Hg, diastolic blood pressure >90 mm Hg, or receiving treatment for these conditions. Hyperlipidemia was defined as total cholesterol >200 mg/dL, low-density lipoprotein cholesterol >130 mg/dL, or receiving lipid-lowering treatment. Obesity was defined as body mass index >30 kg/m².

Table 2
Single-Study Frequencies and Results of Single and Pooled Analyses for rs1333049

Study	Cases, n (%)				Controls, n (%)				OR Per 1 Allele (95% CI)	OR Per 2 Alleles (95% CI)	P	
	CC	CG	GG	CC	CG	GG	RE	FE				
GerMIFS II	206 (30.1)	330 (48.2)	149 (21.8)	204 (23.2)	448 (51.0)	226 (25.7)	1.24 (1.08–1.43)	1.54 (1.16–2.05)	0.0028			
UK MI study	201 (26.6)	381 (50.4)	174 (23.0)	164 (22.6)	356 (49.0)	207 (28.5)	1.21 (1.05–1.40)	1.46 (1.09–1.95)	0.0102			
AtheroGene	98 (26.5)	193 (52.2)	79 (21.4)	71 (20.6)	178 (51.6)	96 (27.8)	1.30 (1.05–1.60)	1.68 (1.10–2.57)	0.0175			
Left main disease study	141 (29.2)	252 (52.2)	90 (18.6)	95 (21.5)	238 (53.8)	109 (24.7)	1.34 (1.11–1.62)	1.80 (1.23–2.64)	0.0024			
MONICA/ KORA study	168 (29.6)	284 (50.1)	115 (20.3)	215 (21.4)	522 (52.0)	266 (26.5)	1.35 (1.16–1.57)	1.82 (1.35–2.45)	8.79×10 ⁻⁵			
PopGen	312 (29.2)	512 (47.9)	246 (23.0)	205 (20.5)	502 (50.3)	292 (29.2)	1.34 (1.19–1.52)	1.80 (1.41–2.30)	2.38×10 ⁻⁶			
PRIME	171 (32.6)	261 (49.7)	93 (17.7)	140 (27.0)	257 (49.3)	123 (23.7)	1.26 (1.06–1.50)	1.60 (1.13–2.26)	0.0081			
Pooled CAD without MI	321 (30.7)	528 (50.5)	196 (18.8)	511 (22.2)	1175 (50.9)	620 (26.9)	1.38 (1.24–1.54)	1.91 (1.55–2.37)	0.0095	3.08×10 ⁻⁹		
Pooled CAD only MI	976 (28.6)	1685 (49.4)	750 (22.0)	1094 (22.3)	2501 (50.9)	1319 (26.8)	1.27 (1.19–1.35)	1.60 (1.41–1.82)	0.0003	1.72×10 ⁻¹³		
Pooled total	1297 (29.1)	2213 (49.7)	946 (21.2)	1094 (22.3)	2501 (50.9)	1319 (26.8)	1.29 (1.22–1.37)	1.67 (1.48–1.87)	0.0001	1.17×10 ⁻¹⁷		

The FE models of the meta-analyses provide more impressive probability values based on the assumption that the ORs of all different studies are identical. In contrast to FE models, RE models are more conservative in that they allow for between-study variabilities. Displayed are genotype frequencies, ORs, 95% CIs, and probability values from Cochrane-Armitage trend test. ORs, 95% CIs, and P values for the pooled data from RE and FE logistic regression models are under the assumption of an additive model.

Table 3

Subgroup Analysis of rs1333049

Risk Factor	Cases, n	Controls, n	OR (95% CI)	P for Subgroup	P for Interaction
Gender					0.672
Male	3849	3515	1.29 (1.21–1.38)	3.07×10 ⁻⁴	...
Female	796	1662	1.32 (1.14–1.52)	1.24×10 ⁻²	...
Age, y					0.0708
>55	2640	2884	1.36 (1.26–1.48)	6.89×10 ⁻⁶	...
55	1874	2290	1.22 (1.11–.35)	1.51×10 ⁻³	...
Body mass index, kg/m ²					0.2963
>30	1136	893	1.39 (1.21–1.60)	5.98×10 ⁻⁵	...
30	3294	4278	1.29 (1.20–1.39)	9.90×10 ⁻⁸	...
Smoking					0.7536
Yes	3022	2679	1.31 (1.21–1.43)	3.26×10 ⁻⁹	...
No	1261	2011	1.33 (1.18–1.50)	8.48×10 ⁻⁶	...
Hypertension					0.5512
Yes	2933	1722	1.33 (1.20–1.47)	7.92×10 ⁻⁸	...
No	1536	3196	1.27 (1.14–1.41)	1.60×10 ⁻⁵	...
Hyperlipidemia					0.7709
Yes	2793	2512	1.32 (1.20–1.46)	1.11×10 ⁻⁷	...
No	1674	2359	1.28 (1.14–1.43)	2.74×10 ⁻⁵	...

Displayed are the number of analyzed individuals in each group, ORs, 95% CIs, and corresponding *P* values from the RE model separately for each of the risk factors.

Table 4

Allelic Effect of the rs1333049-C Allele According to the Haplotypic Backgrounds Derived From rs7044859, rs1292136, and rs7865618

Study	Haplotypic Background		
	TTA	TTG	ACA
GerMIFS II	0.91 (0.51–1.65)	1.05 (0.73–1.50)	1.35 (1.03–1.76)
UK MI study	0.72 (0.39–1.34)	1.31 (0.93–1.85)	1.37 (1.05–1.79)
Left Main Disease study	0.79 (0.38–1.63)	0.93 (0.61–1.42)	1.31 (0.89–1.92)
MONICA/KORA study	1.58 (0.92–2.73)	1.20 (0.82–1.75)	1.27 (0.96–1.68)
PopGen	0.96 (0.55–1.69)	1.12 (0.84–1.48)	1.73 (1.39–2.16)
Test for homogeneity across studies	0.36	0.78	0.39
Pooled OR (95% CI)	0.99 (0.76–1.30)	1.12 (0.96–1.31)	1.44 (1.27–1.62)
	0.94	0.14	4.08×10^{-9}

Allelic OR (95% CI) associated with the rs1333049-C allele according to the 3 haplotypic backgrounds by which it was carried: TTA-, TTG-, and ACA-. These backgrounds are composed of the alleles at the rs7044859 (T/A), rs1292136 (C/T), and rs7865618 (A/G) SNPs.