UC Irvine
UC Irvine Previously Published Works

Title
Genome-wide analysis identifies novel susceptibility loci for myocardial infarction

Permalink
https://escholarship.org/uc/item/30n8729m

Journal
European Heart Journal, 42(9)

ISSN
0195-668X

Authors
Hartiala, Jaana A
Han, Yi
Jia, Qiong
et al.

Publication Date
2021-03-01

DOI
10.1093/eurheartj/ehaa1040

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

ESC European Society of Cardiology

European Heart Journal (2021) **42**, 919–933
doi:10.1093/eurheartj/ehaa1040

CLINICAL RESEARCH
Ischaemic heart disease

Genome-wide analysis identifies novel susceptibility loci for myocardial infarction

Jaana A. Hartiala [1†], Yi Han [1,2†], Qiong Jia [1,2], James R. Hilser[1,2], Pin Huang[1,2], Janet Gukasyan[1,2], William S. Schwartzman [1,2], Zhiheng Cai[1,2], Subarna Biswas [1], David-Alexandre Trégouët[3], Nicholas L. Smith [4,5,6], The INVENT Consortium[‡], The CHARGE Consortium Hemostasis Working Group[‡], The GENIUS-CHD Consortium[‡], Marcus Seldin [7], Calvin Pan [8], Margarete Mehrabian [9], Aldons J. Lusis[8,9,10], Peter Bazeley[11,12,13], Yan V. Sun [14,15], Chang Liu [14,16], Arshed A. Quyyumi [16], Markus Scholz [17,18], Joachim Thiery[18,19], Graciela E. Delgado[20], Marcus E. Kleber[20], Winfried März[20,21,22], Laurence J. Howe [23], Folkert W. Asselbergs [23,24,25], Marion van Vugt [25], Georgios J. Vlachojannis [25], Riyaz S. Patel [23,26], Leo-Pekka Lyytikäinen [27,28,29], Mika Kähönen[30,31], Terho Lehtimäki[27,28], Tuomo V.M. Nieminen [32], Pekka Kuukasjärvi[33], Jari O. Laurikka [33,34], Xuling Chang [35,36], Chew-Kiat Heng [35,36], Rong Jiang[37], William E. Kraus [38,39], Elizabeth R. Hauser [38,40], Jane F. Ferguson [41], Muredach P. Reilly[42,43], Kaoru Ito [44], Satoshi Koyama [44], Yoichiro Kamatani[45,46,47], Issei Komuro[48], Biobank Japan, Lindsey K. Stolze[50], Casey E. Romanoski [50], Mohammad Daud Khan[51,52,53,54], Adam W. Turner [51,52,53,54], Clint L. Miller[51,52,53,54], Redouane Aherrahrou [51,52], Mete Civelek [51,52], Lijiang Ma[55], Johan L.M. Björkegren[55,56], S. Ram Kumar [57], W.H. Wilson Tang[11,12,13], Stanley L. Hazen [12,13], and Hooman Allayee [1,2*]

[1]Department of Preventive Medicine, Keck School of Medicine, University of Southern California, 2250 Alcazar Street, CSC202, Los Angeles, CA 90033, USA; [2]Department of Biochemistry & Molecular Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA 90033, USA; [3]Institut National pour la Santé et la Recherche Médicale (INSERM) UMR_S 1219, Bordeaux Population Health Research Center, University of Bordeaux, 33076 Bordeaux, France; [4]Department of Epidemiology, University of Washington, Seattle, WA 98101, USA; [5]Department of Veterans Affairs, Seattle Epidemiologic Research and Information Center, Office of Research and Development, Seattle, WA 98108, USA; [6]Kaiser Permanente Washington Health Research Institute, Kaiser Permanente Washington, Seattle, WA 98101, USA; [7]Department of Biological Chemistry and Center for Epigenetics and Metabolism, UC Irvine School of Medicine, Irvine, CA 92697, USA; [8]Department of Human Genetics, David Geffen School of Medicine of UCLA, Los Angeles, CA 90095, USA; [9]Department of Medicine, David Geffen School of Medicine of UCLA, Los Angeles, CA 90095, USA; [10]Department of Microbiology, Immunology, & Molecular Genetics, David Geffen School of Medicine of UCLA, Los Angeles, CA 90095, USA; [11]Center for Clinical Genomics, Cleveland Clinic, Cleveland, OH 44195, USA; [12]Department of Cardiovascular Medicine, Cleveland Clinic, Cleveland, OH 44195, USA; [13]Department of Cardiovascular & Metabolic Sciences, Cleveland Clinic, Cleveland, OH 44195, USA; [14]Department of Epidemiology, Emory University Rollins School of Public Health, 1518 Clifton Rd. NE, Atlanta, GA 30322, USA; [15]Department of Biomedical Informatics, Emory University School of Medicine, Atlanta, GA 30322, USA; [16]Division of Cardiology, Department of Medicine, Emory Clinical Cardiovascular Research Institute, Emory University School of Medicine, 1462 Clifton Rd NE, Suite # 507, Atlanta, GA 30322, USA; [17]Institute for Medical Informatics, Statistics and Epidemiology, University of Leipzig, 04107 Leipzig, Germany; [18]LIFE Research Center for Civilization Diseases, University of Leipzig, 04103 Leipzig, Germany; [19]Institute of Laboratory Medicine, Clinical Chemistry and Molecular Diagnostics, University Hospital, 04103 Leipzig, Germany; [20]Vth Department of Medicine, Medical Faculty Mannheim, Heidelberg University, Theodor-Kutzer-Ufer 1-3, 68167 Mannheim, Germany; [21]SYNLAB Academy, SYNLAB Holding Deutschland GmbH, 86156 Augsburg, Germany; [22]Clinical Institute of Medical and Chemical Laboratory Diagnostics, Medical University of Graz, 8036 Graz, Austria; [23]Institute of Cardiovascular Science, Faculty of Population Health Sciences, University College

* Corresponding author. Tel: +1 323 442 1736, Fax: +1 323 442 2764, Email: hallayee@usc.edu
[†] The first two authors contributed equally to this work.
[‡] 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 material online.

Published on behalf of the European Society of Cardiology. All rights reserved. © The Author(s) 2021. For permissions, please email: journals.permissions@oup.com.

London, London WC1E 6HX, UK; [24]Health Data Research UK and Institute of Health Informatics, University College London, London NW1 2DA, UK; [25]Division Heart & Lungs, Department of Cardiology, University Medical Center Utrecht, Utrecht University, 3584 CX Utrecht, the Netherlands; [26]Bart's Heart Centre, St Bartholomew's Hospital, London EC1A 2DA, UK; [27]Department of Clinical Chemistry, Fimlab Laboratories, Tampere 33520, Finland; [28]Department of Clinical Chemistry, Finnish Cardiovascular Research Center—Tampere, Faculty of Medicine and Health Technology, Tampere University, Tampere 33014, Finland; [29]Department of Cardiology, Heart Center, Tampere University Hospital, Tampere 33521, Finland; [30]Department of Clinical Physiology, Tampere University Hospital, Tampere 33521, Finland; [31]Department of Clinical Physiology, Finnish Cardiovascular Research Center—Tampere, Faculty of Medicine and Health Technology, Tampere University, Tampere 33014, Finland; [32]Department of Internal Medicine, Päijät-Häme Central Hospital, Lahti 15850, Finland; [33]Department of Cardio-Thoracic Surgery, Finnish Cardiovascular Research Center—Tampere, Faculty of Medicine and Health Technology, Tampere University, Tampere 33014, Finland; [34]Department of Cardio-Thoracic Surgery, Heart Center, Tampere University Hospital, Tampere 33521, Finland; [35]Department of Paediatrics, Yong Loo Lin School of Medicine, National University of Singapore, Singapore 119228, Singapore; [36]Khoo Teck Puat—National University Children's Medical Institute, National University Health System, Singapore 119074, Singapore; [37]Department of Psychiatry & Behavioral Sciences, Duke University School of Medicine Durham, NC 27710, USA; [38]Duke Molecular Physiology Institute, Duke University School of Medicine Durham, NC 27710, USA; [39]Department of Medicine, Duke University School of Medicine Durham, NC 27710, USA; [40]Department of Biostatistics & Bioinformatics, Duke University School of Medicine Durham, NC 27710, USA; [41]Department of Medicine, Vanderbilt University Medical Center, Nashville, TN 37232, USA; [42]Cardiometabolic Genomics Program, Division of Cardiology, Department of Medicine, Columbia University Irving Medical Center, New York, NY 10032, USA; [43]Irving Institute for Clinical and Translational Research, Columbia University, New York, NY 10032, USA; [44]Laboratory for Cardiovascular Genomics and Informatics, RIKEN Center for Integrative Medical Sciences, Kanagawa 230-0045, Japan; [45]Laboratory for Statistical and Translational Genetics, RIKEN Center for Integrative Medical Sciences, Kanagawa 230-0045, Japan; [46]Human Disease Genomics, Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto 606-8501, Japan; [47]Laboratory of Complex Trait Genomics, Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Tokyo 108-0071, Japan; [48]Department of Cardiovascular Medicine, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan; [49]Institute of Medical Science, The University of Tokyo, Tokyo, Japan; [50]Department of Cellular and Molecular Medicine, University of Arizona College of Medicine, Tucson, AZ 85721, USA; [51]Center for Public Health Genomics, University of Virginia, Charlottesville University of Virginia, Charlottesville, VA 22904, USA; [52]Department of Biomedical Engineering, University of Virginia, Charlottesville University of Virginia, Charlottesville, VA 22904, USA; [53]Department of Biochemistry & Molecular Genetics, University of Virginia, Charlottesville University of Virginia, Charlottesville, VA 22904, USA; [54]Department of Public Health Sciences, University of Virginia, Charlottesville University of Virginia, Charlottesville, VA 22904, USA; [55]Department of Genetics & Genomic Sciences, Institute of Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA; [56]Integrated Cardio Metabolic Centre, Department of Medicine, Karolinska Institutet, Karolinska Universitetssjukhuset, 141 57 Huddinge, Sweden; and [57]Department of Surgery, Keck School of Medicine, University of Southern California, Los Angeles, CA 90033, USA

Received 5 May 2020; revised 18 September 2020; editorial decision 7 December 2020; accepted 7 December 2020; online publish-ahead-of-print 3 February 2021

See page 934 for the editorial comment on this article (doi: 10.1093/eurheartj/ehaa1089)

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\sim$472 000), followed by a meta-analysis with summary statistics from the CARDIoGRAMplusC4D Consortium ($n\sim$167 000). Multiple independent replication analyses and functional approaches were used to prioritize loci and evaluate positional candidate genes. Eight novel regions were identified for MI at the genome wide significance level, of which effect sizes at six loci were more robust for MI than for CAD without the presence of MI. Confirmatory evidence for association of a locus on chromosome 1p21.3 harbouring choline-like transporter 3 (*SLC44A3*) with MI in the context of CAD, but not with coronary atherosclerosis itself, was obtained in Biobank Japan ($n\sim$165 000) and 16 independent angiography-based cohorts ($n\sim$27 000). Follow-up analyses did not reveal association of the *SLC44A3* locus with CAD risk factors, biomarkers of coagulation, other thrombotic diseases, or plasma levels of a broad array of metabolites, including choline, trimethylamine *N*-oxide, and betaine. However, aortic expression of *SLC44A3* was increased in carriers of the MI risk allele at chromosome 1p21.3, increased in ischaemic (vs. non-diseased) coronary arteries, up-regulated in human aortic endothelial cells treated with interleukin-1β (vs. vehicle), and associated with smooth muscle cell migration *in vitro*.

Conclusions A large-scale analysis comprising ∼831 000 subjects revealed novel genetic determinants of MI and implicated *SLC44A3* in the pathophysiology of vulnerable plaques.

Graphical Abstract

Normal artery

Blood flow

Known genetic risk factors for CAD and MI (*9p21*, *LPA*, *SORT1*)

Plaque initiation and progression

Blood flow

Distinct genetic risk factors for MI (*SLC44A3*)

New genetic risk factors for MI (and CAD) (*AHDC1*, *IL1F10*, *FHL5*)

Plaque rupture and thrombus formation

Blood flow

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, 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 and multi-ethnic genome-wide association studies (GWAS) have identified >200 loci that influence risk of MI and CAD via perturbations of lipid metabolism, blood pressure regulation, inflammation, and platelet function,[4–12] as well as through mechanisms that still remain unknown. However, the susceptibility alleles, most of which are common in the population, still only explain a small fraction of the overall heritability for CAD and MI. 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. Again, genetic studies support this concept. For example, *9p21* is one of the most strongly associated loci for CAD but is not specifically associated with MI when comparing CAD-positive/MI-positive (CAD^+/MI^+) individuals to those who are CAD-positive/MI-negative (CAD^+/MI^-).[13,14] In 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.[13] 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 strongly 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

Detailed methods are provided in the Supplementary material online.

Results

Identification of 8 novel loci for MI

To further expand our understanding of the genetic architecture of MI, we first carried out a GWAS for MI with 17 505 cases and 454 212 controls from the UK Biobank (*Figure 1* and Supplementary material online, *Table S1*). This analysis identified 1966 single-nucleotide polymorphisms (SNPs) at 31 loci that were associated with MI at the genome-wide significance threshold of $P = 5.0 \times 10^{-8}$ (Supplementary material online, *Figure S1* and *Table S2*). Twenty-eight of the 31 loci were previously reported for an all-inclusive CAD phenotype that included MI.[6] When MI was defined according to the algorithm provided by the UK Biobank, virtually identical results were obtained (Supplementary material online, *Table S2*). We next combined our results in the UK Biobank with summary statistics from CARDIoGRAMplusC4D[6] 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 data sets (*Figure 1* and Supplementary material online, *Table S1*). This analysis revealed 4419 significantly associated variants at 80 loci (*Figure 2* and Supplementary material online, *Figure S2*), eight of which were novel and associated with MI (or CAD) for the first time herein (*Table 1* and *Figure 3*). The

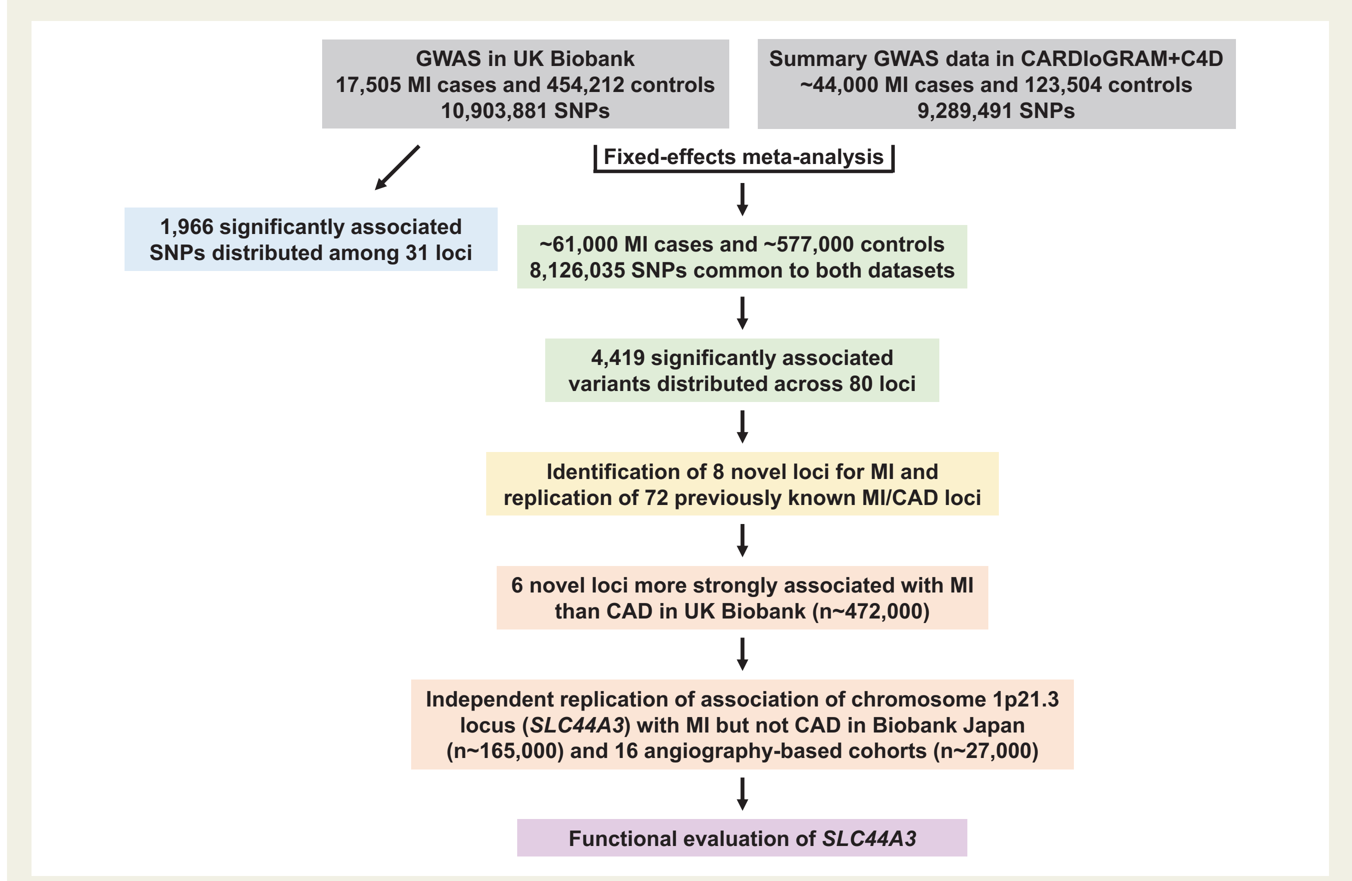

Figure 1 Overview of genetic and functional analyses. A genome-wide association study was first carried out for myocardial infarction using primary-level data in the UK Biobank with ~11 million single-nucleotide polymorphisms. These results were then combined with summary genome-wide association study data from the CARDIoGRAMplusC4D Consortium in a fixed-effects meta-analysis that included a total of ~61 000 myocardial infarction cases and ~577 000 controls, and 8 126 035 single-nucleotide polymorphisms common to both data sets. The meta-analysis identified eight novel loci for myocardial infarction, six of which exhibited stronger association signals for myocardial infarction compared to coronary artery disease. Follow-up analyses and independent replication in Biobank Japan and 16 angiography-based cohorts, encompassing a total of ~831 000 subjects, provided confirmatory evidence for association of the chromosome 1p21.3 locus with myocardial infarction. Bioinformatics and eQTL analyses prioritized *SLC44A3* as one positional candidate on chromosome 1p21.3 for functional evaluation.

other 72 genome-wide significant loci in our MI meta-analysis overlapped with the 205 previously identified CAD regions[7–12] (Supplementary material online, *Table S3*). We also obtained evidence for association of the 133 remaining known CAD loci at $P < 2.5 \times 10^{-3}$, although 12 signals would not be considered significant at the Bonferroni-corrected threshold for testing 205 regions ($P = 0.05/205 = 2.4 \times 10^{-4}$) (Supplementary material online, *Table S3*). Thus, our meta-analysis with the UK Biobank and CARDIoGRAMplusC4D replicated nearly all 205 known CAD loci and, together with the eight novel regions, brings the total number of MI/CAD susceptibility loci to 213 at the time of this analysis (Supplementary material online, *Table S3*).

Prioritization of positional candidate genes and follow-up analyses with novel MI loci

To identify candidate causal genes at the new loci, we first used multi-tissue gene expression data from the GTEx Project,[15] the eQTLgen Consortium, or previously published studies available through the Phenoscanner database.[16] For each locus, at least one candidate gene could be prioritized based on the lead SNP yielding a *cis* eQTL in one or more tissues relevant to MI (Supplementary material online, *Table S4*). Candidate causal genes were prioritized further using co-localization analysis with summary statistics from our meta-analysis and eQTL data from the STARNET cohort[17] in blood, atherosclerotic aortic artery, internal mammary artery, visceral and subcutaneous adipose, liver, and skeletal muscle. Based on posterior probabilities of ≥75%, we obtained evidence for *SLC44A3*, *TMEM87B*, and *FHL5* as being causal positional candidate genes on chromosomes 1p21.3, 2q13, and 6q16.1, respectively (Supplementary material online, *Table S5*). To explore the biological relevance of the MI loci, we also evaluated the lead variants for association with CAD risk factors in the UK Biobank and other disease phenotypes using the PhenoScanner database.[16] Five loci yielded genome-wide significant associations with blood pressure, lipid levels, body mass index, and/or type 2 diabetes in the UK Biobank (Supplementary material online, *Table S6*). The other three loci on chromosomes 1p21.3 (*SLC44A3*), 1p36.11,

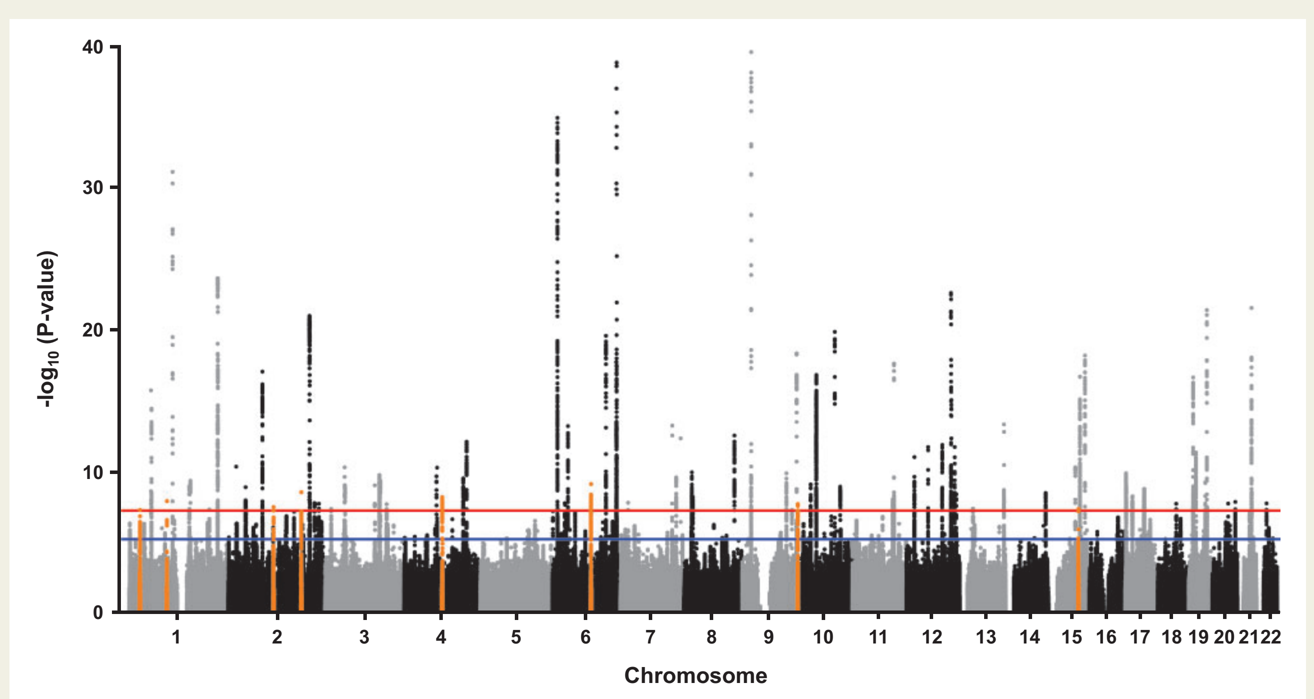

Figure 2 Manhattan plot of results from genome-wide association study meta-analysis for myocardial infarction. (*A*) Eight novel loci on chromosomes 1p36.11, 1p21.3, 2q13, 2q32.1, 4q22.3, 6q16.1, 9q34.3, and 15q24.2 (orange dots) were significantly associated with myocardial infarction. Genome-wide thresholds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

Table 1 Novel loci identified for MI through GWAS meta-analysis of the UK Biobank and CARDIoGRAMplusC4D

						MI		CAD	
SNP	Chr	Pos	Nearest gene(s)	EA/OA	EAF	OR (95% CI)	*P*-value	OR (95% CI)	*P*-value
rs113716316	1p36.11	27 928 640	*AHDC1*	G/A	0.93	1.09 (1.06–1.13)	4.4×10^{-8}	1.07 (1.05–1.10)	5.0×10^{-8}
rs12743267	1p21.3	95 249 306	*SLC44A3*	C/T	0.77	1.05 (1.03–1.07)	1.1×10^{-8}	1.03 (1.01–1.04)	2.0×10^{-4}
rs6761276	2q13	113 832 312	*IL1F10*	T/C	0.43	1.04 (1.03–1.06)	2.8×10^{-8}	1.03 (1.01–1.04)	2.2×10^{-5}
rs12693302	2q32.1	183 211 443	*PDE1A*	G/A	0.39	1.05 (1.03–1.06)	2.5×10^{-9}	1.03 (1.01–1.04)	2.5×10^{-5}
rs2452009	4q22.3	95 495 908	*PDLIM5*	A/G	0.70	1.05 (1.03–1.07)	5.8×10^{-9}	1.03 (1.02–1.05)	9.4×10^{-7}
rs9486719	6q16.1	97 060 124	*FHL5*	G/A	0.80	1.06 (1.04–1.08)	6.8×10^{-10}	1.04 (1.03–1.06)	1.1×10^{-8}
rs28429551	9q34.3	139 243 334	*GPSM1*	A/T	0.76	1.06 (1.04–1.08)	1.7×10^{-8}	1.04 (1.02–1.05)	4.0×10^{-6}
rs8037798	15q24.2	75 240 030	*COX5A-RPP25*	G/T	0.23	1.05 (1.03–1.07)	3.8×10^{-8}	1.02 (1.01–1.04)	1.6×10^{-3}

Chr, chromosome; CI, confidence interval; EA, effect allele; EAF, effect allele frequency; OA, other allele; OR, odds ratio; *P*, *P*-value obtained from meta-analysis of the UK Biobank and CARDIoGRAMplusC4D; Pos, base-pair position (hg19).

(*AHDC1*), and 4q22.3 (*PDLIM5*) were either not associated with any CAD risk factor or only yielded suggestive associations (Supplementary material online, *Table S6*). Based on Phenoscanner, the loci on chromosomes 1p21.3 (*SLC44A3*) and 1p36.11 (*AHDC1*) have also not been associated with other disease-related phenotypes, whereas the lead variants (or tightly linked proxies) at the remaining MI loci have been suggestively or significantly associated with other complex traits, including inflammatory cytokines, circulating leucocytes, prostate cancer, and migraine (Supplementary material online, *Table S7*).

Comparison of association signals for MI and CAD phenotypes at novel loci

We next investigated the phenotypic specificity of the association signals for MI and CAD 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 online Methods for details). This was followed by a meta-analysis of the results with summary statistics for CAD provided by the CARDIoGRAMplusC4D Consortium. Compared to MI, all eight loci yielded some degree of association with CAD in our meta-analysis with the UK Biobank and CARDIoGRAMplusC4D, with two loci on chromosomes 1p36.11 and 6q16.1 exhibiting genome-wide significance (*Table 1* and Supplementary material online, *Table S8*). These latter observations suggest that the association signals on chromosomes 1p36.11 and 6q16.1 may not be specific to MI. The associations between the eight novel loci and CAD were also consistent with another recent meta-analysis for CAD using the UK Biobank and CARDIoGRAMplusC4D Consortium[11] (Supplementary material online, *Table S3*).

Since CARDIoGRAMplusC4D used an all-inclusive definition of CAD that incorporated MI,[6] it was not possible to determine the true specificity of the associations for MI vs. CAD using our meta-analysis results 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 and a restricted CAD-only phenotype that excluded subjects with MI. As a positive control locus, we also included *ABO* in these analyses. Consistent with previous studies,[13] our lead SNP (rs9411377) at the *ABO* locus in the UK Biobank was strongly associated with MI, but not the restricted CAD-only phenotype (*Table 2*), thus validating this analytical approach. Seven of the eight novel loci identified for MI 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*). We also evaluated association at the eight novel loci in the UK Biobank in analyses comparing cases defined as having both CAD and MI (CAD^+/MI^+) to controls defined as CAD-only subjects (CAD^+/MI^-). In addition to the expected association with *ABO*, six of the eight loci were nominally associated ($P < 0.05$) with MI among subjects with CAD (*Table 2*). Taken together, these results suggest that the association signals at some of the novel eight loci are either specific to or more robust for MI than with a CAD-only phenotype.

We next carried out the same analyses in the UK Biobank for 15 previously identified loci that have been suggested to modulate risk

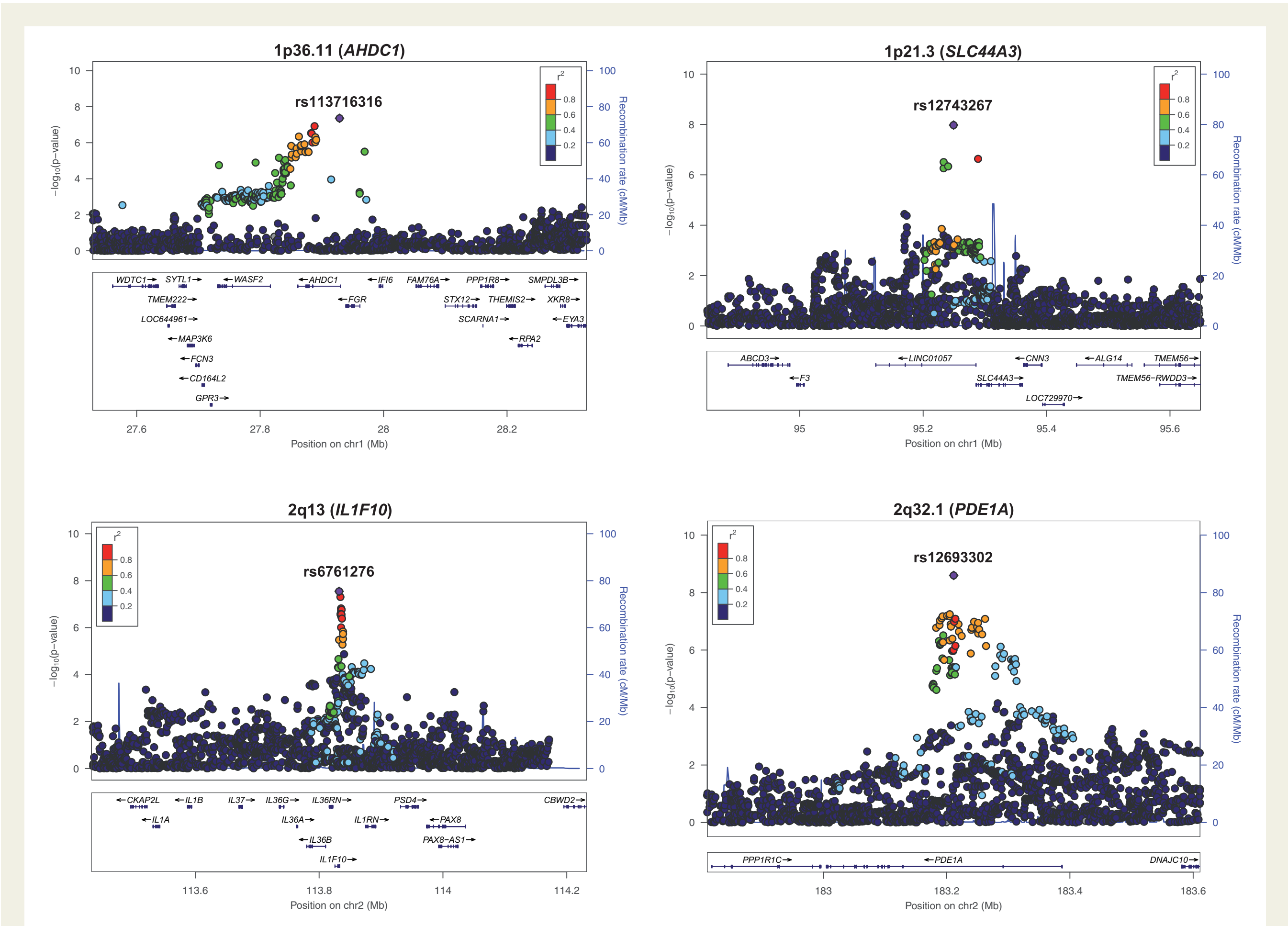

Figure 3 Regional plots of eight novel loci for myocardial infarction. The chromosome band and nearest gene (in parentheses) is indicated for each locus. Each region is centred on the lead single-nucleotide polymorphism (purple diamond) and the genes in the interval are indicated in the bottom panel. The degree of linkage disequilibrium between the lead single-nucleotide polymorphism and other variants is shown as r^2 values according to the colour-coded legend in the box.

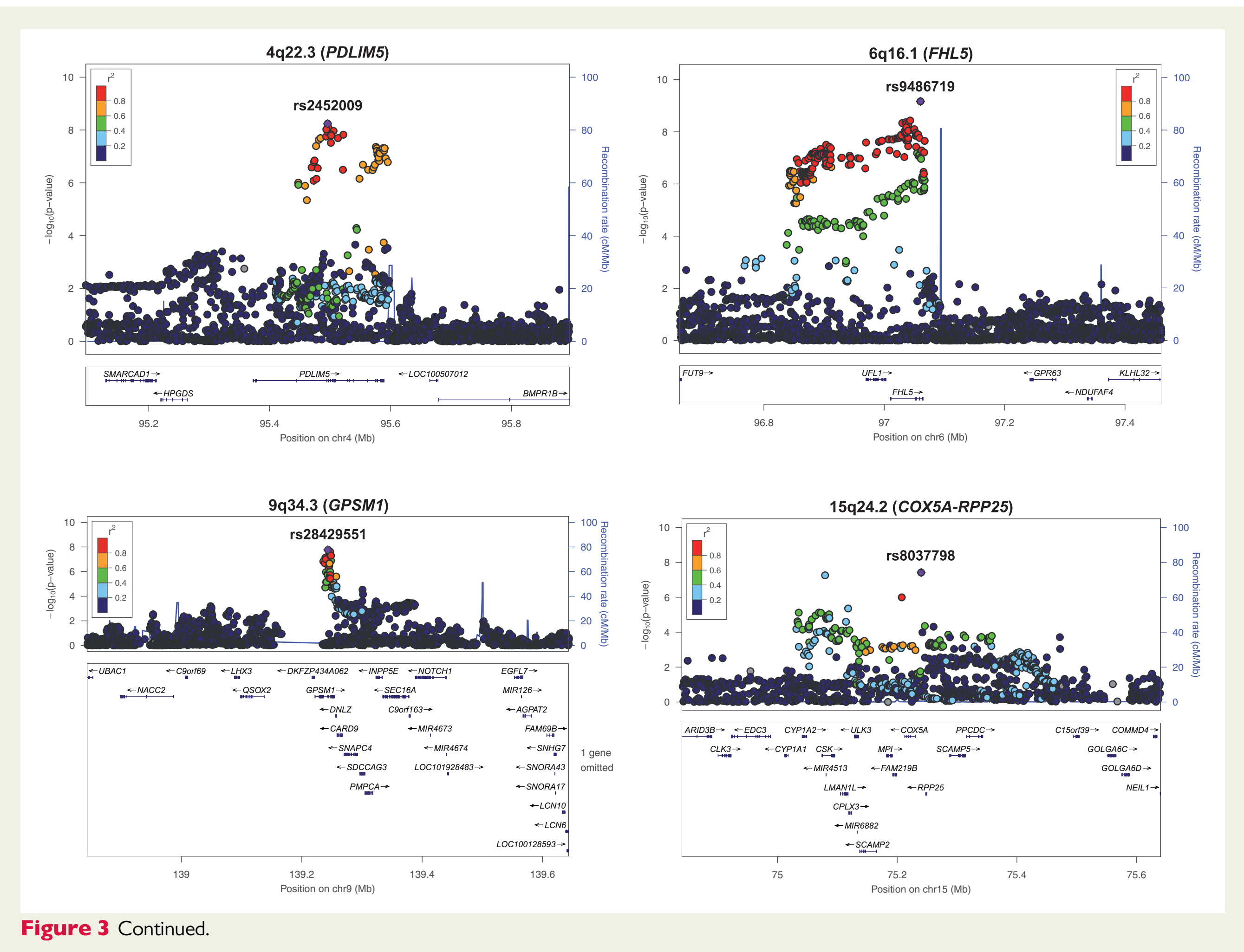

Figure 3 Continued.

of CAD through thrombotic mechanisms.[5–11] At a Bonferroni-corrected significance threshold for testing 15 SNPs ($P = 0.05/15 = 3.3 \times 10^{-3}$), the lead variants from our MI meta-analysis at seven of these loci were associated with MI, but not the CAD-only phenotype, in the UK Biobank (Supplementary material online, *Table S9*). Four of these seven loci were also associated with MI among individuals with CAD (CAD^+/MI^+ vs. CAD^+/MI^-) at $P < 0.05$, but none were associated with the CAD-only phenotype (Supplementary material online, *Table S9*). The remaining eight thrombosis-related loci were associated with both MI and CAD but not with MI in the context of CAD (Supplementary material online, *Table S9*). Thus, some, but not all, of the 15 previously identified CAD/MI loci related to thrombosis exhibited association patterns in the UK Biobank that were similar to those observed at the *ABO* locus and several of the novel MI loci (*Table 2*).

To determine whether the novel MI loci were associated with other CAD phenotypes and whether the association signals differed by ancestry, we carried out sensitivity analyses in the UK Biobank. As shown in Supplementary material online, *Table S10*, there was no evidence for association with 'soft' endpoints, such as angina and death due to CAD, which may have been due to decreased sample size. Although the *P*-values for MI in subjects of non-European ancestry did not reach significance either, presumably also due to decreased power, the effect sizes were all directionally consistent with those in European ancestry subjects (Supplementary material online, *Table S10*) and still contributed to the overall increased significance observed at the MI loci in analyses that included all subjects from the UK Biobank (*Table 2*).

Replication of comparative association signals for MI and CAD in Biobank Japan

To replicate the association signals at the novel loci in a large non-European ancestry population, we carried out the same comparative analyses for MI vs. CAD only in Biobank Japan ($n \sim 165\,000$). Since the restricted CAD phenotype in Biobank Japan could only be defined based on a diagnosis of stable angina, we first evaluated the lead SNP at *9p21* (rs2891168) as a positive control CAD locus. This analysis yielded the expected strong association with CAD only [odds ratio (OR) = 1.14, 95% confidence interval (CI) 1.11–1.17; $P = 7.3 \times 10^{-21}$]. Similar to the UK Biobank, the *ABO* locus was also strongly associated with MI in Biobank Japan but not CAD-only (Supplementary material

Table 2 Comparative associations of the 8 novel loci and the *ABO* locus with MI and CAD in the UK Biobank

SNP	Chr	Pos	Nearest gene(s)	EA/OA	EAF	MI vs. Control (17 505/454 212)		CAD only vs. Control (15 580/454 212)		CAD+/MI+ vs. CAD+/MI− (17 505/15 580)	
						OR (95% CI)	*P*-value	OR (95% CI)	*P*-value	OR (95% CI)	*P*-value
rs113716316	1p36.11	27 928 640	*AHDC1*	G/A	0.93	1.11 (1.07–1.16)	7.2×10^{-7}	1.07 (1.02–1.12)	4.1×10^{-3}	1.04 (0.98–1.11)	0.21
rs12743267	1p21.3	95 249 306	*SLC44A3*	C/T	0.76	1.04 (1.01–1.06)	3.1×10^{-3}	1.00 (0.97–1.03)	0.98	1.04 (1.01–1.08)	0.02
rs6761276	2q13	113 832 312	*IL1F10*	T/C	0.42	1.03 (1.01–1.06)	1.9×10^{-3}	1.01 (0.99–1.03)	0.44	1.03 (0.99–1.06)	0.11
rs12693302	2q32.1	183 211 443	*PDE1A*	G/A	0.36	1.06 (1.03–1.08)	1.3×10^{-6}	0.98 (0.96–1.01)	0.19	1.07 (1.04–1.10)	2.9×10^{-5}
rs2452009	4q22.3	95 495 908	*PDLIM5*	A/G	0.70	1.04 (1.02–1.07)	2.6×10^{-4}	1.01 (0.98–1.03)	0.68	1.03 (1.001–1.07)	0.04
rs28429551	9q34.3	139 243 334	*GPSM1*	A/T	0.76	1.07 (1.04–1.10)	4.8×10^{-8}	1.01 (0.98–1.03)	0.54	1.07 (1.03–1.11)	2.8×10^{-4}
rs8037798	15q24.2	75 240 030	*COX5A-RPP25*	G/T	0.24	1.05 (1.03–1.08)	3.2×10^{-5}	1.00 (0.97–1.02)	0.85	1.06 (1.02–1.10)	1.7×10^{-3}
rs9411377	9q34.2	136 145 404	*ABO*	A/C	0.30	1.06 (1.04–1.09)	3.3×10^{-7}	0.99 (0.97–1.02)	0.67	1.07 (1.03–1.10)	1.3×10^{-4}

Number of cases and controls for each phenotype defined in the UK Biobank are shown in parentheses.
For CAD$^+$/MI$^+$ vs. CAD$^+$/MI$^-$ analyses, cases were defined as subjects positive for both CAD and MI; controls were defined as CAD positive subjects without MI.
Chr, chromosome; CI, confidence interval; EA, effect allele; EAF, effect allele frequency; OA, other allele; OR, odds ratio; *P*, *P*-value obtained from linear mixed model analysis in UK Biobank; Pos, base-pair position (hg19).

online, *Table S11*). Based on these results further validating this comparative strategy and its applicability to Biobank Japan, we tested the novel regions for association with MI vs. CAD-only. 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 CARDIoGRAMplusC4D (*Table 1*), they were not considered in these analyses. None of the six remaining newly identified loci were associated with the CAD-only phenotype, whereas three regions (1p21.3, 2q32.1, and 15q24.2) yielded nominal ($P < 0.05$) associations with MI in Biobank Japan (Supplementary material online, *Table S11*) that were directionally consistent with the UK Biobank (*Table 2*). However, only the lead SNP (rs12743267) at the chromosome 1p21.3 locus harbouring *SLC44A3* was also associated with MI among CAD cases (Supplementary material online, *Table S11*).

Preferential association of the *SLC44A3* locus with MI in the presence of atherosclerosis

We next sought to replicate the association signals for MI at the novel loci using independent cohorts in which the presence of CAD was more directly assessed by angiography. Case–control analyses were carried out in a first set of six cohorts with ∼14 000 angiographically documented CAD patients with MI (CAD$^+$/MI$^+$ cases; $n = 6514$) and without MI (CAD$^+$/MI$^-$ controls; $n = 7411$) (Supplementary material online, *Table S12*). A fixed-effects meta-analysis with these six cohorts revealed consistent and strong association of the *SLC44A3* locus on chromosome 1p21.3 with risk of MI among individuals with CAD (OR = 1.16, 95% CI 1.09–1.23; $P = 3.3 \times 10^{-6}$) (*Table 3*), with no significant evidence for heterogeneity (*P*-het = 0.10) (Supplementary material online, *Table S12*). Exclusion of the Emory cohort, which itself exhibited a very strong effect size with large variation, did not appreciably change the direction or significance level of the overall association between the *SLC44A3* locus and MI (OR = 1.15, 95% CI 1.08–1.22; $P = 6.2 \times 10^{-6}$) (Supplementary material online, *Table S12*).

As another replication study, we evaluated association of the newly identified MI loci in 10 additional angiography-based cohorts comprising 7412 CAD$^+$/MI$^+$ cases and 5542 CAD$^+$/MI$^-$ controls (Supplementary material online, *Table S13*). These analyses also yielded evidence for association of the *SLC44A3* locus with MI in the context of CAD (OR = 1.09, 95% CI 1.03–1.16; $P = 2.1 \times 10^{-3}$) but not the remaining five loci. When all 16 angiography-based cohorts were meta-analysed together (n∼27 000), association of the *SLC44A3* locus with MI in the presence of coronary atherosclerosis increased in significance by several fold (OR = 1.12, 95% CI 1.08–1.17; $P = 5.6 \times 10^{-8}$) (*Table 3*). Notably, the *SLC44A3* locus was highly significantly associated with MI in an all-inclusive meta-analysis with UK Biobank, Biobank Japan, and the 16 angiography-based cohorts ($n = 41\ 336$ CAD$^+$/MI$^+$ cases and 40 363 CAD$^+$/MI$^-$ controls) and exceeded the threshold for genome-wide significance (OR = 1.07, 95% CI 1.05–1.10; $P = 5.4 \times 10^{-11}$). Taken together with the weak associations observed with CAD in the meta-analyses with CARDIoGRAMplusC4D and UK Biobank and the comparative analyses in the UK Biobank and Biobank Japan, these results provide compelling evidence for the *SLC44A3* locus being preferentially associated

Table 3 Association of novel loci with MI in the presence of CAD in angiography-based cohorts

SNP	Chr	Pos	Nearest Gene(s)	EA/OA	EAF	Angiography cohorts I (6514/7411)		Angiography cohorts II (7412/5542)		Meta-analysis (13 926/12 953)	
						OR (95% CI)	*P*-value[a]	OR (95% CI)	*P*-value[b]	OR (95% CI)	*P*-value
rs12743267	1p21.3	95 249 306	*SLC44A3*	C/T	0.77	1.16 (1.09–1.23)	3.3×10^{-6}	1.09 (1.03–1.16)	2.1×10^{-3}	1.12 (1.08–1.17)	5.6×10^{-8}
rs6761276	2q13	113 832 312	*IL1F10*	T/C	0.43	1.03 (0.98–1.08)	0.31	1.03 (0.97–1.08)	0.34	1.03 (0.99–1.06)	0.16
rs12693302	2q32.1	183 211 443	*PDE1A*	G/A	0.35	0.99 (0.93–1.04)	0.63	1.07 (1.004–1.13)	0.04	1.02 (0.98–1.06)	0.28
rs2452009	4q22.3	95 495 908	*PDLIM5*	A/G	0.69	1.01 (0.95–1.06)	0.83	1.04 (0.99–1.10)	0.13	1.03 (0.99–1.07)	0.21
rs28429551	9q34.3	139 243 334	*GPSM1*	A/T	0.76	1.02 (0.95–1.09)	0.66	1.04 (0.95–1.13)	0.37	1.03 (0.97–1.08)	0.36
rs8037798	15q24.2	75 240 030	*COX5A-RPP25*	G/T	0.26	1.004 (0.93–1.08)	0.91	1.03 (0.97–1.10)	0.31	1.02 (0.98–1.06)	0.38

Number of cases, defined as subjects positive for MI and CAD based on angiographic data (CAD^+/MI^+), and controls, defined as CAD positive subjects without MI (CAD^+/MI^-), are shown in parentheses.
Chr, chromosome; CI, confidence interval; EA, effect allele; EAF, effect allele frequency in European ancestry subjects; OA, other allele; OR, odds ratio; Pos, base-pair position (hg19).
[a]*P*, *P*-value from meta-analysis of the GeneBank, Emory Cardiovascular Biobank, ANGES/FINCAVAS, LURIC, LIFE-Heart, and UCORBIO cohorts.
[b]*P*, *P*-Value from meta-analysis of the SMART, SCADGENS, PennCath, MedStar, OHGS, CADomics, ADVANCE, WTCCC, and CATHGEN cohorts.

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 data from the MEGASTROKE Consortium,[18] there was no evidence for association of rs12743267 with most forms of stroke except for nominal associations with cardioembolic and small vessel stroke in subjects of European ancestry that would not be considered significant at a Bonferroni corrected *P*-value of 0.01 for testing five forms of stroke (0.05/5 = 0.01) (Supplementary material online, *Table S14*). Second, variants at the chromosome 1p21.3 locus had been previously associated with circulating levels of D-dimer,[19] which is produced when cross-linked fibrin is degraded by plasmin and the most widely used clinical marker of activated blood coagulation.[20] However, rs12743267 was not associated with D-dimer levels (beta = -0.011; SE = 0.007; *P* = 0.12) based on a GWAS carried out by the CHARGE Consortium[19] and the lead SNP for D-dimer (rs12029080) was not associated with MI in our meta-analysis with the UK Biobank and CARDIoGRAMplusC4D Consortium (OR = 0.99, 95% CI 0.98–1.01; *P* = 0.32) or in Biobank Japan (OR = 0.98, 95% CI 0.96–1.01; *P* = 0.12). Lastly, *SLC44A2*, a member of the solute carrier family of membrane transporters that includes *SLC44A3*, has been associated with venous thromboembolism (VTE),[21] 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 GWAS carried out by the INVENT Consortium.[21] By comparison, the lead VTE SNP in *SLC44A2* (rs2288904) was associated with CAD (OR = 1.04, 95% CI 1.03–1.05; $P = 7.0 \times 10^{-8}$) and MI (OR = 1.04, 95% CI 1.02–1.06; $P = 1.5 \times 10^{-5}$) in our meta-analyses, as well as with CAD in Biobank Japan (OR = 1.03, 95% CI 1.01–1.06; $P = 1.1 \times 10^{-3}$).

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.[22] In humans, elevated plasma levels of choline and products of its metabolism have been linked to risk of MI-related outcomes.[23–25] However, we did not obtain evidence in the Genebank cohort for association of the *SLC44A3* locus with plasma levels of these metabolites or a panel of choline-related small molecule amines that have also been associated with CAD and MI[26–31] (Supplementary material online, *Table S15*). Based on data from three metabolomics and proteomics studies,[32–34] the *SLC44A3* locus did yield associations with small molecules in plasma or urine, but these would not be considered significant at Bonferroni-corrected thresholds for the number of analytes tested in each data set (Supplementary material online, *Table S16*).

Functional analysis of *SLC44A3*

We next used functional studies to evaluate *SLC44A3* as a candidate causal gene at the chromosome 1p21.3 locus. Among 600 CAD patients in the STARNET study,[17] *SLC44A3* was expressed at

relatively high levels in several MI-relevant tissues, such as atherosclerotic aortic root, adipose tissue, mammary artery, and liver (*Figure 4A*). In addition, the lead SNP on chromosome 1p21.3 yielded *cis* eQTLs for *SLC44A3* in atherosclerotic aorta and mammary artery, where the MI risk allele (C) was associated with increased expression (*Figure 4B*). In the GTEx Project, similar eQTLs were observed in aorta and coronary artery (*Figure 4C*), as well as in whole blood and various components of the gastrointestinal tract (Supplementary material online, *Table S4*). These findings were consistent with mRNA levels of *SLC44A3* being significantly higher in ischaemic coronary arteries compared to non-diseased coronary arteries in another independent data set (*Figure 4D*). To explore the vascular cell type in which *SLC44A3* could mediate its biological effects on MI, we used RNAseq and functional data from two additional independent data sets of human aortic endothelial cells (HAECs) and smooth muscle cells (SMCs), respectively. Compared to vehicle control, *SLC44A3* expression was significantly up-regulated in HAECs treated with the pro-atherogenic inflammatory cytokine interleukin (IL)-1β (*Figure 4E*). *SLC44A3* expression in SMCs was also modestly, but significantly, inversely correlated with migration towards platelet-derived growth factor-BB *in vitro* (*Figure 4F*). Taken together, these data provide supportive functional evidence that *SLC44A3* is 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 that in total incorporated ∼831 000 subjects from the UK Biobank, CARDIoGRAMplusC4D Consortium, Biobank Japan, and over a dozen angiography-based cohorts. Based on our own meta-analyses with CARDIoGRAMplusC4D and the UK Biobank and another recent comparable analysis,[11] the strength of the associations at the eight loci was, 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 provided further evidence that six of the novel loci were either specifically or more strongly associated with MI than with CAD. However, only one of these loci yielded independent association with MI among subjects with CAD in replication analyses. Thus, it is possible that some of the novel loci may also influence risk of CAD and are therefore not truly specific for MI. Nevertheless, our collective analyses led to the identification of eight novel genetic determinants of cardiovascular outcomes, bringing the total number of loci associated with atherosclerosis-related outcomes to 213.

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 were also supported by association signals in the comparably sized Biobank Japan that were equivalent in magnitude and significance to those in the UK Biobank. Further and consistent association of the *SLC44A3* locus with MI was also observed in an initial set of 6 followed by another 10 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 (OR = 1.12) was stronger than the ORs obtained in the GWAS meta-analysis, UK Biobank, or Biobank Japan (OR∼1.05), and equivalent to some of the most significantly associated loci identified to date for CAD.[11] Taken together, these results support the notion that the biological mechanism(s) underlying the association of the *SLC44A3* locus may be related to plaque rupture rather than plaque progression per se. In this regard, *ABO* was similarly identified as being only associated with MI in the original study by Reilly *et al.*,[13] which we replicated in our analogous comparative analyses with the UK Biobank and Biobank Japan. Thus, to our knowledge, the *SLC44A3* locus represents the second and only other genetic risk factor 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.[35] However, it should be noted that our meta-analyses for MI had approximately 10-fold higher numbers of subjects than the VTE GWAS.[21] 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 of variants around its gene with circulating D-dimer levels,[19] *F3* would be considered a more biologically plausible candidate gene for a thrombosis-related phenotype such as MI. However, we did not obtain any evidence that would prioritize *F3* as a candidate causal gene 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. 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 observations and the presence of *cis* eQTLs for *SLC44A3* in multiple tissues and independent data sets, we focused on *SLC44A3* as a candidate causal gene for MI. *SLC44A3* is one of five members of the SLC44 family of solute carriers (*SLC44A1-5*) that have been proposed to function as choline transporters.[22] However, *SLC44A1* is the only member of this transporter family for which a role in transporting choline across both the plasma and mitochondrial membranes has been demonstrated by direct experimentation.[36,37] In addition, the *SLC44A3* locus was not associated with plasma levels of choline, pro-atherogenic choline-derived small molecule amines, such as trimethylamine *N*-oxide and betaine,[24,25] or with a large panel of metabolomic and proteomic targets in plasma and urine.[32–35] Thus, additional functional studies will be needed to demonstrate whether *SLC44A3* encodes a transporter for choline or other molecules and whether such activity would modulate levels of metabolites that influence risk of MI.

Several lines of evidence from our functional and bioinformatics analyses further pointed to *SLC44A3* as one causal positional candidate on chromosome 1p21.3 and suggested that putative biological

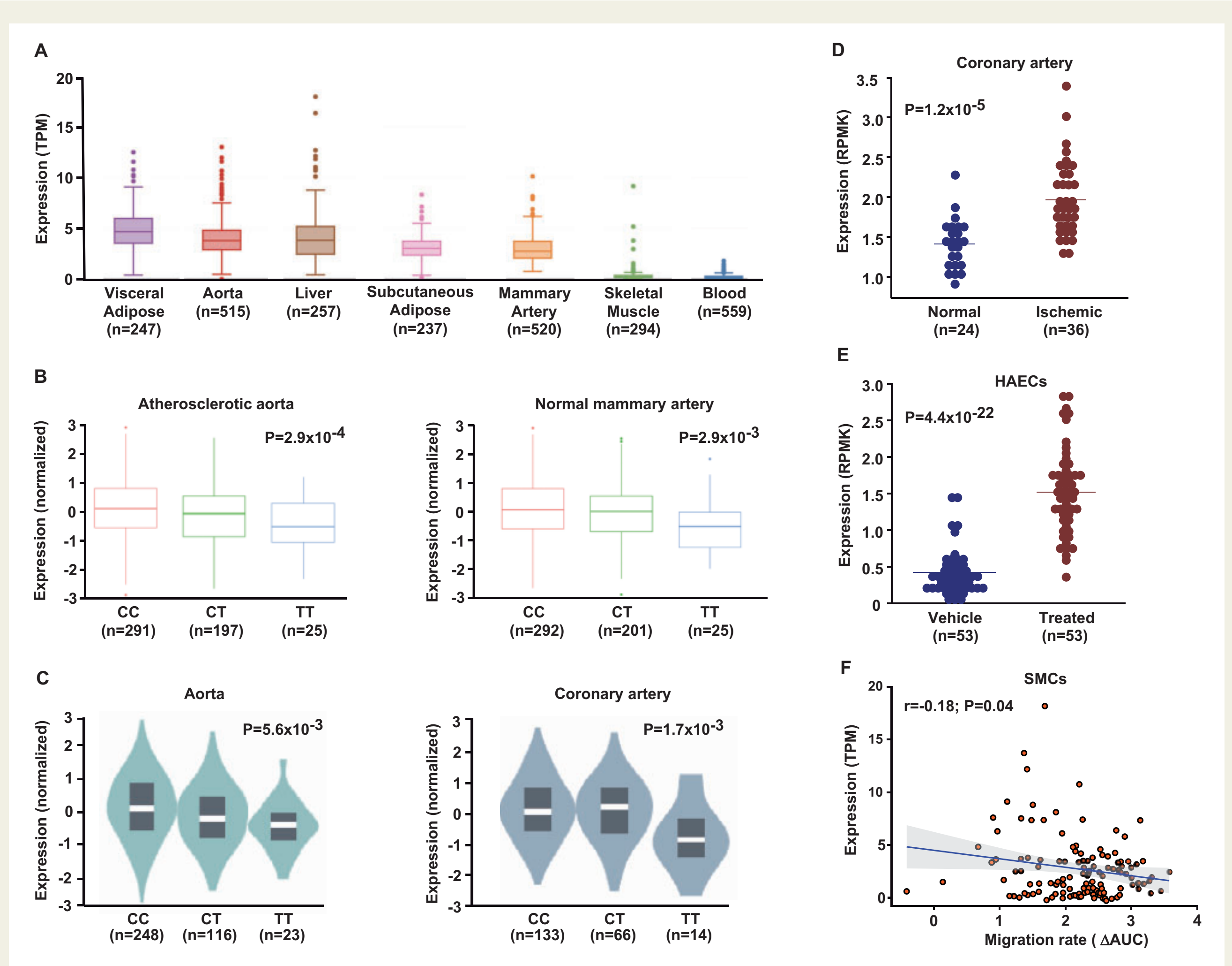

Figure 4 Functional analyses of *SLC44A3* in myocardial infarction-relevant tissues. (*A*) In the STARNET cohort, *SLC44A3* was expressed at relatively high levels in tissues relevant to myocardial infarction, including atherosclerotic aortic root (aorta), visceral adipose, mammary artery, and liver. (*B*) The lead single-nucleotide polymorphism 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 myocardial infarction risk allele (C) was associated with higher mRNA levels. (*C*) A similar pattern of *cis* eQTLs was also independently observed with the *SLC44A3* locus in aorta and coronary artery based on data from the GTEx Project. (*D*) In another independent human data set, *SLC44A3* expression was increased in ischaemic coronary arteries (n = 36) from heart donors with coronary artery disease compared to normal coronary arteries from non-diseased donors (n = 24). (*E*) Incubation of human aortic endothelial cells isolated from a different and independent set of anonymous heart donors (n = 53) with interleukin-1β for 4 h up-regulated *SLC44A3* expression ∼3-fold compared to paired vehicle-treated human aortic endothelial cells. (*F*) Using a fourth independent human data set (n = 151), *SLC44A3* expression was also observed in smooth muscle cells and inversely correlated with migration rate towards platelet-derived growth factor-BB *in vitro*.

mechanisms through which this gene could influence plaque rupture and/or thrombosis may be through direct effects at the level of the vessel wall. First, *SLC44A3* was expressed in MI-relevant vascular tissues, such as the aorta and mammary artery. Second, co-localization analyses carried out in atherosclerotic aorta yielded a strong posterior probability for *SLC44A3*, but not the other genes at the chromosome 1p21.3 locus (i.e. *F3*), as being causal for MI. Third, carriers of the MI risk allele had significantly higher *SLC44A3* mRNA levels than non-carriers, with a stronger effect size observed in atherosclerotic aortic root than mammary artery. The same *cis* eQTLs for *SLC44A3* were independently observed in aorta and coronary artery in the GTEx Project. Fourth, expression analyses in two independent heart donor data sets demonstrated up-regulation of *SLC44A3* in ischaemic coronary arteries by ∼50% 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.[38] Lastly, although we did not detect an eQTL for *SLC44A3* in SMCs (or HAECs), possibly due to insufficient power, 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 promote secretion of extra cellular matrix proteins and the formation of a protective fibrous cap that renders a lesion less prone to rupture.[39] Taken together, 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 data from the GTEx Project, expression was also high in kidney, pancreas, the small intestine, and colon. Moreover, the eQTLs in GTEx for *SLC44A3* in aorta and coronary artery were modest relative to those observed in whole blood, heart, pancreas, liver, and colon. In some of these tissues, such as liver, the allelic association of rs12743267 with *SLC44A3* mRNA levels was also opposite to that observed in arterial tissues. Although these observations suggest that *SLC44A3* could influence risk of MI through mechanisms related to metabolism, the *SLC44A3* locus was not associated with traditional CAD risk factors, such as lipid levels and type 2 diabetes. Nonetheless, we still cannot rule out the possibility that *SLC44A3* could also increase risk of plaque rupture 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, the majority of 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 *SLC44A3* locus yielded an equivalent association with MI in Biobank Japan and exhibited directionally consistent effect sizes in other Asian populations, suggesting that at least a subset of the association signals identified herein may also be relevant in other ethnicities as well. Second, it is possible, albeit unlikely, that some subjects in the UK Biobank and CARDIoGRAMplusC4D Consortium overlapped, which could have been a confounding factor in the meta-analysis. However, a recent analysis concluded that duplicate samples between CARDIoGRAMplusC4D and the UK Biobank were minimal (<0.1%) and would not significantly influence test statistics.[11] Third, 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.[11] There could also have been misclassification in our analyses since, for example, MI and CAD may not have been defined in exactly the same in CARDIoGRAMplusC4D, the UK Biobank, and Biobank Japan. We note that if such misclassifications had occurred, they would have most likely been non-differential and biased the results towards the null. Finally, even though SNPs with minor allele frequencies as low as 0.5% were included in our 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, along with vascular cell-specific eQTL analyses, will require additional investigation.

In summary, our results identify several previously unrecognized loci for MI and provide new avenues for exploring the pathophysiology of vulnerable atherosclerotic lesions. Most importantly, 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 potential genetic susceptibility factor. Future studies will be needed to explore the clinical relevance of these findings for patients at risk of MI.

URLs

The UK Biobank (https://www.ukbiobank.ac.uk/); CARDIoGRAMplusC4D, http://www.cardiogramplusc4d.org/; Biobank Japan, https://biobankjp.org/english/index.html; GWAMA, https://www.geenivaramu.ee/en/tools/gwama/; Genotype-Tissue Expression Project, http://gtexportal.org/; Phenoscanner, http://www.phenoscanner.medschl.cam.ac.uk/phenoscanner; R statistical software, http://www.R-project.org/.

Supplementary material

Supplementary material is available at *European Heart Journal* online.

Data availability

Full summary statistics relating to the GWAS analysis in the UK Biobank and the meta-analysis with CARDIoGRAMplusC4D 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.

Acknowledgements

We gratefully acknowledge the UK Biobank Resource for providing access to their data under Application Number 33307 and the CARDIoGRAMplusC4D 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.

Funding

This work was supported, in part, by National Institutes of Health [R01HL133169, R01HL148110, R01ES021801, R01ES025786, R01MD010358, P01ES022845, R01HL103866, R01HL128300, R01HL135920, R01HL147883, R01HL144651, R00HL138193, R21HL135230, R01HL147187, R01HL148239, R01HL125863, R01HL142856, R01HL113147, R01HL150359, and K24HL107643]; American Heart Association [18POST33990046, 19TPA34910021, and A14SFRN20840000]; Transatlantic Networks of Excellence Awards from Foundation Leducq [12CVD02, 18CVD02]; the EPIDEMIOM-VT Senior Chair from the University of Bordeaux Initiative of Excellence IdEX [to D.-A.T.]; British Heart Foundation Intermediate Fellowship [FS/14/76/30933 to R.S.P.]; the National Institute for Health Research University College London Hospitals Biomedical Research Centre; the Swedish Research Council [2018-02529]; and Astra-Zeneca through ICMC, Karolinska Institutet, Sweden [to J.L.M.B.]. The GeneBank study was supported in part by National Institutes of Health grants [P01HL098055, P01HL076491, P01HL147823, R01HL126827, 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 [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 [755320 for TAXINOMISIS and 848146 for To Aition]. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflict of interest: 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 has 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. S.L.H. also reports having been paid as a consultant from Procter & Gamble Company and having received research funds from Procter & Gamble Company and Roche. M.S. receives funding from Pfizer Inc. for a project not related to this research. 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. G.J.V. reports grants from MicroPort Medical Shanghai, grants from Daiichi Sankyo, and personal fees from AstraZeneca. All other authors have no conflict of interests to declare.

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 Committee and Stroke Statistics Subcommittee. 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; JUPITER Study Group. Rosuvastatin to prevent vascular events in men and women with elevated C-reactive protein. *N Engl J Med* 2008;**359**:2195–2207.
3. Hansson GK. Inflammation, atherosclerosis, and coronary artery disease. *N Engl J Med* 2005;**352**:1685–1695.
4. Schunkert H, Erdmann J, Samani NJ. Genetics of myocardial infarction: a progress report. *Eur Heart J* 2010;**31**:918–925.
5. Erdmann J, Stark K, Esslinger UB, Rumpf PM, Koesling D, de Wit C, Kaiser FJ, Braunholz D, Medack A, Fischer M, Zimmermann ME, Tennstedt S, Graf E, Eck S, Aherrahrou Z, Nahrstaedt J, Willenborg C, Bruse P, Brænne I, Nöthen MM, Hofmann P, Braund PS, Mergia E, Reinhard W, Burgdorf C, Schreiber S, Balmforth AJ, Hall AS, Bertram L, Steinhagen-Thiessen E, Li S-C, März W, Reilly M, Kathiresan S, McPherson R, Walter U; CARDIoGRAM, Ott J, Samani NJ, Strom TM, Meitinger T, Hengstenberg C, Schunkert H. Dysfunctional nitric oxide signalling increases risk of myocardial infarction. *Nature* 2013;**504**:432–436.
6. 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. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet* 2015;**47**:1121–1130.
7. Verweij N, Eppinga RN, Hagemeijer Y, 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.
8. 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, 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, Saleheen D. Fifteen new risk loci for coronary artery disease highlight arterial-wall-specific mechanisms. *Nat Genet* 2017;**49**:1113–1119.
9. 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, Kathiresan S. Genetic analysis in UK Biobank links insulin resistance and transendothelial migration pathways to coronary artery disease. *Nat Genet* 2017;**49**:1392–1397.
10. 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, Schunkert H, Farrall M, Danesh J, Samani NJ, Watkins H, Deloukas P. Association analyses based on false discovery rate implicate new loci for coronary artery disease. *Nat Genet* 2017;**49**:1385–1391.
11. van der Harst P, 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.
12. Koyama S, Ito K, Terao C, Akiyama M, Horikoshi M, Momozawa Y, Matsunaga H, Ieki H, Ozaki K, Onouchi Y, Takahashi A, Nomura S, Morita H, Akazawa H, Kim C, Seo JS, Higasa K, Iwasaki M, Yamaji T, Sawada N, Tsugane S, Koyama T, Ikezaki H, Takashima N, Tanaka K, Arisawa K, Kuriki K, Naito M, Wakai K, Suna S, Sakata Y, Sato H, Hori M, Sakata Y, Matsuda K, Murakami Y, Aburatani H, Kubo M, Matsuda F, Kamatani Y, Komuro I. Population-specific and trans-ancestry genome-wide analyses identify distinct and shared genetic risk loci for coronary artery disease. *Nat Genet* 2020;**52**:1169–1177.
13. 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, 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**:383–392.
14. Beutner F, Teupser D, Gielen S, Holdt LM, Scholz M, Boudriot E, Schuler G, 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.
15. GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* 2013;**45**:580–585.
16. Staley JR, Blackshaw J, Kamat MA, Ellis S, Surendran P, Sun BB, Paul DS, Freitag D, Burgess S, Danesh J, Young R, Butterworth AS. PhenoScanner: a database of human genotype-phenotype associations. *Bioinformatics* 2016;**32**:3207–3209.
17. 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, Bjorkegren JL. Cardiometabolic risk loci share downstream cis- and trans-gene regulation across tissues and diseases. *Science* 2016; **353**:827–830.
18. Malik R, Chauhan G, Traylor M, Sargurupremraj M, Okada Y, Mishra A, Rutten-Jacobs L, Giese A-K, van der Laan SW, Gretarsdottir S, Anderson CD, Chong M, Adams HHH, Ago T, Almgren P, Amouyel P, Ay H, Bartz TM, Benavente OR, Bevan S, Boncoraglio GB, Brown RD, Butterworth AS, Carrera C, Carty CL, Chasman DI, Chen W-M, Cole JW, Correa A, Cotlarciuc I, Cruchaga C, Danesh J, de Bakker PIW, DeStefano AL, den Hoed M, Duan Q, Engelter ST, Falcone GJ, Gottesman RF, Grewal RP, Gudnason V, Gustafsson S, Haessler J, Harris TB, Hassan A, Havulinna AS, Heckbert SR, Holliday EG, Howard G, Hsu F-C, Hyacinth HI, Ikram MA, Ingelsson E, Irvin MR, Jian X, Jiménez-Conde J, Johnson JA, Jukema JW, Kanai M, Keene KL, Kissela BM, Kleindorfer DO, Kooperberg C, Kubo M, Lange LA, Langefeld CD, Langenberg C, Launer LJ, Lee J-M, Lemmens R, Leys D, Lewis CM, Lin W-Y, Lindgren AG, Lorentzen E, Magnusson PK, Maguire J, Manichaikul A, McArdle PF, Meschia JF, Mitchell BD, Mosley TH, Nalls MA, Ninomiya T, O'Donnell MJ, Psaty BM, Pulit SL, Rannikmäe K, Reiner AP, Rexrode KM, Rice K, Rich SS, Ridker PM, Rost NS, Rothwell PM, Rotter JI, Rundek T, Sacco RL, Sakaue S, Sale MM, Salomaa V, Sapkota BR, Schmidt R, Schmidt CO, Schminke U, Sharma P, Slowik A, Sudlow CLM, Tanislav C, Tatlisumak T, Taylor KD, Thijs VNS, Thorleifsson G, Thorsteinsdottir U, Tiedt S, Trompet S, Tzourio C, van Duijn CM, Walters M, Wareham NJ, Wassertheil-Smoller S, Wilson JG, Wiggins KL, Yang Q, Yusuf S, Bis JC, Pastinen T, Ruusalepp A, Schadt EE, Koplev S, Björkegren JLM, Codoni V, Civelek M, Smith NL, Trégouët DA, Christophersen IE, Roselli C, Lubitz SA, Ellinor PT, Tai ES, Kooner JS, Kato N, He J, van der Harst P, Elliott P, Chambers JC, Takeuchi F, Johnson AD, Sanghera DK, Melander O, Jern C, Strbian D, Fernandez-Cadenas I, Longstreth WT, Rolfs A, Hata J, Woo D, Rosand J, Pare G, Hopewell JC, Saleheen D, Stefansson K, Worrall BB, Kittner SJ, Seshadri S, Fornage M, Markus HS, Howson JMM, Kamatani Y, Debette S, Dichgans M. Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. *Nat Genet* 2018;**50**:524–537.
19. 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, Hayward C. Genetic predictors of fibrin D-dimer levels in healthy adults. *Circulation* 2011;**123**:1864–1872.
20. Adam SS, Key NS, Greenberg CS. D-dimer antigen: current concepts and future prospects. *Blood* 2009;**113**:2878–2887.
21. 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, 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**:532–542.
22. Traiffort E, O'Regan S, Ruat M. The choline transporter-like family SLC44: properties and roles in human diseases. *Mol Aspects Med* 2013;**34**:646–654.
23. 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, Hazen SL. Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature* 2011;**472**:57–63.
24. Tang WH, Wang Z, Levison BS, Koeth RA, Britt EB, Fu X, Wu Y, Hazen SL. Intestinal microbial metabolism of phosphatidylcholine and cardiovascular risk. *N Engl J Med* 2013;**368**:1575–1584.
25. Wang Z, Tang WH, Buffa JA, Fu X, Britt EB, Koeth RA, Levison BS, Fan Y, Wu Y, Hazen SL. Prognostic value of choline and betaine depends on intestinal microbiota-generated metabolite trimethylamine-N-oxide. *Eur Heart J* 2014;**35**: 904–910.
26. 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, Hazen SL. Intestinal microbiota metabolism of L-carnitine, a nutrient in red meat, promotes atherosclerosis. *Nat Med* 2013;**19**:576–585.
27. 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, Hazen SL. gamma-Butyrobetaine is a proatherogenic intermediate in gut microbial metabolism of L-carnitine to TMAO. *Cell Metab* 2014;**20**:799–812.
28. Hartiala J, Bennett B J, Tang WW, Wang Z, Stewart AF, Roberts R, Mcpherson R, Lusis AJ, Hazen SL, Allayee H, 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**:1307–1313.
29. Hartiala JA, Tang WH, Wang Z, Crow AL, Stewart AF, Roberts R, McPherson R, Erdmann J, Willenborg C, Hazen SL, Allayee H. Genome-wide association study and targeted metabolomics identifies sex-specific association of CPS1 with coronary artery disease. Nat Commun 2016;7:10558.
30. 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, Hazen SL. Untargeted metabolomics identifies trimethyllysine, a TMAO-producing nutrient precursor, as a predictor of incident cardiovascular disease risk. *JCI Insight* 2018;**3**:e99096.
31. 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, Allayee H. Genetic determinants of circulating glycine levels and risk of coronary artery disease. *J Am Heart Assoc* 2019;**8**:e011922.
32. Shin SY, Fauman EB, Petersen AK, Krumsiek J, Santos R, Huang J, Arnold M, Erte I, Forgetta V, Yang TP, Walter K, Menni C, Chen L, Vasquez L, Valdes AM, Hyde CL, Wang V, Ziemek D, Roberts P, Xi L, Grundberg E, Waldenberger M, Richards JB, Mohney RP, Milburn MV, John SL, Trimmer J, Theis FJ, Overington JP, Suhre K, Brosnan MJ, Gieger C, Kastenmuller G, Spector TD, Soranzo N. An atlas of genetic influences on human blood metabolites. *Nat Genet* 2014;**46**: 543–550.
33. Sun BB, Maranville JC, Peters JE, Stacey D, Staley JR, Blackshaw J, Burgess S, Jiang T, Paige E, Surendran P, Oliver-Williams C, Kamat MA, Prins BP, Wilcox SK, Zimmerman ES, Chi A, Bansal N, Spain SL, Wood AM, Morrell NW, Bradley JR, Janjic N, Roberts DJ, Ouwehand WH, Todd JA, Soranzo N, Suhre K, Paul DS, Fox CS, Plenge RM, Danesh J, Runz H, Butterworth AS. Genomic atlas of the human plasma proteome. *Nature* 2018;**558**:73–79.
34. Schlosser P, Li Y, Sekula P, Raffler J, Grundner-Culemann F, Pietzner M, Cheng Y, Wuttke M, Steinbrenner I, Schultheiss UT, Kotsis F, Kacprowski T, Forer L, Hausknecht B, Ekici AB, Nauck M, Volker U, Investigators G, Walz G, Oefner PJ, Kronenberg F, Mohney RP, Kottgen M, Suhre K, Eckardt KU, Kastenmuller G, Kottgen A. Genetic studies of urinary metabolites illuminate mechanisms of detoxification and excretion in humans. *Nat Genet* 2020;**52**:167–176.
35. Dichgans M, Malik R, König 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, März 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, Gschwendtner A, Bevan S, Chen Y-C, DeStefano AL, Parati EA, Quertermous T, Ziegler A, Boerwinkle E, Holm H, Fischer M, Kessler T, Willenborg C, Laaksonen R, Voight BF, Stewart AFR, Rader DJ, Hall AS, Kooner JS; 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.

36. O'Regan S, Traiffort E, Ruat M, Cha N, Compaore D, Meunier FM. An electric lobe suppressor for a yeast choline transport mutation belongs to a new family of transporter-like proteins. *Proc Natl Acad Sci USA* 2000;**97**:1835–1840.
37. Michel V, Bakovic M. The solute carrier 44A1 is a mitochondrial protein and mediates choline transport. *FASEB J* 2009;**23**:2749–2758.
38. Gimbrone MA, García-Cardeña G. Endothelial cell dysfunction and the pathobiology of atherosclerosis. *Circ Res* 2016;**118**:620–636.
39. 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, Owens GK. Activation of the pluripotency factor OCT4 in smooth muscle cells is atheroprotective. *Nat Med* 2016;**22**:657–665.