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[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], Marion van Vugt[26], Georgios J. Vlachojannis[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], Xuling Chang[36,37], Chew-Kiat Heng[36,37], Rong Jiang[38], William E. Kraus[39,40], Elizabeth R. Hauser[39, 41], Jane F. Ferguson[42], Muredach P. Reilly[43,44], Kaoru Ito[45], Satoshi Koyama[45], Yoichiro Kamatani[46,47,48], Issei Komuro[49], Biobank Japan[50], Lindsey K. Stolze[51], Casey E. Romanoski[51], Mohammad Daud Khan[52,53,54,55], Adam W. Turner[52,53,54,55], Clint L. Miller[52,53,54,55], Redouane Aherrahrou[52,53], Mete Civelek[52,53,], Lijiang Ma[56], Johan L.M. Björkegren[56,57], S. Ram Kumar[58], 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, 33076 Bordeaux, France; [4]Department of Epidemiology, University of Washington, Seattle, WA 98101; [5]Seattle Epidemiologic Research and Information Center, Office of Research and Development, Department of Veterans Affairs, Seattle WA 98108; [6]Kaiser Permanente Washington Health Research Institute, Kaiser Permanente Washington, Seattle WA 98101; [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 Material**; [8]Department of Biological Chemistry and Center for Epigenetics and Metabolism, UC Irvine School of Medicine, Irvine, CA 92697; Departments of [9]Human Genetics, [10]Medicine, and [11]Microbiology, Immunology, & Molecular Genetics, 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 # 507, Atlanta, GA 30322; [18]Institute for Medical Informatics, Statistics and Epidemiology, University of Leipzig, 04107 Leipzig, Germany; [19]LIFE Research Center for Civilization Diseases, University of Leipzig, 04103 Leipzig, Germany; [20]Institute of Laboratory Medicine, Clinical Chemistry and Molecular Diagnostics, University Hospital, 04103 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, 86156 Augsburg, Germany; [23]Clinical Institute of Medical and Chemical Laboratory Diagnostics, Medical University of Graz, 8036 Graz, Austria; [24]Institute of Cardiovascular Science, Faculty of Population Health Sciences, University College London, London WC1E 6HX, UK; [25]Health Data Research UK and Institute of Health Informatics, University College London, London NW1 2DA, UK; [26]Department of Cardiology, Division Heart & Lungs, University Medical Center Utrecht, Utrecht University, 3584 CX Utrecht, the Netherlands; [27]Bart's Heart Centre, St Bartholomew's Hospital, London EC1A 2DA, 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 15850, 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 Paediatrics, Yong Loo Lin School of Medicine, National University of Singapore, Singapore 119228, Singapore; [37]Khoo Teck Puat - National University Children's Medical Institute, National University Health System, Singapore 119074, Singapore; [38]Department of Psychiatry &

Behavioral Sciences, Duke University School of Medicine Durham, NC 27710; [39]Duke Molecular Physiology Institute and Departments of [40]Medicine and [41]Biostatistics & Bioinformatics, Duke University School of Medicine Durham, NC 27710; [42]Department of Medicine, Vanderbilt University Medical Center, Nashville TN 37232; [43]Cardiometabolic Genomics Program, Division of Cardiology, Department of Medicine, Columbia University Irving Medical Center, New York, NY 10032; [44]Irving Institute for Clinical and Translational Research, Columbia University, New York, NY 10032; [45]Laboratory for Cardiovascular Genomics and Informatics, RIKEN Center for Integrative Medical Sciences, Kanagawa 230-0045, Japan; [46]Laboratory for Statistical and Translational Genetics, RIKEN Center for Integrative Medical Sciences, Kanagawa 230-0045, Japan; [47]Human Disease Genomics, Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto 606-8501, Japan; [48]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; [49]Department of Cardiovascular Medicine, Graduate School of Medicine, The University of Tokyo, Tokyo 113-0033, Japan; [50]Institute of Medical Science, The University of Tokyo, Tokyo 108-8639, Japan; [51]Department of Cellular and Molecular Medicine, University of Arizona College of Medicine, Tucson, AZ 85721; [52]Center for Public Health Genomics and Departments of [53]Biomedical Engineering, [54]Biochemistry & Molecular Genetics, and [55]Public Health Sciences, University of Virginia, Charlottesville University of Virginia, Charlottesville, VA 22904; [56]Department of Genetics & Genomic Sciences, Institute of Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York, NY, 10029; [57]Integrated Cardio Metabolic Centre, Department of Medicine, Karolinska Institutet, Karolinska

Universitetssjukhuset, 141 57 Huddinge, Sweden; [58]Department of Surgery, Keck School of

Medicine, University of Southern California, Los Angeles, CA 90033.

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: (323) 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 the CARDIoGRAMplusC4D Consortium (n~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 harboring choline-like transporter 3 (*SLC44A3*) with MI in the context of CAD, but not with coronary atherosclerosis itself, was obtained in Biobank Japan (n~165,000) and 16 independent angiography-based cohorts (n~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 ischemic (vs. non-diseased) coronary arteries, upregulated 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.

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]. By contrast, the same analytical approach initially identified *ABO*, which

1   defines the common ABO blood group system, as being associated with MI among individuals

2   with CAD but not necessarily with the presence of coronary atherosclerosis itself[13]. Thus, even

3   though nearly all loci identified to date for CAD are also associated with MI, it is likely that

4   additional genetic factors predisposing more strongly or specifically to plaque rupture and

5   thrombotic phenotypes exist as well. However, with the exception of *ABO*, no other such locus

6   has been identified. In the present study, we sought to further explore the genetic architecture of

7   MI and address the hypothesis that distinct genetic risk factors may underlie susceptibility to MI

8   and CAD.

9

Methods

Detailed methods are provided in the online Supplemental Materials.

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 1,966 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 CARDIoGRAM+C4D[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 datasets (**Figure 1** and **Supplementary material online,**

Table S1). This analysis revealed 4,419 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 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**).

1   <u>Prioritization of Positional Candidate Genes and Follow-up Analyses with Novel MI Loci:</u> To

2   identify candidate causal genes at the new loci, we first used multi-tissue gene expression data

3   from the GTEx Project[15], the eQTLgen Consortium, or previously published studies available

4   through the Phenoscanner database[16]. For each locus, at least one candidate gene could be

5   prioritized based on the lead SNP yielding a *cis* eQTL in one or more tissues relevant to MI

6   (**Supplementary material online, Table S4**). Candidate causal genes were prioritized further

7   using colocalization analysis with summary statistics from our meta-analysis and eQTL data

8   from the STARNET cohort[17] in blood, atherosclerotic aortic artery, internal mammary artery,

9   visceral and subcutaneous adipose, liver, and skeletal muscle. Based on posterior probabilities of

10 ≥75%, we obtained evidence for *SLC44A3*, *TMEM87B*, and *FHL5* as being causal positional

11 candidate genes on chromosomes 1p21.3, 2q13, and 6q16.1, respectively (**Supplementary**

12 **material online, Table S5**). To explore the biological relevance of the MI loci, we also

13 evaluated the lead variants for association with CAD risk factors in the UK Biobank and other

14 disease phenotypes using the PhenoScanner database[16]. Five loci yielded genome-wide

15 significant associations with blood pressure, lipid levels, BMI, and/or type 2 diabetes in the UK

16 Biobank (**Supplementary material online, Table S6**). The other three loci on chromosomes

17 1p21.3 (*SLC44A3*), 1p36.11, (*AHDC1*) and 4q22.3 (*PDLIM5*) were either not associated with

18 any CAD risk factor or only yielded suggestive associations (**Supplementary material online,**

19 **Table S6**). Based on Phenoscanner, the loci on chromosomes 1p21.3 (*SLC44A3*) and 1p36.11

20 (*AHDC1*) have also not been associated with other disease-related phenotypes whereas the lead

21 variants (or tightly linked proxies) at the remaining MI loci have been suggestively or

22 significantly associated with other complex traits, including inflammatory cytokines, circulating

23 leukocytes, prostate cancer, and migraine (**Supplementary material online, Table S7**).

1   Comparison of Association Signals for MI and CAD Phenotypes at Novel Loci: We next

2   investigated the phenotypic specificity of the association signals for MI and CAD using various

3   analytical strategies. In the first approach, we carried out association analyses with the eight

4   novel loci in the UK Biobank using an all-inclusive definition of CAD (see online Methods for

5   details). This was followed by a meta-analysis of the results with summary statistics for CAD

6   provided by the CARDIoGRAMplusC4D Consortium. Compared to MI, all eight loci yielded

7   some degree of association with CAD in our meta-analysis with the UK Biobank and

8   CARDIoGRAM+C4D, with two loci on chromosomes 1p36.11 and 6q16.1 exhibiting genome-

9   wide significance (**Table 1** and **Supplementary material online, Table S8**). These latter

10   observations suggest that the association signals on chromosomes 1p36.11 and 6q16.1 may not

11   be specific to MI. The associations between the eight novel loci and CAD were also consistent

12   with another recent meta-analysis for CAD using the UK Biobank and CARDIoGRAMplusC4D

13   Consortium[11] (**Supplementary material online, Table S3**).

14   Since CARDIoGRAMplusC4D used an all-inclusive definition of CAD that incorporated

15   MI[6], it was not possible to determine the true specificity of the associations for MI vs. CAD

16   using our meta-analysis results for CAD. Therefore, as a second approach, we used primary

17   level data in the UK Biobank to compare association of the eight novel loci with MI and a

18   restricted CAD only phenotype that excluded subjects with MI. As a positive control locus, we

19   also included *ABO* in these analyses. Consistent with previous studies[13], our lead SNP

20   (rs9411377) at the *ABO* locus in the UK Biobank was strongly associated with MI but not the

21   restricted CAD only phenotype (**Table 2**), thus validating this analytical approach. Seven of the

22   eight novel loci identified for MI were not associated with CAD in the comparative analyses

23   using the UK Biobank (**Table 2**). The only exception was the *AHDC1* locus on chromosome

1   1p36.11, although the effect size and significance level were weaker for CAD than with MI

2   (**Table 2**). We also evaluated association at the eight novel loci in the UK Biobank in analyses

3   comparing cases defined as having both CAD and MI (CAD$^+$/MI$^+$) to controls defined as CAD

4   only subjects (CAD$^+$/MI$^-$). In addition to the expected association with *ABO*, six of the eight loci

5   were nominally associated (P<0.05) with MI among subjects with CAD (**Table 2**). Taken

6   together, these results suggest that the association signals at some of the novel eight loci are

7   either specific to or more robust for MI than with a CAD only phenotype.

8   We next carried out the same analyses in the UK Biobank for 15 previously identified

9   loci that have been suggested to modulate risk of CAD through thrombotic mechanisms[5-11]. At a

10   Bonferroni-corrected significance threshold for testing 15 SNPs (P=0.05/15=3.3x10^{-3}), the lead

11   variants from our MI meta-analysis at seven of these loci were associated with MI, but not the

12   CAD only phenotype, in the UK Biobank (**Supplementary material online, Table S9**). Four of

13   these seven loci were also associated with MI among individuals with CAD (CAD$^+$/MI$^+$ vs.

14   CAD$^+$/MI$^-$) at P<0.05, but none were associated with the CAD only phenotype (**Supplementary**

15   **material online, Table S9**). The remaining eight thrombosis-related loci were associated with

16   both MI and CAD but not with MI in the context of CAD (**Supplementary material online,**

17   **Table S9**). Thus, some, but not all, of the 15 previously identified CAD/MI loci related to

18   thrombosis exhibited association patterns in the UK Biobank that were similar to those observed

19   at the *ABO* locus and several of the novel MI loci (**Table 2**).

20   To determine whether the novel MI loci were associated with other CAD phenotypes and

21   whether the association signals differed by ancestry, we carried out sensitivity analyses in the

22   UK Biobank. As shown in **Supplementary material online, Table S10**, there was no evidence

23   for association with "soft" endpoints, such as angina and death due to CAD, which may have

1   been due to decreased sample size. Although the P-values for MI in subjects of non-European

2   ancestry did not reach significance either, presumably also due to decreased power, the effect

3   sizes were all directionally consistent with those in European ancestry subjects (**Supplementary**

4   **material online, Table S10**) and still contributed to the overall increased significance observed

5   at the MI loci in analyses that included all subjects from the UK Biobank (**Table 2**).

6

7   <u>Replication of Comparative Association Signals for MI and CAD in Biobank Japan:</u> To replicate

8   the association signals at the novel loci in a large non-European ancestry population, we carried

9   out the same comparative analyses for MI vs. CAD only in Biobank Japan (n~165,000). Since

10   the restricted CAD phenotype in Biobank Japan could only be defined based on a diagnosis of

11   stable angina, we first evaluated the lead SNP at *9p21* (rs2891168) as a positive control CAD

12   locus. This analysis yielded the expected strong association with CAD only (OR=1.14, 95% CI

13   1.11-1.17; P=7.3×10^{-21}). Similar to the UK Biobank, the *ABO* locus was also strongly associated

14   with MI in Biobank Japan but not CAD only (**Supplementary material online, Table S11**).

15   Based on these results further validating this comparative strategy and its applicability to

16   Biobank Japan, we tested the novel regions for association with MI vs. CAD only. Since the loci

17   on chromosomes 1p36.11 (*AHDC1*) and 6q16.1 (*FHL5*) yielded genome-wide significant

18   association with CAD in the meta-analysis with the UK Biobank and CARDIoGRAM+C4D

19   (**Table 1**), they were not considered in these analyses. None of the six remaining newly

20   identified loci were associated with the CAD only phenotype whereas three regions (1p21.3,

21   2q32.1, and 15q24.2) yielded nominal (P<0.05) associations with MI in Biobank Japan

22   (**Supplementary material online, Table S11**) that were directionally consistent with the UK

23   Biobank (**Table 2**). However, only the lead SNP (rs12743267) at the chromosome 1p21.3 locus

15

1   harboring *SLC44A3* was also associated with MI among CAD cases (**Supplementary material**

2   **online, Table S11**).

3

4   <u>Preferential Association of the *SLC44A3* Locus with MI in the Presence of Atherosclerosis:</u> We

5   next sought to replicate the association signals for MI at the novel loci using independent cohorts

6   in which the presence of CAD was more directly assessed by angiography. Case-control

7   analyses were carried out in a first set of six cohorts with ~14,000 angiographically-documented

8   CAD patients with MI (CAD$^+$/MI$^+$ cases; n=6,514) and without MI (CAD$^+$/MI$^-$ controls;

9   n=7,411) (**Supplementary material online, Table S12**). A fixed-effects meta-analysis with

10   these six cohorts revealed consistent and strong association of the *SLC44A3* locus on

11   chromosome 1p21.3 with risk of MI among individuals with CAD (OR=1.16, 95% CI 1.09-1.23;

12   P=3.3x10^{-6}) (**Table 3**), with no significant evidence for heterogeneity (P-het=0.10)

13   (**Supplementary material online, Table S12**). Exclusion of the Emory cohort, which itself

14   exhibited a very strong effect size with large variation, did not appreciably change the direction

15   or significance level of the overall association between the *SLC44A3* locus and MI (OR=1.15,

16   95% CI 1.08-1.22; P=6.2x10^{-6}) (**Supplementary material online, Table S12**).

17   As another replication study, we evaluated association of the newly identified MI loci in

18   10 additional angiography-based cohorts comprised of 7,412 CAD$^+$/MI$^+$ cases and 5,542

19   CAD$^+$/MI$^-$ controls (**Supplementary material online, Table S13**). These analyses also yielded

20   evidence for association of the *SLC44A3* locus with MI in the context of CAD (OR=1.09, 95%

21   CI 1.03-1.16; P=2.1x10^{-3}) but not the remaining five loci. When all 16 angiography-based

22   cohorts were meta-analyzed together (n~27,000), association of the *SLC44A3* locus with MI in

23   the presence of coronary atherosclerosis increased in significance by several fold (OR=1.12, 95%

1   CI 1.08-1.17; P=5.6x10^{-8}) (**Table 3**). Notably, the *SLC44A3* locus was highly significantly

2   associated with MI in an all-inclusive meta-analysis with UK Biobank, Biobank Japan, and the

3   16 angiography-based cohorts (n=41,336 CAD^{+}/MI^{+} cases and 40,363 CAD^{+}/MI^{-} controls) and

4   exceeded the threshold for genome-wide significance (OR=1.07, 95% CI 1.05-1.10; P=5.4x10^{-}

5   ^{11}). Taken together with the weak associations observed with CAD in the meta-analyses with

6   CARDIoGRAMplusC4D and UK Biobank and the comparative analyses in the UK Biobank and

7   Biobank Japan, these results provide compelling evidence for the *SLC44A3* locus being

8   preferentially associated with plaque instability and/or rupture in the presence of coronary

9   atherosclerosis but not atherosclerotic CAD itself.

10

11   Association of the *SLC44A3* Locus with Other Thrombotic Phenotypes: We next explored

12   whether the *SLC44A3* locus was associated with other thrombotic and coagulation phenotypes

13   related to MI. Based on data from the MEGASTROKE Consortium[18], there was no evidence for

14   association of rs12743267 with most forms of stroke except for nominal associations with

15   cardioembolic and small vessel stroke in subjects of European ancestry that would not be

16   considered significant at a Bonferroni corrected P-value of 0.01 for testing five forms of stroke

17   (0.05/5=0.01) (**Supplementary material online, Table S14**). Second, variants at the

18   chromosome 1p21.3 locus had been previously associated with circulating levels of D-dimer[19],

19   which is produced when crosslinked fibrin is degraded by plasmin and the most widely used

20   clinical marker of activated blood coagulation[20]. However, rs12743267 was not associated with

21   D-dimer levels (beta=-0.011; SE=0.007; P=0.12) based on a GWAS carried out by the CHARGE

22   Consortium[19] and the lead SNP for D-dimer (rs12029080) was not associated with MI in our

23   meta-analysis with the UK Biobank and CARDIoGRAMplusC4D Consortium (OR=0.99, 95%

1   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,

2   *SLC44A2*, a member of the solute carrier family of membrane transporters that includes

3   *SLC44A3*, has been associated with venous thromboembolism (VTE)[21], another coagulation and

4   thrombotic phenotype relevant to MI. However, there was no association of rs12743267 with

5   VTE (OR=0.97, 95% CI 0.92-1.02; P=0.23) in a GWAS carried out by the INVENT

6   Consortium[21]. By comparison, the lead VTE SNP in *SLC44A2* (rs2288904) was associated with

7   CAD (OR=1.04, 95% CI 1.03-1.05; P=7.0×10^{-8}) and MI (OR=1.04, 95% CI 1.02-1.06;

8   P=1.5×10^{-5}) in our meta-analyses, as well as with CAD in Biobank Japan (OR=1.03, 95% CI

9   1.01-1.06; P=1.1×10^{-3}).

10

11   Association of the *SLC44A3* Locus with Choline-related Metabolites: While the function of

12   *SLC44A3* as a solute carrier is not entirely known, it has been reported to encode a putative

13   choline-like transporter[22]. In humans, elevated plasma levels of choline and products of its

14   metabolism have been linked to risk of MI-related outcomes[23-25]. However, we did not obtain

15   evidence in the Genebank cohort for association of the *SLC44A3* locus with plasma levels of

16   these metabolites or a panel of choline-related small molecule amines that have also been

17   associated with CAD and MI[26-31] (**Supplementary material online, Table S15**). Based on data

18   from three metabolomics and proteomics studies[32-34], the *SLC44A3* locus did yield associations

19   with small molecules in plasma or urine, but these would not be considered significant at

20   Bonferroni-corrected thresholds for the number of analytes tested in each dataset

21   (**Supplementary material online, Table S16**).

22

1   <u>Functional Analysis of *SLC44A3*</u>: We next used functional studies to evaluate *SLC44A3* as a

2   candidate causal gene at the chromosome 1p21.3 locus. Among 600 CAD patients in the

3   STARNET study [17], *SLC44A3* was expressed at relatively high levels in several MI-relevant

4   tissues, such as atherosclerotic aortic root, adipose tissue, mammary artery, and liver (**Figure**

5   **4A**). In addition, the lead SNP on chromosome 1p21.3 yielded *cis* eQTLs for *SLC44A3* in

6   atherosclerotic aorta and mammary artery, where the MI risk allele (C) was associated with

7   increased expression (**Figure 4B**). In the GTEx Project, similar eQTLs were observed in aorta

8   and coronary artery (**Figure 4C**), as well in whole blood and various components of the

9   gastrointestinal tract (**Supplementary material online, Table S4**). These findings were

10   consistent with mRNA levels of *SLC44A3* being significantly higher in ischemic coronary

11   arteries compared to non-diseased coronary arteries in another independent dataset (**Figure 4D**).

12   To explore the vascular cell type in which *SLC44A3* could mediate its biological effects on MI,

13   we used RNAseq and functional data from two additional independent datasets of human aortic

14   endothelial cells (HAECs) and smooth muscle cells (SMCs), respectively. Compared to vehicle

15   control, *SLC44A3* expression was significantly upregulated in HAECs treated with the pro-

16   atherogenic inflammatory cytokine IL-1β (**Figure 4E**). *SLC44A3* expression in SMCs was also

17   modestly, but significantly, inversely correlated with migration towards platelet-derived growth

18   factor (PDGF)-BB *in vitro* (**Figure 4F**). Taken together, these data provide supportive

19   functional evidence that *SLC44A3* is at least one candidate causal at the novel MI locus on

20   chromosome 1p21.3 locus and suggest that this putative solute carrier could promote increased

21   risk of plaque rupture and thrombosis through mechanisms at the level of the artery wall.

22

23

Discussion

1   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 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 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.

16 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,

20

1   the magnitude of the effect size of the *SLC44A3* locus on MI in the context of coronary

2   atherosclerosis (OR=1.12) was stronger than the ORs obtained in the GWAS meta-analysis, UK

3   Biobank, or Biobank Japan (OR~1.05), and equivalent to some of the most significantly

4   associated loci identified to date for CAD[11]. Taken together, these results support the notion that

5   the biological mechanism(s) underlying the association of the *SLC44A3* locus may be related to

6   plaque rupture rather than plaque progression per se. In this regard, *ABO* was similarly

7   identified as being only associated with MI in the original study by Reilly *et al*.[13], which we

8   replicated in our analogous comparative analyses with the UK Biobank and Biobank Japan.

9   Thus, to our knowledge, the *SLC44A3* locus represents the second and only other genetic risk

10   factor that is specifically associated with MI but not with CAD. We also did not obtain evidence

11   for association of the *SLC44A3* locus with other thrombotic phenotypes, such as stroke or VTE.

12   This observation is not entirely surprising since the genetic determinants of CAD and stroke,

13   while shared, do not completely overlap[35]. However, it should be noted that our meta-analyses

14   for MI had approximately 10-fold higher numbers of subjects than the VTE GWAS[21]. Thus, it is

15   possible that power was insufficient in the INVENT Consortium to detect an association of the

16   *SLC44A3* locus with VTE.

17   The lead SNP on chromosome 1p21.3 (rs12743267) is located ~36kb upstream of the

18   transcriptional start site for *SLC44A3* and ~250kb away from the gene encoding tissue factor or

19   coagulation factor III (*F3*). Given the known role of tissue factor in the blood coagulation

20   cascade and the association of variants around its gene with circulating D-dimer levels[19], *F3*

21   would be considered a more biologically plausible candidate gene for a thrombosis-related

22   phenotype such as MI. However, we did not obtain any evidence that would prioritize *F3* as a

23   candidate causal gene since our lead SNP was not associated with D-dimer levels and the lead

21

1 SNP for D-dimer (rs12029080) showed no evidence for association with MI. Furthermore, *cis*

2 eQTLs for *F3* were not observed with our lead SNP or proxy variants in any available tissue in

3 STARNET or the GTEx Project. Given these observations and the presence of *cis* eQTLs for

4 *SLC44A3* in multiple tissues and independent datasets, we focused on *SLC44A3* as a candidate

5 causal gene for MI. *SLC44A3* is one of five members of the SLC44 family of solute carriers

6 (*SLC44A1-5*) that have been proposed to function as choline transporters[22]. However, *SLC44A1*

7 is the only member of this transporter family for which a role in transporting choline across both

8 the plasma and mitochondrial membranes has been demonstrated by direct experimentation[36, 37].

9 In addition, the *SLC44A3* locus was not associated with plasma levels of choline, proatherogenic

10 choline-derived small molecule amines, such as trimethylamine *N*-oxide and betaine[24, 25], or with

11 a large panel of metabolomic and proteomic targets in plasma and urine[32-34]. Thus, additional

12 functional studies will be needed to demonstrate whether *SLC44A3* encodes a transporter for

13 choline or other molecules and whether such activity would modulate levels of metabolites that

14 influence risk of MI.

15   Several lines of evidence from our functional and bioinformatics analyses further pointed

16 to *SLC44A3* as one causal positional candidate on chromosome 1p21.3 and suggested that

17 putative biological mechanisms through which this gene could influence plaque rupture and/or

18 thrombosis may be through direct effects at the level of the vessel wall. First, *SLC44A3* was

19 expressed in MI-relevant vascular tissues, such as the aorta and mammary artery. Second,

20 colocalization analyses carried out in atherosclerotic aorta yielded a strong posterior probability

21 for *SLC44A3*, but not the other genes at the chromosome 1p21.3 locus (i.e. *F3*), as being causal

22 for MI. Third, carriers of the MI risk allele had significantly higher *SLC44A3* mRNA levels than

23 non-carriers, with a stronger effect size observed in atherosclerotic aortic root than mammary

1   artery. The same *cis* eQTLs for *SLC44A3* were independently observed in aorta and coronary

2   artery in the GTEx Project. Fourth, expression analyses in two independent heart donor datasets

3   demonstrated upregulation of *SLC44A3* in ischemic coronary arteries by ~50% compared to

4   normal arteries and by ~3-fold in HAECs incubated with the pro-atherogenic cytokine IL-1β.

5   This latter observation suggests that *SLC44A3* might be involved in the response of HAECs to

6   inflammatory stimuli that increase expression and secretion of various pro-atherogenic genes,

7   such as adhesion molecules and chemokines[38]. Lastly, although we did not detect an eQTL for

8   *SLC44A3* in SMCs (or HAECs), possibly due to insufficient power, an *in vitro* assay

9   demonstrated that *SLC44A3* expression was inversely correlated with SMC migration. In this

10   regard, previous studies have shown that SMC proliferation and migration can promote secretion

11   of extra cellular matrix proteins and the formation of a protective fibrous cap that renders a

12   lesion less prone to rupture[39]. Taken together, these functional data and the results of our genetic

13   analyses collectively implicate *SLC44A3* as at least one candidate causal gene on chromosome

14   1p21.3 and suggest that its expression is positively associated with MI-promoting characteristics

15   of various vascular cell types. However, in STARNET, *SLC44A3* mRNA levels in adipose and

16   liver were equivalent to those observed in aorta, and based on data from the GTEx Project,

17   expression was also high in kidney, pancreas, the small intestine, and colon. Moreover, the

18   eQTLs in GTEx for *SLC44A3* in aorta and coronary artery were modest relative to those

19   observed in whole blood, heart, pancreas, liver, and colon. In some of these tissues, such as

20   liver, the allelic association of rs12743267 with *SLC44A3* mRNA levels was also opposite to that

21   observed in arterial tissues. Although these observations suggest that *SLC44A3* could influence

22   risk of MI through mechanisms related to metabolism, the *SLC44A3* locus was not associated

23   with traditional CAD risk factors, such as lipid levels and type 2 diabetes. Nonetheless, we still

1   cannot rule out the possibility that *SLC44A3* could also increase risk of plaque rupture via a role

2   in other MI-relevant tissues.

3   While our results point to novel and distinct genetic determinants of MI, certain

4   limitations of our study should still be taken into consideration. First, the majority of subjects in

5   our analyses were of European ancestry and it is possible that some of the genetic associations

6   may not be generalizable to other populations. However, the *SLC44A3* locus yielded an

7   equivalent association with MI in Biobank Japan and exhibited directionally consistent effect

8   sizes in other Asian populations, suggesting that at least a subset of the association signals

9   identified herein may also be relevant in other ethnicities as well. Second, it is possible, albeit

10   unlikely, that some subjects in the UK Biobank and CARDIoGRAMplusC4D Consortium

11   overlapped, which could have been a confounding factor in the meta-analysis. However, a

12   recent analysis concluded that duplicate samples between CARDIoGRAMplusC4D and the UK

13   Biobank were minimal (<0.1%) and would not significantly influence test statistics[11]. Third, we

14   did not exclude subjects with a positive family history of CAD from the control group in the UK

15   Biobank as was done in another recent GWAS meta-analysis for CAD[11]. There could also have

16   been misclassification in our analyses since, for example, MI and CAD may not have been

17   defined in exactly the same in CARDIoGRAMplusC4D, the UK Biobank, and Biobank Japan.

18   We note that if such misclassifications had occurred, they would have most likely been non-

19   differential and biased the results towards the null. Finally, even though SNPs with MAFs as

20   low as 0.5% were included in our analyses, our study was primarily focused on discovery of

21   main effects with common susceptibility alleles. However, rare variants or GxE interactions still

22   likely play important roles in modulating risk of MI, which, along with vascular cell-specific

23   eQTL analyses, will require additional investigation.

1   In summary, our results identify several previously unrecognized loci for MI and provide

2   new avenues for exploring the pathophysiology of vulnerable atherosclerotic lesions. Most

3   importantly, our data support the concept that some of the heritable determinants of plaque

4   rupture and thrombus formation are distinct from those that contribute to development of

5   coronary atherosclerosis, with *SLC44A3* emerging as one such potential genetic susceptibility

6   factor. Future studies will be needed to explore the clinical relevance of these findings for

7   patients at risk of MI.

8

9

1   **URLs.** The UK Biobank, (https://www.ukbiobank.ac.uk/); CARDIoGRAMplusC4D,

2   http://www.cardiogramplusc4d.org/; Biobank Japan, https://biobankjp.org/english/index.html;

3   GWAMA, https://www.geenivaramu.ee/en/tools/gwama/; Genotype-Tissue Expression Project,

4   http://gtexportal.org/; Phenoscanner, http://www.phenoscanner.medschl.cam.ac.uk/phenoscanner;

5   R statistical software, http://www.R-project.org/.

6

Data Availability Statement

8   Full summary statistics relating to the GWAS analysis in the UK Biobank and the meta-analysis

9   with CARDIoGRAMplusC4D will be deposited with The NHGRI-EBI Catalog of published

10 genome-wide association studies (https://www.ebi.ac.uk/gwas/docs/about). All other relevant

11 data are available upon request from the authors.

12

Funding

11

Acknowledgments

19

Conflict of Interest

21 Dr. Hazen (S.L.H.) is named as co-inventor on pending and issued patents held by the Cleveland

22 Clinic relating to cardiovascular diagnostics and therapeutics and have the right to receive

23 royalty payment for inventions or discoveries related to cardiovascular diagnostics or

Author Contributions

Concept and design: J.H., Y.H., and H.A. Acquisition, analysis, or interpretation of data: J.H.,

Y.H., Q.J., P.H., J.R.H., 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., M.V.V., G.J.V.,

R.S.P., L-P.L., M.K., T.L., T.V.M.N., P.K., J.O.L., X.C., C-K.H., R.J., W.E.K., E.R.H, J.F.F.,

M.P.R., K.I., S.K., Y.K., I.K., L.K.S., C.E.R., M.D.K., A.W.T., C.L.M., R.A., M.C., L.M.,

1   J.L.M.B., S.R.K., W.H.W.T., S.L.H., and H.A. Drafting of the manuscript: J.H., Y.H., and H.A.

2   Critical revision of the manuscript for important intellectual content: All authors.

3

4

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, Stroke Statistics S. Heart disease and stroke statistics--2014 update: a report from the American Heart Association. *Circulation* 2014;**129**(3):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, Group JS. Rosuvastatin to prevent vascular events in men and women with elevated C-reactive protein. *N Engl J Med* 2008;**359**(21):2195-207.
3. Hansson GK. Inflammation, atherosclerosis, and coronary artery disease. *N Engl J Med* 2005;**352**(16):1685-95.
4. Schunkert H, Erdmann J, Samani NJ. Genetics of myocardial infarction: a progress report. *Eur Heart J* 2010;**31**(8):918-25.
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, Braenne I, Nothen MM, Hofmann P, Braund PS, Mergia E, Reinhard W, Burgdorf C, Schreiber S, Balmforth AJ, Hall AS, Bertram L, Steinhagen-Thiessen E, Li SC, Marz 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**(7480):432-6.
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, Consortium

CD. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet* 2015;**47**(10):1121-30.

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**(1):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, 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, Saleheen D. Fifteen new risk loci for coronary artery disease highlight arterial-wall-specific mechanisms. *Nat Genet* 2017;**49**(7):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, Consortium CD, Kathiresan S. Genetic analysis in UK Biobank links insulin resistance and transendothelial migration pathways to coronary artery disease. *Nat Genet* 2017;**49**(9):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, Consortium E-C, CardioGramplusC4D, group UKBCCCw, 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**(9):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**(3):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**(11):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**(9763):383-92.

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**(12):e29070.

15. Consortium G. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* 2013;**45**(6):580-5.

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**(20):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**(6301):827-30.

18. Malik R, Chauhan G, Traylor M, Sargurupremraj M, Okada Y, Mishra A, Rutten-Jacobs L, Giese AK, 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, Jr., Butterworth AS, Carrera C, Carty CL, Chasman DI, Chen WM, 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 FC, Hyacinth HI, Ikram MA, Ingelsson E, Irvin MR, Jian X, Jimenez-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 JM, Lemmens R, Leys D, Lewis CM, Lin WY, 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, Rannikmae 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, Consortium AF, Cohorts for H, Aging Research in Genomic Epidemiology C, International Genomics of Blood Pressure C, Consortium I, Starnet, Bis JC, Pastinen T, Ruusalepp A, Schadt EE, Koplev S, Bjorkegren JLM, Codoni V, Civelek M, Smith NL, Tregouet 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, BioBank Japan Cooperative Hospital G, Consortium C, Consortium E-C, Consortium EP-I, International Stroke Genetics C, Consortium M, Neurology Working Group of the CC, Network NSG, Study UKYLD, Consortium M, Sanghera DK, Melander O, Jern C, Strbian D, Fernandez-Cadenas I, Longstreth WT, Jr., 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**(4):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**(17):1864-72.

20. Adam SS, Key NS, Greenberg CS. D-dimer antigen: current concepts and future prospects. *Blood* 2009;**113**(13):2878-87.

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**(4):532-42.

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**(2-3):646-54.

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**(7341):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**(17):1575-84.

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**(14):904-10.

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**(5):576-85.

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**(5):799-812.

28. Hartiala J, Bennett BJ, Tang WH, Wang Z, Stewart AF, Roberts R, McPherson R, Lusis AJ, Hazen SL, Allayee H, Consortium C. 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**(6):1307-13.

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**(6).

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**(10):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, Multiple Tissue Human Expression Resource C, 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**(6):543-50.

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**(7708):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**(2):167-176.

35. 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, Consortium M, Consortium CA, Consortium CD, International Stroke Genetics C. Shared genetic susceptibility to ischemic stroke and coronary artery disease: a genome-wide analysis of common variants. *Stroke* 2014;**45**(1):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. *Proceedings of the National Academy of Sciences of the United States of America* 2000;**97**(4):1835-1840.

37. Michel V, Bakovic M. The solute carrier 44A1 is a mitochondrial protein and mediates choline transport. *FASEB J* 2009;**23**(8):2749-58.

38. Gimbrone MA, Jr., Garcia-Cardena G. Endothelial cell dysfunction and the pathobiology of atherosclerosis. *Circ Res* 2016;**118**(4):620-36.

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**(6):657-65.

1   **Figure Legends**

2
3   **Figure 1. Overview of genetic and functional analyses.** A GWAS was first carried out for MI

4   using primary level data in the UK Biobank with ~11 million SNPs. These results were then

5   combined with summary GWAS data from the CARDIoGRAMplusC4D Consortium in a fixed-

6   effects meta-analysis that included a total of ~61,000 MI cases and ~577,000 controls, and

7   8,126,035 SNPs common to both datasets. The meta-analysis identified eight novel loci for MI,

8   6 of which exhibited stronger association signals for MI compared to CAD. Follow-up analyses

9   and independent replication in Biobank Japan and 16 angiography-based cohorts, encompassing

10   a total of ~831,000 subjects, provided confirmatory evidence for association of the chromosome

11   1p21.3 locus with MI. Bioinformatics and eQTL analyses prioritized *SLC44A3* as one positional

12   candidate on chromosome 1p21.3 for functional evaluation.

13

14   **Figure 2. Manhattan plot of results from GWAS meta-analysis for MI. (A)** Eight novel

15   loci on chromosomes 1p36.11, 1p21.3, 2q13, 2q32.1, 4q22.3, 6q16.1, 9q34.3, and 15q24.2

16   (orange dots) were significantly associated with MI. Genome-wide thresholds for significant

17   (P=5.0×10^{-8}) and suggestive (P=5.0×10^{-6}) association are indicated by the horizontal red and blue

18   lines, respectively. P-values are truncated at $-\log_{10}(P)=40$.

19

20   **Figure 3. Regional plots of eight novel loci for MI.** The chromosome band and nearest gene

21   (in parentheses) is indicated for each locus. Each region is centered on the lead SNP (purple

22   diamond) and the genes in the interval are indicated in the bottom panel. The degree of linkage

23   disequilibrium (LD) between the lead SNP and other variants is shown as r^2 values according to

24   the color-coded legend in the box.

36

1

2 **Figure 4. Functional Analyses of *SLC44A3* in MI-relevant Tissues.** (**A**) In the STARNET

3 cohort, *SLC44A3* was expressed at relatively high levels in tissues relevant to MI, including

4 atherosclerotic aortic root (aorta), visceral adipose, mammary artery, and liver. (**B**) The lead

5 SNP at the chromosome 1p21.3 locus yielded significant *cis* eQTLs for *SLC44A3* in

6 atherosclerotic aortic root and normal mammary artery among subjects from the STARNET

7 cohort, where the MI risk allele (C) was associated with significantly higher mRNA levels. (**C**)

8 A similar pattern of *cis* eQTLs was also independently observed with the *SLC44A3* locus in aorta

9 and coronary artery based on data from the GTEx Project. (**D**) In another independent human

10 dataset, *SLC44A3* expression was increased in ischemic coronary arteries (n=36) from heart

11 donors with CAD compared to normal coronary arteries from non-diseased donors (n=24). (**E**)

12 Incubation of human aortic endothelial cells (HAECs) isolated from a different and independent

13 set of anonymous heart donors (n=53) with IL-1β for 4 hours upregulated *SLC44A3* expression

14 ~3-fold compared to paired vehicle-treated HAECs. (**F**) Using a fourth independent human

15 dataset (n=151), *SLC44A3* expression was also observed in smooth muscle cells (SMCs) and

16 inversely correlated with migration rate towards platelet-derived growth factor (PDGF)-BB *in*

17 *vitro*.

18

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

SNP	Chr	Pos	Nearest Gene(s)	EA/OA	EAF	MI OR (95% CI)	MI P	CAD OR (95% CI)	CAD P
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; Pos, base-pair position (hg19), EA, effect allele; OA, other allele; EAF, effect allele frequency; OR, odds ratio; CI, confidence interval; P, P-value obtained from meta-analysis of the UK Biobank and CARDIoGRAM+C4D.

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	OR (95% CI)	P	OR (95% CI)	P
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; Pos, base-pair position (hg19), EA, effect allele; OA, other allele; EAF, effect allele frequency; OR, odds ratio; CI, confidence interval; P, P-value obtained from linear mixed model analysis in UK Biobank.

Table 3. Association of Novel Loci with MI in the Presence of CAD in Angiography-based Cohorts.

						Angiography Cohorts I (6,514/7,411)		Angiography Cohorts II (7,412/5,542)		Meta-Analysis (13,926/12,953)	
SNP	Chr	Pos	Nearest Gene(s)	EA/OA	EAF	OR (95% CI)	[a]P	OR (95% CI)	[b]P	OR (95% CI)	P
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; Pos, base-pair position (hg19), EA, effect allele; OA, other allele; EAF, effect allele frequency in European ancestry subjects; OR, odds ratio; CI, confidence interval.

[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.
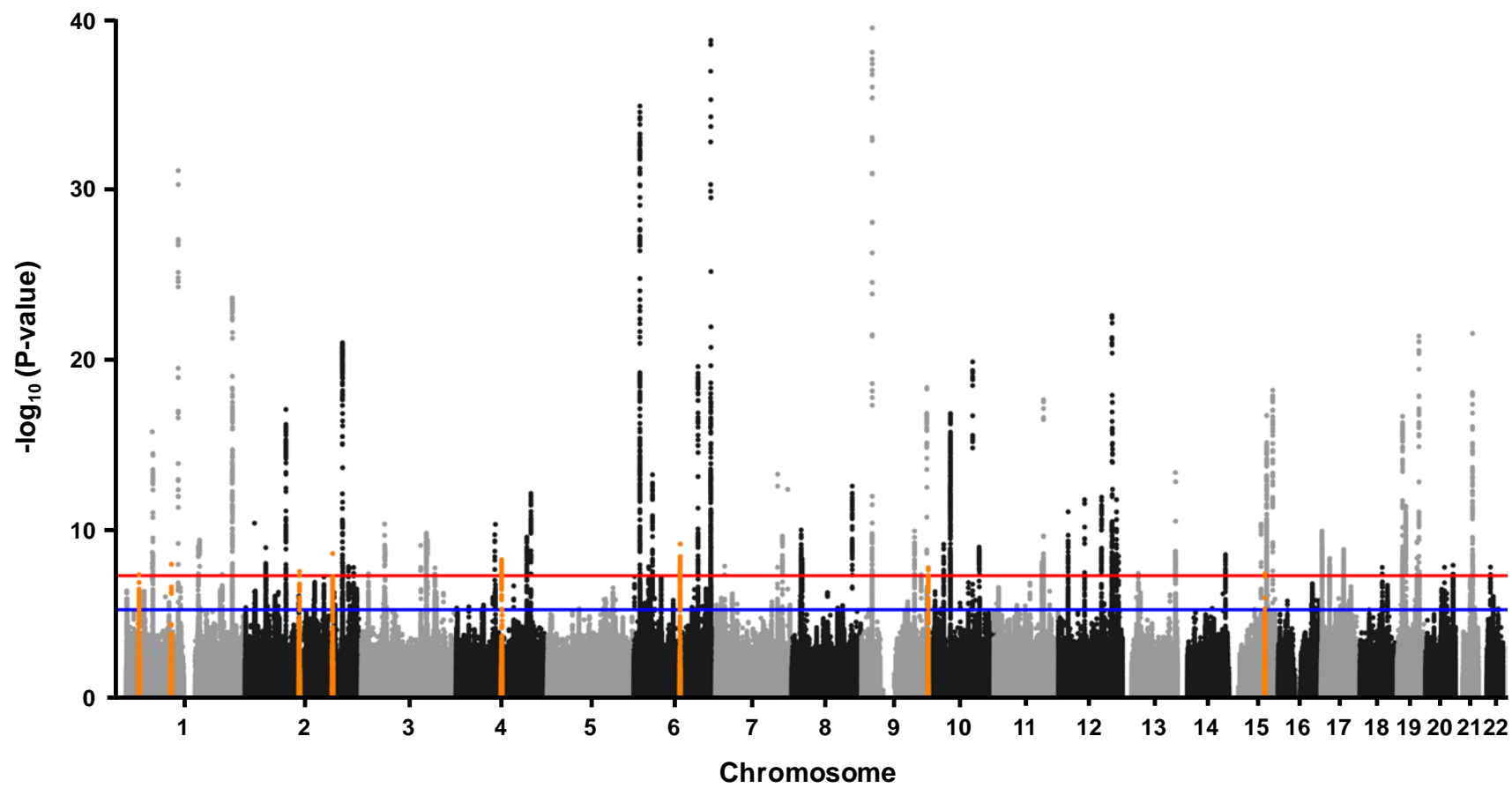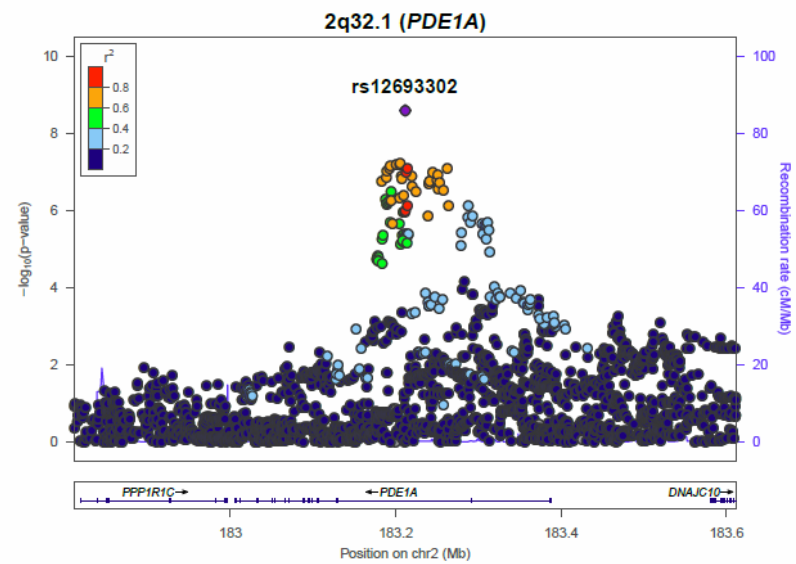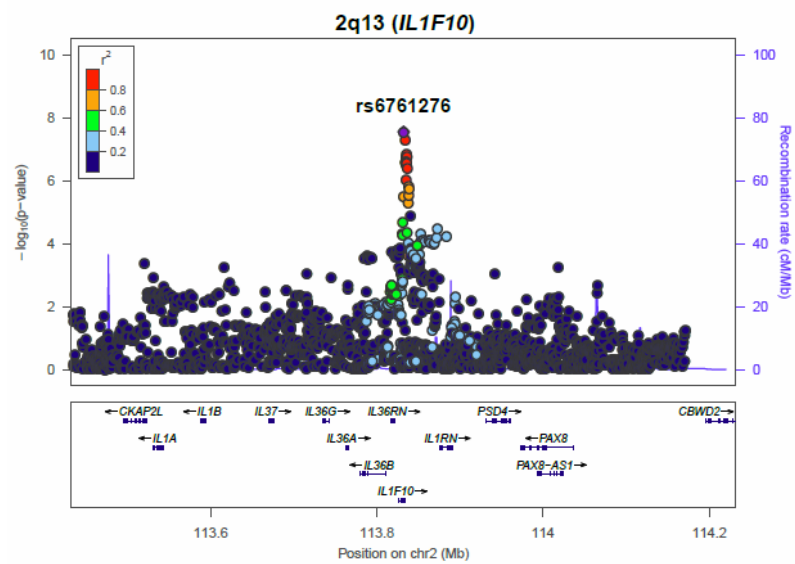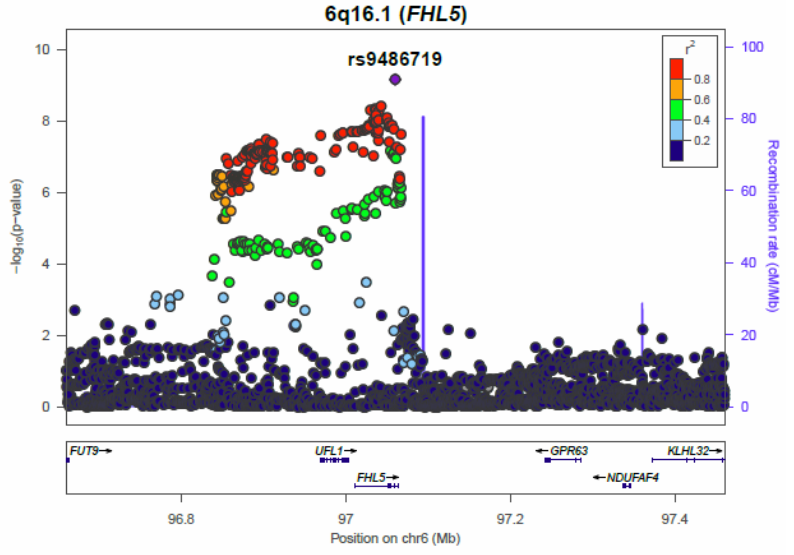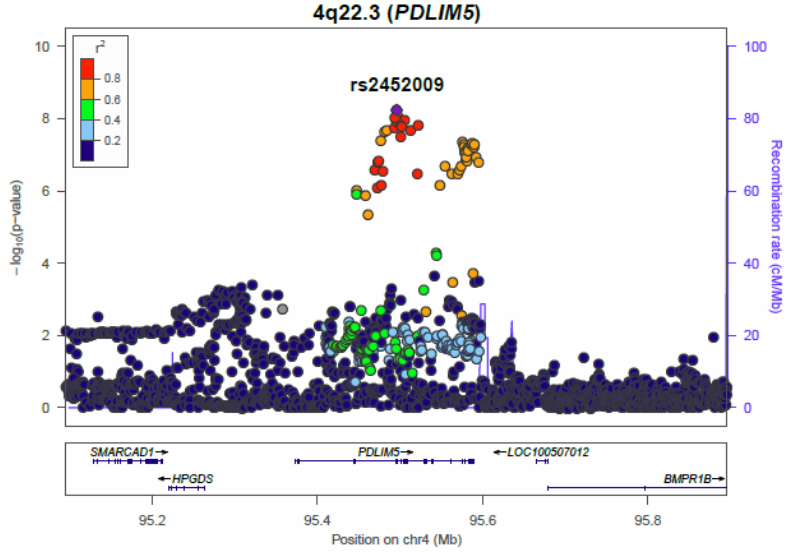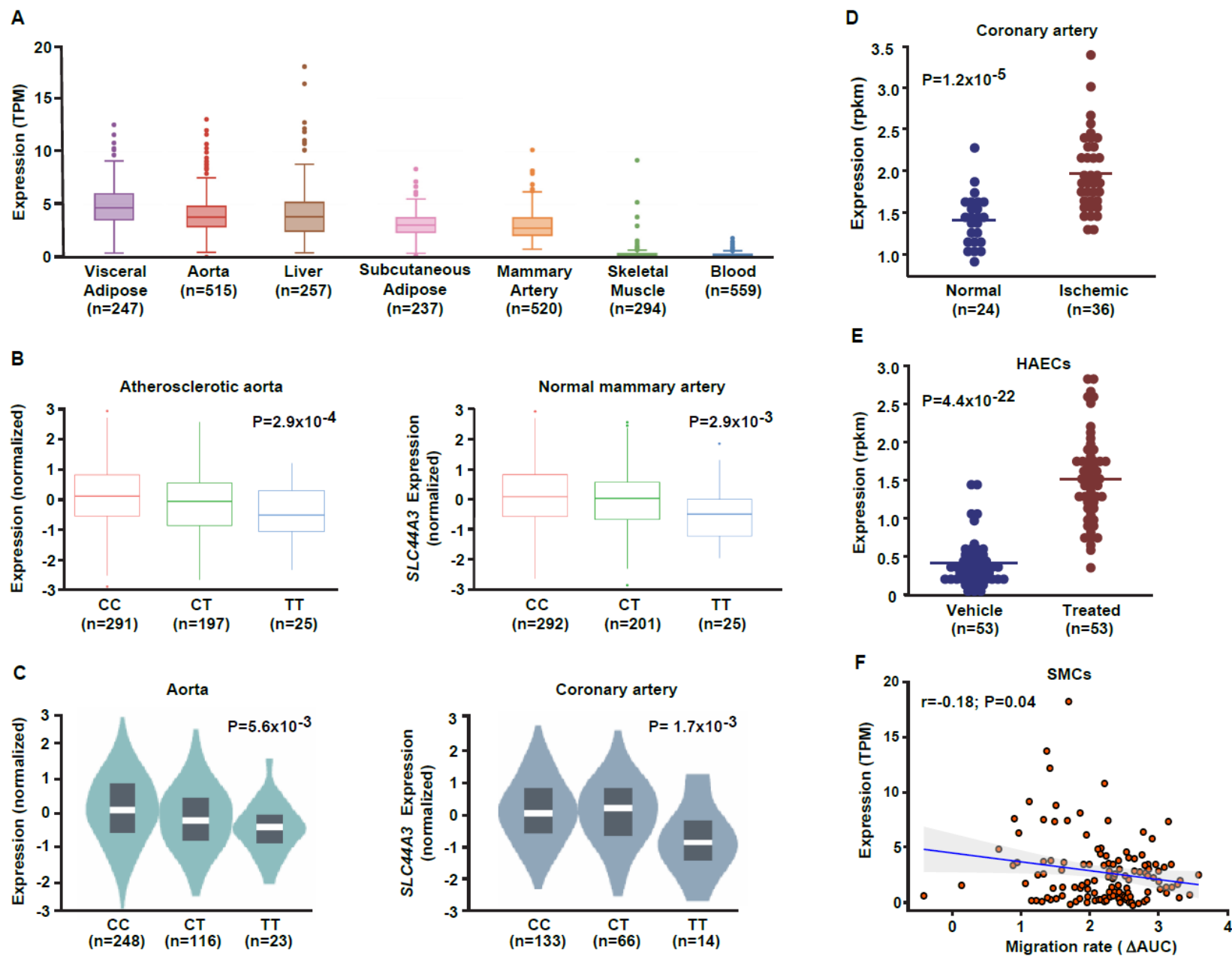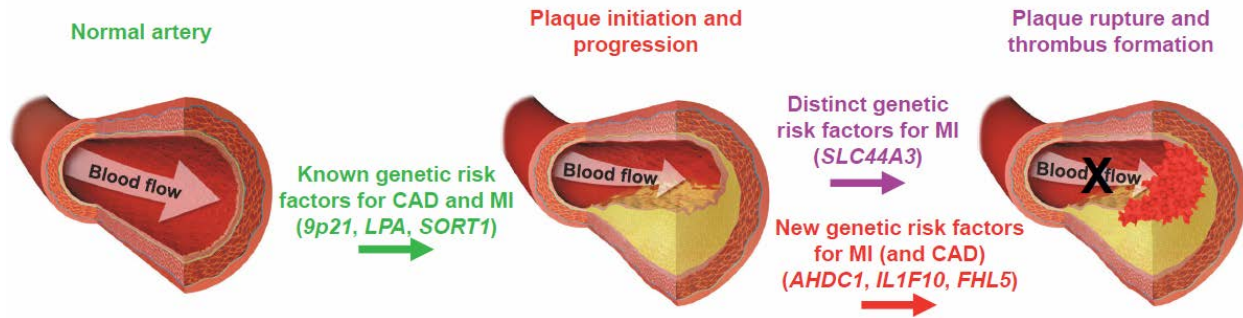
Figure 1
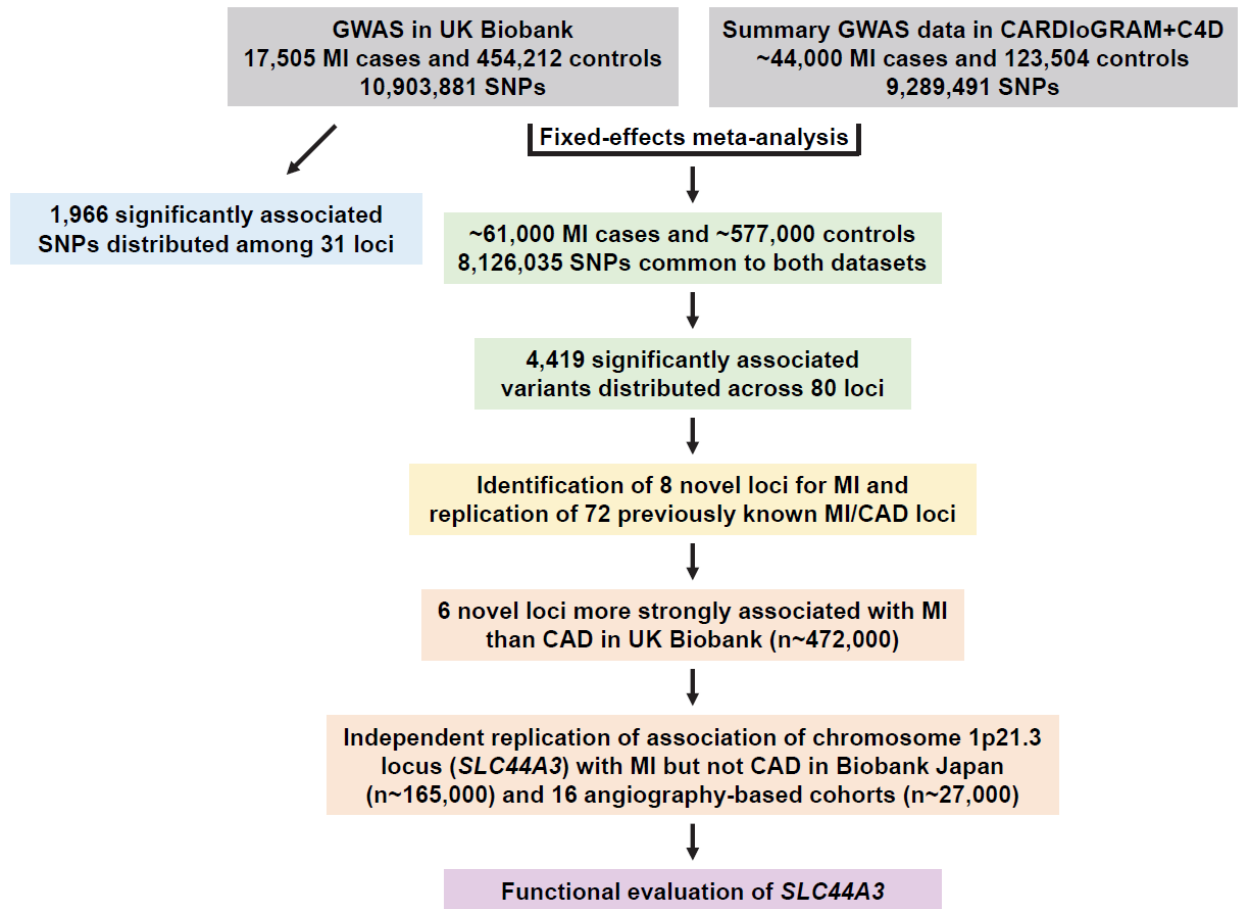
Figure 2

Figure 3
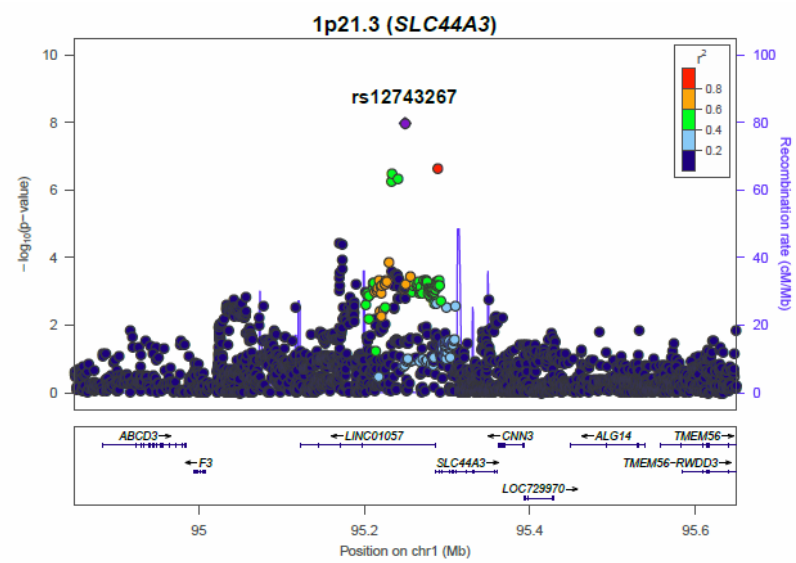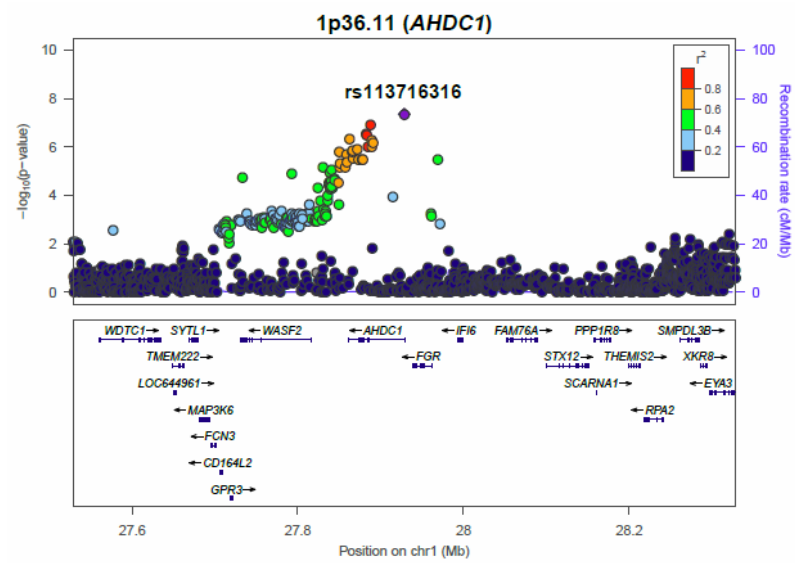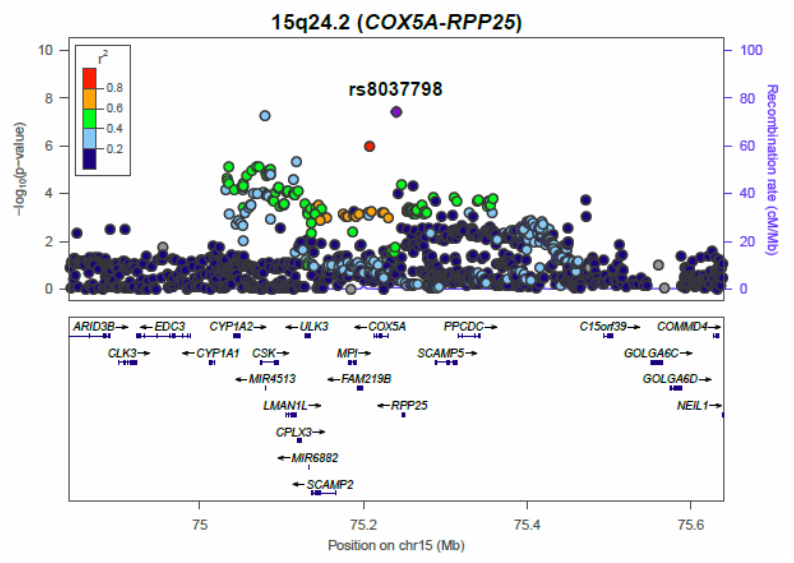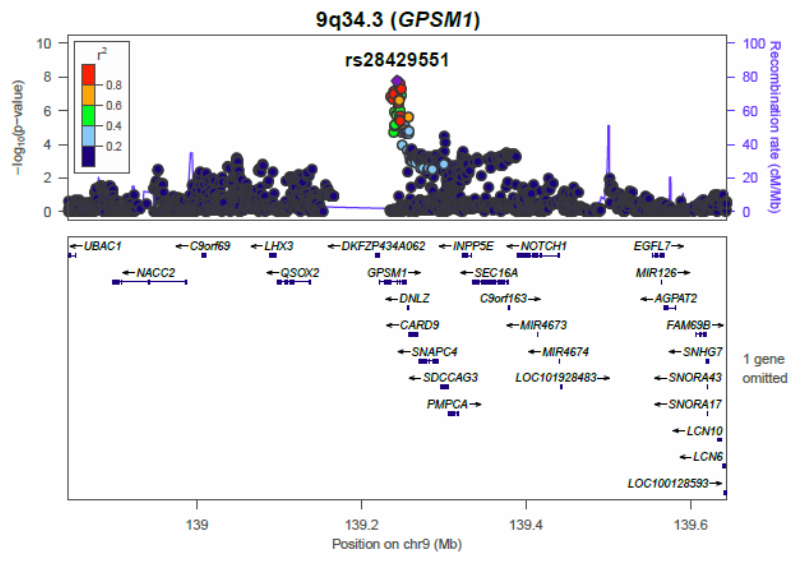
Figure 4

Graphical Abstract. Large-scale analyses in ~831,000 subjects identified eight novel susceptibility factors for myocardial infarction (MI) and demonstrated that some of genetic determinants of plaque rupture and thrombus formation, such as *SLC44A3*, are distinct from those that contribute to development of coronary atherosclerosis. These findings may provide new avenues for exploring the pathophysiology of vulnerable lesions and development of MI.