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 Braun, 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. Ghoroi, 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; Niles J. Samani, FMedSci; for the Cardiogenics Consortium

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 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=6.04×10−10 (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. (Circulation. 2008;117:1675-1684.)

Key Words: 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/H11005/3544). The remaining cases had evidence of CAD based on either a revascularization procedure or anginal symptoms with a positive stress test (n/H11005/1101).

#### Genotyping

In the GerMIFS II, the UK MI study, AtheroGene, the MONICA/KORA study, PopGen, and PRIME samples, genotyping was per-

---

### Table 1. Baseline Characteristics of the Study Populations Stratified by Affection Status

<table>
<thead>
<tr>
<th></th>
<th>GerMIFS II Cases</th>
<th>GerMIFS II Controls</th>
<th>UK MI Study Cases</th>
<th>UK MI Study Controls</th>
<th>AtheroGene Cases</th>
<th>AtheroGene Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Subject characteristic</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Individuals, n (%)</td>
<td>707 (100)</td>
<td>911 (100)</td>
<td>772 (100)</td>
<td>734 (100)</td>
<td>392 (100)</td>
<td>355 (100)</td>
</tr>
<tr>
<td><strong>Main phenotype</strong></td>
<td>MI</td>
<td></td>
<td>MI</td>
<td>MI</td>
<td>CAD</td>
<td></td>
</tr>
<tr>
<td><strong>Ancestry</strong></td>
<td>Germany</td>
<td></td>
<td>Leicester, Great Britain</td>
<td>Mainz, Germany</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Phenotype, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MI</td>
<td>707 (100)</td>
<td></td>
<td>751 (100)</td>
<td>278 (71.9)</td>
<td>141 (71.9)</td>
<td></td>
</tr>
<tr>
<td>Family history of CAD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ever smoked</td>
<td>357 (50.5)</td>
<td>463 (50.8)</td>
<td>468 (60.6)</td>
<td>266 (36.2)</td>
<td>268 (68.4)</td>
<td>86 (24.2)</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>129 (18.2)</td>
<td>53 (5.8)</td>
<td>70 (9.1)</td>
<td>14 (1.9)</td>
<td>57 (14.5)</td>
<td>10 (2.8)</td>
</tr>
<tr>
<td>Hypertension</td>
<td>565 (79.9)</td>
<td>455 (49.9)</td>
<td>229 (30.6)</td>
<td>106 (14.5)</td>
<td>244 (62.2)</td>
<td>89 (25.1)</td>
</tr>
<tr>
<td>Hyperlipidemia</td>
<td>479 (67.8)</td>
<td>79 (8.7)</td>
<td>558 (89.1)</td>
<td>511 (72.2)</td>
<td>212 (54.1)</td>
<td>109 (30.7)</td>
</tr>
<tr>
<td>Obesity</td>
<td>168 (23.8)</td>
<td>166 (18.2)</td>
<td>224 (30.9)</td>
<td>85 (11.6)</td>
<td>83 (21.2)</td>
<td>79 (22.3)</td>
</tr>
</tbody>
</table>

### Clinical Perspective p 1684

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.
formed 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-μL 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 al13 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 use of the 2-sided asymptotic Cochrane-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.14 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.15 Fixed-effect (FE) logistic regression models were used as the sensitivity analysis.

**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...
The recruitment strategy, main phenotypes, and source of the data are shown. In addition, we estimated the population-attributable fractions for the marker with highest linkage disequilibrium (LD; r² = 0.87 for rs238206 [used in studies from McPherson et al5]; r² = 0.9 by another SNP already selected a total of 31 SNPs (online-only Data Supplement Table I) for fine-map genotyping if a) they had higher r² values with regional association and b) they were not covered at r²>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 Adjustments of population-attributable fractions were made on these cardiovascular risk factors.

### Combination of Cardiovascular Risk Factors

For all studies, a multiple RE logistic regression was estimated 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.

### 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 an additive model. For all studies, a multiple RE logistic regression was estimated 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.

### 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, CHS, DSH, and OHS-3, ie, replication studies used in McPherson et al4; 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; r² = 0.87 for rs238206 [used in studies from McPherson et al4]; r² = 0.9 by another SNP already selected a total of 31 SNPs (online-only Data Supplement Table I) for fine-map genotyping if a) they had higher r² values with regional association and b) they were not covered at r²>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 Adjustments of population-attributable fractions were made on these cardiovascular risk factors.

### 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 > 2) with SNPs showing CAD association (P < 10⁻⁵) in the Welcome 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 r² values with regional association and b) they were not covered at r²>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
data\textsuperscript{24} is shown in Figure 1. The lead SNP, rs1333049, identified by Samani et al\textsuperscript{3} 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,\textsuperscript{3,4} including those identified by McPherson et al\textsuperscript{5} (rs10757274 and rs2383206) and Helgadottir et al\textsuperscript{6} (rs10116277, rs1333040, rs2383207, and rs10757278). In moderate LD (average $|D'| = 0.60$) with this main block is a neighboring LD block that harbors further SNPs associated with CAD.\textsuperscript{3} The SNPs in this second block generate 5 haplotypes with a frequency $> 2\%$ that explain $\approx 97\%$ of its haplotypic diversity.\textsuperscript{3} Three SNPs (rs7044859, rs1292136, and rs7865618) are sufficient to characterize the main haplotypes of this second block.\textsuperscript{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 model 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/m$^2$ ($P=0.29$), 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 al$^{19}$ 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 ($P=3.38\times10^{-16}$). 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).

**Table 3. Subgroup Analysis of rs1333049**

<table>
<thead>
<tr>
<th>Risk Factor</th>
<th>Cases, n</th>
<th>Controls, n</th>
<th>OR (95% CI)</th>
<th>$P$ for Subgroup</th>
<th>$P$ for Interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>3849</td>
<td>3515</td>
<td>1.29 (1.21–1.38)</td>
<td>$3.07\times10^{-4}$</td>
<td>...</td>
</tr>
<tr>
<td>Female</td>
<td>796</td>
<td>1662</td>
<td>1.32 (1.14–1.52)</td>
<td>$1.24\times10^{-2}$</td>
<td>...</td>
</tr>
<tr>
<td>Age, y</td>
<td></td>
<td></td>
<td></td>
<td>$0.0708$</td>
<td></td>
</tr>
<tr>
<td>$&gt;55$</td>
<td>2640</td>
<td>2884</td>
<td>1.36 (1.26–1.48)</td>
<td>$6.89\times10^{-6}$</td>
<td>...</td>
</tr>
<tr>
<td>$\leq55$</td>
<td>1874</td>
<td>2290</td>
<td>1.22 (1.11–1.35)</td>
<td>$1.51\times10^{-3}$</td>
<td>...</td>
</tr>
<tr>
<td>Body mass index, kg/m$^2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$&gt;30$</td>
<td>1136</td>
<td>893</td>
<td>1.39 (1.21–1.60)</td>
<td>$5.98\times10^{-5}$</td>
<td>...</td>
</tr>
<tr>
<td>$\leq30$</td>
<td>3294</td>
<td>4278</td>
<td>1.29 (1.20–1.39)</td>
<td>$9.90\times10^{-5}$</td>
<td>...</td>
</tr>
<tr>
<td>Smoking</td>
<td></td>
<td></td>
<td></td>
<td>$0.2963$</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>3022</td>
<td>2679</td>
<td>1.31 (1.21–1.43)</td>
<td>$3.26\times10^{-9}$</td>
<td>...</td>
</tr>
<tr>
<td>No</td>
<td>1261</td>
<td>2011</td>
<td>1.33 (1.18–1.50)</td>
<td>$8.48\times10^{-6}$</td>
<td>...</td>
</tr>
<tr>
<td>Hypertension</td>
<td></td>
<td></td>
<td></td>
<td>$0.7536$</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>2933</td>
<td>1722</td>
<td>1.33 (1.20–1.47)</td>
<td>$7.92\times10^{-8}$</td>
<td>...</td>
</tr>
<tr>
<td>No</td>
<td>1536</td>
<td>3196</td>
<td>1.27 (1.14–1.41)</td>
<td>$1.60\times10^{-5}$</td>
<td>...</td>
</tr>
<tr>
<td>Hyperlipidemia</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>2793</td>
<td>2512</td>
<td>1.32 (1.20–1.46)</td>
<td>$1.11\times10^{-7}$</td>
<td>...</td>
</tr>
<tr>
<td>No</td>
<td>1674</td>
<td>2359</td>
<td>1.28 (1.14–1.43)</td>
<td>$2.74\times10^{-5}$</td>
<td>...</td>
</tr>
</tbody>
</table>

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.

![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.](http://circ.ahajournals.org/doi/suppl/10.1161/CIRCULATIONAHA.107.7044859-A)
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; \( P = 4.08 \times 10^{-4} \)) 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 al, McPherson et al, and Helgadottir et al). We analyzed rs1333049, rs2383206, and rs10757278 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 \( 1.62 \) (95% CI, 1.57 to 1.66) in the RE model and \( 1.20 \) to \( 1.29 \) in the FE model, respectively. Including additional association data from the initial 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 is also 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 for this finding. The region on chromosome 9p21.3 associated with CAD is located in 2 blocks of strong intrablock and
interblock LD. 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.

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 involve-

---

**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: rs1333049; OHS-2, ARIC, CCHS, DHS, OHS-3: rs2383206; Iceland B, Atlanta, Philadelphia, Durham: rs10757278).
ment 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. 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. 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 phenotypes 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.

Acknowledgments

We thank Petra Bruse, Kristin Blankenberg, Viviane Nicaud, and Maylis de Sureau 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.

Disclosures

None.

References

CLINICAL PERSPECTIVE

Recent studies exploring the entire genome identified the chromosomal locus 9p21.3 as being associated with the risk of coronary artery disease and myocardial infarction. To use this information for clinical risk prediction, the present study investigated close to 10,000 individuals. The findings consistently replicated the significant association between the 9p21.3 locus and coronary artery disease in 7 case-control studies. The chromosomal locus was further investigated by genotyping of 31 additional single-nucleotide polymorphisms and detailed haplotypes that further characterized the genomic structure of the chromosomal locus and its association with coronary artery disease. Moreover, the present study revealed that an autosomal-additive mode of inheritance best explained the underlying association. This information is essential for appropriate incorporation of the genotypic information into algorithms for risk prediction. Indeed, the present study can be considered an important step for using genomic information to better predict risk of coronary artery disease.
Repeated Replication and a Prospective Meta-Analysis of the Association Between Chromosome 9p21.3 and Coronary Artery Disease

Circulation. 2008;117:1675-1684; originally published online March 24, 2008; doi: 10.1161/CIRCULATIONAHA.107.730614
Circulation is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2008 American Heart Association, Inc. All rights reserved.
Print ISSN: 0009-7322. Online ISSN: 1524-4539

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circ.ahajournals.org/content/117/13/1675

Data Supplement (unedited) at:
http://circ.ahajournals.org/content/suppl/2008/03/21/CIRCULATIONAHA.107.730614.DC1

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation is online at:
http://circ.ahajournals.org/subscriptions/