Published in final edited form as:
Lancet. 2011 January 29; 377(9763): 383–392. doi:10.1016/S0140-6736(10)61996-4.

Identification of *ADAMTS7* as a novel locus for coronary atherosclerosis and association of *ABO* with myocardial infarction in the presence of coronary atherosclerosis: two genome-wide association studies

Muredach P Reilly, MB[*], **Mingyao Li, PhD**[*], **Jing He, PhD**, **Jane F Ferguson, PhD**, **Ioannis M Stylianou, PhD**, **Nehal N Mehta, MD**, **Mary Susan Burnett, PhD**, **Joseph M Devaney, PhD**, **Christopher W Knouff, MD**, **Prof John R Thompson, PhD**, **Benjamin D Horne, PhD**, **Alexandre F R Stewart, PhD**, **Themistocles L Assimes, MD**, **Philipp S Wild, MD**, **Hooman Allayee, PhD**, **Patrick Linsel Nitschke, MD**, **Riyaz S Patel, MD**, **Myocardial Infarction Genetics Consortium**[†], **Wellcome Trust Case Control Consortium**[†], **Nicola Martinelli, MD**, **Prof Domenico Girelli, MD**, **Prof Arshed A Quyyumi, MD**, **Prof Jeffrey L Anderson, MD**, **Prof Jeanette Erdmann, PhD**, **Alistair S Hall, MD**, **Prof Heribert Schunkert, MD**, **Prof Thomas Quertermous, MD**, **Prof Stefan Blankenberg, MD**, **Prof Stanley L Hazen, PhD**, **Prof Robert Roberts, MD**, **Sekar Kathiresan, MD**, **Prof Nilesh J Samani, MD**, **Prof Stephen E Epstein, MD**, and **Prof Daniel J Rader, MD**

Cardiovascular Institute (M P Reilly MB, J F Ferguson PhD, N N Mehta MD, Prof D J Rader MD), Institute for Translational Medicine and Therapeutics, School of Medicine (M P Reilly, I M Stylianou PhD, N N Mehta, Prof D J Rader), and Biostatistics and Epidemiology (M Li PhD, J He PhD), University of Pennsylvania, Philadelphia, PA, USA; Cardiovascular Research Institute, MedStar Health Research Institute, Washington Hospital Center, Washington, DC, USA (M S Burnett PhD, J M Devaney PhD, Prof S E Epstein MD); Genetics Division and Drug Discovery, GlaxoSmithKline, King of Prussia, PA, USA (C W Knouff MD); Department of Health Sciences (Prof J R Thompson PhD), and Department of Cardiovascular Sciences (Prof N J Samani MD), University of Leicester, Leicester, UK; Cardiovascular Department, Intermountain Medical Center, Salt Lake City, UT, USA (B D Horne PhD, Prof J L Anderson MD); Genetic Epidemiology Division (B D Horne), and Cardiology Division (Prof J L Anderson), University of Utah, Salt Lake City, UT, USA; Department of Medicine, University of Ottawa Heart Institute, Ottawa, ON, Canada (A F R Stewart PhD, Prof R Roberts MD); Department of Medicine, Stanford University School of Medicine, Stanford, CA, USA (T L Assimes MD, Prof T Quertermous MD); Department of Medicine II, University Medical Centre Mainz, Mainz, Germany (P S Wild MD, Prof S Blankenberg MD); Department of Preventive Medicine, Institute for Genetic Medicine, Keck School of

Correspondence to: Dr Muredach P Reilly, Cardiovascular Institute, University of Pennsylvania Medical Center, 609 BRB 2/3, 421 Curie Blvd, Philadelphia, PA 19104-6160, USA muredach@mail.med.upenn.edu.

[*]Authors contributed equally

[†]Webappendix shows full listing of participants from the Wellcome Trust Case Control Consortium (WTCCC) and the Myocardial Infarction Genetics (MI-GEN) Consortium

Contributions

MPR, ML, SEE, SK, NJS, and DJR contributed to study design. JFF, IMS, NNM, MSB, JMD, BDH, AFRS, TLA, PSW, RSP, PLN, NM, DG, AAQ, JLA, JE, ASH, HS, TQ, SB, SLH, RR, DJR, and NJS contributed to data collection and performed research. ML, JH, MPR, JFF, JRT, HA, NM, CWK, BDH, AFRS, TLA, JE, TQ, SB, SLH, RR, SK, DJR, and NJS contributed to data analysis and interpreted results. MPR, ML, JFF, and DJR wrote drafts of the manuscript. MPR, ML, JFF, MSB, JMD, CWK, BDH, AFRS, TLA, PSW, HA, PLN, RSP, DG, AAQ, JLA, JE, ASH, HS, TQ, SB, SLH, RR, SK, NJS, SEE, and DJR revised and reviewed the final manuscript.

Conflicts of interest

MPR, MSB, JMD, SEE, and DJR received research grant support from GlaxoSmithKline. CWK was an employee of GlaxoSmithKline at the time of the study. ML, JH, JFF, IMS, NNM, JRT, BDH, AFRS, TLA, PSW, HA, PLN, RSP, NM, DG, AAQ, JLA, JE, ASH, HS, TQ, SB, SLH, RR, SK, and NJS declare that they have no conflicts of interest.

Medicine, University of Southern California, Los Angeles, CA, USA (H Allayee PhD); Medizinische Klinik II, University of Lübeck, Lübeck, Germany (P L Nitschke MD, Prof J Erdmann PhD, Prof H Schunkert MD); Division of Cardiology, Emory University School of Medicine, Atlanta, GA, USA (R S Patel MD, Prof A A Quyyumi MD); Department of Medicine, University of Verona, Verona, Italy (N Martinelli MD, Prof D Girelli MD); Leeds Institute of Genetics Health and Therapeutics (LIGHT), and Multidisciplinary Cardiovascular Research Centre (MCRC), University of Leeds, Leeds, UK (A S Hall MD); The Center for Cardiovascular Diagnostics & Prevention, Cleveland Clinic, Cleveland, OH, USA (Prof S L Hazen PhD); Cardiovascular Research Center and Cardiology Division, Massachusetts General Hospital, Boston, MA, USA (S Kathiresan MD); Center for Human Genetic Research, Massachusetts General Hospital, Boston, MA, USA (S Kathiresan); Department of Medicine, Harvard Medical School, Boston, MA, USA (S Kathiresan); Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA, USA (S Kathiresan); and Leicester National Institute for Health Research Biomedical Research Unit in Cardiovascular Disease, Glenfield Hospital, Leicester, UK (Prof N J Samani)

Summary

Background—We tested whether genetic factors distinctly contribute to either development of coronary atherosclerosis or, specifically, to myocardial infarction in existing coronary atherosclerosis.

Methods—We did two genome-wide association studies (GWAS) with coronary angiographic phenotyping in participants of European ancestry. To identify loci that predispose to angiographic coronary artery disease (CAD), we compared individuals who had this disorder (n=12 393) with those who did not (controls, n=7383). To identify loci that predispose to myocardial infarction, we compared patients who had angiographic CAD and myocardial infarction (n=5783) with those who had angiographic CAD but no myocardial infarction (n=3644).

Findings—In the comparison of patients with angiographic CAD versus controls, we identified a novel locus, *ADAMTS7* ($p=4{\cdot}98\times10^{-13}$). In the comparison of patients with angiographic CAD who had myocardial infarction versus those with angiographic CAD but no myocardial infarction, we identified a novel association at the *ABO* locus ($p=7{\cdot}62\times10^{-9}$). The *ABO* association was attributable to the glycotransferase-deficient enzyme that encodes the ABO blood group O phenotype previously proposed to protect against myocardial infarction.

Interpretation—Our findings indicate that specific genetic predispositions promote the development of coronary atherosclerosis whereas others lead to myocardial infarction in the presence of coronary atherosclerosis. The relation to specific CAD phenotypes might modify how novel loci are applied in personalised risk assessment and used in the development of novel therapies for CAD.

Funding—The PennCath and MedStar studies were supported by the Cardiovascular Institute of the University of Pennsylvania, by the MedStar Health Research Institute at Washington Hospital Center and by a research grant from GlaxoSmithKline. The funding and support for the other cohorts contributing to the paper are described in the webappendix.

Introduction

Definition of the genetic architecture of coronary artery disease (CAD) and myocardial infarction can provide substantial benefit through improved risk prediction and development of novel therapies. Recent genome-wide association studies (GWAS) provide promise, with identification of several novel loci for these disorders.[1–6] However, only a small proportion of the inherited component has been identified.[5] Atherosclerotic plaque rupture is the most common cause of myocardial infarction.[7] Since all patients with plaque rupture or myocardial infarction have coronary atherosclerosis but only a few with coronary

atherosclerosis develop myocardial infarction, unique factors—some genetic—are likely to predispose to plaque rupture or myocardial infarction in coronary atherosclerosis. Within clinically defined myocardial infarction, however, the mechanisms that drive events are unknown, such as those that cause progression of atherosclerosis, those that modulate plaque vulnerability, or factors that lead to arterial thrombosis.[7] In fact, it is yet to be determined if identified loci for myocardial infarction[5] contribute to initiation and progression of atherosclerosis or to plaque rupture and thrombosis in leading to myocardial infarction.

We report two GWAS of CAD designed to address the hypothesis that genetic factors predisposing to myocardial infarction in patients with coronary atherosclerosis are distinct from those that associate with the presence of coronary atherosclerosis. Unlike previous GWAS in this disease, we used coronary angiography in primary ascertainment of CAD phenotypes. This approach allows discrimination of risk alleles for plaque rupture and myocardial infarction from those for coronary atherosclerosis.

Methods

Participants

The webappendix shows a detailed description of design and clinical characteristics for study samples (pp 2–4). Our primary focus was on studies of angiographic CAD in patients of European ancestry. Discovery studies were PennCath and MedStar (webappendix p 2), which recruited patients before coronary angiography at the University of Pennsylvania Medical Center (Penn) and Washington Hospital Center (WHC), respectively. Selection of younger patients with angiographic CAD (mean age about 51±7 years) and older controls (mean age about 61±9 years) was aligned across these two studies, which used similar angiographic criteria and had similar numbers of cases of angiographic CAD with and without myocardial infarction.

PennCath is a hospital-based study of genes and biomarkers for CAD.[1,5,8] Briefly, between July 1, 1998, and March 31, 2003, PennCath recruited 3815 consecutive patients undergoing cardiac catheterisation. All patients gave written informed consent. Clinical information was extracted from the medical records. Coronary angiograms were scored at the time of procedures. Blood was taken from patients in a fasting state, DNA and plasma were isolated, and lipoproteins and glucose were assayed. A nested case-control GWAS study of angiographic CAD was done in PennCath (n=1401 white patients). Controls (n=468) had no or minimum CAD (<10% stenosis of any vessel) on angiography and were aged older than 40 years for men and 45 years for women. Patients with angiographic CAD (n=933) had at least one coronary vessel with 50% or more stenosis, and men were aged 60 years or younger and women 65 years or younger. About half of patients with angiographic CAD presented with or had history of myocardial infarction (n=469).

MedStar is a hospital-based study designed for biomarker and genetic studies of CAD. Briefly, between Aug 1, 2004, and March 31, 2007, a case-control sample of patients with and without angiographic CAD undergoing cardiac catheterisation were recruited.[5] All patients gave written informed consent and demographic, clinical, risk factor, and angiographic data were recorded. Study design of the case-control GWAS was similar to that for PennCath, with 1322 white patients composed of controls (n=447) and those with angiographic CAD (n=875). The patients with angiographic CAD were divided into those with myocardial infarction (n=421) and those without (n=454). Patients' age at diagnosis of CAD was 55 years or younger for men and 60 years or younger for women. Controls were older than 45 years.

The webappendix (pp 2–4, 8, and 9) shows additional studies, including case-control subsets from the Wellcome Trust Case-Control Consortium (WTCCC) Study,[3,4] the Ottawa Heart Genomics Study (OHGS),[2,9] the Atherosclerotic Disease Vascular Function and Genetic Epidemiology study (ADVANCE),[10] the Cadomics study,[6,11] the Emory Genebank study,[1] the Intermountain Heart Collaborative Study (IHCS),[12] the Verona Heart Study (VHS),[5,13] the Cleveland Clinic GeneBank Study,[9] the Myocardial Infarction Genetics Consortium (MI-GEN),[5] and the German myocardial infarction Family Studies.[6]

Procedures and study design

Detailed descriptions of genotyping, imputation, and quality control are shown in the webappendix (pp 4, 5, 10, and 11). Figure 1 shows the study design of the two multistage, case-control GWAS studies of phenotypes of angiographic CAD. A GWAS is a study in which genetic information is gathered for individuals with known phenotypes using a large array of genetic markers that represent common variation in the human genome. The aim is to map susceptibility genes through the detection of associations between these genetic markers and the phenotype of interest.[14] First, we did a GWAS of patients with angiographic CAD compared with controls to identify loci for coronary atherosclerosis (study A, figure 1). We used a four-stage strategy similar to that used by Kathiresan and colleagues[5] in their studies of myocardial infarction (study B, figure 1). Second, we did a GWAS study within patients with angiographic CAD comparing those who had myocardial infarction with those who did not to identify loci that predispose specifically to myocardial infarction. With a modification of a two-stage GWAS approach,[15] we restricted the number of single nucleotide polymorphisms (SNPs) examined in stage 2, in view of modest power (webappendix p 5).

In stage 1 of study A, more than 2·4 million genotyped and imputed SNPs were studied. The union of top SNPs in PennCath, MedStar, and their meta-analysis (5425 SNPs with additive trend test $p<1\times10^{-3}$) were selected for stage-2 replication in analysis of existing genome-wide data for the WTCCC,[3,4] OHGS,[2,9] and ADVANCE[10] studies (figure 1, webappendix pp 2, 3, and 8). After meta-analysis of stages 1 and 2, 25 novel SNPs ($p <1\times10^{-5}$) and ten SNPs from previous GWAS[4,5] were selected for stage-3 wet-lab genotyping in the Emory GeneBank Study,[1] Utah IHCS,[12] and the VHS.[5,13] The criteria for SNP selection at each stage were based on the approach of Kathiresan and colleagues.[5] A significant finding was defined a priori by published criteria[5,14] as $p <5\times10^{-8}$ in stages 1–3 combined as well as independent replication within stage-3 corrected for the number of SNP tests (n=35). For SNPs that reached genome-wide significance in stages 1–3, we undertook additional replication (figure 1, stage 4).[9] Separately, we did an exploratory analysis of novel findings from study A in early-myocardial infarction studies consisting of MIGEN[5] and GerMIFS[6] (n=6256 myocardial infarction cases and 7711 controls).

For study B we assembled a six-study GWAS for an a-priori stage-1 meta-analysis of patients with angiographic CAD, comparing those with myocardial infarction with those without (figure 1, webappendix pp 2–4, and 9). Participants had evidence of angiographic CAD (by angiography in PennCath, MedStar, OHGS, CADomics and ADVANCE or by history of coronary revascularisation in WTCCC) and were classified for myocardial infarction as described.[4–6] In most studies, patients with myocardial infarction were younger than those without and had broadly similar risk factors (webappendix p 9). In the stage-1 meta-analysis, between 1·2 and 2·4 million SNPs were examined. In stage 2, SNPs with $p<5\times10^{-6}$ in stage 1 (n=49) were interrogated in two Cleveland Clinic GeneBank studies[9] with existing genome-wide data. A significant finding was defined a priori as $p<5\times10^{-8}$ in stages 1 and 2 combined as well as independent replication within stage 2 after Bonferroni correction for independent SNP tests (in stage 2, multiple SNPs were in strong linkage disequilibrium; therefore, we corrected for independent tests with the Nyholt method[16]). We

also did a meta-analysis of all eight studies. Finally, we examined the association of published loci for myocardial infarction[5] with angiographic CAD (in study A) as well as myocardial infarction in patients with angiographic CAD (in study B).

Statistical analysis

For genotyped SNPs, association was tested by logistic regression with the assumption of additive genetic effects with PLINK.[17] For imputed SNPs, association was examined with SNPTEST, which can account for imputation uncertainty.[18] We adjusted the effects of sex and age at diagnosis of CAD in all analyses. The estimated genomic control inflation factor lambda (λ),[19] an indicator of potential population stratification, was calculated as the median of the test statistics divided by the median of χ^2 distribution with one degree of freedom. For top SNPs identified in studies A and B, we also explored in PennCath the association with traditional risk factors such as diabetes, smoking, hypertension, hyperlipidaemia, and assessed top SNP associations with phenotypes of angiographic CAD after controlling for these risk factors. Finally, for top SNPs identified in study B, we used PennCath samples to determine the associations of these SNPs with myocardial infarction in patients with angiographic CAD after adjustment for Gensini score, a semi-quantitative estimate of the burden of angiographic CAD.[20]

Meta-analysis was done by a weighted *Z*-score method with METAL,[21] which accounts for the direction of association relative to a consistent reference allele. In this method, p values for each study are first converted to a *Z* score. Then a weighted sum of *Z* scores is calculated, in which each statistic is weighted by the square root of the effective sample size for each study. The resulting sum is divided by the square root of the total effective sample size to obtain an overall *Z* statistic, which is used to assess the overall evidence for association. The reason for use of the effective sample size is to adjust for asymmetric case-control sample sizes. In accordance with de Bakker and colleagues,[22] we used the non-centrality parameter for the given asymmetric case-control sample size, and then iteratively determined the effective (symmetric) case-control sample size that returns the same non-centrality parameter. To test for consistency of allelic effects across studies at the same SNP, we calculated two summary statistics of heterogeneity, Cochran's *Q* statistic, which provides a test of heterogeneity of allelic effects at the *test* SNP, and an alternative I^2 index, which quantifies heterogeneity in allelic effects across studies, over that expected by chance. The webappendix shows the power calculations (p 5).

In study B, SNPs at the *ABO* locus had genome-wide significant associations with myocardial infarction in patients with angiographic CAD. Therefore, in PennCath we inferred ABO blood groups and analysed blood group association with angiographic CAD phenotypes (webappendix pp 5 and 6). Stratified and conditional analyses were done in PennCath to examine whether the top *ABO* SNPs for myocardial infarction in patients with angiographic CAD were independent of ABO blood groups and distinct *ABO* SNPs.

Role of the funding source

Employees of GlaxoSmithKline contributed to study design, data interpretation, and editing of the report. The corresponding author had full access to the data and final responsibility for decision to submit for publication.

Results

Study A: GWAS of angiographic CAD patients versus controls

We examined loci for myocardial infarction, established through published GWAS,[5] for their association with angiographic CAD in studies in which all cases and controls were

defined by angiography (6886 patients with angiographic CAD and 3226 controls). The direction and strength of association with angiographic CAD for risk alleles (table 1) was largely consistent with published findings.[4,5]

Meta-analysis of stage-1–4 studies identified a novel genotyped SNP, rs1994016, on 15q25.1 that exceeded genome-wide significance for angiographic CAD (figure 2, table 2). This finding had modest heterogeneity across studies (Cochran's Q=17·7, p=0·013; I^2=0·60) with strong effects in initial studies, consistent with a so-called winners curse[24] (figure 2, table 2). Within stages 1–3, rs1994016C reached genome-wide significance (table 2). In the same analysis, the published 9p21 rs4977574G allele (frequency about 0·56) had an OR of 1·35, p=$2{\cdot}98\times10^{-27}$. Within the stage-3 wet-lab replications, rs1994016C had an OR of 1·14, p=$4{\cdot}97\times10^{-4}$ (p=0·0174 after Bonferroni correction). In the stage-4 replication, the effect size was close to that seen in the other stages although the p value was 0·076, which probably indicates the small effective sample. In PennCath, rs1994016 was not associated with any risk factors for cardiovascular disease (data not shown) and its association with angiographic CAD was not attenuated after adjustment for risk factors (OR 1·32–1·35 before and after adjustment). The webappendix shows findings for all SNPs analysed in stages 1–3 (pp 12 and 13).

In separate analysis of MI-GEN and GerMIFS (early-onset myocardial infarction studies), the rs1994016C allele was not associated with myocardial infarction (OR 1·02, p=0·81; figure 2). This finding was unlikely to be due to low power because the analysed sample had more than 80% power to detect genotype relative-risk as low as 1·08.

The variant rs1994016 maps within intron 8 of *ADAMTS7*, a member of the family of disintegrin and metalloproteinase with thrombospondin motifs proteins.[25] The webappendix shows the gene annotation and the linkage disequilibrium structure at this 15q25.1 region (p 23). The linkage disequilibrium is weak (r^2<0·2) between rs1994016 and SNPs in any genes within 500 kb 3′ or 5′ of *ADAMTS7*, which suggests that *ADAMTS7* is probably the atherosclerosis gene at this locus.

Study B: GWAS of myocardial infarction in patients with angiographic CAD

By contrast with angiographic CAD versus control, none of the published GWAS SNPs for myocardial infarction[5] were significant (p<0·05) for myocardial infarction in patients with angiographic CAD (table 1); see for example, allele 9p21 rs4977574G). Our findings are unlikely to be attributable to low power, because this analysed sample had more than 80% power to detect genotype-relative risks from 1·07 to 1·20 for the reported effect sizes and allele frequencies at these loci.

In stage 1, 49 SNPs showed p<5×10^{-6} for myocardial infarction in patients with angiographic CAD. The top 11 SNPs mapped to one region on 9q34.2 within *ABO*, the blood group locus. Combined with stage-2 data, the association of multiple *ABO* SNPs, all in strong linkage disequilibrium (r^2 >0·85), exceeded genome-wide significance (table 3; webappendix pp 14 and 24). For example, the OR of the G allele of the top genotyped SNP, rs612169, was 1·20 (p=$3{\cdot}66\times10^{-8}$) for patients with angiographic CAD who had myocardial infarction whereas for the C allele at rs514659, the top imputed SNP, it was 1·21 (p=$7{\cdot}62\times10^{-9}$, figure 3). Findings for rs514659 were consistent across studies in stages 1 and 2 (Cochran's Q=4·8, p=0·69; I^2=0, figure 3). Within stage-2 replication, the OR of the rs514659C allele was 1·24 (p=0·00145; p=0·0174 after Bonferroni correction for 12 independent tests). The webappendix shows results of the analysis of the combined stages 1 and 2 for all 49 SNPs selected in stage 1 (p 14).

We tested the association of risk alleles with specific angiographic CAD phenotypes in PennCath. Relative to controls, *ABO* rs514659C was related to angiographic CAD with myocardial infarction, but not to angiographic CAD without myocardial infarction (table 4), which suggests the *ABO* locus is related to myocardial infarction but not to coronary atherosclerosis. In PennCath, *ABO* SNPs were not associated with traditional risk factors (data not shown) and *ABO* associations with myocardial infarction in patients with angiographic CAD were not attenuated after risk factor adjustment (eg, rs514659, OR 1·38 before and after adjustment). Furthermore, in PennCath the signal for myocardial infarction in patients with angiographic CAD was not attenuated after adjusting for Gensini scores (eg, rs514659, OR 1·38 before and after adjustment). These findings suggest that results are not affected by slight differences in the burden of coronary atherosclerosis between angiographic CAD patients with with myocardial infarction and angiographic CAD patients without myocardial infarction.

Figure 4 and p 25 of webappendix show the gene annotation and linkage-disequilibrium structure of the GWAS region mapping to 9q34.2 as well as meta-analysis p values and recombination rates. The SNPs with strongest association lie in a linkage-disequilibrium block in intron 1 of the *ABO* locus. There is only modest linkage disequilibrium with SNPs (all r^2<0·2) in other genes within 500 kb 5′ and 3′, suggesting that *ABO* is the myocardial infarction gene at this locus. The A allele of rs612169, our top genotyped SNP, tags the O blood group. Therefore, we inferred the main ABO blood groups using SNPs that tag *ABO* alleles[26] (webappendix pp 5, 6, and 16) and examined their associations with CAD phenotypes in PennCath. Blood group A, B, and AB genotypes had greater odds of myocardial infarction in patients with angiographic CAD than did blood group O (table 4 and webappendix p 16). Subgroup analysis in PennCath revealed that, relative to controls, non-O blood groups had higher odds of angiographic CAD with myocardial infarction but not of angiographic CAD without myocardial infarction (table 4). This result suggests that ABO blood group O is protective against myocardial infarction in patients with angiographic CAD but is not related to coronary atherosclerosis. In a subgroup analysis of PennCath patients restricted to those with non-O blood groups, rs514659 was not associated with myocardial infarction in patients with angiographic CAD (OR 0·90 [0·59–1·38], p=0·63), suggesting that rs514659 has no effect independent of the ABO blood group allele.

Discussion

We identified two loci for distinct CAD phenotypes, *ADAMTS7*, a novel locus for angiographic CAD but not myocardial infarction, and *ABO*, a gene for myocardial infarction in patients with angiographic CAD, but not for angiographic CAD itself. Further, our data suggest that the *ABO* GWAS signal for myocardial infarction in patients with angiographic CAD is mediated by the glycotransferase-deficient isoform that encodes the ABO blood group O phenotype.

Clinical CAD phenotypes are heritable but highly complex. The association of several published loci for myocardial infarction[5] might be mediated by diverse pathological processes including those that promote atherosclerosis, precipitate plaque rupture, or facilitate arterial thrombosis. Our use of coronary angiography reduced heterogeneity in coronary atherosclerosis within patients with CAD while allowing discrimination of risk alleles for plaque rupture or myocardial infarction from those for atherosclerosis. Although most published loci for myocardial infarction had significant signals for angiographic CAD compared with controls, none were associated with myocardial infarction in patients with angiographic CAD. This finding suggests that these loci relate to myocardial infarction indirectly via coronary atherosclerosis rather than having a specific role in vulnerable plaque and myocardial infarction. In fact, independent studies support this concept for the 9p21

locus. Consistent with our data, Horne and colleagues[12] showed that this locus did not predict incident or prevalent myocardial infarction in patients with CAD but was strongly associated with the presence of angiographic CAD versus controls.

Our discovery of *ADAMTS7* as a novel locus for CAD might have been facilitated by use of coronary angiography because, unlike clinically defined cases, the definition of angiographic CAD required a pre-specified burden of coronary atherosclerosis. All *ADAMTS* genes have a similar domain structure, consisting of a preproregion, a reprolysin-type catalytic domain, a disintegrin-like domain, a thrombospondin type-1 module, a cysteine-rich domain, a spacer domain, and a COOH-terminal thrombospondin type-1 module. ADAMTS7 degrades cartilage oligomeric matrix protein and has been implicated in inflammatory arthritis and bone growth. Overexpression of ADAMTS7 accelerates migration of vascular smooth muscle cells in vitro and exacerbates neointimal thickening after carotid artery injury in vivo, perhaps through degradation of cartilage oligomeric matrix protein.[27] These data implicate ADAMTS7 in the proliferative response to vascular injury, a process that has parallels to the progressive phase of atherosclerosis.[7] These mechanistic findings coincide with the lack of association of *ADAMTS7* SNPs with early-onset myocardial infarction. Together they raise the provocative possibility that some proteins, such as ADAMTS7, could increase plaque size but not affect plaque stability. Overall, *ADAMTS7* might be a novel therapeutic target for progression of atherosclerosis but seems less likely to be one for prevention of myocardial infarction in high-risk patients.

Discovery of *ABO* as the top locus for myocardial infarction in patients with angiographic CAD is notable, in view of decades of work suggesting a relation between ABO blood groups and both thrombosis and coronary heart disease.[28,29] The *ABO* gene encodes proteins (transferase A, a 1-3-N-acetylgalactosaminyltransferase; transferase B, a 1-3-galactosyltransferase) related to the ABO blood group system.[30] Blood group O is caused by a deletion of guanine-258 near the N-terminus of the protein. This deletion causes a frameshift, which results in translation of a protein with no glycosyltransferase activity.[30] In a meta-analysis,[29] Wu and colleagues reported ORs for non-O relative to the O blood group of 1·79 (1·56–2·05) for venous thrombo embolism, 1·25 (1·14–1·36) for myocardial infarction, but only 1·03 (0·89–1·19) for angina.[29] Ketch and colleagues[31] reported that patients with non-O blood groups had higher thrombus burden despite less extensive coronary atherosclerosis at the time of acute myocardial infarction. These data, coupled with our genetic findings, strongly suggest a primary relation of non-O ABO glycotransferase activity with coronary thrombosis rather than atherosclerosis. ABO-related thrombosis is thought to be mediated by ABO carbohydrate-modification of von Willebrand Factor (VWF) resulting in impaired proteolysis and higher circulating von Willebrand Factor and Factor VIII.[32]

Tregouet and co-workers identified *ABO* as the most significant locus in a GWAS of venous thrombo-embolism.[33] The top *ABO* SNPs for venous thromboembolism, rs657152 and rs505922, have strong associations with myocardial infarction in patients with angiographic CAD in our data (eg, rs505922 $p=1{\cdot}032\times10^{-8}$) and are in strong linkage disequilibrium with our top *ABO* SNP signals for myocardial infarction (eg, r^2 1·0 with rs514659 (figure 4, table 3, webappendix p 6). We note that the top SNPs (rs687621 and rs687289) in a GWAS[34] of plasma von Willebrand Factor and Factor VIII are in complete linkage disequilibrium with rs514659 and blood group O and also reach genome-wide significance for myocardial infarction in patients with angiographic CAD (table 3). Thus, common *ABO* genetic variation, linked to blood group O, reduced glycotransferase activity and lower circulating von Willebrand Factor and Factor VIII, lowers risk of myocardial infarction in the setting of angiographic CAD, while also protecting against venous thromboembolism.

The relation between *ABO* and atherosclerotic cardiovascular disease, however, might be more complex than modulation of thrombosis. Other GWAS also identified *ABO* as a locus for low-density lipoprotein (LDL-C),[35] type-2 diabetes,[26] inflammatory risk biomarkers E-selectin, P-selectin, and sol-ICAM1[26,36–38] as well as angiotensin-converting enzyme[39] (figure 4). Indeed, ABO blood group associations with plasma cholesterol were described several decades ago. In PennCath, however, *ABO* SNP associations with myocardial infarction in patients with angiographic CAD were not attenuated by adjustment for *ABO* SNPs related to LDL-C, ICAM-1, and E-selectin (webappendix p 6). Taken together, these factors suggest that *ABO* might modulate various distinct pathways related to cardiovascular risk factors, atherosclerosis, and thrombosis.

Our study has potential limitations. The multistudy design might have introduced selection bias and confounding. The absence of genomic control inflation, however, argues against confounding caused by genetic differences in source populations. Angiography cannot detect early subclinical atherosclerosis in controls resulting in misclassification of controls as free of CAD. Misclassification is potentially a greater drawback in our study within patients with angiographic CAD, in which those without myocardial infarction might subsequently develop myocardial infarction. Patients with angiographic CAD who had myocardial infarction, however, tended to be younger than those who did not have myocardial infarction, despite having broadly similar risk factors. This finding suggests that additional factors beyond age and traditional risk must contribute to myocardial infarction among patients with angiographic CAD. Overall, heterogeneity and misclassification would tend to bias to wards the null, would not affect our novel findings, but would limit the power for additional discoveries.

Our results indicate that specific genetic variants predispose to the development of coronary atherosclerosis whereas others predispose to subsequent plaque rupture and acute myocardial infarction. Further, many published loci for myocardial infarction are likely to relate to the initiation and progression of coronary atherosclerosis rather than having a specific role in vulnerable plaque and myocardial infarction. Translation of GWAS discoveries for CAD into prognostic and therapeutic benefit will need greater insights into the relation between each locus and the phenotypes of atherosclerosis, plaque rupture, and thrombosis.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

All acknowledgements are provided online in pp 19–22 of the webappendix.

References

1. Helgadottir A, Thorleifsson G, Manolescu A, et al. A common variant on chromosome 9p21 affects the risk of myocardial infarction. Science. 2007; 316:1491–93. [PubMed: 17478679]
2. McPherson R, Pertsemlidis A, Kavaslar N, et al. A common allele on chromosome 9 associated with coronary heart disease. Science. 2007; 316:1488–91. [PubMed: 17478681]
3. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature. 2007; 447:661–78. [PubMed: 17554300]
4. Samani NJ, Erdmann J, Hall AS, et al. Genomewide association analysis of coronary artery disease. N Engl J Med. 2007; 357:443–53. [PubMed: 17634449]

5. Kathiresan S, Voight BF, Purcell S, et al. Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. Nat Genet. 2009; 41:334–41. [PubMed: 19198609]
6. Erdmann J, Grosshennig A, Braund PS, et al. New susceptibility locus for coronary artery disease on chromosome 3q22.3. Nat Genet. 2009; 41:280–82. [PubMed: 19198612]
7. Hansson GK. Inflammation, atherosclerosis, and coronary artery disease. N Engl J Med. 2005; 352:1685–95. [PubMed: 15843671]
8. Teslovich TM, Musunuru K, Smith AV, et al. Biological, clinical and population relevance of 95 loci for blood lipids. Nature. 2010; 466:707–13. [PubMed: 20686565]
9. Dandona S, Chen L, Fan M, et al. The transcription factor GATA-2 does not associate with angiographic coronary artery disease in the Ottawa Heart Genomics and Cleveland Clinic GeneBank Studies. Hum Genet. 2010; 127:101–05. [PubMed: 19885677]
10. Assimes TL, Knowles JW, Basu A, et al. Susceptibility locus for clinical and subclinical coronary artery disease at chromosome 9p21 in the multi-ethnic ADVANCE study. Hum Mol Genet. 2008; 17:2320–28. [PubMed: 18443000]
11. Tiret L, Godefroy T, Lubos E, et al. Genetic analysis of the interleukin-18 system highlights the role of the interleukin-18 gene in cardiovascular disease. Circulation. 2005; 112:643–50. [PubMed: 16043644]
12. Horne BD, Carlquist JF, Muhlestein JB, et al. Association of variation in the chromosome 9p21 locus with myocardial infarction versus chronic coronary artery disease. Circ Cardiovasc Genet. 2008; 1:85–92. [PubMed: 19956784]
13. Martinelli N, Olivieri O, Shen GQ, et al. Additive effect of LRP8/ APOER2 R952Q variant to APOE epsilon2/epsilon3/epsilon4 genotype in modulating apolipoprotein E concentration and the risk of myocardial infarction: a case-control study. BMC Med Genet. 2009; 10:41. [PubMed: 19439088]
14. McCarthy MI, Abecasis GR, Cardon LR, et al. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. Nat Rev Genet. 2008; 9:356–69. [PubMed: 18398418]
15. Skol AD, Scott LJ, Abecasis GR, et al. Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. Nat Genet. 2006; 38:209–13. [PubMed: 16415888]
16. Nyholt DR. A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. Am J Hum Genet. 2004; 74:765–69. [PubMed: 14997420]
17. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007; 81:559–75. [PubMed: 17701901]
18. Marchini J, Howie B, Myers S, et al. A new multipoint method for genome-wide association studies by imputation of genotypes. Nat Genet. 2007; 39:906–13. [PubMed: 17572673]
19. Devlin B, Roeder K. Genomic control for association studies. Biometrics. 1999; 55:997–1004. [PubMed: 11315092]
20. Gensini GG. A more meaningful scoring system for determining the severity of coronary heart disease. Am J Cardiol. 1983; 51:606. [PubMed: 6823874]
21. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics. 2010; 26:2190–91. [PubMed: 20616382]
22. de Bakker PI, Ferreira MA, Jia X, et al. Practical aspects of imputation-driven meta-analysis of genome-wide association studies. Hum Mol Genet. 2008; 17(R2):R122–28. [PubMed: 18852200]
23. Purcell S, Cherny SS, Sham PC. Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. Bioinformatics. 2003; 19:149–50. [PubMed: 12499305]
24. Zollner S, Pritchard JK. Overcoming the winner's curse: estimating penetrance parameters from case-control data. Am J Hum Genet. 2007; 80:605–15. [PubMed: 17357068]
25. Somerville RP, Longpre JM, Apel ED, et al. ADAMTS7B, the full-length product of the ADAMTS7 gene, is a chondroitin sulfate proteoglycan containing a mucin domain. J Biol Chem. 2004; 279:35159–75. [PubMed: 15192113]

26. Qi L, Cornelis MC, Kraft P, et al. Genetic variants in ABO blood group region, plasma soluble E-selectin levels and risk of type 2 diabetes. Hum Mol Genet. 2010; 19:1856–62. [PubMed: 20147318]
27. Wang L, Zheng J, Bai X, et al. ADAMTS-7 mediates vascular smooth muscle cell migration and neointima formation in balloon-injured rat arteries. Circ Res. 2009; 104:688–98. [PubMed: 19168437]
28. Bronte-Stewart B, Botha MC, Krut LH. ABO blood groups in relation to ischaemic heart disease. BMJ. 1962; 1:1646–50. [PubMed: 13873429]
29. Wu O, Bayoumi N, Vickers MA, et al. ABO(H) blood groups and vascular disease: a systematic review and meta-analysis. J Thromb Haemost. 2008; 6:62–69. [PubMed: 17973651]
30. Storry JR, Olsson ML. The ABO blood group system revisited: a review and update. Immunohematology. 2009; 25:48–59. [PubMed: 19927620]
31. Ketch TR, Turner SJ, Sacrinty MT, et al. ABO blood types: influence on infarct size, procedural characteristics and prognosis. Thromb Res. 2008; 123:200–05. [PubMed: 18378281]
32. Jenkins PV, O'Donnell JS. ABO blood group determines plasma von Willebrand factor levels: a biologic function after all? Transfusion. 2006; 46:1836–44. [PubMed: 17002642]
33. Tregouet DA, Heath S, Saut N, et al. Common susceptibility alleles are unlikely to contribute as strongly as the FV and ABO loci to VTE risk: results from a GWAS approach. Blood. 2009; 113:5298–303. [PubMed: 19278955]
34. Smith NL, Chen MH, Dehghan A, et al. Novel associations of multiple genetic loci with plasma levels of factor VII, factor VIII, and von Willebrand factor: The CHARGE (Cohorts for Heart and Aging Research in Genome Epidemiology) Consortium. Circulation. 2010; 121:1382–92. [PubMed: 20231535]
35. Chasman DI, Pare G, Mora S, et al. Forty-three loci associated with plasma lipoprotein size, concentration, and cholesterol content in genome-wide analysis. PLoS Genet. 2009; 5:e1000730. [PubMed: 19936222]
36. Pare G, Chasman DI, Kellogg M, et al. Novel association of ABO histo-blood group antigen with soluble ICAM-1: results of a genome-wide association study of 6,578 women. PLoS Genet. 2008; 4:e1000118. [PubMed: 18604267]
37. Paterson AD, Lopes-Virella MF, Waggott D, et al. Genome-wide association identifies the ABO blood group as a major locus associated with serum levels of soluble E-selectin. Arterioscler Thromb Vasc Biol. 2009; 29:1958–67. [PubMed: 19729612]
38. Barbalic M, Dupuis J, Dehghan A, et al. Large-scale genomic studies reveal central role of ABO in sP-selectin and sICAM-1 levels. Hum Mol Genet. 2010; 19:1863–72. [PubMed: 20167578]
39. Chung CM, Wang RY, Chen JW, et al. A genome-wide association study identifies new loci for ACE activity: potential implications for response to ACE inhibitor. Pharmacogenomics J. 2010 published online Jan 12. 10.1038/tpj.2009.70

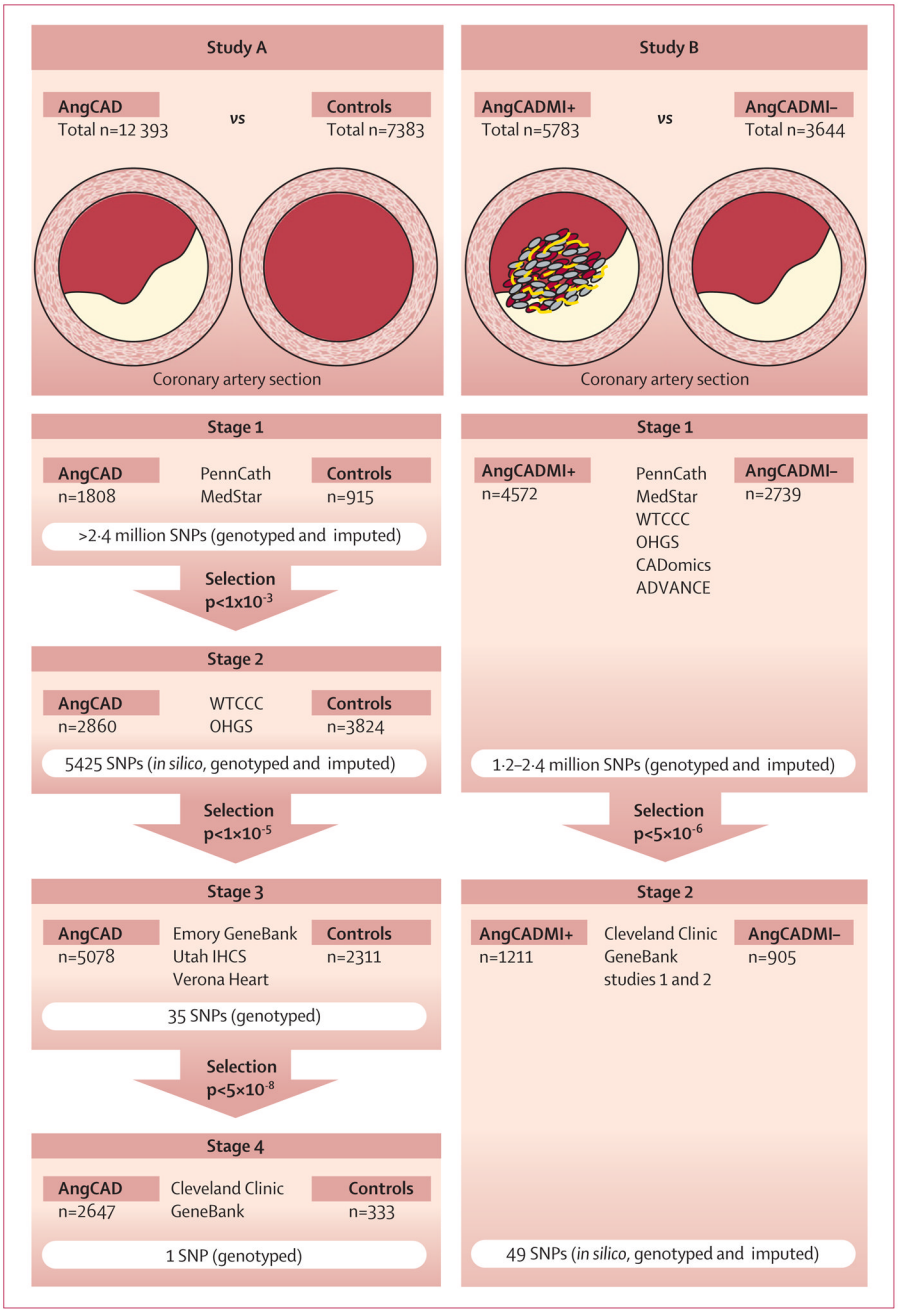

Figure 1. Study design

Cleveland Clinic GeneBank Study 1 was derived from a GWAS of CAD cases with in-stent restenosis. Cleveland Clinic GeneBank Study 2 was derived from a GWAS study of angiographic CAD cases without diabetes mellitus. GWAS=genome-wide association study. CAD=coronary artery disease. AngCADMI+=angiographic CAD with myocardial infacrtion. AngCADMI–=angiographic CAD without myocardial infarction. WTCCC=Wellcome Trust Case-Control Consortium. OHGS=Ottawa Heart Genomics Study. CADomics=Coronary Artery Disease and genomics. ADVANCE=Atherosclerotic Disease Vascular Function and Genetic Epidemiology study. Cleveland GB=Cleveland

Clinic GeneBank. Emory GB=Emory GeneBank Study. Utah IHCS=Utah Intermountain Heart Collaborative Study.

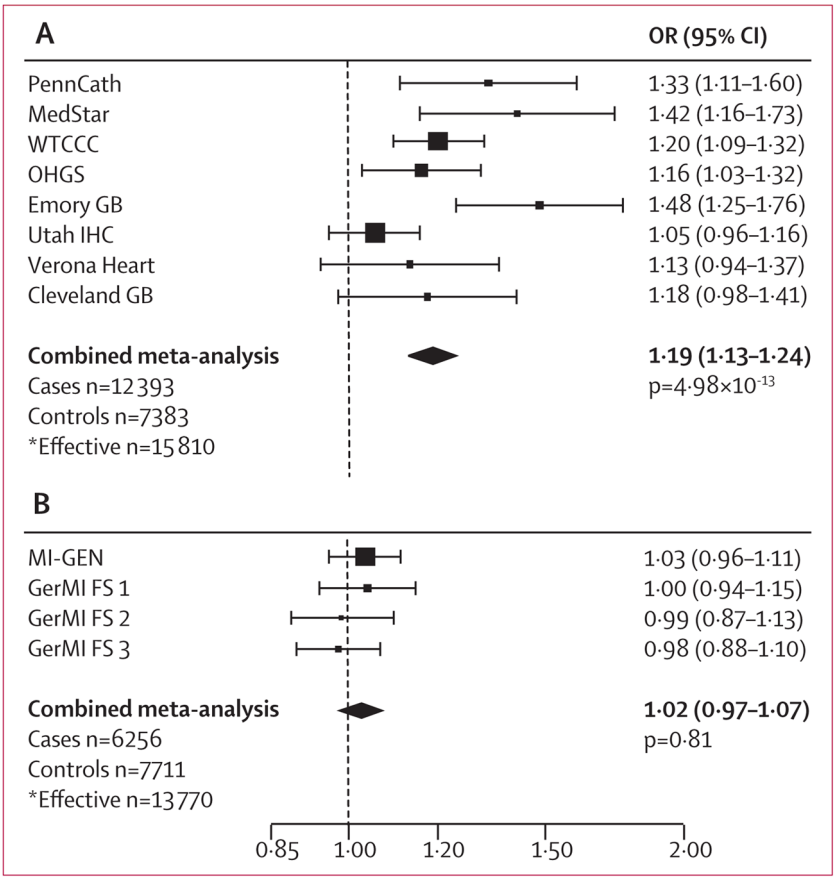

Figure 2. Association of rs1994016 at the ADAMTS7 locus with (A) angiographic CAD (AngCAD *vs* control) and (B) early-onset myocardial infarction (myocardial infarction *vs* control)

Forest plots are presented for the association of rs1994016 with (A) angiographic CAD versus controls in each individual study and overall in meta-analysis of eight studies, and (B) myocardial infarction versus control in two separate studies of early-onset myocardial infarction and in their meta-analysis. WTCCC=Wellcome Trust Case-Control Consortium. OHGS=Ottawa Heart Genomics Study. Emory GB=Emory GeneBank Study. Utah IHCS=Utah Intermountain Heart Collaborative Study. Cleveland GB=Cleveland Clinic GeneBank. MI-GEN=Myocardial Infarction Genetics Consortium. GerMI FS=German Myocardial Infarction Family Study. *We used Genetic Power Calculator[23] to estimate the non-centrality parameter for the given asymmetric case-control sample size, and then determined the effective (symmetric) case-control sample size that returns the same non-centrality parameter.

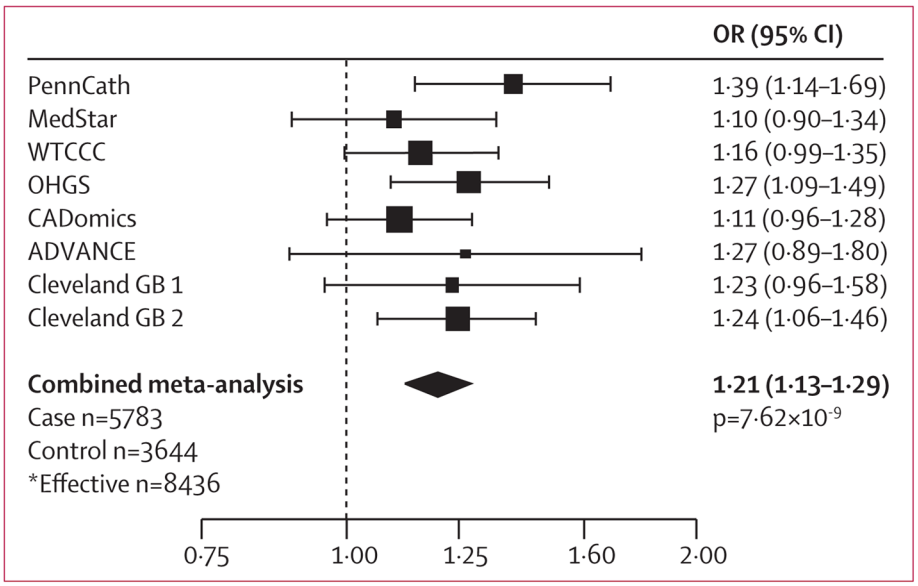

Figure 3. Association of rs514659 at the 9q34.2 ABO locus with myocardial infarction in patients with angiographic CAD (myocardial infarction versus no myocardial infarction)

Forest plot for the association of rs514659 with myocardial infarction in each individual study and overall in meta-analysis of all eight studies: PennCath, MedStar, WTCC, OHGS, CADomics, ADVANCE, and Cleveland GB 1 and 2. CAD=coronary artery disease. WTCCC=Wellcome Trust Case-Control Consortium. OHGS=Ottawa Heart Genomics Study. CADomics=Coronary Artery Disease and genomics. ADVANCE=Atherosclerotic Disease VAscular functioN and genetiC Epidemiology study. GB=GeneBank. *We used Genetic Power Calculator[23] to estimate the non-centrality parameter for the given asymmetric case-control sample size, and then determined the effective (symmetric) case-control sample size that returns the same non-centrality parameter.

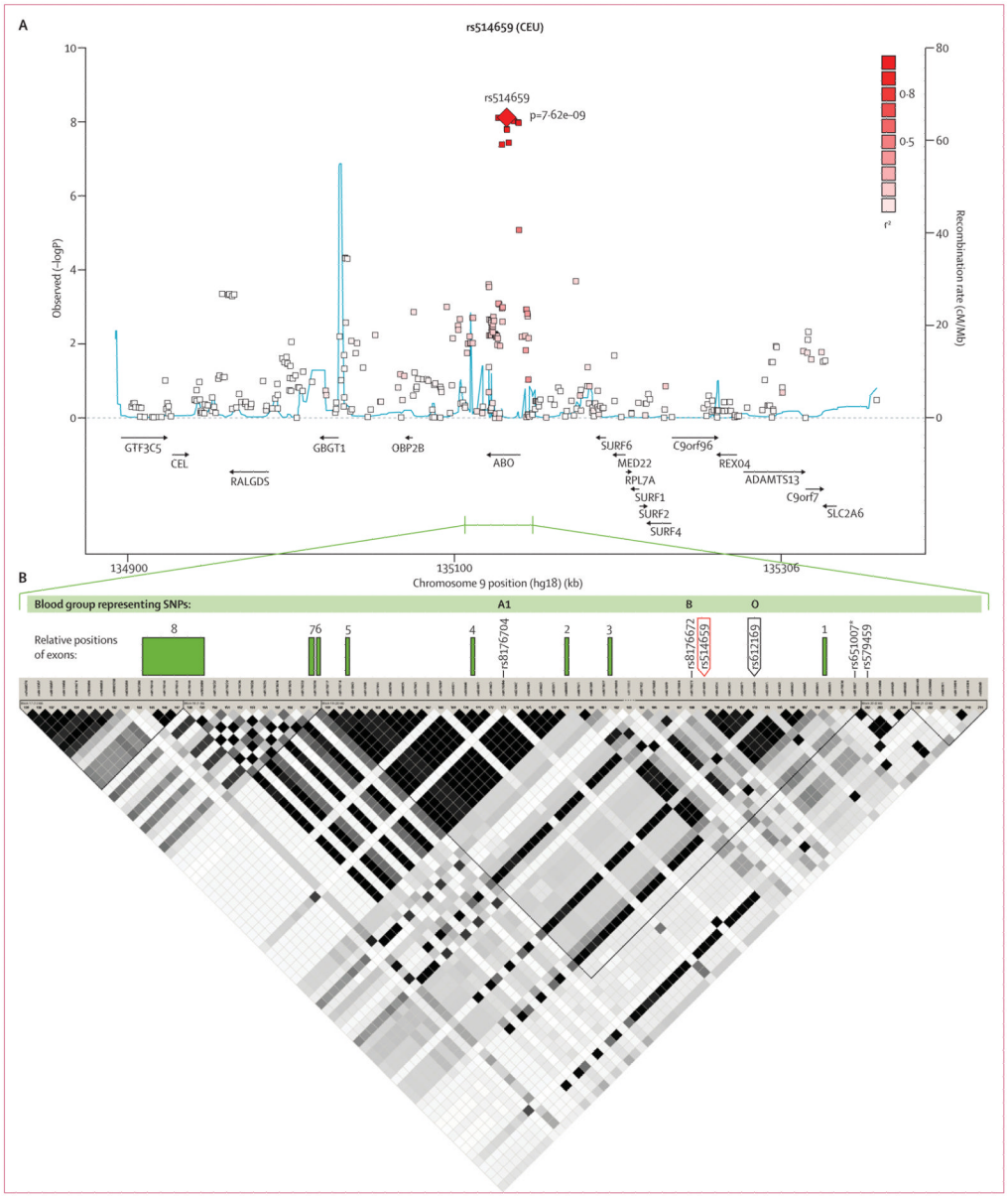

Figure 4. Association between the 9q34.2 ABO locus and myocardial infarction in patients with angiographic CAD

(A) Recombination rate, linkage disequilibrium (in r^2; colour coding of the squares correspond to strength of linkage disequilibrium), and p values for SNPs in the ABO region based on meta-analysis data from all eight studies. (B) Selected SNPs and haplotype blocks at the ABO locus. The top imputed SNP, rs514659, shown in the red box. The top genotyped SNP, rs612169, is shown in the black box; the rs612169A allele fully tags the allele of blood group O. Top SNPs for myocardial infarction in patients with angiographic CAD, including rs514659 and rs612169, lie in intron 1 of ABO and are all in tight linkage disequilibrium with the inferred blood group O allele. The rs8176672A allele tags blood group B whereas the rs8176704A allele tags blood group A1. Additional symbols mark rs651007 (*) and rs579459 (†). The rs651007 SNP tags (r^2=0·88) rs507666, which was the top SNP in GWAS of circulating concentrations of LDL-cholesterol[34] and ICAM-1.[35] The rs579459 SNP was a top SNP in GWAS of circulating levels of E-selectin and P-

selectin.[36,37] These SNPs are in moderate linkage disequilibrium with rs514659 (r^2=0·53 for rs651007, r^2=0·36 for rs579459) and had modest associations with AngCADMI+ (p=0·053 for rs651007, p=0·001 for rs579459). CAD=coronary artery disease. AngCADMI +=angiographic CAD with myocardial infarction. AngCADMI–=angiographic CAD without myocardial infarction. ICAM-1=Inter-cellular adhesion molecule 1).

Table 1

Association of published GWAS loci for myocardial infarction with angiographic CAD (study A: angiographic CAD versus control) and myocardial infarction in patients with angiographic CAD (study B: myocardial infarction versus no myocardial infarction)

	Chr	Location	Position (NCBI36)	Effect allele*	Effect allele frequency†	Gene(s)	Study A: angiographic CAD *vs* control			Study B: myocardial infarction *vs* no myocardial infarction		
							OR (95% CI)	p	Effective sample size‡	OR (95% CI)	p	Effective sample size‡
rs646776	1	1p13	109620053	T	0·79	*SORT3*	1·33 (1·22–1·44)	$7{\cdot}88\times10^{-11}$	7790	1·02 (0·94–1·11)	0·69	6916
rs17465637	1	1q41	220890152	C	0·72	*MIA3*	1·09 (1·01–1·18)	0·02708	7790	1·03 (0·95–1·11)	0·55	6360
rs11206510	1	1p32	55268627	T	0·82	*PCSK3*	1·03 (0·94–1·13)	0·4857	7790	1·02 (0·94–1·11)	0·65	8436
rs6725887	2	2q33	203454130	C	0·13	*WDR33*	1·28 (1·15–1·45)	$1{\cdot}74\times10^{-5}$	6454	0·99 (0·91–1·08)	0·87	8436
rs9818870	3	3q22	139604812	T	0·16	*MRAS*	1·15 (1·03–1·27)	0·009773	6656	0·96 (0·89–1·04)	0·25	8436
rs12526453	6	6q24	13035530	C	0·66	*PHACTR3*	1·17 (1·09–1·26)	$1{\cdot}45\times10^{-5}$	7790	1·05 (0·98–1·12)	0·20	8196
rs49775·74	9	9p21	22088574	G	0·53	*ANRIL, CDKN3A, CDKN3B*	1·3 (1·22–1·39)	$4{\cdot}78\times10^{-14}$	7790	1·00 (0·94–1·06)	0·96	8436
rs1746048	10	10q11	44095830	C	0·86	*CXCL33*	1·08 (0·97–1·18)	0·1681	7790	1·04 (0·95–1·14)	0·42	8436
rs11066301	12	12q24.3	111355755	G	0·46	*ATXN3, SH3B3, PTPN33*	§	§	§	1·03 (0·96–1·10)	0·42	6676
rs1122608	19	19p13	11024601	G	0·75	*LDLR*	1·18 (1·09–1·28)	0·0001568	6454	0·98 (0·91–1·05)	0·62	8436
rs9982601	21	21q22	34520998	T	0·13	*SLC3A3, MRPS3, KCNE3*	1·18 (1·07–1·30)	0·001106	7790	1·07 (0·97–1·18)	0·15	6676

(A) The meta-analysis of angiographic CAD versus control was restricted to those studies that ascertained both cases and controls via coronary angiography. These were PennCath, MedStar, Emory GeneBank, the Intermountain Heart Collaborative Study, and Verona Heart Study. (B) The meta-analysis of myocardial infarction versus no myocardial infarction in patients with angiographic CAD was done with PennCath, MedStar, Wellcome Trust Case Control Consortium, Ottawa Heart Genomics Study, Coronary Artery Disease and genomics, Atherosclerotic Disease, Vascular Function, and Genetic Epidemiology, and Cleveland Clinic GeneBank Studies 1 and 2. CAD=coronary artery disease. Chr=chromosome. OR=odds ratio.

*Effect allele was based on risk allele in previously published studies where available.[5]

†Effect allele frequency was based on their frequencies in PennCath.

‡Effective sample size: we used Genetic Power Calculator[39] to estimate the non-centrality parameter (NCP) for the specific asymmetric case-control sample size, and then determined the effective (symmetric) case-control sample size that returns the same non-centrality parameter and used this in meta-analysis.

§rs11066301 was not genotyped in most studies of angiographic CAD versus controls.

Table 2

Study A: Association of ADAMTS7 rs1994016 with angiographic CAD (angiographic CAD versus control)

	Chr	Position (NCBI-36)	Risk allele	Risk allele frequency	Gene	Stage 1	Stage 2	Stage 3	Stages 1, 2, 3	Stage 4	All stages combined
						OR (95% CI) p	OR (95% CI) p	OR (95% CI) p	OR (95% CI) p	OR (95% CI) p	OR (95% CI) p
rs1994016	15	76867289	C	0·60	ADAMTS7	1·37 (1·20–1·57) $3·94\times10^{-6}$	1·19 (1·10–1·28) $7·01\times10^{-6}$	1·14 (1·06–1·23) 0·000497	1·19 (1·13–1·25) $2·41\times10^{-12}$	1·18 (0·98–1·41) 0·0767	1·19 (1·13–1·24) $4·98\times10^{-13}$
Sample size											
Number of cases	..	..	..	..	..	1808	2860	5078	9746	2647	12 393
Number of controls	..	..	..	..	..	915	3824	2311	7050	333	7383
Effective sample number*	..	..	..	..	..	2380	5910	6350	14 640	1170	15 810

Chr = chromosome. OR = odds ratio. CI = confidence interval. Stage 1 corresponds to PennCath and MedStar; stage 2 corresponds to Wellcome Trust Case Control Consortium and Ottawa Heart Genomics Study; stage 3 corresponds to Emory Genebank, Utah Intermountain Heart Collaborative Study, and Verona Heart Study; and stage 4 corresponds to Cleveland Clinic GeneBank.

*We used Genetic Power Calculator[39] to estimate the non-centrality parameter for the given asymmetric case-control sample size, and then determined the effective (symmetric) case-control sample size that returns the same non-centrality parameter and used this in meta-analysis.

Table 3

Study B. Association of *ABO* SNPs with myocardial infarction in patients with angiographic CAD (myocardial infarction versus no myocardial infarction)

	Position NCBI 36	r^2 with rs514659	r^2 with allele O	r^2 with allele A	r^2 with allele B	Non-risk allele	Risk allele	Risk allele frequency	Stage 1 (PennCath, MedStar, WTCCC, OHGS, CADomics, ADVANCE)		Stage 2 (Cleveland Clinic GeneBank 1 and 2)		Combined stage 1 and 2	
									OR (95% CI)	p value	OR (95% CI)	p value	OR (95% CI)	p value
rs514659*	135,132,024	..	0·97	0·64	0·18	A	C	0·37	1·20 (1·11–1·29)	$1·23\times10^{-6}$	1·24 (1·09–1·42)	$1·45\times10^{-3}$	1·21 (1·13–1·29)	$7·62\times10^{-9}$
rs687289*	135,126,927	1·0	0·97	0·64	0·18	G	A	0·37	1·20 (1·11–1·29)	$1·11\times10^{-6}$	1·23 (1·08–1·41)	$1·69\times10^{-3}$	1·20 (1·13–1·29)	$7·75\times10^{-9}$
rs687621*	135,126,886	1·0	0·97	0·64	0·18	A	G	0·37	1·20 (1·11–1·29)	$1·11\times10^{-6}$	1·24 (1·08–1·41)	$1·70\times10^{-3}$	1·20 (1·13–1·28)	$7·81\times10^{-9}$
rs545971*	135,133,193	1·0	0·97	0·64	0·18	C	T	0·37	1·20 (1·11–1·29)	$1·33\times10^{-6}$	1·24 (1·09–1·42)	$1·60\times10^{-3}$	1·20 (1·13–1·29)	$8·96\times10^{-9}$
rs674302*	135,136,485	1·0	0·97	0·64	0·18	T	A	0·37	1·20 (1·11–1·28)	$1·33\times10^{-6}$	1·24 (1·08–1·42)	$1·69\times10^{-3}$	1·20 (1·13–1·28)	$9·59\times10^{-9}$
rs505922*	135,139,050	1·0	0·97	0·64	0·18	T	C	0·37	1·19 (1·11–1·28)	$1·45\times10^{-6}$	1·24 (1·08–1·42)	$1·70\times10^{-3}$	1·20 (1·13–1·28)	$1·03\times10^{-8}$
rs529565*	135,139,321	1·0	0·97	0·64	0·18	T	C	0·37	1·19 (1·11–1·28)	$1·49\times10^{-6}$	1·24 (1·08–1·42)	$1·71\times10^{-3}$	1·20 (1·12–1·28)	$1·07\times10^{-8}$
rs644234*	135,132,038	0·91	0·89	0·58	0·16	T	G	0·39	1·20 (1·12–1·29)	$6·13\times10^{-7}$	1·20 (1·05–1·36)	$7·73\times10^{-3}$	1·20 (1·13–1·28)	$1·57\times10^{-8}$
rs643434*	135,132,176	0·91	0·89	0·58	0·18	G	A	0·39	1·20 (1·12–1·29)	$6·30\times10^{-7}$	1·20 (1·05–1·36)	$7·87\times10^{-3}$	1·20 (1·13–1·28)	$1·64\times10^{-8}$
rs612169	135,133,263	0·98	0·98	0·64	0·16	A	G	0·37	1·19 (1·10–1·28)	$4·88\times10^{-6}$	1·24 (1·08–1·42)	$1·69\times10^{-3}$	1·20 (1·12–1·28)	$3·66\times10^{-8}$
rs657152	135,129,086	0·89	0·88	0·58	0·16	C	A	0·40	1·19 (1·11–1·28)	$1·50\times10^{-6}$	1·19 (1·04–1·36)	$8·33\times10^{-3}$	1·19 (1·12–1·27)	$4·11\times10^{-8}$
Sample size														
Number of cases	..	..	..	..	..	..	..	..	4572		1211		5783	
Number of controls	..	..	..	..	..	..	..	..	2739		905		3644	
Effective sample number*	..	..	..	..	..	..	..	..	6600		1836		8436	

WTCCC = Wellcome Trust Case Control Consortium. OHGS=Ottawa Heart Genomics Study. CADomics=Coronary Artery Disease and genomics. ADVANCE=Atherosclerotic Disease, Vascular function, and Genetic Epidemiology. OR=Odds ratio; CI=confidence interval. SNPs=single-nucleotide polymorphism. Cleveland Clinic Genebank Study 1 was derived from a GWAS of CAD cases with in-stent restenosis. Cleveland Clinic Genebank Study 2 was derived from a GWAS study of angiographic CAD cases without diabetes mellitus.

*We used Genetic Power Calculator[39] to estimate the non-centrality parameter for the given asymmetric case-control sample size, and then determined the effective (symmetric) case-control sample size that returns the same non-centrality parameter and used this in meta-analysis.

Table 4

Association of (A) the ABO rs514659C allele and (B) ABO blood groups with CAD phenotypes in PennCath

	AngCADMI+ *vs* angCADMI–*		AngCADMI+ *vs* control*		AngCADMI– *vs* control*		All AngCAD *vs* control*	
	OR (95% CI)	**p**	**OR (95% CI)**	**p**	**OR (95% CI)**	**p**	**OR (95% CI)**	**p**
(A) rs514659								
C allele	1·39 (1·14–1·69)	0·001	1·33 (1·06–1·66)	0·01	1·00 (0·81–1·25)	0·97	1·13 (0·93–1·37)	0·22
(B) ABO Blood group†								
A *vs* O	1·49 (1·11–2·00)	0·01	1·37 (1·02–1·84)	0·04	0·92 (0·69–1·23)	0·57	1·11 (0·86–1·43)	0·42
B *vs* O	2·15 (1·37–3·40)	0·0005	1·78 (1·15–2·78)	0·0081	0·83 (0·51–1·34)	0·42	1·24 (0·84–1·85)	0·30
AB *vs* O	1·65 (0·85–3·25)	0·12	1·35 (0·71–2·60)	0·36	0·82 (0·41–1·62)	0·63	1·05 (0·60–1·88)	0·89
A/B/AB *vs* O	1·62 (1·23–2·13)	0·0004	1·44 (1·10–1·90)	0·01	0·89 (0·68–1·17)	0·43	1·13 (0·90–1·43)	0·30

CAD=coronary artery disease. OR=odds ratio. AngCADMI+=angiographic CAD with myocardial infarction.

AngCADMI–=angiographic CAD without myocardial infarction.

*N=933 for all cases of angiographic CAD; N=470 for cases of angiographic CAD with myocardial infarction; N=463 for cases of angiographic CAD without myocardial infarction; N=468 for controls.

†The number of individuals with blood groups A, B, AB, or O is 571, 162, 68, and 578, respectively. Association analysis is conducted assuming an additive model, and allele A for rs514659 is coded as the reference allele for OR calculation.