Genome-wide Analysis Identifies Novel Susceptibility Loci for Myocardial Infarction

Jaana A. Hartiala[1,2]*, Yi Han[1,2]*, Qiong Jia[1,2], Pin Huang[1,2], Janet Gukasyan[1,2], William S. Schwartzman[1,2], David-Alexandre Trégouët[3], Nicholas L. Smith[4,5,6], The INVENT Consortium[7], The CHARGE Consortium Hemostasis Working Group[7], The GENIUS-CHD Consortium[7], Marcus Seldin[8], Calvin Pan[9], Margarete Mehrabian[10], Aldons J. Lusis[9,10,11], Peter Bazeley[12,13,14], Yan V. Sun[15,16], Chang Liu[15,17], Arshed A. Quyyumi[17], Markus Scholz[18,19], Joachim Thiery[19,20], Graciela E. Delgado[21], Marcus E. Kleber[21], Winfried März[21,22,23], Laurence J. Howe[24], Folkert W. Asselbergs[24,25,26], Riyaz S. Patel[24,27], Leo-Pekka Lyytikäinen[28,29,30], Mika Kähönen[31,32], Terho Lehtimäki[28,29], Tuomo V.M. Nieminen[33], Pekka Kuukasjärvi[34], Jari O. Laurikka[34,35], Lindsey K. Stolze[36], Casey E. Romanoski[36], Mohammad Daud Khan[37,38], Adam W. Turner[37,38], Clint Miller[37,38], Redouane Aherrahrou[37,38], Mete Civelek[37,38], Lijiang Ma[39], Johan L.M. Bjorkegren[39,40], W.H. Wilson Tang[12,13,14], Stanley L. Hazen[13,14], and Hooman Allayee[1,2]

*These authors contributed equally to this work.

Departments of [1]Preventive Medicine and [2]Biochemistry & Molecular Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA 90033; [3]Institut National pour la Santé et la Recherche Médicale (INSERM) UMR_S 1219, Bordeaux Population Health Research Center, University of Bordeaux, France; [4]Department of Epidemiology, University of Washington, Seattle, WA; [5]Seattle Epidemiologic Research and Information Center, Office of Research and Development, Department of Veterans Affairs, Seattle WA; [6]Kaiser Permanente Washington Health Research Institute, Kaiser Permanente Washington, Seattle WA; [7]A full list

of members in the INVENT Consortium, the CHARGE Consortium Hemostasis Working Group, and the GENIUS-CHD Consortium is provided in the supplementary materials; [8]Department of Biological Chemistry and Center for Epigenetics and Metabolism, UC Irvine School of Medicine, Irvine, CA 92697; Departments of Human Genetics[9], Medicine[10], and Microbiology, Immunology, & Molecular Genetics[11], David Geffen School of Medicine of UCLA, Los Angeles, CA 90095; Center for Clinical Genomics[12] and Departments of Cardiovascular Medicine[13] and Cardiovascular & Metabolic Sciences[14], Cleveland Clinic, Cleveland, OH 44195; [15]Department of Epidemiology, Emory University Rollins School of Public Health, 1518 Clifton Rd. NE, Atlanta, GA 30322; [16]Department of Biomedical Informatics, Emory University School of Medicine, Atlanta, GA 30322; [17]Division of Cardiology, Department of Medicine, Emory Clinical Cardiovascular Research Institute, Emory University School of Medicine. 1462 Clifton Rd NE, Suite # 513, Atlanta, GA 30322; [18]Institute for Medical Informatics, Statistics and Epidemiology, University of Leipzig, Leipzig, Germany; [19]LIFE Research Center for Civilization Diseases, University of Leipzig, Leipzig, Germany; [20]Institute of Laboratory Medicine, Clinical Chemistry and Molecular Diagnostics, University Hospital, Leipzig, Germany; [21]Vth Department of Medicine, Medical Faculty Mannheim, Heidelberg University, Theodor-Kutzer-Ufer 1-3, 68167 Mannheim, Germany; [22]SYNLAB Academy, SYNLAB Holding Deutschland GmbH, Mannheim, Germany; [23]Clinical Institute of Medical and Chemical Laboratory Diagnostics, Medical University of Graz, Graz, Austria; [24]Institute of Cardiovascular Science, Faculty of Population Health Sciences, University College London, London, UK; [25]Department of Cardiology, Division Heart & Lungs, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands; [26]Health Data Research UK and Institute of Health Informatics, University College London, London, UK; [27]Bart's Heart Centre, St Bartholomew's

Hospital, London, EC1A2DA, UK; [28]Department of Clinical Chemistry, Fimlab Laboratories, Tampere 33520, Finland; [29]Department of Clinical Chemistry, Finnish Cardiovascular Research Center - Tampere, Faculty of Medicine and Health Technology, Tampere University, Tampere 33014, Finland; [30]Department of Cardiology, Heart Center, Tampere University Hospital, Tampere 33521, Finland; [31]Department of Clinical Physiology, Tampere University Hospital, Tampere 33521, Finland; [32]Department of Clinical Physiology, Finnish Cardiovascular Research Center - Tampere, Faculty of Medicine and Health Technology, Tampere University, Tampere 33014, Finland; [33]Department of Internal Medicine, Päijät-Häme Central Hospital, Lahti, Finland; [34]Department of Cardio-Thoracic Surgery, Finnish Cardiovascular Research Center - Tampere, Faculty of Medicine and Health Technology, Tampere University, Tampere 33014, Finland; [35]Department of Cardio-Thoracic Surgery, Heart Center, Tampere University Hospital, Tampere 33521, Finland; [36]Department of Cellular and Molecular Medicine, University of Arizona College of Medicine, Tucson, AZ 85721; [37]Center for Public Health Genomics and [38]Department of Biomedical Engineering, University of Virginia, Charlottesville, VA ; [39]Department of Genetics & Genomic Sciences, Institute of Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York, NY, 10029; [40]Integrated Cardio Metabolic Centre, Department of Medicine, Karolinska Institutet, Karolinska Universitetssjukhuset, Huddinge, Sweden.

Address correspondence and reprint requests to:

Hooman Allayee, PhD
Department of Preventive Medicine
Keck School of Medicine of USC
2250 Alcazar Street, CSC202
Los Angeles, CA 90033
Phone: (323) 442-1736/Fax: (332) 442-2764
Email: hallayee@usc.edu

Abstract

Aims: While most patients with myocardial infarction (MI) have underlying coronary atherosclerosis, not all patients with coronary artery disease (CAD) develop MI. We sought to address the hypothesis that some of the genetic factors which establish atherosclerosis may be distinct from those that predispose to vulnerable plaques and thrombus formation.

Methods and Results: We carried out a genome-wide association study for MI in the UK Biobank (n~472,000), followed by a meta-analysis with summary statistics from CARDIoGRAM+C4D (total n~639,000). Multiple independent replication analyses and functional approaches were then used to prioritize loci and evaluate positional candidate genes. Eight novel regions were identified for MI at the genome wide significance level, of which effect sizes at six loci were more robust for MI than for CAD in the absence of MI. Confirmatory evidence for association of the chromosome 1p21.3 locus harboring choline-like transporter 3 (*SLC44A3*) with MI in the context of CAD, but not with coronary atherosclerosis itself, was obtained in six independent angiography-based cohorts (n~14,000). Follow-up analyses did not reveal association of the *SLC44A3* locus with other thrombotic diseases, biomarkers of coagulation, or plasma choline and trimethylamine *N*-oxide levels. However, *SLC44A3* expression was increased in carriers of the risk allele, increased in ischemic (vs. non-diseased) coronary arteries, upregulated in human aortic endothelial cells treated with interleukin-1β (vs. vehicle), and associated with the migration rate of smooth muscle cells.

Conclusions: The present findings implicate *SLC44A3* as an MI-specific genetic risk factor and suggest its role in the pathophysiology of vulnerable plaques.

Keywords: myocardial infarction; genetic factors; genome-wide association study; meta-analysis; *SLC44A3*.

Introduction

Myocardial infarction (MI) and coronary artery disease (CAD) are the leading causes of death in Western societies[1], even in the contemporary era of high-potency statin therapy[2]. Individuals with CAD are typically asymptomatic, with the first manifestations often being major adverse clinical events, such as MI, stroke, or sudden death, due to the rupture of an atherosclerotic plaque[3]. Thus, understanding the biological mechanisms that precipitate plaque rupture and thrombosis could have important clinical implications since it may lead to earlier detection or better prediction of the transition from a stable lesion to a vulnerable plaque.

It is generally accepted that common forms of MI and CAD are characterized by heritable susceptibility factors in the context of lifetime exposure to an atherogenic environment. Consistent with this notion, large-scale studies in humans have identified >160 loci that influence risk of MI and CAD through unknown mechanisms and via perturbations of lipid metabolism, blood pressure regulation, inflammation, and platelet function[4-9]. However, the susceptibility alleles identified to date, most of which are common in the population, still only explain a small fraction of the overall heritability for these cardiovascular phenotypes. Furthermore, even though the vast majority of patients with MI have underlying coronary atherosclerosis, not all patients with coronary atherosclerosis develop MI. This observation suggests that some of the mechanisms that establish atherosclerosis or drive its progression may be distinct from those that predispose to plaque vulnerability and thrombus formation. Genetic studies support this concept as well. For example, *9p21* is one of the most strongly associated loci for CAD but is not directly associated with MI when comparing CAD-positive/MI-negative individuals with those who are CAD-positive/MI-positive[10, 11]. By contrast, the same analytical approach initially identified *ABO*, which defines the common ABO blood group system, as being

associated with MI among individuals with CAD but not necessarily with the presence of coronary atherosclerosis itself[10]. Thus, even though nearly all loci identified to date for CAD are also associated with MI, it is likely that additional genetic factors predisposing more robustly or specifically to plaque rupture and thrombotic phenotypes exist as well. However, with the exception of *ABO*, no other such locus has been identified. In the present study, we sought to further explore the genetic architecture of MI and address the hypothesis that distinct genetic risk factors may underlie susceptibility to MI and CAD.

Methods

Study Populations. The UK Biobank recruited participants between 40-69 years of age who were registered with a general practitioner of the UK National Health Service (NHS). From 2006-2010, a total of 503,325 individuals were included. All study participants provided informed consent and the study was approved by the North West Multi-centre Research Ethics Committee. Detailed methods used by the UK Biobank have been described elsewhere[12]. Individual cohorts in the CARDIoGRAM+C4D Consortium 1000 Genomes GWAS studies have been previously described in detail[4]. All subjects gave written consent for participation in genetic studies, and the protocol of each study was approved by the corresponding local research ethics committee or institutional review board. The present study was approved by the Institutional Review Boards of the Cleveland Clinic and USC Keck School of Medicine.

Clinical Definitions in the UK Biobank. MI cases were defined as positive for International Classification of Diseases version-10 (ICD10) codes: I21, I22, I23, I25.2, which included myocardial infarction (MI), and complications following acute MI. Doctor-diagnosed and self-reported MI were also included in the definition. CAD cases were defined in two ways. The first was an all-inclusive definition that included subjects positive for MI according to the criteria described above, as well as positive for other ICD-10 codes I24.0, I24.8, I24.9, I25.0, I25.1, I25.4, I25.8, and I25.9, which included ischemic heart diseases. Office of Population Censuses and Surveys Classification of Interventions and Procedures, version 4 (OPCS-4) codes K40-K46, K49, K50 and K75, covering replacement, transluminal balloon angioplasty, other therapeutic transluminal operations on coronary artery and percutaneous transluminal balloon angioplasty and insertion of stent into coronary artery. The second definition of CAD was more restricted and included subjects who were only positive for CAD and not for MI. To avoid any

misclassification of cases, we also excluded 59 samples who were only positive for ICD10 codes I24.1 (Dressler's syndrome), I25.3 (aneurysm of heart), I25.5 (ischaemic cardiomyopathy), or I25.6 (silent myocardial ischaemia). This strategy resulted in the classification of 17,505 MI cases, 15,580 CAD only cases, 33,085 all-inclusive CAD cases, and 454,212 controls who were not positive for any of the clinical designations used to define CAD and MI.

GWAS Analyses in the UK Biobank. Quality control of samples and DNA variants and imputation were performed by the Wellcome Trust Centre for Human Genetics, as described in detail elsewhere[12]. Briefly, ~90 million single nucleotide polymorphisms (SNPs) imputed from the Haplotype Reference Consortium, UK10K, and 1000 Genomes imputation were available in the UK Biobank. Of these, 10,903,881 variants were used for GWAS analysis after filtering on autosomal SNPs with INFO scores >0.8 (directly from the UK Biobank) and with minor allele frequencies (MAF) >0.5% in the 487,379 individuals with imputed genotypes. A GWAS analysis for MI was performed with BOLT-LMM V2.3.2 using a standard (infinitesimal) mixed model to correct for structure due to relatedness, ancestral heterogeneity, with adjustment for age, sex, the first 20 principal components, and genotyping array[13]. The genome-wide significance threshold was set at $P=5.0x10^{-8}$. Since BOLT-LMM relies on linear models even for qualitative traits, SNP effect size estimates on the quantitative scale were transformed to obtain odds ratios (ORs) and standard errors (SEs) using the following formula: β or $SE/(\mu * (1 - \mu))$, where μ = case fraction[13].

Meta-analysis for MI. Publicly available summary statistics for MI with 9,289,491 SNPs from the CARDIoGRAM+C4D Consortium[4] were combined with our GWAS results for MI in the UK Biobank. We carried out a fixed-effects meta-analysis with 8,126,035 SNPs common to both

datasets assuming an additive model, as implemented in GWAMA[14]. The genome-wide threshold for significant association was set at $P=5.0x10^{-8}$. A locus was defined as novel if our lead SNP was >1Mb away or in weak or no linkage disequilibrium ($r^2 \leq 0.1$) with the lead variants at the 162 previously reported loci for CAD (which also included MI as a criteria for case status)[9]. Replication of the 162 known CAD loci was considered significant at a Bonferroni-corrected threshold of $P=3.0x10^{-4}$ (0.05/162).

Comparison of Association Signals at Novel Loci. To compare the association signals at the eight novel loci, we carried out several different analyses. First, we determined the association of the lead SNPs at each MI locus with CAD in the UK Biobank. These analyses were carried out with BOLT-LMM V2.3.2 using a standard (infinitesimal) mixed model to correct for structure due to relatedness, ancestral heterogeneity, with adjustment for age, sex, the first 20 principal components, and genotyping array[13]. A fixed-effects meta-analysis was then carried out for the eight novel loci with the CAD association results from the UK Biobank and those from CARDIoGRAM+C4D under an additive model, as implemented in GWAMA[14]. As another approach, we also used primary level data in the UK Biobank alone to carry out additional case-control analyses. The first compared individuals without MI or CAD (controls, n=454,212) to CAD only cases (n=15,580). The second cases-control analysis tested the eight novel loci for association in subjects with both MI and CAD (cases, n=17,505) compared to subjects who only had CAD ("controls", n=15,580).

Replication Analyses for MI Loci. To replicate the association of the novel MI loci and determine their phenotypic specificity, we carried out a case-control analysis in several cohorts with angiographically-documented CAD patients of northern European ancestry with (cases) and

without (controls) adjudicated prior diagnoses of MI. Detailed descriptions of the angiography-based cohorts are provided in the **Supplemental Materials**. Association of lead SNPs at the identified loci with MI in the presence of coronary atherosclerosis was carried out in each cohort individually using logistic regression with adjustment for age and sex. Analyses in the ANGES/FINCAVAS cohort additionally adjusted for smoking status, hyperlipidemia, statin use, hypertension, and diabetes status. This was followed by a fixed-effects meta-analysis with all angiography-based cohorts using the 'meta' package[15] in R (v3.6.0).

Metabolomics Analyses with the *SCL44A3* Locus. Association of the *SCL44A3* locus with plasma levels of proatherogenic metabolites was evaluated in the GeneBank cohort[16-22]. Briefly, metabolites were quantified using stable isotope dilution high performance liquid chromatography (HPLC) with online electrospray ionization tandem mass spectrometry. Stable isotope labeled internal standards for each monitored analyte were added to plasma samples prior to protein precipitation and monitored at the appropriate transitions in in multiple reaction monitoring mode using characteristic parent-daughter ion transitions at m/z ratios for each metabolite. Genotypes for rs12743267 were obtained in subjects profiled on either the Affymetrix 6.0 GeneChip or Illumina Infinium Global Screening Array. Linear regression analyses for each metabolite was carried out using natural-log transformed values, with adjustment for age and sex (STATA v15.0, StataCorp LP, Texas, USA).

Expression Analyses of *SCL44A3* in the STARNET cohort. The Stockholm-Tartu Atherosclerosis Reverse Networks Engineering Task study (STARNET) recruited 600 patients with CAD who were eligible for open-thorax surgery at the Department of Cardiac Surgery,

Tartu University Hospital in Estonia[23]. After providing informed consent and at the time of surgery, venous blood (n=559) and tissue biopsies were obtained from atherosclerotic aortic artery (n=515), mammary artery (n=520), visceral (n=247) and subcutaneous (n=237) adipose, liver (n=257), and skeletal muscle (n=294). RNA sequencing was performed on the Illumina TruSeq platform and genotyping was performed using the Illumina Infinium assay with the human OmniExpressExome-8v1 bead chip[23]. Association of the lead SNP (rs12743267) at the chromosome 1p21.3 locus with *SLC44A3* expression levels in aorta and mammary artery was computed by one-way ANOVA in R (v3.6.0) among all subjects and separately in subjects with and without MI.

Expression Analyses of *SLC44A3* in Human Coronary Artery and Aortic Endothelial Cells. De-identified coronary artery tissue samples were obtained from heart transplant donors with CAD or from non-diseased donor hearts rejected for orthotopic heart transplantation at the Stanford University School of Medicine. All subjects provided consent for participation in research studies under an IRB-approved protocol. Proximal coronary artery segments from main branches of the left anterior descending, circumflex or right coronary arteries were dissected from all donor hearts. Clinical and histopathology information from diseased hearts was used to classify ischemic and non-ischemic arteries. All normal arteries originated from non-diseased hearts with left ventricular ejection fraction (LVEF) greater than 50%. Total RNA was extracted from frozen coronary artery segments using the miRNeasy Mini RNA Extraction kit (Qiagen, catalog # 217004, Germantown, MD) and subjected to 150bp paired-end sequencing on an Illumina NovaSeq S4 Flowcell. Expression levels of *SLC44A3* were compared between ischemic (n=36) and normal (n=24) coronary arteries using DEseq2 (v3.1)[24], after correcting for age, sex, RIN Score, ethnicity, and hidden confounding variables.

Human aortic endothelial cells (HAECs) were isolated from aortic explants of anonymous donors (n=149) through the UCLA heart transplant program[25, 26]. Cells were grown to 90% confluence and incubated for 4hrs in M-199 medium (catalog # MT10–060-CV, ThermoFisher Scientific, Waltham, MA) supplemented with 1.2% sodium pyruvate (catalog # 11360070, ThermoFisher Scientific, Waltham, MA), 1% 100X pen strep glutamine (catalog # 10378016, ThermoFisher Scientific, Waltham, MA), 20% fetal bovine serum (FBS, GE Healthcare, Hyclone, Pittsburgh, PA), 1.6% endothelial cell growth serum (Product #356006, Corning, Corning, NY), 1.6% heparin, and 10µl/50ml amphotericin B (catalog #15290018, ThermoFisher Scientific, Waltham, MA). Donor HAECs from up to 53 individuals were expanded at 5% CO_2 at 37ºC and subsequently treated for 4 hours in media containing 1% FBS with and without 10ng/ml human recombinant interleukin (IL)-1β protein (catalog # 201-LB-005/CF, R&D Systems, Minneapolis, MN). RNA was extracted from approximately $5x10^6$ cells using the Quick-RNA MicroPrep kit (catalog # R1051, Zymo Research, Irvine, CA) for preparation of libraries, as described previously[27], and subjected to sequencing on an Illumina HiSeq 4000 according to the manufacturer's specifications. Expression levels of *SLC44A3* in HAECs with and without IL-1β treatment were compared using DEseq[24].

***In Vitro* Smooth Muscle Cell Migration Assay:** Vascular smooth muscle cells (SMCs) were isolated from aortic explants of anonymous donors (n=151) through the UCLA heart transplant program. A portion of these donors overlapped with those from whom HAECs were isolated. The cells were grown in Smooth Muscle Cell Basal Medium (catalog # CC-3181, Lonza, Basel, Switzerland) supplemented with the Smooth Muscle Growth Medium-2 SingleQuots Kit (SmGM-2, catalog # CC-4149, Lonza, Basel, Switzerland). Migration of SMCs was monitored continuously for 24 hours using electronically integrated 16-well Boyden chamber plates (CIM-

plate 16, catalog # 5665817001, ACEA Biosciences, San Diego, CA) with 8μm pores in an xCelligence Real-Time Cell Analysis Instrument (ACEA Biosciences, San Diego, CA) placed in a 5% CO_2 humidified incubator and maintained at 37°C. The upper chamber of each well was seeded with $3x10^5$ cells in serum-free media whereas the lower chamber contained either serum-free media (control) or serum-free media supplemented with 100ng/mL PDGF-BB as the chemoattractant. Changes in electrical impedance as the cells migrated to the lower chamber was monitored and translated to an index that was proportional to the number of migrating cells. All experiments were performed in quadruplicate. For each donor, the response to PDGF-BB was calculated by estimating the difference in the area-under-the-curve (AUC) between PDGF-BB and control media (migration response difference). The difference in migration rate was also determined by calculating the difference in the slopes between PDGF-BB and control media before the cells reached maximum migration.

Results

Identification of 8 Novel Loci for MI: We sought to expand our understanding of the genetic architecture of MI through a large-scale genome-wide analysis with the UK Biobank and the CARDIoGRAM+C4D Consortium[4] (**Figure 1** and **Supplementary Table 1**). We first carried out a genome-wide association study (GWAS) for MI in the UK Biobank with 17,505 cases and 454,212 controls (**Supplemental material, Extended Data Table 1**). This analysis identified 1,966 significantly associated SNPs distributed among 31 loci (**Supplemental material, Extended Data Figure 1** and **Extended Data Table 2**). Three of these loci were not known to be associated with MI or CAD whereas the remaining 28 loci were previously reported for an all-inclusive CAD phenotype that included MI[4] (**Supplemental material, Extended Data Table 2**). We next combined our GWAS results in the UK Biobank with summary statistics for MI from CARDIoGRAM+C4D[4] in a fixed-effects meta-analysis that included a total of ~61,000 MI cases and ~578,000 controls and 8,126,035 SNPs common to both datasets (**Supplemental material, Extended Data Table 1**). This analysis revealed 4,419 significantly associated variants distributed across 80 loci (**Figure 2** and **Supplemental material, Extended Data Figure 2**). Of these loci, eight were novel and associated with MI (or CAD) for the first time herein (**Table 1** and **Figure 3**). The other 72 genome-wide significant loci in our MI meta-analysis overlapped with the 162 previously known CAD regions identified in other meta-analyses combining CARDIoGRAM+C4D with either the UK Biobank or samples genotyped on the MetaboChip (**Supplementary Table 3**)[5-9]. We also obtained evidence for association of 87 of the remaining 90 known CAD loci at the Bonferroni-corrected threshold for testing 162 regions ($P=0.05/162=3.1\times10^{-4}$) (**Supplemental material, Extended Data Table 3**). Thus, with the exception of three regions, our meta-analyses with the UK Biobank and CARDIoGRAM+C4D

replicated association signals at all known 162 MI/CAD loci. All together with the 8 novel regions we identified, our results bring the total number of MI/CAD susceptibility loci to 170 at the time of this analysis (**Supplemental material, Extended Data Table 3**).

Expression quantitative trait locus (eQTL) Analysis Prioritizes Positional Candidate Genes: To identify candidate causal genes at the newly identified loci, we used multi-tissue gene expression data from The GTEx Project[28]. For each locus, at least one candidate causal gene could be prioritized based on the lead SNP yielding a *cis* expression quantitative trait locus (eQTL) in one or more tissues relevant to MI, such as aorta, blood, heart, and visceral adipose (**Supplemental material, Extended Data Table 4**). We next carried out a PheWAS analysis with our lead variants (or tightly linked proxies), which also revealed suggestive and genome-wide significant associations at several of the novel loci for other complex traits, including inflammatory cytokines, circulating leukocytes, prostate cancer, and migraine (**Supplemental material, Extended Data Table 5**).

Comparison of Association Signals for MI and CAD at Eight Novel Loci: We next investigated the phenotypic specificity of the association signals for MI vs. CAD at the eight novel loci using various analytical strategies. In the first approach, we carried out association analyses with the eight novel loci in the UK Biobank using an all-inclusive definition of CAD (see Methods for details). This was followed by a meta-analysis of the results with summary statistics for CAD provided by CARDIoGRAM+C4D. Compared to MI, all eight loci yielded some degree of association with CAD in our meta-analysis with the UK Biobank and CARDIoGRAM+C4D, with two loci on chromosomes 1p36.11 and 6q16.1 exhibiting genome-wide significance (**Table 1** and **Supplemental material, Extended Data Table 6**). These latter observations suggest that

the association signals on chromosomes 1p36.11 and 6q16.1 may not be specific to MI. The CAD associations we obtained with the eight novel loci were also highly consistent with the results of another recent meta-analysis for CAD using data from the UK Biobank and CARDIoGRAM+C4D[9] (**Supplemental material, Extended Data Table 3**). However, since CARDIoGRAM+C4D used an all-inclusive definition of CAD that incorporated MI[4], it was not possible to determine the true specificity of the associations for MI vs. CAD using the results of our meta-analysis for CAD. Therefore, as a second approach, we used primary level data in the UK Biobank to compare association of the eight novel loci with MI but not the restricted CAD only phenotype that excluded subjects positive for MI. As a positive control locus, we also included *ABO* in these analyses. As expected and consistent with previous studies[10], our lead SNP at the *ABO* locus (rs9411377) in the UK Biobank was strongly associated with MI but not the restricted CAD only phenotype (**Table 2**), thus validating this analytical approach. With respect to the eight novel loci identified for MI, seven were not associated with CAD in the comparative analyses using the UK Biobank (**Table 2**). The only exception was the *AHDC1* locus on chromosome 1p36.11, although the effect size and significance level were weaker for CAD than with MI (**Table 2**). Finally, we also evaluated association at the eight novel loci in the UK Biobank in an analysis comparing cases defined as subjects with both CAD and MI (CAD^+/MI^+) to controls defined as CAD only subjects (CAD^+/MI^-). These analyses revealed nominal ($p<0.05$) association of six of the eight loci with MI among subjects with CAD as well as the expected association at the *ABO* locus (**Table 2**). Taken together with the effect sizes for MI and CAD that were observed in our meta-analyses, these results suggest that the association signals at some of the novel loci are either specific to or more robust for MI than with a CAD only phenotype. U

Preferential Association of the *SLC44A3* Locus with MI in the Presence of Atherosclerosis: We next sought to independently replicate the stronger association signals for MI at the novel loci using independent cohorts (GeneBank, Emory Cardiovascular Biobank, ANGES/FINCAVAS, LURIC, LIFE-Heart and UCORBIO) in which the presence of CAD in subjects was more directly assessed by angiography. For this purpose, case-control analyses were carried out among angiographically-documented CAD patients with (cases) and without (controls) adjudicated prior diagnoses of MI. Since the loci on chromosomes 1p36.11 (*AHDC1*) and 6q16.1 (*FHL5*) yielded genome-wide significant association with CAD in the meta-analysis with the UK Biobank and CARDIoGRAM+C4D, they were not considered in these analyses. Five of the six remaining newly identified regions did not yield evidence for association with MI among angiographically documented CAD cases (**Supplemental material, Extended Data Table 7** and **Table 3**). However, the C allele of the lead SNP (rs12743267) at the chromosome 1p21.3 locus harboring *SLC44A3* was associated with increased risk of MI among individuals with CAD in several of the angiography-based cohorts (**Supplemental material, Extended Data Table 7**). Association of rs12743267 with MI also became highly significant in a fixed-effects meta-analysis (OR=1.16, 95% CI 1.09-1.23; $P=3.3x10^{-6}$) with all ~14,000 subjects from these cohorts (**Table 3**) with no significant evidence for heterogeneity (P-het=0.10). Lastly, we used publicly available results from a Japanese GWAS for MI with 1,666 cases and 3,198 controls[29] to carry out another *in silico* replication study. The risk allele (C) of rs12743267 yielded directionally consistent association with MI in the Japanese cohort as well (OR=1.09, 95% CI 1.004-1.19; P=0.04). When these results were combined with our meta-analysis results for MI from the UK Biobank and CARDIoGRAM+C4D (total n~644,000), association of the *SLC44A3* locus with MI became even more significant (OR=1.05, 95% CI 1.04-1.07; $P=1.7x10^{-9}$). Taken together

with the weak associations observed with CAD in the meta-analyses and the comparative analyses in the UK Biobank described above, these results provide consistent evidence that the *SLC44A3* locus preferentially associates with plaque instability and/or rupture in the presence of coronary atherosclerosis but not atherosclerotic CAD itself.

Association of the *SLC44A3* Locus with Other Thrombotic Phenotypes: We next explored whether the *SLC44A3* locus was associated with other thrombotic and coagulation phenotypes related to MI. Based on publicly available data from the MEGASTROKE Consortium[30], there was no evidence for association of rs12743267 with most forms of stroke except for nominally significant associations with cardioembolic and small vessel stroke in subjects of European ancestry (**Supplemental material, Extended Data Table 8**). However, these associations would not be considered significant after a Bonferroni correction for testing five forms of stroke (p=0.05/5=0.01) and they were not significant in analyses with larger numbers of multi-ethnic subjects (**Supplemental material, Extended Data Table 8**). Second, variants at the chromosome 1p21.3 locus had been previously associated with circulating levels of D-dimer[31], one of several peptides produced when crosslinked fibrin is degraded by plasmin and the most widely used clinical marker of activated blood coagulation[32]. However, rs12743267 was not associated with D-dimer levels (beta=-0.011; SE=0.007; P=0.12) in a large GWAS carried out by the CHARGE Consortium[31] and the lead SNP for D-dimer (rs12029080) showed no evidence for association with MI in our meta-analysis for MI in the UK Biobank and CARDIoGRAM+C4D (OR=0.99, 95% CI 0.98-1.01; P=0.32). Lastly, *SLC44A2*, a member of the solute carrier family of membrane transporters that includes *SLC44A3*, has been associated with venous thromboembolism (VTE)[33], another coagulation and thrombotic phenotype relevant to MI. However, there was no association of rs12743267 with VTE (OR=0.97, 95% CI 0.92-1.02;

P=0.23) in a large GWAS carried out by the INVENT Consortium[33]. By comparison, we did obtain suggestive evidence for association of the lead VTE SNP in *SLC44A2* (rs2288904) with CAD (OR=1.04, 95% CI 1.03-1.05; $P=7.0x10^{-8}$) and MI (OR=1.04, 95% CI 1.02-1.06; $P=1.5x10^{-5}$) in our meta-analyses.

Association of the *SLC44A3* Locus with Choline-related Metabolites: While the function of *SLC44A3* as a solute carrier is not entirely known, it has been reported to encode a putative choline-like transporter[34]. In humans, elevated plasma levels of choline and two choline-derived metabolites, trimethylamine *N*-oxide (TMAO) and betaine, have been linked to increased risk of CAD and prospective risk of major adverse cardiac events, such as MI, stroke, and all-cause mortality[16, 35]. Therefore, we used the Genebank cohort to investigate whether association of rs12743267 with MI could be explained by genetics effects on choline-related metabolites. However, plasma levels of choline, TMAO, betaine, and homocysteine did not vary as a function of rs12743267 genotype (**Supplemental material, Extended Data Table 9**). Evaluation of a panel of related small molecule pro-atherogenic amines also did not reveal any associations with rs12743267 (**Supplemental material, Extended Data Table 9**).

Functional Analysis of *SLC44A3*: Based on the genetic association and eQTL results described above, we focused *SLC44A3* as a positional candidate gene at the chromosome 1p21.3 locus. We first examined the tissue distribution pattern of *SLC44A3* expression in the STARNET study, a cohort of 600 CAD patients in whom genome-wide genotyping and multi-tissue whole transcriptome analysis was carried out[23]. As shown in **Figure 4A**, *SLC44A3* was expressed at relatively high levels in several tissues relevant to MI, including atherosclerotic aortic root, adipose tissue, mammary artery, and liver. In addition, the lead SNP at chromosome 1p21.3

locus yielded *cis* eQTLs in atherosclerotic aorta and mammary artery, where the risk allele for MI (C) was associated with increased *SLC44A3* expression (**Figure 4B**). In GTEx, highly similar genotype-mRNA associations were independently observed in aorta and coronary artery tissue as well (**Figure 4C**). Consistent with these observations, mRNA levels of *SLC44A3* were significantly upregulated in an independent dataset of ischemic coronary arteries compared to non-diseased coronary arteries (**Figure 4D**). To explore the vascular cell type in which *SLC44A3* could mediate its biological effects on MI risk, we used RNAseq and functional data from two additional independent datasets of human aortic endothelial cells (HAECs) and smooth muscle cells (SMCs), respectively. Compared to untreated cells, *SLC44A3* expression was significantly upregulated in HAECs treated with the inflammatory and pro-atherogenic cytokine IL-1β (**Figure 4E**). An *in vitro* assay also demonstrated that *SLC44A3* mRNA levels in SMCs were modestly, but significantly, inversely correlated with their migration rate towards platelet-derived growth factor (PDGF)-BB (**Figure 4F**). Taken together, these data provide supportive functional evidence that *SLC44A3* could be at least one candidate causal at the novel MI locus on chromosome 1p21.3 locus and suggest that this putative solute carrier could promote increased risk of plaque rupture and thrombosis through mechanisms at the level of the artery wall.

Discussion

In the present study, we identified eight novel loci for MI through a large-scale gene discovery effort with ~639,000 subjects from the UK Biobank and CARDIoGRAM+C4D. Based on our own meta-analysis for CAD using data from CARDIoGRAM+C4D and the UK Biobank and another recent comparable analysis[9], the strength of the associations at the eight loci were, for the most part, stronger with MI than with CAD. This pattern of association signals is not entirely surprising since our primary meta-analysis was specifically for a plaque rupture phenotype. Various follow-up analyses with individual level data in the UK Biobank provided further evidence that six of the novel loci were either specifically or more strongly associated with MI than with CAD. However, of these six loci, only the chromosome 1p21.3 locus yielded independent association with MI among subjects with angiographically documented CAD. Although the relatively smaller sample size in the angiography-based replication cohorts may not have provided sufficient power to detect associations with the other five loci, it is possible that some of the novel loci we identified may also influence risk of CAD and are therefore not truly specific to MI. Even so, our collective analyses still led to the identification of eight novel genetic determinants of cardiovascular outcomes, bringing the total number of loci associated with atherosclerosis-related traits to 170 at the time of this analysis.

Of the loci identified, multiple independent analytical approaches provided evidence that the *SLC44A3* locus was specifically associated with MI but not CAD. This association was revealed not only by our initial meta-analysis and subsequent comparative analyses in the UK Biobank, but also supported by similar results in an independent Japanese case-control dataset. Further and consistent association of the *SLC44A3* locus with MI was also observed in six additional independent cohorts in which associations were tested specifically with MI among

individuals with angiographically documented CAD. Importantly, the magnitude of the effect size of the *SLC44A3* locus on MI in the context of coronary atherosclerosis was stronger than that observed in the meta-analysis (OR=1.16 vs. OR=1.05, respectively) and equivalent to effect sizes for some of the most significantly associated loci identified to date for CAD traits, such as *LPA* or chromosome 9p21[9]. Taken together, these results suggest that the biological mechanism(s) underlying the association of this locus may be related to thrombosis and/or plaque rupture rather than plaque progression per se. In this regard, *ABO* was similarly identified as being associated only with MI in the original study by Reilly *et al.*[10], which we replicated in our analogous comparative analyses with the UK Biobank. Thus, to our knowledge, the *SLC44A3* locus may be the only other genetic risk factor out of the 170 known susceptibility loci that is specifically associated with MI but not with CAD. We also did not obtain evidence for association of the *SLC44A3* locus with other thrombotic phenotypes, such as stroke or VTE. This observation is not entirely surprising since the genetic determinants of CAD and stroke, while shared, do not completely overlap[36]. However, it should be noted that the number of subjects in our meta-analyses for MI was approximately 10-fold greater than that in the GWAS for VTE[33]. Thus, it is possible that power was insufficient in the INVENT Consortium to detect an association of the *SLC44A3* locus with VTE.

The lead SNP on chromosome 1p21.3 (rs12743267) is located ~36kb upstream of the transcriptional start site for *SLC44A3* and ~250kb away from the gene encoding tissue factor or coagulation factor III (*F3*). Given the known role of tissue factor in the blood coagulation cascade and the association variants around its gene with circulating D-dimer levels[31], *F3* would be considered a more biologically plausible candidate gene for MI as a thrombosis-related phenotype that mapped to this locus. However, we did not obtain any evidence that would

prioritize *F3* as a candidate causal gene for MI since our lead SNP was not associated with D-dimer levels and the lead SNP for D-dimer (rs12029080) showed no evidence for association with MI in our study. Furthermore, *cis* eQTLs for *F3* were not observed with our lead SNP or proxy variants in any available tissue in STARNET or the GTEx Project. Given these results and the presence of *cis* eQTLs for *SLC44A3* in multiple tissues, we focused on *SLC44A3* as a candidate causal gene for MI.

In addition to the genetic evidence provided by our study, our functional experiments suggested that putative biological mechanisms through which *SLC44A3* could influences plaque rupture and/or thrombosis may be through direct effects at the level of the vessel wall. For example, in the STARNET cohort, *SLC44A3* was expressed in MI-relevant vascular tissues, such as the aorta and mammary artery, in which the MI risk allele of rs12743267 significantly higher *SLC44A3* mRNA levels, with a stronger effect size observed in atherosclerotic aortic root. The same allelic associations were also independently observed with *SLC44A3* expression in aorta and coronary artery using GTEx data. Furthermore, in two independent heart donor datasets, *SLC44A3* expression was upregulated ~50% in ischemic coronary arteries compared to normal arteries and by ~3-fold in HAECs incubated with the pro-atherogenic cytokine IL-1β. This latter observation suggests that *SLC44A3* might be involved in the response of HAECs to inflammatory stimuli that increase expression and secretion of various pro-atherogenic genes, such as adhesion molecules and chemokines[37]. Lastly, an *in vitro* assay demonstrated that *SLC44A3* expression was inversely correlated with SMC migration. In this regard, previous studies have shown that SMC proliferation and migration can be protective by promoting secretion of extra cellular matrix proteins and the formation of a protective fibrous cap that renders a lesion less prone to rupture[38]. Thus, these functional data and the results of our genetic

analyses collectively implicate *SLC44A3* as at least one candidate causal gene on chromosome 1p21.3 and suggest that its expression is positively associated with MI-promoting characteristics of various vascular cell types. However, in STARNET, *SLC44A3* mRNA levels in adipose and liver were equivalent to those observed in aorta, and based on GTEx data, expression was also high in kidney, pancreas, the small intestine, and colon. Thus, we cannot rule out the possibility that *SLC44A3* could also increase risk of plaque rupture and thrombosis via a role in other MI-relevant tissues.

While our results point to novel and distinct genetic determinants of MI, certain limitations of our study should still be taken into consideration. First, nearly all subjects in our analyses were of European ancestry and it is possible that some of the genetic associations may not be generalizable to other populations. However, the chromosome 1p21.3 locus was associated with MI in Japanese subjects, suggesting that at least a subset of the association signals identified herein may also be relevant in other ethnicities. Second, it is possible, albeit unlikely, that there was overlap of samples between CARDIoGRAM+C4D and the UK Biobank, which could have been a confounding factor in the meta-analysis. However, a recent analysis concluded that the presence and influence of duplicate samples between CARDIoGRAM+C4D and the UK Biobank was minimal (<0.1%) and that any duplications did not significantly influence test statistics[9]. Lastly, we did not exclude subjects with a positive family history of CAD from the control group in the UK Biobank as was done in another recent GWAS meta-analysis for CAD[9]. Furthermore, there could have been misclassification in our analyses since, for example, MI and CAD may not have been defined similarly in CARDIoGRAM+C4D and the UK Biobank. We note that if such misclassifications had occurred, they would most likely have been non-differential and biased the results towards the null. Finally, even though SNPs with

MAFs as low as 0.5% were included in the genetic analyses, our study was primarily focused on discovery of main effects with common susceptibility alleles. However, rare variants or GxE interactions still likely play important roles in modulating risk of MI, which will need to be addressed in appropriately designed future studies.

In summary, our results identify several previously unrecognized loci for MI and provide new avenues for exploring the pathophysiology of vulnerable atherosclerotic lesions. Furthermore, our data support the concept that some of the heritable determinants of plaque rupture and thrombus formation are distinct from those that contribute to development of coronary atherosclerosis, with *SLC44A3* emerging as one such genetic susceptibility factor.

URLs. The UK Biobank, (https://www.ukbiobank.ac.uk/); CARDIoGRAMplusC4D, http://www.cardiogramplusc4d.org/; Japanese National Bioscience Database Center, https://humandbs.biosciencedbc.jp/en/; GWAMA, https://www.geenivaramu.ee/en/tools/gwama Phenoscanner, http://www.phenoscanner.medschl.cam.ac.uk/phenoscanner; Genotype-Tissue Expression Project, http://gtexportal.org/; R statistical software, http://www.R-project.org/.

Data Availability

Full summary statistics relating to the GWAS analysis in the UK Biobank and the meta-analysis with CARDIoGRAM+C4D will be deposited with The NHGRI-EBI Catalog of published genome-wide association studies (https://www.ebi.ac.uk/gwas/docs/about). All other relevant data are available upon request from the authors.

Acknowledgments

We gratefully acknowledge the UK Biobank Resource for providing access to their data under Application Number 33307 and the CARDIoGRAM+C4D Consortium for making their data publicly available. We also thank the LURIC study team who were either temporarily or permanently involved in patient recruitment as well as sample and data handling, in addition to the laboratory staff at the Ludwigshafen General Hospital and the Universities of Freiburg and Ulm, Germany.

Sources of Funding

This work was supported, in part, by National Institutes of Health grants [R01HL133169, R01ES021801, R01ES025786, MD010358, P01ES022845, R01HL103866, P20HL113452, R21HL135230, R01HL147187], American Heart Association grants [18POST33990046 and 19TPA34910021], and Transatlantic Networks of Excellence Awards from Foundation Leducq. The GeneBank study was supported in part by National Institutes of Health grants [P01HL098055, P01HL076491, and R01HL103931]. The Angiography and Genes Study (ANGES) and Finnish Cardiovascular Study (FINCAVAS) have been financially supported by the Competitive Research Funding of the Tampere University Hospital [Grant 9M048 and 9N035], the Finnish Cultural Foundation, the Finnish Foundation for Cardiovascular Research, the Emil Aaltonen Foundation, Finland, the Tampere Tuberculosis Foundation, and EU Horizon 2020 [grant 755320 for TAXINOMISIS]. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Disclosures

Dr. Hazen (S.L.H.) is named as co-inventor on pending and issued patents held by the Cleveland Clinic relating to cardiovascular diagnostics and therapeutics and have the right to receive royalty payment for inventions or discoveries related to cardiovascular diagnostics or therapeutics from Cleveland Heart Lab, Quest Diagnostics, and Procter & Gamble Company. Dr. Hazen also reports having been paid as a consultant from Procter & Gamble Company and having received research funds from Procter & Gamble Company and Roche. Dr. Scholz (M.Sc.) receives funding from Pfizer Inc. for a project not related to this research. Dr. März (W.M.) reports grants from Siemens Healthineers, grants and personal fees from Aegerion

Pharmaceuticals, grants and personal fees from AMGEN, grants from Astrazeneca, grants and personal fees from Sanofi, grants and personal fees from Alexion Pharmaceuticals, grants and personal fees from BASF, grants and personal fees from Abbott Diagnostics, grants and personal fees from Numares AG, grants and personal fees from Berlin-Chemie, grants and personal fees from Akzea Therapeutics, grants from Bayer Vital GmbH , grants from bestbion dx GmbH, grants from Boehringer Ingelheim Pharma GmbH Co KG, grants from Immundiagnostik GmbH, grants from Merck Chemicals GmbH, grants from MSD Sharp and Dohme GmbH, grants from Novartis Pharma GmbH, grants from Olink Proteomics; and other support from SYNLAB Holding Deutschland GmbH, all outside the submitted work. The other authors report no conflicts.

Author Contributions

Concept and design: J.H., Y.H., and H.A. Acquisition, analysis, and interpretation of data: Y.H., Q.J., P.H., J.G., W.S.S., D.A.T, N.L.S., M.Se., C.P., M.M., A.J.L., P.B., Y.V.S., C.L., A.A.Q., M.Sc., J.T., G.E.D., M.E.K., W.M., L.J.H., F.W.A., R.S.P., L-P.L., M.K., T.L., T.V.M.N., P.K., J.O.L., L.K.S., C.E.R., M.D.K., A.W.T., C.M., R. A., M.C., L.M., J.L.M.B., W.H.W.T., S.L.H., and H.A. Drafting of the manuscript: J.H., Y.H., and H.A. Critical revision of the manuscript for important intellectual content: All authors.

References

1. Go AS, Mozaffarian D, Roger VL, Benjamin EJ, Berry JD, Blaha MJ, Dai S, Ford ES, Fox CS, Franco S, Fullerton HJ, Gillespie C, Hailpern SM, Heit JA, Howard VJ, Huffman MD, Judd SE, Kissela BM, Kittner SJ, Lackland DT, Lichtman JH, Lisabeth LD, Mackey RH, Magid DJ, Marcus GM, Marelli A, Matchar DB, McGuire DK, Mohler ER, 3rd, Moy CS, Mussolino ME, Neumar RW, Nichol G, Pandey DK, Paynter NP, Reeves MJ, Sorlie PD, Stein J, Towfighi A, Turan TN, Virani SS, Wong ND, Woo D, Turner MB, American Heart Association Statistics C and Stroke Statistics S. Heart disease and stroke statistics--2014 update: a report from the American Heart Association. *Circulation* 2014;**129**:e28-e292.
2. Ridker PM, Danielson E, Fonseca FA, Genest J, Gotto AM, Jr., Kastelein JJ, Koenig W, Libby P, Lorenzatti AJ, MacFadyen JG, Nordestgaard BG, Shepherd J, Willerson JT, Glynn RJ and Group JS. Rosuvastatin to prevent vascular events in men and women with elevated C-reactive protein. *N Engl J Med* 2008;**359**:2195-207.
3. Hansson GK. Inflammation, atherosclerosis, and coronary artery disease. *N Engl J Med* 2005;**352**:1685-95.
4. Nikpay M, Goel A, Won HH, Hall LM, Willenborg C, Kanoni S, Saleheen D, Kyriakou T, Nelson CP, Hopewell JC, Webb TR, Zeng L, Dehghan A, Alver M, Armasu SM, Auro K, Bjonnes A, Chasman DI, Chen S, Ford I, Franceschini N, Gieger C, Grace C, Gustafsson S, Huang J, Hwang SJ, Kim YK, Kleber ME, Lau KW, Lu X, Lu Y, Lyytikainen LP, Mihailov E, Morrison AC, Pervjakova N, Qu L, Rose LM, Salfati E, Saxena R, Scholz M, Smith AV, Tikkanen E, Uitterlinden A, Yang X, Zhang W, Zhao W, de Andrade M, de Vries PS, van Zuydam NR, Anand SS, Bertram L, Beutner F, Dedoussis G, Frossard P, Gauguier D, Goodall AH, Gottesman O, Haber M, Han BG, Huang J, Jalilzadeh S, Kessler T, Konig IR, Lannfelt L, Lieb W, Lind L, Lindgren CM, Lokki ML, Magnusson PK, Mallick NH, Mehra N, Meitinger T, Memon FU, Morris AP, Nieminen MS, Pedersen NL, Peters A, Rallidis LS, Rasheed A, Samuel M, Shah SH, Sinisalo J, Stirrups KE, Trompet S, Wang L, Zaman KS, Ardissino D, Boerwinkle E, Borecki IB, Bottinger EP, Buring JE, Chambers JC, Collins R, Cupples LA, Danesh J, Demuth I, Elosua R, Epstein SE, Esko T, Feitosa MF, Franco OH, Franzosi MG, Granger CB, Gu D, Gudnason V, Hall AS, Hamsten A, Harris TB, Hazen SL, Hengstenberg C, Hofman A, Ingelsson E, Iribarren C, Jukema JW, Karhunen PJ, Kim BJ, Kooner JS, Kullo IJ, Lehtimaki T, Loos RJ, Melander O, Metspalu A, Marz W, Palmer CN, Perola M, Quertermous T, Rader DJ, Ridker PM, Ripatti S, Roberts R, Salomaa V, Sanghera DK, Schwartz SM, Seedorf U, Stewart AF, Stott DJ, Thiery J, Zalloua PA, O'Donnell CJ, Reilly MP, Assimes TL, Thompson JR, Erdmann J, Clarke R, Watkins H, Kathiresan S, McPherson R, Deloukas P, Schunkert H, Samani NJ, Farrall M and CARDIoGRAMplusC4D Consortium. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet* 2015;**47**:1121-30.
5. Verweij N, Eppinga RN, Hagemeijer Y and van der Harst P. Identification of 15 novel risk loci for coronary artery disease and genetic risk of recurrent events, atrial fibrillation and heart failure. *Sci Rep* 2017;**7**:2761.
6. Howson JMM, Zhao W, Barnes DR, Ho WK, Young R, Paul DS, Waite LL, Freitag DF, Fauman EB, Salfati EL, Sun BB, Eicher JD, Johnson AD, Sheu WHH, Nielsen SF, Lin WY, Surendran P, Malarstig A, Wilk JB, Tybjaerg-Hansen A, Rasmussen KL, Kamstrup PR, Deloukas P, Erdmann J, Kathiresan S, Samani NJ, Schunkert H, Watkins H, CardioGramplusC4D, Do R, Rader DJ, Johnson JA, Hazen SL, Quyyumi AA, Spertus JA,

Pepine CJ, Franceschini N, Justice A, Reiner AP, Buyske S, Hindorff LA, Carty CL, North KE, Kooperberg C, Boerwinkle E, Young K, Graff M, Peters U, Absher D, Hsiung CA, Lee WJ, Taylor KD, Chen YH, Lee IT, Guo X, Chung RH, Hung YJ, Rotter JI, Juang JJ, Quertermous T, Wang TD, Rasheed A, Frossard P, Alam DS, Majumder AAS, Di Angelantonio E, Chowdhury R, Epic CVD, Chen YI, Nordestgaard BG, Assimes TL, Danesh J, Butterworth AS and Saleheen D. Fifteen new risk loci for coronary artery disease highlight arterial-wall-specific mechanisms. *Nat Genet* 2017;**49**:1113-1119.

7. Klarin D, Zhu QM, Emdin CA, Chaffin M, Horner S, McMillan BJ, Leed A, Weale ME, Spencer CCA, Aguet F, Segre AV, Ardlie KG, Khera AV, Kaushik VK, Natarajan P, CARDIoGRAMplusC4D Consortium, and Kathiresan S. Genetic analysis in UK Biobank links insulin resistance and transendothelial migration pathways to coronary artery disease. *Nat Genet* 2017;**49**:1392-1397.
8. Nelson CP, Goel A, Butterworth AS, Kanoni S, Webb TR, Marouli E, Zeng L, Ntalla I, Lai FY, Hopewell JC, Giannakopoulou O, Jiang T, Hamby SE, Di Angelantonio E, Assimes TL, Bottinger EP, Chambers JC, Clarke R, Palmer CNA, Cubbon RM, Ellinor P, Ermel R, Evangelou E, Franks PW, Grace C, Gu D, Hingorani AD, Howson JMM, Ingelsson E, Kastrati A, Kessler T, Kyriakou T, Lehtimaki T, Lu X, Lu Y, Marz W, McPherson R, Metspalu A, Pujades-Rodriguez M, Ruusalepp A, Schadt EE, Schmidt AF, Sweeting MJ, Zalloua PA, AlGhalayini K, Keavney BD, Kooner JS, Loos RJF, Patel RS, Rutter MK, Tomaszewski M, Tzoulaki I, Zeggini E, Erdmann J, Dedoussis G, Bjorkegren JLM, Epic-CVD Consortium, CARDIoGRAMplusC4D, U. K. Biobank CardioMetabolic Consortium CHD working group, Schunkert H, Farrall M, Danesh J, Samani NJ, Watkins H and Deloukas P. Association analyses based on false discovery rate implicate new loci for coronary artery disease. *Nat Genet* 2017;**49**:1385-1391.
9. van der Harst P and Verweij N. Identification of 64 novel genetic loci provides an expanded view on the genetic architecture of coronary artery disease. *Circ Res* 2018;**122**:433-443.
10. Reilly MP, Li M, He J, Ferguson JF, Stylianou IM, Mehta NN, Burnett MS, Devaney JM, Knouff CW, Thompson JR, Horne BD, Stewart AF, Assimes TL, Wild PS, Allayee H, Nitschke PL, Patel RS, Martinelli N, Girelli D, Quyyumi AA, Anderson JL, Erdmann J, Hall AS, Schunkert H, Quertermous T, Blankenberg S, Hazen SL, Roberts R, Kathiresan S, Samani NJ, Epstein SE, Rader DJ, Qasim AN, DerOhannessian SL, Qu L, Cappola TP, Chen Z, Matthai W, Hakonarson HH, Wilensky R, Kent KM, Lindsay JM, Pichard AD, Satler L and Waksman R. 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. *Lancet* 2011;**377**:383-92.
11. Beutner F, Teupser D, Gielen S, Holdt LM, Scholz M, Boudriot E, Schuler G and Thiery J. Rationale and design of the Leipzig (LIFE) Heart Study: phenotyping and cardiovascular characteristics of patients with coronary artery disease. *PLoS One* 2011;**6**:e29070.
12. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, Motyer A, Vukcevic D, Delaneau O, O'Connell J, Cortes A, Welsh S, Young A, Effingham M, McVean G, Leslie S, Allen N, Donnelly P and Marchini J. The UK Biobank resource with deep phenotyping and genomic data. *Nature* 2018;**562**:203-209.
13. Loh P-R, Tucker G, Bulik-Sullivan BK, Vilhjalmsson BJ, Finucane HK, Salem RM, Chasman DI, Ridker PM, Neale BM and Berger BJNg. Efficient Bayesian mixed-model analysis increases association power in large cohorts. 2015;**47**:284.

14. Magi R and Morris AP. GWAMA: software for genome-wide association meta-analysis. *BMC Bioinformatics* 2010;**11**:288.
15. Schwarzer G, Carpenter JR and Rücker G. *Meta-analysis with R*. Cham: Springer; 2015.
16. Wang Z, Klipfell E, Bennett BJ, Koeth R, Levison BS, Dugar B, Feldstein AE, Britt EB, Fu X, Chung YM, Wu Y, Schauer P, Smith JD, Allayee H, Tang WH, DiDonato JA, Lusis AJ and Hazen SL. Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature* 2011;**472**:57-63.
17. Koeth RA, Wang Z, Levison BS, Buffa JA, Org E, Sheehy BT, Britt EB, Fu X, Wu Y, Li L, Smith JD, DiDonato JA, Chen J, Li H, Wu GD, Lewis JD, Warrier M, Brown JM, Krauss RM, Tang WH, Bushman FD, Lusis AJ and Hazen SL. Intestinal microbiota metabolism of L-carnitine, a nutrient in red meat, promotes atherosclerosis. *Nat Med* 2013;**19**:576-85.
18. Hartiala J, Bennett BJ, Tang WH, Wang Z, Stewart AF, Roberts R, McPherson R, Lusis AJ, Hazen SL, Allayee H and CARDIoGRAM Consortium. Comparative genome-wide association studies in mice and humans for trimethylamine N-oxide, a proatherogenic metabolite of choline and L-carnitine. *Arterioscler Thromb Vasc Biol* 2014;**34**:1307-13.
19. Koeth RA, Levison BS, Culley MK, Buffa JA, Wang Z, Gregory JC, Org E, Wu Y, Li L, Smith JD, Tang WH, DiDonato JA, Lusis AJ and Hazen SL. gamma-Butyrobetaine is a proatherogenic intermediate in gut microbial metabolism of L-carnitine to TMAO. *Cell Metab* 2014;**20**:799-812.
20. Hartiala JA, Tang WH, Wang Z, Crow AL, Stewart AF, Roberts R, McPherson R, Erdmann J, Willenborg C, Hazen SL and Allayee H. Genome-wide association study and targeted metabolomics identifies sex-specific association of CPS1 with coronary artery disease. *Nat Commun* 2016;**7**:10558.
21. Li XS, Wang Z, Cajka T, Buffa JA, Nemet I, Hurd AG, Gu X, Skye SM, Roberts AB, Wu Y, Li L, Shahen CJ, Wagner MA, Hartiala JA, Kerby RL, Romano KA, Han Y, Obeid S, Luscher TF, Allayee H, Rey FE, DiDonato JA, Fiehn O, Tang WHW and Hazen SL. Untargeted metabolomics identifies trimethyllysine, a TMAO-producing nutrient precursor, as a predictor of incident cardiovascular disease risk. *JCI Insight* 2018;**3**.
22. Jia Q, Han Y, Huang P, Woodward NC, Gukasyan J, Kettunen J, Ala-Korpela M, Anufrieva O, Wang Q, Perola M, Raitakari O, Lehtimaki T, Viikari J, Jarvelin MR, Boehnke M, Laakso M, Mohlke KL, Fiehn O, Wang Z, Tang WHW, Hazen SL, Hartiala JA and Allayee H. Genetic Determinants of Circulating Glycine Levels and Risk of Coronary Artery Disease. *J Am Heart Assoc* 2019;**8**:e011922.
23. Franzen O, Ermel R, Cohain A, Akers NK, Di Narzo A, Talukdar HA, Foroughi-Asl H, Giambartolomei C, Fullard JF, Sukhavasi K, Koks S, Gan LM, Giannarelli C, Kovacic JC, Betsholtz C, Losic B, Michoel T, Hao K, Roussos P, Skogsberg J, Ruusalepp A, Schadt EE and Bjorkegren JL. Cardiometabolic risk loci share downstream cis- and trans-gene regulation across tissues and diseases. *Science* 2016;**353**:827-30.
24. Anders S and Huber W. Differential expression analysis for sequence count data. *Genome biology* 2010;**11**:R106.
25. Romanoski CE, Lee S, Kim MJ, Ingram-Drake L, Plaisier CL, Yordanova R, Tilford C, Guan B, He A, Gargalovic PS, Kirchgessner TG, Berliner JA and Lusis AJ. Systems genetics analysis of gene-by-environment interactions in human cells. *Am J Hum Genet* 2010;**86**:399-410.
26. Romanoski CE, Che N, Yin F, Mai N, Pouldar D, Civelek M, Pan C, Lee S, Vakili L, Yang WP, Kayne P, Mungrue IN, Araujo JA, Berliner JA and Lusis AJ. Network for activation of

human endothelial cells by oxidized phospholipids: a critical role of heme oxygenase 1. *Circ Res* 2011;**109**:e27-41.
27. Hogan NT, Whalen MB, Stolze LK, Hadeli NK, Lam MT, Springstead JR, Glass CK and Romanoski CE. Transcriptional networks specifying homeostatic and inflammatory programs of gene expression in human aortic endothelial cells. *Elife* 2017;**6**.
28. GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* 2013;**45**:580-5.
29. Hirokawa M, Morita H, Tajima T, Takahashi A, Ashikawa K, Miya F, Shigemizu D, Ozaki K, Sakata Y, Nakatani D, Suna S, Imai Y, Tanaka T, Tsunoda T, Matsuda K, Kadowaki T, Nakamura Y, Nagai R, Komuro I and Kubo M. A genome-wide association study identifies PLCL2 and AP3D1-DOT1L-SF3A2 as new susceptibility loci for myocardial infarction in Japanese. *Eur J Hum Genet* 2015;**23**:374-80.
30. Malik R, Traylor M, Pulit SL, Bevan S, Hopewell JC, Holliday EG, Zhao W, Abrantes P, Amouyel P, Attia JR, Battey TW, Berger K, Boncoraglio GB, Chauhan G, Cheng YC, Chen WM, Clarke R, Cotlarciuc I, Debette S, Falcone GJ, Ferro JM, Gamble DM, Ilinca A, Kittner SJ, Kourkoulis CE, Lemmens R, Levi CR, Lichtner P, Lindgren A, Liu J, Meschia JF, Mitchell BD, Oliveira SA, Pera J, Reiner AP, Rothwell PM, Sharma P, Slowik A, Sudlow CL, Tatlisumak T, Thijs V, Vicente AM, Woo D, Seshadri S, Saleheen D, Rosand J, Markus HS, Worrall BB, Dichgans M, Group IA, collaboration M, Wellcome Trust Case Control C and Network NSG. Low-frequency and common genetic variation in ischemic stroke: The METASTROKE collaboration. *Neurology* 2016;**86**:1217-26.
31. Smith NL, Huffman JE, Strachan DP, Huang J, Dehghan A, Trompet S, Lopez LM, Shin SY, Baumert J, Vitart V, Bis JC, Wild SH, Rumley A, Yang Q, Uitterlinden AG, Stott DJ, Davies G, Carter AM, Thorand B, Polasek O, McKnight B, Campbell H, Rudnicka AR, Chen MH, Buckley BM, Harris SE, Peters A, Pulanic D, Lumley T, de Craen AJ, Liewald DC, Gieger C, Campbell S, Ford I, Gow AJ, Luciano M, Porteous DJ, Guo X, Sattar N, Tenesa A, Cushman M, Slagboom PE, Visscher PM, Spector TD, Illig T, Rudan I, Bovill EG, Wright AF, McArdle WL, Tofler G, Hofman A, Westendorp RG, Starr JM, Grant PJ, Karakas M, Hastie ND, Psaty BM, Wilson JF, Lowe GD, O'Donnell CJ, Witteman JC, Jukema JW, Deary IJ, Soranzo N, Koenig W and Hayward C. Genetic predictors of fibrin D-dimer levels in healthy adults. *Circulation* 2011;**123**:1864-72.
32. Adam SS, Key NS and Greenberg CS. D-dimer antigen: current concepts and future prospects. *Blood* 2009;**113**:2878-87.
33. Germain M, Chasman DI, de Haan H, Tang W, Lindstrom S, Weng LC, de Andrade M, de Visser MC, Wiggins KL, Suchon P, Saut N, Smadja DM, Le Gal G, van Hylckama Vlieg A, Di Narzo A, Hao K, Nelson CP, Rocanin-Arjo A, Folkersen L, Monajemi R, Rose LM, Brody JA, Slagboom E, Aissi D, Gagnon F, Deleuze JF, Deloukas P, Tzourio C, Dartigues JF, Berr C, Taylor KD, Civelek M, Eriksson P, Cardiogenics C, Psaty BM, Houwing-Duitermaat J, Goodall AH, Cambien F, Kraft P, Amouyel P, Samani NJ, Basu S, Ridker PM, Rosendaal FR, Kabrhel C, Folsom AR, Heit J, Reitsma PH, Tregouet DA, Smith NL and Morange PE. Meta-analysis of 65,734 individuals identifies TSPAN15 and SLC44A2 as two susceptibility loci for venous thromboembolism. *Am J Hum Genet* 2015;**96**:532-42.
34. Traiffort E, O'Regan S and Ruat M. The choline transporter-like family SLC44: properties and roles in human diseases. *Mol Aspects Med* 2013;**34**:646-54.

35. Tang WH, Wang Z, Levison BS, Koeth RA, Britt EB, Fu X, Wu Y and Hazen SL. Intestinal microbial metabolism of phosphatidylcholine and cardiovascular risk. *N Engl J Med* 2013;**368**:1575-84.
36. Dichgans M, Malik R, Konig IR, Rosand J, Clarke R, Gretarsdottir S, Thorleifsson G, Mitchell BD, Assimes TL, Levi C, O'Donnell CJ, Fornage M, Thorsteinsdottir U, Psaty BM, Hengstenberg C, Seshadri S, Erdmann J, Bis JC, Peters A, Boncoraglio GB, Marz W, Meschia JF, Kathiresan S, Ikram MA, McPherson R, Stefansson K, Sudlow C, Reilly MP, Thompson JR, Sharma P, Hopewell JC, Chambers JC, Watkins H, Rothwell PM, Roberts R, Markus HS, Samani NJ, Farrall M, Schunkert H, Metastroke Consortium, CARDIoGRAM Consortium, C4D Consortium, International Stroke Genetics Consortium. Shared genetic susceptibility to ischemic stroke and coronary artery disease: a genome-wide analysis of common variants. *Stroke* 2014;**45**:24-36.
37. Gimbrone MA, Jr. and Garcia-Cardena G. Endothelial Cell Dysfunction and the Pathobiology of Atherosclerosis. *Circ Res* 2016;**118**:620-36.
38. Cherepanova OA, Gomez D, Shankman LS, Swiatlowska P, Williams J, Sarmento OF, Alencar GF, Hess DL, Bevard MH, Greene ES, Murgai M, Turner SD, Geng YJ, Bekiranov S, Connelly JJ, Tomilin A and Owens GK. Activation of the pluripotency factor OCT4 in smooth muscle cells is atheroprotective. *Nat Med* 2016;**22**:657-65.

Figure Legends

Figure 1. Overview of genetic and functional analyses. A GWAS was first carried out for MI using primary level data in the UK Biobank with ~11 million SNPs. These results were then combined with summary GWAS data from the CARDIoGRAM+C4D Consortium in a fixed-effects meta-analysis that included a total of ~61,000 MI cases and ~577,000 controls 8,126,035 SNPs common to both datasets. The meta-analysis identified eight novel loci for MI, 6 of which exhibited stronger association signals for MI compared to CAD. Follow-up analyses and replication in independent cohorts provided confirmatory evidence for association of the chromosome 1p21.3 locus with MI. Bioinformatics and eQTL analyses prioritized *SLC44A3* as one positional candidate on chromosome 1p21.3 for functional evaluation.

Figure 2. Manhattan plot of results from GWAS meta-analysis for MI. (A) Eight novel loci on chromosomes 1p36.11, 1p21.3, 2q13, 2q32.1, 4q22.3, 6q16.1, 9q34.3, and 15q24.2 (orange dots) were significantly associated with MI. Genome-wide thresholds for significant ($P=5.0x10^{-8}$) and suggestive ($P=5.0x10^{-6}$) association are indicated by the horizontal red and blue lines, respectively. P-values are truncated at $-\log_{10}(P)=40$.

Figure 3. Regional plots of eight novel loci for MI. The chromosome band and nearest gene (in parentheses) is indicated for each locus. Each region is centered on the lead SNP (purple diamond) and the genes in the interval are indicated in the bottom panel. The degree of linkage disequilibrium (LD) between the lead SNP and other variants is shown as r2 values according to the color-coded legend in the box.

Figure 4. Functional Analyses of *SLC44A3* in MI-relevant Tissues. (**A**) In the STARNET cohort, *SLC44A3* was expressed at relatively high levels in tissues relevant to MI, including atherosclerotic aortic root (aorta) , visceral adipose, mammary artery, and liver. (**B**) The lead SNP at the chromosome 1p21.3 locus yielded *cis* eQTLs for *SLC44A3* in atherosclerotic aortic root and normal mammary artery among subjects from the STARNET cohort, where the MI risk allele (C) was associated with significantly higher mRNA levels. (**C**) In an independent human dataset, *SLC44A3* expression was increased in ischemic coronary arteries (n=36) from heart donors with CAD compared to normal coronary arteries from non-diseased donors (n=24). (**D**) Incubation of human aortic endothelial cells (HAECs) isolated from a different and independent set of anonymous heart donors (n=53) with IL-1β for 4 hours upregulated *SLC44A3* expressed ~3-fold compared to untreated paired HAECs. (**E**) Using a fourth independent human dataset (n=151), *SLC44A3* expression was also observed in smooth muscle cells (SMCs) and inversely correlated with migration rate *in vitro*.

Table 1. Novel Loci Identified for MI through GWAS Meta-Analysis of the UK Biobank and CARDIoGRAM+C4D.

						MI		CAD	
SNP	**Chr**	**Pos**	**Nearest Gene(s)**	**EA/OA**	**EAF**	**OR (95% CI)**	**P**	**OR (95% CI)**	**P**
rs113716316	1p36.11	27,928,640	*AHDC1*	G/A	0.93	1.09 (1.06-1.13)	$4.4x10^{-08}$	1.07 (1.05-1.10)	$5.0x10^{-08}$
rs12743267	1p21.3	95,249,306	*SLC44A3*	C/T	0.77	1.05 (1.03-1.07)	$1.1x10^{-08}$	1.03 (1.01-1.04)	$2.0x10^{-04}$
rs6761276	2q13	113,832,312	*IL1F10*	T/C	0.43	1.04 (1.03-1.06)	$2.8x10^{-08}$	1.03 (1.01-1.04)	$2.2x10^{-05}$
rs12693302	2q32.1	183,211,443	*PDE1A*	G/A	0.39	1.05 (1.03-1.06)	$2.5x10^{-09}$	1.03 (1.01-1.04)	$2.5x10^{-05}$
rs2452009	4q22.3	95,495,908	*PDLIM5*	A/G	0.70	1.05 (1.03-1.07)	$5.8x10^{-09}$	1.03 (1.02-1.05)	$9.4x10^{-07}$
rs9486719	6q16.1	97,060,124	*FHL5*	G/A	0.80	1.06 (1.04-1.08)	$6.8x10^{-10}$	1.04 (1.03-1.06)	$1.1x10^{-08}$
rs28429551	9q34.3	139,243,334	*GPSM1*	A/T	0.76	1.06 (1.04-1.08)	$1.7x10^{-08}$	1.04 (1.02-1.05)	$4.0x10^{-06}$
rs8037798	15q24.2	75,240,030	*COX5A-RPP25*	G/T	0.23	1.05 (1.03-1.07)	$3.8x10^{-08}$	1.02 (1.01-1.04)	$1.6x10^{-03}$

Chr, chromosome; Pos, base-pair position (hg19), EA, effect allele; OA, other allele; EAF, effect allele frequency; OR, odds ratio; CI, confidence interval; P, p-value.

Table 2. Comparison of Associations with MI and CAD in the UK Biobank for 8 Novel Loci and *ABO* Locus.

					MI vs. Control (17,505/454,212)		CAD only vs. Control (15,580/454,212)		CAD^+/MI^+ vs. CAD^+/MI^- (17,505/15,580)	
SNP	**Chr**	**Pos**	**Nearest Gene(s)**	**EAF**	**OR (95% CI)**	**P**	**OR (95% CI)**	**P**	**OR (95% CI)**	**P**
rs113716316	1p36.11	27,928,640	*AHDC1*	0.93	1.11 (1.07-1.16)	$7.2x10^{-07}$	1.07 (1.02-1.12)	$4.1x10^{-03}$	1.04 (0.98-1.11)	0.21
rs12743267	1p21.3	95,249,306	*SLC44A3*	0.76	1.04 (1.01-1.06)	$3.1x10^{-03}$	1.00 (0.97-1.03)	0.98	1.04 (1.01-1.08)	0.02
rs6761276	2q13	113,832,312	*IL1F10*	0.42	1.03 (1.01-1.06)	$1.9x10^{-03}$	1.01 (0.99-1.03)	0.44	1.03 (0.99-1.06)	0.11
rs12693302	2q32.1	183,211,443	*PDE1A*	0.36	1.06 (1.03-1.08)	$1.3x10^{-06}$	0.98 (0.96-1.01)	0.19	1.07 (1.04-1.10)	$2.9x10^{-05}$
rs2452009	4q22.3	95,495,908	*PDLIM5*	0.69	1.04 (1.02-1.07)	$2.6x10^{-04}$	1.01 (0.98-1.03)	0.68	1.03 (1.001-1.07)	0.04
rs9486719	6q16.1	97,060,124	*FHL5*	0.80	1.05 (1.03-1.08)	$9.2x10^{-05}$	1.01 (0.99-1.04)	0.36	1.04 (1.004-1.08)	0.03
rs28429551	9q34.3	139,243,334	*GPSM1*	0.76	1.07 (1.04-1.10)	$4.8x10^{-08}$	1.01 (0.98-1.03)	0.54	1.07 (1.03-1.11)	$2.8x10^{-04}$
rs8037798	15q24.2	75,240,030	*COX5A-RPP25*	0.23	1.05 (1.03-1.08)	$3.2x10^{-05}$	1.00 (0.97-1.02)	0.85	1.06 (1.02-1.10)	$1.7x10^{-03}$
rs9411377	9q34.2	136,145,404	*ABO*	0.30	1.06 (1.04-1.09)	$3.3x10^{-07}$	0.99 (0.97-1.02)	0.67	1.07 (1.03-1.10)	$1.3x10^{-04}$

Number of cases and controls for each phenotype are shown in parentheses.
For CAD^+/MI^+ vs. CAD^+/MI^- analyses, cases were defined as CAD positive subjects with MI and controls were defined as CAD positive subjects without MI.
Chr, chromosome; Pos, base-pair position (hg19), EA, effect allele; OA, other allele; EAF, effect allele frequency; OR, odds ratio; CI, confidence interval; P, p-value.

Table 3. Meta-analysis Results of Novel Loci with MI in the Presence of CAD Using Angiography-based Cohorts.

					CAD^+/MI^+ (6,514) vs. CAD^+/MI^+ (7,411)	
SNP	**Chr**	**Pos**	**Nearest Gene(s)**	**EA/OA**	**OR (95% CI)**	**P**
rs12743267	1p21.3	95,249,306	*SLC44A3*	C/T	1.16 (1.09-1.23)	$3.3x10^{-06}$
rs6761276	2q13	113,832,312	*IL1F10*	T/C	1.03 (0.98-1.08)	0.31
rs12693302	2q32.1	183,211,443	*PDE1A*	G/A	0.99 (0.93-1.04)	0.63
rs2452009	4q22.3	95,495,908	*PDLIM5*	A/G	1.01 (0.95-1.06)	0.83
rs28429551	9q34.3	139,243,334	*GPSM1*	A/T	1.02 (0.95-1.09)	0.66
rs8037798	15q24.2	75,240,030	*COX5A-RPP25*	G/T	1.00 (0.93-1.08)	0.91

Number of cases and controls for each phenotype are shown in parentheses.
EAF, effect allele frequency; OR, odds ratio; CI, confidence interval; P, p-value.
Cases were defined CAD positive subjects with MI (CAD^+/MI^+) controls were defined as CAD only subjects without MI (CAD^+/MI^-).
See Methods for details.

Figure 1

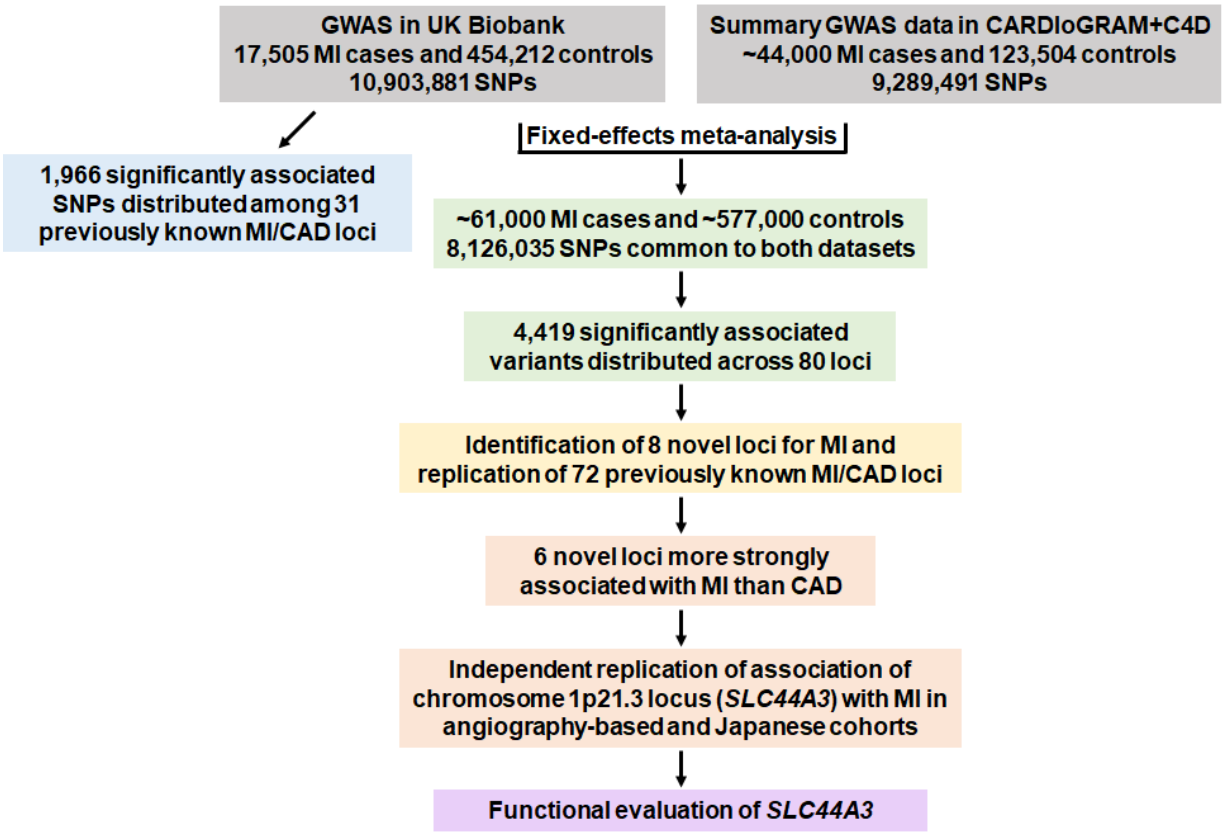

GWAS in UK Biobank
17,505 MI cases and 454,212 controls
10,903,881 SNPs
Summary GWAS data in CARDIoGRAM+C4D
~44,000 MI cases and 123,504 controls
9,289,491 SNPs
1,966 significantly associated SNPs distributed among 31 previously known MI/CAD loci
Fixed-effects meta-analysis
~61,000 MI cases and ~577,000 controls
8,126,035 SNPs common to both datasets
4,419 significantly associated variants distributed across 80 loci
Identification of 8 novel loci for MI and replication of 72 previously known MI/CAD loci
6 novel loci more strongly associated with MI than CAD
Independent replication of association of chromosome 1p21.3 locus (SLC44A3) with MI in angiography-based and Japanese cohorts
Functional evaluation of SLC44A3

Figure 2

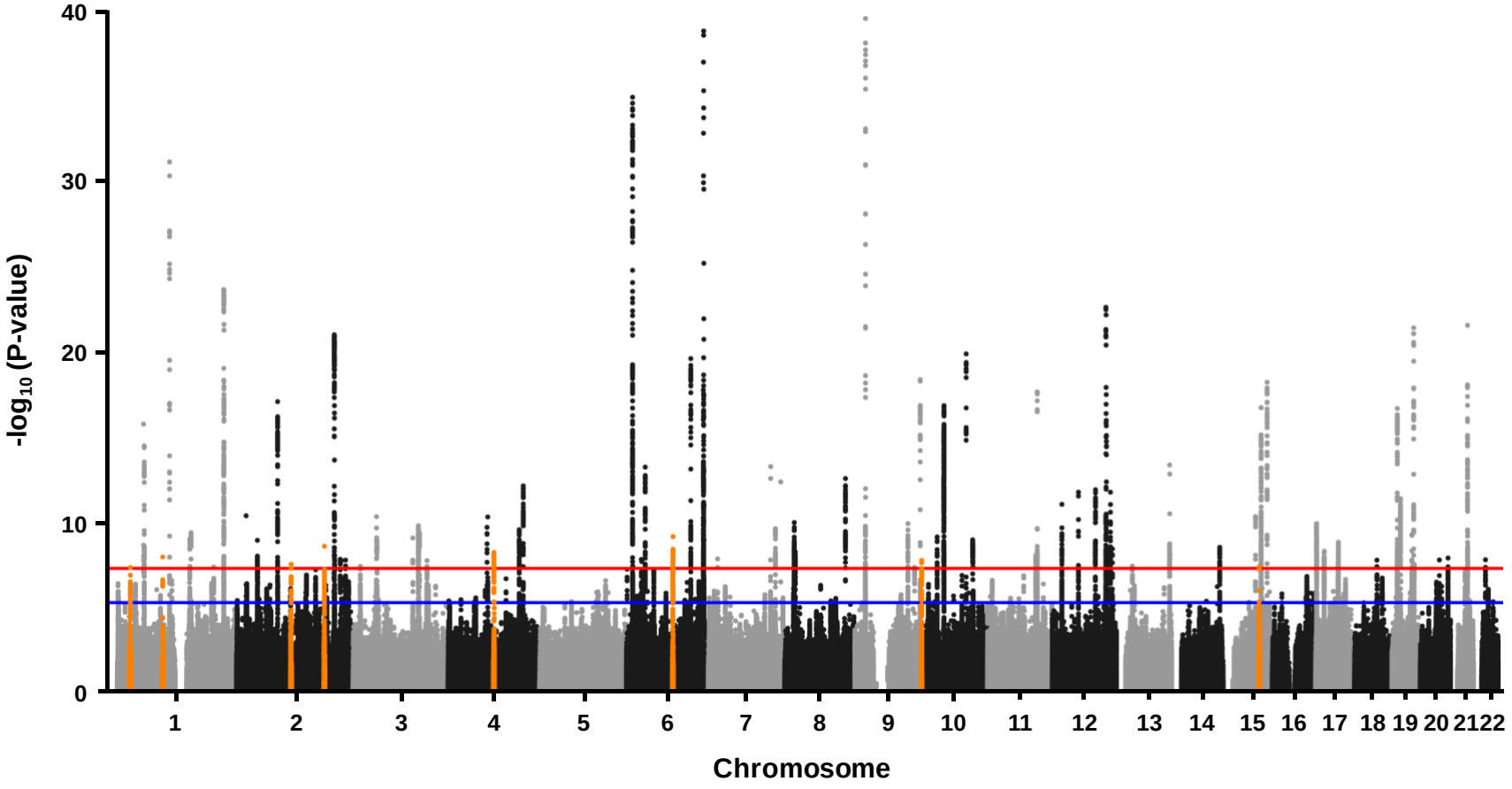
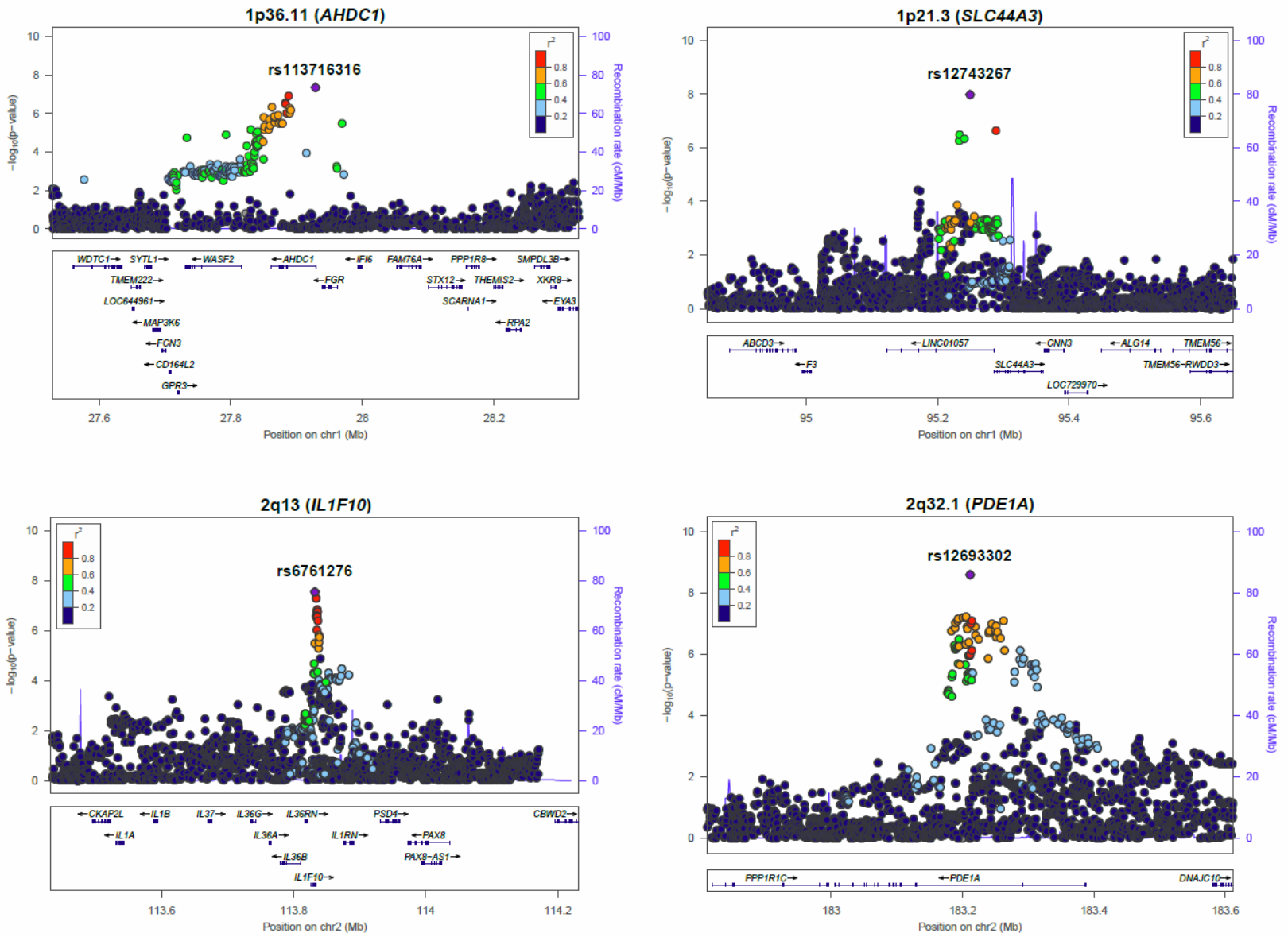

40
30
20
10
0
-log10 (P-value)
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
Chromosome

Figure 3


1p36.11 (AHDC1)
rs113716316
1p21.3 (SLC44A3)
rs12743267
2q13 (IL1F10)
rs6761276
2q32.1 (PDE1A)
rs12693302
Position on chr1 (Mb)
Position on chr2 (Mb)
Recombination rate (cM/Mb)

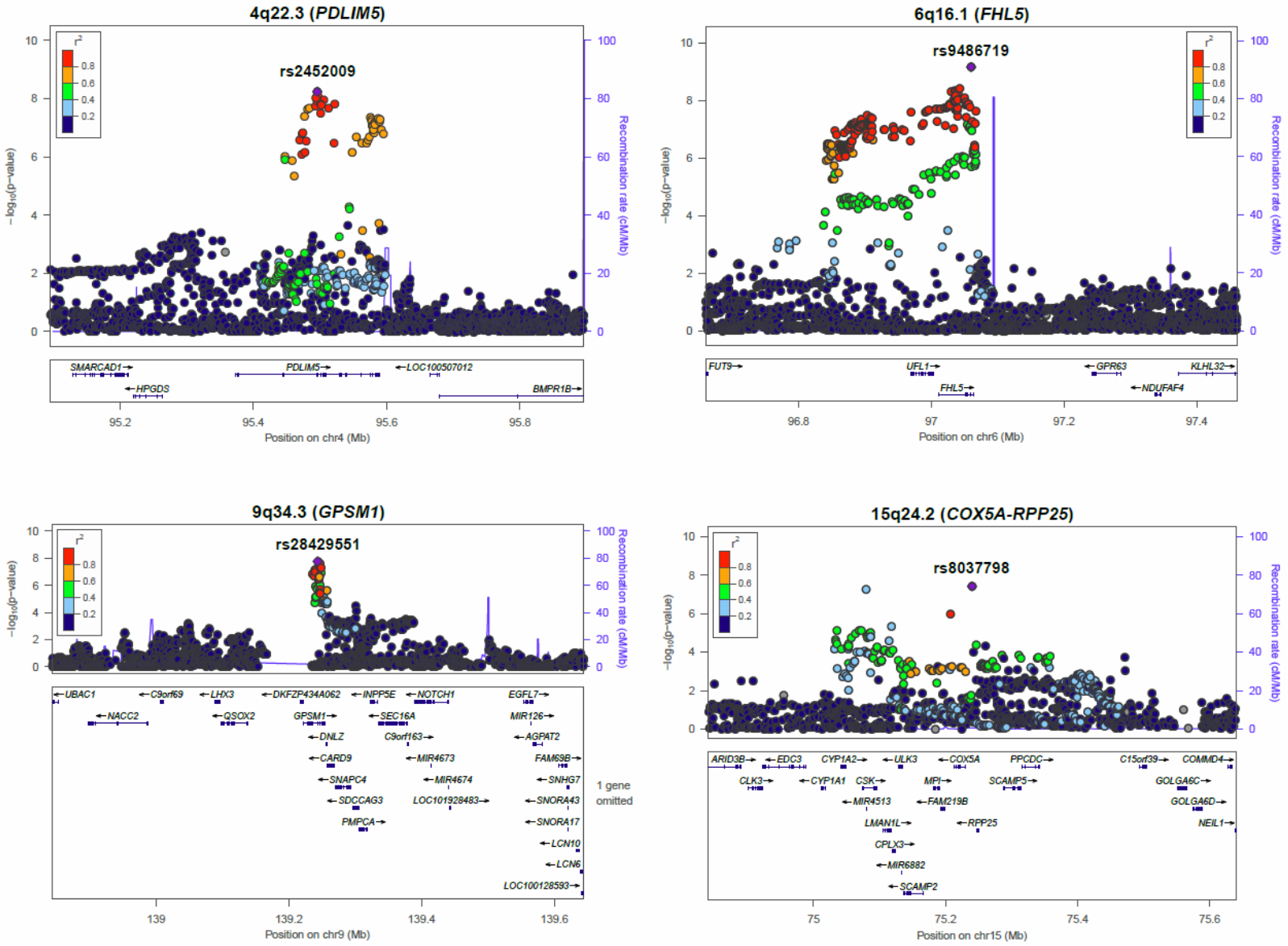
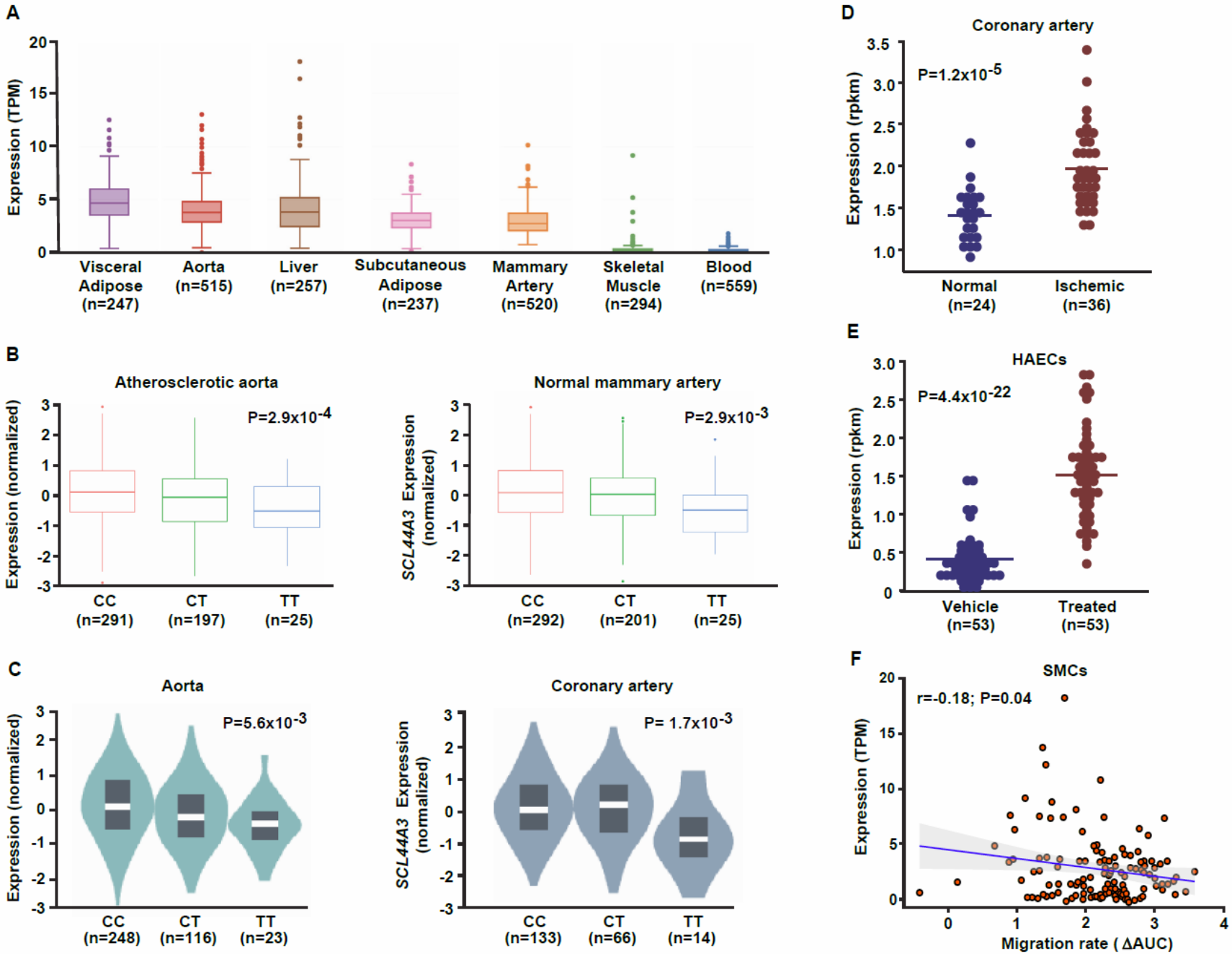

4q22.3 (PDLIM5)
rs2452009
6q16.1 (FHL5)
rs9486719
9q34.3 (GPSM1)
rs28429551
1 gene omitted
15q24.2 (COX5A-RPP25)
rs8037798
Position on chr4 (Mb)
Position on chr6 (Mb)
Position on chr9 (Mb)
Position on chr15 (Mb)
Recombination rate (cM/Mb)

Figure 4


A
Expression (TPM)
Visceral Adipose (n=247)
Aorta (n=515)
Liver (n=257)
Subcutaneous Adipose (n=237)
Mammary Artery (n=520)
Skeletal Muscle (n=294)
Blood (n=559)
B
Atherosclerotic aorta
P=2.9x10-4
Expression (normalized)
CC (n=291)
CT (n=197)
TT (n=25)
Normal mammary artery
P=2.9x10-3
SCL44A3 Expression (normalized)
CC (n=292)
CT (n=201)
TT (n=25)
C
Aorta
P=5.6x10-3
Expression (normalized)
CC (n=248)
CT (n=116)
TT (n=23)
Coronary artery
P= 1.7x10-3
SCL44A3 Expression (normalized)
CC (n=133)
CT (n=66)
TT (n=14)
D
Coronary artery
P=1.2x10-5
Expression (rpkm)
Normal (n=24)
Ischemic (n=36)
E
HAECs
P=4.4x10-22
Expression (rpkm)
Vehicle (n=53)
Treated (n=53)
F
SMCs
r=-0.18; P=0.04
Expression (TPM)
Migration rate (ΔAUC)