Systematic evaluation of pleiotropy identifies six further loci associated with coronary artery disease

Myocardial Infarction Genetics and CARDIoGRAM Exome Consortia Investigators, Thomas R. Webb, PhD[1,2], Jeanette Erdmann, PhD[3,4,5], Kathleen E. Stirrups, PhD[6,7], Nathan O. Stitziel, MD, PhD[8,9,10], Nicholas G.D. Masca, PhD[1,2], Henning Jansen, MD[11,12], Stavroula Kanoni, PhD[6], Christopher P. Nelson, PhD[1,2], Paola G. Ferrario, PhD[4,13], Inke R. König, PhD[4,13], John D. Eicher, PhD[13], Andrew D. Johnson, PhD[13], Stephen E. Hamby[1,2], Christer Betsholtz, PhD[15,16], Arno Ruusalepp, MD, PhD[17,18,19], Oscar Franzén, PhD[19,20], Eric E. Schadt, MD, PhD[20], Johan L. M. Björkegren, MD, PhD[16,17,19,20,] , Peter E. Weeke, MD, PhD[21,22], Paul L. Auer, PhD[23], Ursula M. Schick, PhD[24,25], Yingchang Lu, MD, PhD[25,26], He Zhang, PhD[27], Marie-Pierre Dube, PhD[28,29], Anuj Goel, MSc[30,31], Martin Farrall, FRCPath[30,31], Gina M. Peloso, PhD[32,33,34,35], Hong-Hee Won, PhD[32,33,34,35,36], Ron Do, PhD[26,37,38,39], Erik van Iperen, MSc[40], Jochen Kruppa, PhD[41], Anubha Mahajan, PhD[31], Robert A. Scott, PhD[42], Christina Willenborg, PhD[3], Peter S. Braund, PhD[1,2], Julian C. van Capelleveen, MD[43], Alex S.F. Doney, MD, PhD[44], Louise A. Donnelly, PhD[44], Rosanna Asselta, PhD[45,46], Pier A. Merlini, MD[47], Stefano Duga, PhD[45,46], Nicola Marziliano, PhD[49], Josh C. Denny, MD, MS[21,49], Christian Shaffer, BS[21], Nour Eddine El-Mokhtari, MD[50], Andre Franke, PhD[51], Stefanie Heilmann, PhD[52,53], Christian Hengstenberg, MD[10,11], Per Hoffmann, PhD[52,53,54], Oddgeir L. Holmen, MD[55,56], Kristian Hveem, MD, PhD[55,57], Jan-Håkan Jansson, MD, PhD[58], Karl-Heinz Jöckel, PhD[59], Thorsten Kessler, MD[11,12], Jennifer Kriebel, PhD[60,61,62], Karl L. Laugwitz, MD[12,63], Eirini Marouli, MSc[6], Nicola Martinelli, MD, PhD[64], Mark I. McCarthy, MD[31,65,66], Natalie R. Van Zuydam, PhD[65], Christa Meisinger, MD, MPH[61], Tõnu Esko, PhD[67,68,69,70], Evelin Mihailov, MSc[67], Stefan A. Escher, PhD[71], Maris Alver, MSc[67,72], Susanne Moebus, PhD[59], Andrew D. Morris, MD[73], Jarma Virtamo, MD,PhD[74], Majid Nikpay, PhD[75], Oliviero Olivieri, MD[64], Sylvie Provost, MSc[29], Alaa AlQarawi, BSc[76], Neil R. Robertson, MSc[31,65], Karen O. Akinsansya, PhD[77], Dermot F. Reilly, PhD[77], Thomas F. Vogt, PhD[77], Wu Yin, PhD[77], Folkert W. Asselbergs, MD, PhD[78,79,80] , Charles Kooperberg, PhD[24], Rebecca D. Jackson, MD[81], Eli Stahl, PhD[82],Konstantin Strauch, PhD[83,84], Tibor V. Varga, PhD[71], Melanie Waldenberger, PhD[60,61], Wellcome Trust Case Control Consortium, Lingyao Zeng, MSc[11,12], Rajiv Chowdhury, MD, PhD[85], Veikko Salomaa, MD, PhD[74], Ian Ford, PhD[86], J. Wouter Jukema, MD, PhD[87], Philippe Amouyel, MD, PhD[88], Jukka Kontto, MSSc[74], MORGAM Investigators, Børge G. Nordestgaard, MD, DMSc[89], Jean Ferrières, MD[90], Danish Saleheen, MBBS, PhD[91,92], Naveed Sattar, FRCP, PhD[93], Praveen Surendran, PhD[85], Aline Wagner, MD, PhD[94], Robin Young, PhD[85], Joanna M. M. Howson, PhD[85], Adam S. Butterworth, PhD[85,95], John Danesh, FRCP, DPhil[85,96,97], Diego Ardissino, MD[98], Erwin P. Bottinger, MD[25], Raimund Erbel, MD[59], Paul W. Franks, PhD[71,95,99], Domenico Girelli, MD, PhD[64], Alistair S. Hall, MD, PhD[100], G. Kees Hovingh, MD, PhD[43], Adnan Kastrati, MD[11], Wolfgang Lieb, MD, M.Sc[101], Thomas Meitinger, MD[12,102,103], William E. Kraus, MD[104,105], Svati H. Shah, MD, MPH[104,105], Ruth McPherson, MD, PhD[75], Marju Orho-Melander, PhD[106], Olle Melander, MD, PhD[107], Andres Metspalu, MD, PhD[67,72], Colin N.A. Palmer, PhD[44], Annette Peters, PhD[12,61], Daniel J. Rader, MD[108], Muredach P. Reilly, MB, BCh, MSCE[109], Ruth J.F. Loos, PhD[25,26,110], Alex P. Reiner, MD, MSc[24,111], Dan M. Roden,

MD[21,112], Jean-Claude Tardif, MD[28,29], John R. Thompson, PhD[2,113], Nicholas J. Wareham, MB, BS, PhD[42], Hugh Watkins, MD, PhD[30,31], Cristen J. Willer, PhD[28,114,115], Nilesh J Samani, MD[1,2], Heribert Schunkert, MD[11,12], Panos Deloukas, PhD[6,76], Sekar Kathiresan, MD[32,33,34,35,116].

Word count: 4985

Short title: Novel CAD risk loci and pleiotropy

Author information:

[1] Department of Cardiovascular Sciences, University of Leicester, Leicester, LE3 9QP, UK

[2] NIHR Leicester Cardiovascular Biomedical Research Unit, Glenfield Hospital, Leicester, LE3 9QP, UK

[3] Institute for Integrative and Experimental Genomics, University of Lübeck, 23562 Lübeck, Germany

[4] DZHK (German Research Centre for Cardiovascular Research), partner site Hamburg/Lübeck/Kiel, 23562 Lübeck, Germany

[5] University Heart Center Luebeck, 23562 Lübeck, Germany

[6] William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK

[7] Department of Haematology, University of Cambridge, Cambridge, UK

[8] Cardiovascular Division, Department of Medicine, Washington University School of Medicine, Saint Louis MO 63110 USA

[9] Department of Genetics, Washington University School of Medicine, Saint Louis MO 63110 USA

[10] McDonnell Genome Institute, Washington University School of Medicine, Saint Louis MO 63110 USA

[11] Deutsches Herzzentrum München, Technische Universität München, Lazarettstr. 36, 80636, München, Germany.

[12] DZHK (German Centre for Cardiovascular Research), Partner Site Munich Heart Alliance, Germany.

[13] Institut für Medizinische Biometrie und Statistik, Universität zu Lübeck, Lübeck

[14] Center for Population Studies, National Heart, Lung, and Blood Institute, The Framingham Heart Study, Framingham, Massachusetts, USA.

[15] Department of Immunology, Genetics and Pathology, Rudbeck Laboratory, Uppsala University, Sweden.

[16] Department of Medical Biochemistry and Biophysics, Vascular Biology Unit, Karolinska Institutet, Stockholm, Sweden.

[17] Department of Physiology, Institute of Biomedicine and Translation Medicine, University of Tartu, Estonia.

[18] Department of Cardiac Surgery, Tartu University Hospital, Tartu, Estonia.

[19] Clinical Gene Networks AB, Stockholm, Sweden.

[20] Department of Genetics & Genomic Sciences, Institute of Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York, USA.

[21] Department of Medicine, Vanderbilt University Medical Center, Nashville, TN

[22] Laboratory for Molecular Cardiology, Department of Cardiology, Copenhagen University Hospital Rigshospitalet, Copenhagen, Denmark

[23] School of Public Heath, University of Wisconsin-Milwaukee, Milwaukee, WI, 53205, USA

[24] Fred Hutchinson Cancer Research Center, Seattle WA 98109, USA

[25] The Charles Bronfman Institute for Personalized Medicine, The Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA

[26] The Genetics of Obesity and Related Metabolic Traits Program, The Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA

[27] Department of Internal Medicine, Division of Cardiovascular Medicine, University of Michigan, Ann Arbor, MI, USA

[28] Université de Montréal, Faculté de médecine, Département de médecine, Montreal, QC, Canada

[29] Montreal Heart Institute, Montreal, QC, Canada

[30] Division of Cardiovascular Medicine, Radcliffe Department of Medicine, University of Oxford, Oxford. UK. OX3 9DU

[31] Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK. OX3 7BN

[32] Center for Human Genetic Research, Massachusetts General Hospital, Boston, MA, USA

[33] Cardiovascular Research Center, Massachusetts General Hospital, Boston, MA, USA

[34] Department of Medicine, Harvard Medical School, Boston, Massachusetts, USA

[35] Program in Medical and Population Genetics, Broad Institute, Cambridge, MA, USA

[36] Samsung Advanced Institute for Health Sciences and Technology (SAIHST), Sungkyunkwan University, Samsung Medical Center, Seoul, South Korea

[37] The Center for Statistical Genetics, Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, USA

[38] The Icahn Institute for Genomics and Multiscale Biology, Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, USA

[39] The Zena and Michael A. Weiner Cardiovascular Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA

[40] Department of Biostatistics, Academic Medical Center, Amsterdam, The Netherlands

[41] Institute for Animal Breeding and Genetics, University of Veterinary Medicine Hannover, Bünteweg 17p, Hannover 30559, Germany

[42] MRC Epidemiology Unit, Institute of Metabolic Science, Addenbrooke's Hospital, Box 285, Cambridge, UK

[43] Department of Vascular Medicine, Academic Medical Center, Amsterdam, The Netherlands

[44] Medical Research Institute, University of Dundee, Ninewells Hospital and Medical School, Dundee, DD1 9SY, Scotland, UK.

[45] Department of Biomedical Sciences, Humanitas University, Via Manzoni 113, 20089 Rozzano, Milan, Italy.

[46] Humanitas Clinical and Research Center, Via Manzoni 56, 20089 Rozzano, Milan, Italy

[47] Niguarda Hospital, Milan, Italy

[48] ASL 3 Nuoro, PO San Francesco, Nuoro-IT

[49] Department of Biomedical informatics, Vanderbilt University Medical Center, Nashville, TN

[50] Klinik für Kardiologie, Pneumologie und Innere Medizin, Imland Klinik Rendsburg, Rendsburg, Germany

[51] Institute of Clinical Molecular Biology, Christian-Albrechts-University of Kiel, Kiel, Germany

[52] Institute of Human Genetics, University of Bonn, Bonn, Germany

[53] Department of Genomics, Life & Brain Center, University of Bonn, Bonn, Germany

[54] Division of Medical Genetics, Department of Biomedicine, University of Basel, Basel, Switzerland

[55] HUNT Research Centre, Department of Public Health and General Practice, Norwegian University of Science and Technology, Levanger, Norway

[56] St. Olav Hospital, Trondheim University Hospital, Trondheim, Norway

[57] Department of Medicine, Levanger Hospital, Nord- Trøndelag Health Trust, Levanger, Norway

[58] Department of Public Health and Clinical Medicine, Research Unit Skellefteå, Umeå University, Sweden

[59] Institute for Medical Informatics, Biometry and Epidemiology, University Hospital Essen, Essen, Germany

[60] Research unit of Molecular Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, 85764 Neuherberg, Germany

[61] Institute of Epidemiology II, Helmholtz Zentrum München - German Research Center for Environmental Health, 85764 Neuherberg, Germany

[62] German Center for Diabetes Research, Neuherberg, Germany

[63] I. Medizinische Klinik und Poliklinik, Klinikum rechts der Isar der Technischen Universität München, Munich, Germany

[64] Department of Medicine, Section of Internal Medicine, University of Verona 37134 Italy

[65] Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, OX3 7LJ UK

[66] Oxford National Institute for Health Research Biomedical Research Centre, Churchill Hospital, Old Road Headington, Oxford, OX3 7LJ, UK

[67] Estonian Genome Center, University of Tartu, Tartu, Estonia, 51010

[68] Division of Endocrinology, Boston Children's Hospital, Boston, MA 02115, USA

[69] Department of Genetics, Harvard Medical School, Boston, MA02115, USA

[70] Broad Institute of the Massachusetts Institute of Technology and Harvard University, 140 Cambridge 02142, MA, USA

[71] Genetic and Molecular Epidemiology Unit, Lund University Diabetes Centre, Department of Clinical Sciences, Lund University, SE-20502, Malmö, Sweden

[72] Institute of Molecular and Cell Biology, Tartu, Estonia, 51010

[73] School of Molecular, Genetic and Population Health Sciences,University of Edinburgh,Medical School,Teviot Place,Edinburgh, EH8 9AG, Scotland, UK

[74]), National Institute for Health and Welfare (THL), FI-00271 Helsinki, Finland

[75] Ruddy Canadian Cardiovascular Genetics Centre, University of Ottawa Heart Institute, Ottawa, ON, Canada

[76] Princess Al-Jawhara Al-Brahim Centre of Excellence in Research of Hereditary Disorders (PACER-HD), King Abdulaziz University, Jeddah 21589, Saudi Arabia

[77] Merck Sharp & Dohme Corp., Rahway, NJ 07065, USA

[78] Department of Cardiology, Division Heart & Lungs, UMC Utrecht, 3584 CX, the Netherlands

[79] Durrer Center for Cardiogenetic Research, ICIN-Netherlands Heart Institute, Utrecht, the Netherlands

[80] Institute of Cardiovascular Science, Faculty of Population Health Sciences, University College London, London, United Kingdom

[81] Division of Endocrinology, Diabetes and Metabolism, Department of Medicine, Ohio State University, Columbus OH, 43201, USA

[82] Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA

[83] Institute of Genetic Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, 85764 Neuherberg, Germany.

[84] Institute of Medical Informatics, Biometry and Epidemiology, Chair of Genetic Epidemiology, Ludwig-Maximilians-Universität, Munich, Germany

[85] Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge, CB18RN, UK

[86] Robertson Centre for Biostatistics, University of Glasgow, Glasgow, G12 8QQ, UK

[87] Department of Cardiology, Leiden University Medical Center, Leiden and Interuniversity Cardiology Institute of the Netherlands, Utrecht, The Netherlands

[88] Univ. Lille, Inserm, CHU Lille, Institut Pasteur de Lille, U1167 - RID-AGE - Risk factors and molecular determinants of aging-related diseases, F-59000 Lille, France

[89] Copenhagen University Hospital and Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

[90] Toulouse University School of Medicine, TSA 50032, 31059, Toulouse, France

[91] Department of Biostatistics and Epidemiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104-5127, USA

[92] Center for Noncommunicable Diseases, Karachi, Pakistan

[93] British Heart Foundation, Glasgow Cardiovascular Research Centre, University of Glasgow, Glasgow, G12 8TA, UK

[94] Department or Epîdemiology and Public Health, EA 3430, University of Strasbourg, Strasbourg, France

[95] Department of Nutrition, Harvard School of Public Health, Boston, USA

[96] Wellcome Trust Sanger Institute, Hinxton, Cambridge, CB10 1SA, UK

[97] National Institute of Health Research Blood and Transplant Research Unit in Donor Health and Genomics, University of Cambridge.

[98] Parma University Hospital, Parma, Italy

[99] Department of Public Health & Clinical Medicine, Umeå University Hospital, Umeå, Sweden

[100] Leeds Institute of Genetics, Health and Therapeutics, University of Leeds, Leeds, LS2 9LU, UK

[101] Institute of Epidemiology and Biobank popgen, Christian-Albrechts-University Kiel, Kiel, Germany

[102] Institute of Human Genetics, Helmholtz Zentrum München - German Research Center for Environmental Health, 85764 Neuherberg, Germany

[103] Institute of Human Genetics, Technische Universität München, 81675 Munich, Germany

[104] Duke Molecular Physiology Institute, Duke University, Durham, NC 27701 USA

[105] Division of Cardiology, Department of Medicine, Duke University, Durham, NC 27710 USA

[106] Department of Clinical Sciences in Malmo, Lund University, Clinical Research Center, Malmo, Sweden

[107] Department of Clinical Sciences, Diabetes and Endocrinology, Lund University, University Hospital Malmo, Malmo, Sweden

[108] Department of Genetics, Cardiovascular Institute and Institute of Translational Medicine and Therapeutics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

[109] Division of Cardiology, Department of Medicine and the Irving Institute for Clinical and Translational Research, Columbia University, New York, NY, USA

[110] The Mindich Child Health and Development Institute, The Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA

[111] Department of Epidemiology, University of Washington, Seattle, Washington, 98195, USA

[112] Department of Pharmacology, Vanderbilt University Medical Center, Nashville, TN

[113] Department of Health Sciences, University of Leicester, Leicester, LE1 7RH, UK

[114] Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, USA

[115] Department of Human Genetics, University of Michigan, Ann Arbor, MI, US

[116] Cardiology Division, Massachusetts General Hospital, Boston, MA, USA

Corresponding authors:

Nilesh J Samani MD
Department of Cardiovascular Sciences, University of Leicester
BHF Cardiovascular Research Centre
Glenfield Hospital, Groby Road
Leicester, LE3 9QP
United Kingdom
Email: njs@le.ac.uk
Tel: +44 (0)116 204 4758

Heribert Schunkert MD
Deutsches Herzzentrum München
Technische Universität München
Lazarettstraße 36
80636 Munich

Germany
Email: schunkert@dhm.mhn.de

Panos Deloukas PhD
William Harvey Research Institute
Queen Mary University of London ,
Charterhouse Square
London, EC1M 6BQ
United Kingdom
Email: p.deloukas@qmul.ac.uk

Sekar Kathiresan MD
Cardiovascular Research Center and Center for Human Genetic Research
Massachusetts General Hospital
185 Cambridge Street, CPZN 5.252
Boston, MA 02114
USA
Email: skathiresan@partners.org

Abstract

Background: Genome-wide association studies (GWAS) have so far identified 56 loci associated with risk of coronary artery disease (CAD). Many CAD loci show pleiotropy, i.e. they are also associated with other diseases or traits.

Objectives: To systematically test if genetic variants identified for non-CAD diseases/traits also associate with CAD and to undertake a comprehensive analysis of the extent of pleiotropy of all CAD loci.

Methods: In discovery analyses involving 42,335 CAD cases and 78,240 controls, we tested the association of 29,383 common (minor allele frequency > 5%) single nucleotide polymorphisms (SNPs) available on the Exome array, which included a substantial proportion of known or suspected SNPs associated with common diseases/traits as of 2011. Suggestive association signals were replicated in an additional 30,533 cases and 42,530 controls. To evaluate pleiotropy, we tested CAD loci for association with cardiovascular risk factors (lipid traits, blood pressure phenotypes, body mass index, diabetes and smoking behaviour) and with other diseases/traits through interrogation of currently available GWAS catalogues.

Results: We identified six new loci associated with CAD at genome-wide significance – on 2q37 (*KCNJ13-GIGYF2*), 6p21 (*C2*), 11p15 (*MRVI1-CTR9*), 12q13 (*LRP1*), 12q24 (*SCARB1*) and 16q13 (*CETP*). Risk allele frequencies ranged from 0.15-0.86 and odds ratio per copy of the risk allele ranged from 1.04-1.09. Out of 62 new and known CAD loci, 24 (38.7%) showed statistical association with a traditional cardiovascular risk factor with some loci showing multiple associations and 29 (47%) showed associations at $P<1x10^{-4}$ with a range of other diseases/traits.

Conclusions: We identified six loci associated with CAD at genome-wide significance. Several CAD loci show substantial pleiotropy which may help to understand the mechanisms by which these loci affect CAD risk.

Abbreviations

GWAS	Genome-wide association study/studies
CAD	Coronary artery disease
SNP	Single nucleotide polymorphism
eQTL	Expression quantitative trait loci
LDL	Low-density lipoprotein
HDL	High-density lipoprotein
LD	Linkage disequilibrium
BMI	Body mass index
CETP	Cholesteryl ester transfer protein

Over the past decade GWAS have identified several thousand robust associations ($p<5x10^{-8}$) for a range of human traits and diseases. For CAD, 56 such loci have been identified so far explaining ~15% of the heritability of the disease (1,2). Around a third of the CAD loci also show association with a known or putative cardiovascular risk factor, in particular blood pressure and lipid traits (2). Furthermore, several loci show association with other diseases e.g. the CAD-associated variants in the chromosome 9p21 locus also associate with risk of stroke, abdominal aortic and intracranial aneurysms (3,4). These observations suggest that a comprehensive analysis of variants associated with other diseases and traits may not only identify additional loci associated with risk of CAD but also provide important insights on genetic mechanisms shared by different diseases.

Here, we leveraged the Illumina HumanExome Beadchip to test the contribution of 29,393 common variants (SNPs) (minor-allele frequency > 5%) for association with CAD. The variants included the majority of reported trait/disease associated lead SNPs in the National Human Genome Research Institute (NHGRI) GWAS catalogue as of August 2011 as well as a number of associations for complex diseases unpublished at that time, variants in the HLA region and a scaffold of approximately 5,000 SNPs placed on the array for identity by descent testing. The results of an analysis of rare (minor-allele frequency < 5%) coding sequence ('exome') variants on this array with CAD were recently reported (5).

We identified six new loci associated at genome-wide significance with CAD, annotated these, and undertook a detailed examination of the extent of pleiotropy of these loci as well the previously known CAD loci.

Methods

Study design and participants

The study consisted of discovery and replication phases and is described in more detail elsewhere (5). Briefly, the discovery cohort included 42,335 cases and 78,240 controls from 20 individual studies (**Supplementary Table S1**) while the replication cohort, which was separately assembled and ascertained to have no sample overlap with the discovery cohorts, included 30,533 cases and 42,530 controls from eight studies (**Supplementary Table S2**). With the exception of participants from BRAVE and PROMIS in the replication cohort – who were of South Asian ancestry – all participants were of European ancestry.

Genotyping and Quality Control

Samples were genotyped on the Illumina HumanExome BeadChip v1.0, v1.1, or the Illumina OmniExome (which includes markers from the HumanExome BeadChip) arrays followed by quality control procedures as previously described (5).

Statistical analysis

In discovery samples that passed quality control procedures, we performed individual tests for association of the selected variants with CAD in each study separately, using logistic regression analysis with principal components of ancestry as covariates.[5] We combined evidence across individual studies using an inverse-variance weighted fixed-effects meta-analysis. Heterogeneity was assessed by Cochran's Q statistic (6). In the discovery phase, we defined suggestive novel association as a meta-analysis *P* value $\leq 1x10^{-6}$.

For variants with suggestive association, we performed association analysis in the replication studies (see **Supplementary Methods**). We defined significant novel associations as those nominally significant ($P<0.05$) in the replication study and with an overall (discovery and replication combined) $P<5x10^{-8}$.

Bioinformatics analysis

Annotation of novel loci: To identify any association between the novel loci and gene expression traits we performed a systematic search of cis-eQTLs (described in the **Supplementary methods**). To identify candidate causal SNPs at the new loci we annotated each of the lead variants as well as SNPs in high linkage disequilibrium (LD) ($r^2 > 0.8$) based on position, overlap with regulatory elements and *in silico* SNP prioritisation tools (see **Supplementary methods**).

Assessment of pleiotropy: For both the novel loci and all previously reported CAD loci (1,2), we tested the association of the lead CAD-associated variant (or if unavailable a proxy) with traditional cardiovascular risk factors using publically available GWAS meta-analyses datasets for systolic, diastolic, and pulse pressures (7,8), LDL cholesterol level, HDL cholesterol level, triglycerides level (9,10), type 2 diabetes mellitus (11), BMI (12)), and smoking quantity (13). The maximum sizes of these datasets ranged from 41,150 to 339,224 individuals. For variants available on the Exome array with a known genome-wide association with a risk factor, we also compared the magnitude of the reported association with the risk factor with the observed association with CAD in our analysis.

To identify any associations with other diseases or traits we searched version 2 of the Genome-Wide Repository of Associations between SNPs and Phenotypes (GRASP) database (14) and the NHGRI-EBI GWAS catalogue (accessed 19/11/2015) (15) and collected all associations below $1x10^{-4}$. For all associations we identified the lead variant for that trait or disease and calculated pairwise LD with the lead CAD associated variant using SNAP (16).

Results

Discovery of novel loci

In the discovery cohort, 28 variants not located in a known CAD locus (defined as +/- 300kb from the published lead SNP) showed association with CAD at a *P*-value of $<1x10^{-6}$ (**Supplementary Table S3**). No marked heterogeneity was observed (**Supplementary Table 3**), justifying the use of a fixed-effects model. We then tested these 28 variants for replication and six variants showed both a nominally significant ($p<0.05$) association in the replication cohort and a combined discovery and replication meta-analyses *P*-value exceeding the threshold for genome-wide significance ($p<5x10^{-8}$) (**Table 1**). As typical for GWAS findings, the risk alleles were common (allele frequencies ranging from 15% to 86%) and the risk increase per allele was modest (ranging from 4% to 9%) (**Table 1**).

Annotation of novel loci

Forest and regional association plots for the six novel loci are shown in **Supplementary Figures S1 and S2**, respectively. Interrogation of the 1000G phase 1 EUR data using Haploreg (17) showed that the number of SNPs in high LD ($r^2>0.8$) with the lead variant varied between one (*LRP1* locus and *CETP* locus) and 111 (*KCNJ13-GIGYF2* locus) (**Supplementary Table S4**). Apart from the lead variant at the *KCNJ13-GIGYF2* locus, which is a non-synonymous SNP (see below) none of the other loci had a variant affecting protein sequence in high LD with the lead variant.

Notable *cis* eQTL findings for the new loci are shown in **Supplementary Table S5** and functional annotation of the lead variant and variants in high LD are shown in **Supplementary Figure S3**. The main findings from these analyses are discussed locus by locus below.

16q13: The lead variant, rs1800775, also known as −629C>A, is in the promoter of the *CETP* gene which mediates the transfer of cholesteryl esters from HDL cholesterol to other lipoproteins and was placed on the array because of its association with plasma HDL cholesterol level (9,10). The risk (C) allele is associated with lower HDL cholesterol and modest increases in plasma LDL cholesterol and triglycerides levels (9,10) and previous studies have shown that rs1800775 is itself functional in that the C allele allele disrupts binding of the Sp1 transcription factor resulting in increased promoter activity (18). This is in agreement with our annotation, which predicts this to be more likely to be a functional SNP than the only other SNP in high LD, rs3816117 (**Supplementary Figure S3**). Consistent with this we also found associations between rs1800775 and *CETP* expression (r^2 of 0.77 with the best eSNP) in monocytes and liver (**Supplementary Table S5**) and previous studies have shown that the variant is also associated with plasma CETP level (19,20).

12q24: The lead variant, rs11057830, and all 8 variants in high LD are located in a region of approximately 10kb in intron 1 of *SCARB1* which encodes SR-B1, a receptor for HDL cholesterol. Other variants at this locus have been associated with HDL cholesterol level (9,10). However these HDL cholesterol variants are not in high LD with the CAD-associated variants identified here, which only have a modest association with plasma HDL cholesterol level (**Supplementary Table S6**) but a stronger association with plasma LDL cholesterol and triglycerides levels (**Table 2**). rs11057830 was included on the array because of an association of the A allele (CAD-risk associated allele) with higher levels of vitamin E (**Table 3**) (21). Variants in high LD with the CAD risk allele at rs11057830 have also been associated with increased lipoprotein-associated phospholipase A_2 (Lp-PLA_2) activity (22). eQTL analysis identified an association between rs11057841 (r^2=0.92 with the lead variant), and expression of *SCARB1* in intestine (**Supplementary Table S5**). Functional annotation of the locus did not identify a strong candidate causal SNP, but rs10846744 (r^2=0.94 with the

lead variant) overlaps a DNAse hypersensitivity peak in a region bound by several transcription factors (**Supplementary Figure S3**).

12q13: The lead variant, rs11172113, is in intron 1 of *LRP1* (low density lipoprotein receptor related protein-1) and only has one other adjacent SNP in high LD (**Figure 1** and **Supplementary Table S4**). The risk (C) allele of the lead variant has previously been associated with reduced risk of migraine (23) and there is an association of the alternate (T) allele with reduced lung function (24). There are also associations at this locus for abdominal aortic aneurysm (25) and triglyceride levels (10); however these variants are in modest or low LD to the CAD-associated SNP (r^2 of 0.54 and 0.07, respectively). The lead variant overlaps a region containing peaks in DNAse hypersensitivity in several cells and tissues, including aortic smooth muscle cells, within a predicted enhancer element (**Supplementary Figure S3**). We found associations between the CAD risk allele at rs11172113 and reduced expression of *LRP1* in atherosclerotic and non-atherosclerotic arterial wall as well as eQTLs in omental and subcutaneous adipose tissue (**Supplementary Table S5**).

11p15: The lead variant, rs11042937, at this locus lies in an intergenic region between *MRVI1* (Murine Retrovirus Integration Site 1 Homolog) encoding IRAG (inositol-trisphosphate receptor-associated cGMP kinase substrate) a mediator of smooth muscle tone and *CTR9* which encodes a component of the PAF1 complex with some SNPs in high LD located within intron 1 of *MRVI1* (**Supplementary Figure S3**). The lead variant was included on the array because of a suggestive association with bipolar disorder and schizophrenia (26). There was no association of the locus with any cardiovascular risk factors and we did not identify any eQTLs. Evidence for a regulatory function for either the lead variant or any of the SNPs in high LD was also weak (**Supplementary Figure S3**).

6p21: The lead variant, rs3130683, lies in the HLA complex in intron 1 of *C2*, which encodes the complement C2 protein. There are just 14 SNPs in high LD with the lead variant (**Supplementary Table S4**) but the CAD signal spans a region of approximately 300kb including more than 20 genes (**Supplementary Figure S3**). Apart from a single synonymous variant in *HSPA1A* (Heat Shock 70kDa Protein 1A) the other high LD variants are non-coding with several of the variants showing evidence for regulatory functionality (**Supplementary Figure S3**). Although there are a large number of eQTLs in the HLA region most of these are variants with modest (r^2<0.5) LD with the CAD-associated variants and the only eQTL of note was with *CYP21A2* (Cytochrome P450, Family 21, Subfamily A, Polypeptide 2) expression in whole blood **(Supplementary Table S5)**. rs3869109, another variant at the HLA locus, approximately 700kb away from the new lead variant, has been reported to be associated with CAD (27). In our discovery cohort rs3869109 has a P value of association with CAD of 0.23.

2q37: The lead variant, rs1801251, was included on the array for identity by descent testing. rs1801251 causes a threonine to isoleucine amino acid change at position 95 in KCNJ13, an inwardly rectifying potassium channel protein; however this is not predicted to be functionally important. There is extended linkage at this locus with more than 100 SNPs in high LD with the lead variant in a region of ~170kb also spanning *GIGYF2* (GRB10 Interacting GYF Protein 2) (**Supplementary Figure S3).** *KCJN13* is located entirely within *GIGYF2* and transcribed in the opposite direction. A number of the associated variants are in annotated regulatory regions, with the top scoring candidate by *in silico* prediction, rs11555646, lying in the 5'-UTR of GIGYF2 close to the initiating methionine (**Supplementary Figure S3**). There was no association of the locus with any of the cardiovascular risk factors but we found eQTLs for the lead variant or a variant in high LD for both *GIGFY2* and *KCNJ13* (**Supplementary Table S5**).

CAD loci and pleiotropy

We undertook an updated analysis of the association of all 62 CAD loci (56 published and 6 novel in this report) with traditional cardiovascular risk factors (blood pressure traits, lipid traits, BMI, type 2 diabetes and smoking). The full results are shown in **Supplementary Table S6** and the significant associations summarised in **Table 2**. Of the 62 CAD loci, 24 (38.7%) showed a statistical association at a Bonferroni corrected $P<8.32 \times 10^{-5}$ with a traditional cardiovascular risk factor with some loci showing multiple associations (**Figure 1**). The largest number of associations are with lipid traits (14 with LDL cholesterol, 9 with HDL cholesterol, 7 with triglycerides) followed by blood pressure traits (5 with diastolic blood pressure, 4 with systolic blood pressure and 1 with pulse pressure), BMI (5 associations) and type 2 diabetes (one association). The majority of associations were in the direction consistent from the epidemiological association of these risk factors with CAD, although a few displayed effects in the opposite direction (the risk variants at 2q33 and 12q24 are associated with reduced plasma LDL cholesterol and at 10q24, 12q24 and 19q13 are associated with lower BMI).

To inform the interpretation of these data, in a complementary analysis, for variants available on the array with a known genome-wide association with a risk factor, we also compared the magnitude of the reported association with the risk factor with the observed association with CAD in our data. Except for LDL cholesterol and BMI, the correlations between the two effects were either weak or insignificant (**Supplementary Figure S5**). In a separate analysis conducted in the 150,000 participants in UK Biobank with currently released genotype data (www.**ukbiobank**.ac.**uk**) we confirmed that none of the CAD-associated variants showed a gender difference in allele frequency (data not shown).

We next analysed the association of the 62 CAD loci with other diseases and traits. When restricted to variants with a high LD ($r^2>0.8$) with the lead CAD variant, 29/62 loci (47%)

showed such an association with another disease/trait at a *P*-value of $<1x10^{-4}$. Several loci showed multiple associations (**Table 3**). While in the majority of cases, the CAD associated risk allele was also associated with an increased risk (or level) of the other disease or trait, this was not always the case. Furthermore, in some loci with multiple associations, the direction of association varied between diseases (**Table 3**).

Discussion

This large-scale meta-analysis of common variants including many with prior evidence for association with another complex trait resulted in the identification of six new CAD loci at genome-wide significance. We also show that almost half of the CAD loci identified to date demonstrate pleiotropy, i.e. an association with another disease or trait. The findings add to our understanding of the genetic basis of CAD and may provide clues to the mechanisms by which such loci affect CAD risk.

Our findings of a genome-wide association with CAD of a functional variant in the promoter of the *CETP* gene that is also associated with its expression and plasma activity (18-20) adds to previous evidence linking genetically determined increase in the activity of this gene with higher risk of CAD (20). There has been a long-standing interest in CETP inhibition as a therapeutic target primarily because of the effect on plasma HDL-cholesterol level. However, several CETP inhibitors have recently failed to improve cardiovascular outcomes in large randomised clinical trials (28-30), and in one case caused harm (28), despite markedly increasing plasma HDL cholesterol level. Furthermore, Mendelian randomisation studies have questioned the causal role of lower plasma HDL-cholesterol in increasing CAD risk (31,32). While previous studies have shown that the CETP genetic variant we report here impacts on activity of CETP, the precise mechanism(s) by which this variant affects CAD risk remains uncertain.

A notable finding is the association with CAD of common variants located in the *SCARB1* gene. Association of variants at the *SCARB1* locus with CAD was also reported by the CARDIoGRAMplusC4D consortium but this did not reach genome-wide significance (1). The gene encodes the canonical receptor, SR-BI, responsible for HDL cholesteryl ester uptake in hepatocytes and steroidogenic cells (33). Genetic modulation of SR-BI levels in mice is associated with marked changes in plasma HDL-cholesterol (34, 47). Consistent with

this, a rare loss of function variant in which leucine replaces proline 376 (P376L) in *SCARB1* was recently identified through sequencing of individuals with high plasma HDL-cholesterol (35). Interestingly, despite having higher plasma HDL cholesterol levels, 346L carriers had an increased risk of CAD suggesting that the association of variation at this locus on CAD is not driven primarily through plasma HDL cholesterol level (35). Indeed, there is only a nominal association of the lead CAD variant at this locus (rs11057830) with plasma HDL cholesterol (**Supplementary Table S6**). The variant is also modestly associated with plasma LDL cholesterol and serum triglyceride level (**Table 2**). All three of these lipid associations are directionally consistent with epidemiological associations of these lipids with CAD risk and could in combination explain the association of the locus with CAD. However, the lead variant is more strongly associated with Lp-PLA_2 activity and mass (**Table 3**) which could provide an alternative explanation for its association with CAD (see below). Irrespective of the mechanism, our findings when combined with those of Zanoni et al. (35) suggest that modulating SR-B1 may be therapeutically beneficial.

We found, after adjusting for multiple testing, that just over one third of the CAD loci showed an association with what are regarded as traditional cardiovascular risk factors. Although, the vast majority of the associations were in the direction consistent with the epidemiological association of these risk factors with CAD, as noted above with respect to loci affecting HDL-cholesterol level, this should not be interpreted as implying that these loci affect CAD risk through an impact on the specific risk factor. Indeed, for variants available on the array with a known genome-wide association with these risk factors, we found a poor correlation between the magnitudes of their effect of the risk factor and their association with CAD in our dataset except for LDL cholesterol (**Supplementary Figure S5**). Nonetheless, formal causal inference analyses, using Mendelian randomisation have implicated LDL

cholesterol, triglyceride-rich lipoproteins, blood pressure, type 2 diabetes and BMI as causally involved in CAD (36).

Almost half the CAD loci showed a strong or suggestive association with other diseases or traits with in many cases the identical variant being the lead variant reported for the association with these other conditions (**Table 3**). Some of associations with other traits, for example with coronary calcification (3q22, 6p24, 9p21, 13q34, 15q25) or carotid intima-media thickness (4q31, 19q13), are not surprising as these traits are known to be correlated with CAD. Others, e.g. with risk of stroke (7p21, 9p21), may reflect a shared aetiology. However, the mechanism(s) behind most of the observed pleiotropy is not clear although the findings could provide clues as to how the locus may affect CAD risk. As an example, five loci (12q24, 1p13, 6q25, 11q23, and 19q13) show strong associations with plasma activity and/or mass of lipoprotein-associated phospholipase A_2 (Lp-PLA_2). Lp-PLA_2 is expressed in atherosclerotic plaques where studies have suggested a role in the production of proinflammatory and pro-apoptotic mediators primarily through interaction with oxidised-LDL (37,38). A meta-analysis of prospective studies showed an independent and continuous relationship of plasma Lp-PLA_2 with risk of CAD (39). However, it should be noted that Mendelian randomisation analyses have not supported a causal role of secreted Lp-PLA2 in CHD (40) and Phase III trials of darapladib, a Lp-PLA2 inhibitor, has shown no benefit in patients with stable coronary heart disease (41) or acute coronary syndromes (42) when added to conventional treatments including statins.

Chronic inflammation plays a key role in both the pathogenesis of CAD as well as that of inflammatory bowel disease (IBD). It is therefore interesting to note the association of the same locus at 15q22 with CAD as well as Crohn's disease and ulcerative colitis (**Table 3**). Association of this locus with CAD at genome-wide significance was recently reported by the CARDIoGRAMplusC4D consortium (2) with the lead SNP (rs56062135) showing strong

linkage disequilibrium (r^2=0.9) with the lead SNP (rs17293632) associated with IBD. Both rs56062135 and rs17293632 lie in a region of ~ 30kb within the initial introns of *SMAD3* (SMAD family member 3) gene, a signal transducer in the transforming growth factor beta (TGF-β) pathway. Indeed, rs17293632 was included on the exome array because of its known association with Crohn's disease and showed a significant association with CAD in our combined dataset (*P*-value $1.78x10^{-8}$). Farh et al (434) interrogated ChIP-seq data from ENCODE and found allele specific binding of the AP-1 transcription factor to the major (C) allele in heterozygous cell lines and suggested that the T allele of rs17293632 increases risk of Crohn's disease by disrupting AP-1 regulation of *SMAD3* expression. Interestingly, the direction of effect on CAD risk observed for this variant is in the opposite direction to that for inflammatory disorders with the C allele being the risk allele. Recent analysis of this variant in arterial smooth muscle cells confirmed that the CAD risk allele preserves AP-1 transcription factor binding and increases expression of *SMAD3* (44). Further investigation of the discordant effects of SMAD3 may shed light on the mechanisms of both diseases.

Our study has several limitations. First, in our discovery study we were only able to interrogate common variants associated with other diseases and traits that were known at the time of the creation of the exome array in late 2011 and thus included on the array. On the other hand our interrogation for pleiotropic associations of the new and known CAD has used the latest data available in the GWAS catalogues and other sources. Second, the common variants tested in our study confer statistically robust yet quantitatively modest effects on both CAD and potentially related traits. Thus, we may have missed associations with other traits. However, if such traits were considered as intermediary steps in the aetiology of CAD, exploration of our large GWAS sample sets and respective GWAS catalogues should have detected relevant associations. Third, our discovery analysis is based largely on subjects with Western-European ancestry and any association with CAD of the new loci in other

populations needs further evaluation. Finally, although we used relatively stringent criteria (minimal $r^2 > 0.8$ between the CAD SNP and the lead variant associated with the other disease/trait), the limited content of the exome array and the information available in the GWAS catalogues means that we cannot examine the extent of overlap in the loci in detail.

In summary, through an analysis of selected variants associated with other disease traits we report the discovery of six further loci associated with CAD. Furthermore, in the most comprehensive analysis to date, we show that several of the new and previously established loci show substantial pleiotropy which may help to understand the mechanisms by which these loci affect CAD risk.

References

1. CARDIoGRAMplusC4D Consortium. Large-scale association analysis identifies new risk loci for coronary artery disease. Nat Genet 2013;45:25-33.

2. CARDIoGRAMplusC4D Consortium. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. Nat Genet 2015;47:1121-30.

3. Dichgans M, Malik R, König IR, et al. Shared genetic susceptibility to ischemic stroke and coronary artery disease: a genome-wide analysis of common variants. Stroke 2014;45:24-36.

4. Helgadottir A, Thorleifsson G, Magnusson KP, et al. The same sequence variant on 9p21 associates with myocardial infarction, abdominal aortic aneurysm and intracranial aneurysm. Nat Genet 2008;40(2):217-24.

5. Myocardial Infarction Genetics and CARDIoGRAM Exome Consortia Investigators. Coding Variation in ANGPTL4, LPL, and SVEP1 and the Risk of Coronary Disease. N Engl J Med 2016; 374: 1134-1144.

6. Cochran WG. The combination of estimates from different experiments. Biometrics 1954;10:101-129.

7. International Consortium for Blood Pressure Genome-Wide Association Studies. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. Nature 2011;478:103-9.

8. Wain LV, Verwoert GC, O'Reilly PF, et al. Genome-wide association study identifies six new loci influencing pulse pressure and mean arterial pressure. Nat Genet 2011;43:1005-11.

9. Teslovich TM, Musunuru K, Smith AV, et al. Biological, clinical and population relevance of 95 loci for blood lipids. Nature 2010;466:707-13.

10. Global Lipids Genetics Consortium. Discovery and refinement of loci associated with lipid levels. Nat Genet 2013;45:1274-83.

11. Morris AP, Voight BF, Teslovich TM, et al. Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. Nat Genet 2012;44:981-90.

12. Locke AE, Kahali B, Berndt SI, et al. Genetic studies of body mass index yield new insights for obesity biology. Nature 2015;518: 197-206.

13. Liu JZ, Tozzi F, Waterworth DM, et al. Meta-analysis and imputation refines the association of 15q25 with smoking quantity. Nat Genet 2010;42:436-40.

14. Eicher JD, Landowski C, Stackhouse B, et al. GRASP v2.0: an update on the Genome-Wide Repository of Associations between SNPs and phenotypes. Nucleic Acids Res 2015;43:D799-804.

15. Welter D, MacArthur J, Morales J, et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. Nucleic Acids Research 2014;42:D1001-6.

16. Johnson AD, Handsaker RE, Pulit S, et al. SNAP: A web-based tool for identification and annotation of proxy SNPs using HapMap.Bioinformatics 2008;24:2938-9.

17. Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. Nucleic Acids Res. 2012; 40:D930-4.

18. Dachet C, Poirier O, Cambien F, et al. New functional promoter polymorphism, CETP/-629, in cholesteryl ester transfer protein (CETP) gene related to CETP mass and high density lipoprotein cholesterol levels: role of Sp1/Sp3 in transcriptional regulation. Arterioscler Thromb Vasc Biol 2000;20:507-15.

19. Klerkx AH, Tanck MW, Kastelein JJ, et al. Haplotype analysis of the CETP gene: not TaqIB, but the closely linked -629C-->A polymorphism and a novel promoter variant are independently associated with CETP concentration. Hum Mol Genet. 2003;12:111-23.

20. Thompson A, Di Angelantonio E, Sarwar N, et al. Association of cholesteryl ester transfer protein genotypes with CETP mass and activity, lipid levels, and coronary risk. JAMA 2008;299:2777–88.

21. Major JM, Yu K, Wheeler W, et al. Genome-wide association study identifies common variants associated with circulating vitamin E levels. Hum Mol Genet 2011;20:3876-83.

22. Suchindran S, Rivedal D, Guyton JR, et al. Genome-wide association study of Lp-PLA(2) activity and mass in the Framingham Heart Study. PLoS genet 2010;6:e1000928.

23. Anttila V, Winsvold BS, Gormley P, et al. Genome-wide meta-analysis identifies new susceptibility loci for migraine. Nat Genet 2013;45:912-7.

24. Soler Artigas M, Loth DW, Wain LV, et al. Genome-wide association and large-scale follow up identifies 16 new loci influencing lung function. Nat Genet 2011;43:1082-90.

25. Bown MJ, Jones GT, Harrison SC, et al. Abdominal aortic aneurysm is associated with a variant in low-density lipoprotein receptor-related protein 1. Am J Hum Genet 2011;89:619-

26. Wang KS, Liu XF, Aragam N. A genome-wide meta-analysis identifies novel loci associated with schizophrenia and bipolar disorder. Schizophr Res 2010;124:192-9.

27. Davies RW, Wells GA, Stewart AF, et al. A genome-wide association study for coronary artery disease identifies a novel susceptibility locus in the major histocompatibility complex. Circ Cardiovasc Genet 2012;5:217-25.

28. Barter PJ, Caulfield M, Eriksson M, et al. Effects of torcetrapib in patients at high risk for coronary events. N Engl J Med 2007;357:2109–22.

29. Schwartz, GS, Olsson, AG, Abt M et al. Effects of Dalcetrapib in Patients with a Recent Acute Coronary Syndrome. N Engl J Med 2012;367:2089-99.

30. Nicholls SJ, Lincoff A, Barter P, et al. Impact of the Cholesteryl Ester Transfer Protein Inhibitor Evacetrapib on Cardiovascular Events: Results of the ACCELERATE trial.

Presented at the 2016 American College of Cardiology 65th Annual Scientific Sessions. April 2-4, 2016; Chicago, IL.

31. Voight BF, Peloso GM, Orho-Melander M, et al. Plasma HDL cholesterol and risk of myocardial infarction: a mendelian randomisation study. Lancet. 2012;380:572-80.

32. Holmes MV, Asselbergs FW, Palmer TM et al. Mendelian randomisation of blood lipids for coronary heart disease. Eur Heart J 2015;36:539-50.33. Acton S, Rigotti A, Landschulz KT, et al. Identification of scavenger receptor SR-BI as a high density lipoprotein receptor. Science 1996;271:518–520.

34. Kozarsky KF, Donahee MH, Glick JM, et al. Gene transfer and hepatic overexpression of the HDL receptor SR-BI reduces atherosclerosis in the cholesterol-fed LDL receptor-deficient mouse. Arterioscler Thromb Vasc Biol 2000;20:721-7.

35. Zanoni P, Khetarpal SA, Larach DP et al. Rare variant in scavenger receptor BI raises HDL cholesterol and increases risk of coronary heart disease. Science 2016;351: 1166-71.

36. Jansen H, Samani NJ, Schunkert H. Mendelian randomization studies in coronary artery disease. Eur Heart J 2014;35:1917-1924.

37. Carpenter KL, Dennis IF, Challis IR, et al. Inhibition of lipoprotein-associated phospholipase A2 diminishes the death inducing effects of oxidised LDL on human monocyte-macrophages. FEBS Lett 2001;505:357-63.

38. Shi Y, Zhang P, Zhang L, et al. Role of lipoprotein-associated phospholipase A2 in leukocyte activation and inflammatory responses. Atherosclerosis 2007;191:54-62.

39. Thompson A, Gao P, Orfei L, et al. Lipoprotein-associated phospholipase A(2) and risk of coronary disease, stroke, and mortality: collaborative analysis of 32 prospective studies. Lancet 2010;375:1536-44.

40. Talmud PJ, Holmes MV. Deciphering the Causal Role of sPLA2s and Lp-PLA2 in Coronary Heart Disease. Arterioscler Thromb Vasc Biol 2015;35**:**2281-2289.

41. White HD, Held C, Stewart R, et al. Darapladib for preventing ischemic events in stable coronary heart disease. N Engl J Med 2014;370:1702–1711.

42. O'Donoghue ML, Braunwald E, White HD, et al. Effect of darapladib on major coronary events after an acute coronary syndrome: the SOLID-TIMI 52 randomized clinical trial. JAMA 2014;312:1006–1015.

43. Farh KK, Marson A, Zhu J, et al. Genetic and epigenetic fine mapping of causal autoimmune disease variants. Nature 2015;518:337-43.

44. Turner AW, Martinuk A, Silva A, et al. Functional analysis of a novel genome-wide association study signal in SMAD3 that confers protection from coronary artery disease. Arterioscler Thromb Vasc Biol 2016;36:972-83.

Figures

Figure 1. Chord diagram showing significant associations of CAD loci with selected cardiovascular risk factors.

Table 1. Novel loci showing significant association with coronary artery disease

Lead Variant	Locus	Locus Name	A1/2 (freq)	Discovery			Replication			Meta-analysis		Known Associations with Lead Variant
				Cases/Controls	OR (95% CI)	p value	Cases/Controls	OR (95% CI)	p value	OR (CI 95%)	p value	
rs1801251	2q37	*KCNJ13-GIGYF2*	A/G (0.35)	42332/78229	1.06 (1.04-1.08)	$1.46x10^{-08}$	30528/42521	1.03 (1.01-1.06)	0.007	1.05 (1.03-1.06)	**$1.48x10^{-09}$**	-
rs3130683	6p21	*C2*	T/C (0.86)	39494/72267	1.09 (1.06-1.13)	$7.87x10^{-08}$	30450/42485	1.09 (1.05-1.14)	$2.97x10^{-05}$	1.09 (1.07-1.12)	**$1.04x10^{-11}$**	-
rs11042937	11p15	*MRVI1-CTR9*	T/G (0.49)	42335/78234	1.05 (1.03-1.07)	$3.21x10^{-08}$	30533/42527	1.03 (1.00-1.05)	0.019	1.04 (1.03-1.06)	**$1.18x10^{-08}$**	Bipolar disorder and schizophrenia
rs11172113	12q13	*LRP1*	C/T (0.41)	42335/78234	1.06 (1.04-1.08)	$1.78x10^{-08}$	28503/36433	1.06 (1.03-1.08)	$1.16x10^{-06}$	1.06 (1.04-1.07)	**$9.25\ x10^{-14}$**	Migraine, Pulmonary function
rs11057830	12q24	*SCARB1*	A/G (0.15)	42331/78237	1.09 (1.06-1.11)	$3.69x10^{-10}$	20395/30592	1.07 (1.03-1.11)	0.0003	1.08 (1.06-1.10)	**$4.61x10^{-13}$**	Vitamin E level
rs1800775	16q13	*CETP*	C/A (0.51)	38810/62756	1.06 (1.04-1.08)	$2.21x10^{-08}$	22445/32148	1.03 (1.00-1.05)	0.032	1.04 (1.03-1.06)	**$9.83x10^{-09}$**	HDL Cholesterol

A1/2, allele 1/allele 2; freq, frequency of allele 1; OR, odds ratio for disease for carriers of allele 1; CI, confidence interval.

Table 2. Significant associations of CAD variants with selected cardiovascular risk factors

Locus	Locus Name	Lead Variant	Trait	Effect	p value
New Loci					
6p21	*C2*	rs3130683	T2D	1.12*	$2.7x10^{-5}$
	SCARB1	rs11057830	LDL	0.006	$2.6x10^{-5}$
			TG	0.022	$8.3x10^{-5}$
16q13	*CETP*	rs1800775	LDL	0.041	$8.5x10^{-24}$
			HDL	-0.202	$3.3x10^{-644}$
			TG	0.04	$1.3x10^{-26}$
Known Loci					
1p32	*PCSK9*	rs11206510	LDL	0.083	$2.4x10^{-53}$
1p13	*SORT1*	rs602633	LDL	0.159	$1.5x10^{-261}$
			HDL	-0.033	$3.5x10^{-14}$
2p24	*APOB*	rs515135	LDL	0.139	$1.1x10^{-178}$
2p21	*ABCG5-ABCG8*	rs6544713	LDL	0.081	$4.84x10^{-83}$
4q32	*GUCY1A3*	rs7692387	DBP	0.326	$3.4x10^{-5}$
5q31	*SLC22A4-SLC22A5*	rs273909	LDL	0.022	$2.3x10^{-5}$
2q33	*WDR12*	rs6725887	LDL	-0.026	$1.3x10^{-5}$
6q25	*LPA*	rs3798220	LDL	0.158	$6.1x10^{-11}$
		rs2048327	LDL	0.019	$1.3x10^{-6}$
7q32	*ZC3HC1*	rs11556924	DBP	NA	$1.8x10^{-5}$
			HDL	-0.018	$1.3x10^{-5}$
7q36	*NOS3*	rs3918226	SBP	0.96	1.1x10-6
			DBP	0.81	2.2x10-9
8p21	*LPL*	rs264	HDL	-0.098	$8x10^{-77}$
			TG	0.093	$2.4x10^{-84}$
8q24	*TRIB1*	rs2954029	LDL	0.056	$2.1x10^{-50}$
			HDL	-0.04	$2.7x10^{-29}$
			TG	0.076	$1x10^{-107}$
9q34	*ABO*	rs579459	LDL	0.067	$2.4x10^{-44}$
10q24	*CYP17A1-CNNM2-NT5C2*	rs12413409	SBP	1.034	$2x10^{-9}$
			DBP	0.483	$3.4x10^{-5}$
			PP	0.56	$5.7x10^{-8}$
			BMI	-0.03	$2.2x10^{-8}$
11q23	*ZNF259-APOA5-APOA1*	rs964184	LDL	0.086	$2x10^{-26}$
			HDL	-0.107	$6.1x10^{-48}$
			TG	0.234	$6.6x10^{-244}$
12q24	*SH2B3*	rs3184504	LDL	-0.027	$4.2x10^{-12}$
			HDL	-0.026	$4.1x10^{-12}$
			SBP	0.598	$2x10^{-9}$
			DBP	0.483	$8.8x10^{-6}$
			BMI	-0.131	$9.4x10^{-6}$
15q26	*FURIN-FES*	rs17514846	SBP	0.509	$1.2x10^{-5}$
17p13	*SMG6*	rs2281727	BMI	0.015	$3.64x10^{-6}$
18q21	*PMAIP1-MC4R*	rs663129	HDL	-0.026	$5.5x10^{-9}$
			BMI	0.056	$8.8x10^{-53}$
19p13	*LDLR*	rs1122608	LDL	0.074	$8.5x10^{-57}$
19q13	*APOE-APOC1*	rs2075650	LDL	0.177	$1.7x10^{-214}$
			HDL	-0.055	$9.7x10^{-26}$
			TG	0.044	$2.3x10^{-21}$
			BMI	-0.026	$1.3x10^{-8}$
		rs445925	LDL	0.363	$6.6x10^{-397}$
			HDL	-0.051	$1.9x10^{-10}$
			TG	0.101	$3.6x10^{-39}$

LDL, plasma LDL-cholesterol level; TG, plasma triglycerides level; HDL, plasma HDL-cholesterol level; T2D, type-2 diabetes; DBP, diastolic blood pressure; Effects are either absolute beta estimates of the association of the CAD risk allele on the trait (with a positive association indicating a higher value of the trait per copy of the risk allele) or log odds ratio for diabetes mellitus* per copy of the risk allele. This table only includes associations that passed Bonferroni correction. **Supplementary Table S6** shows association results for all CAD loci.

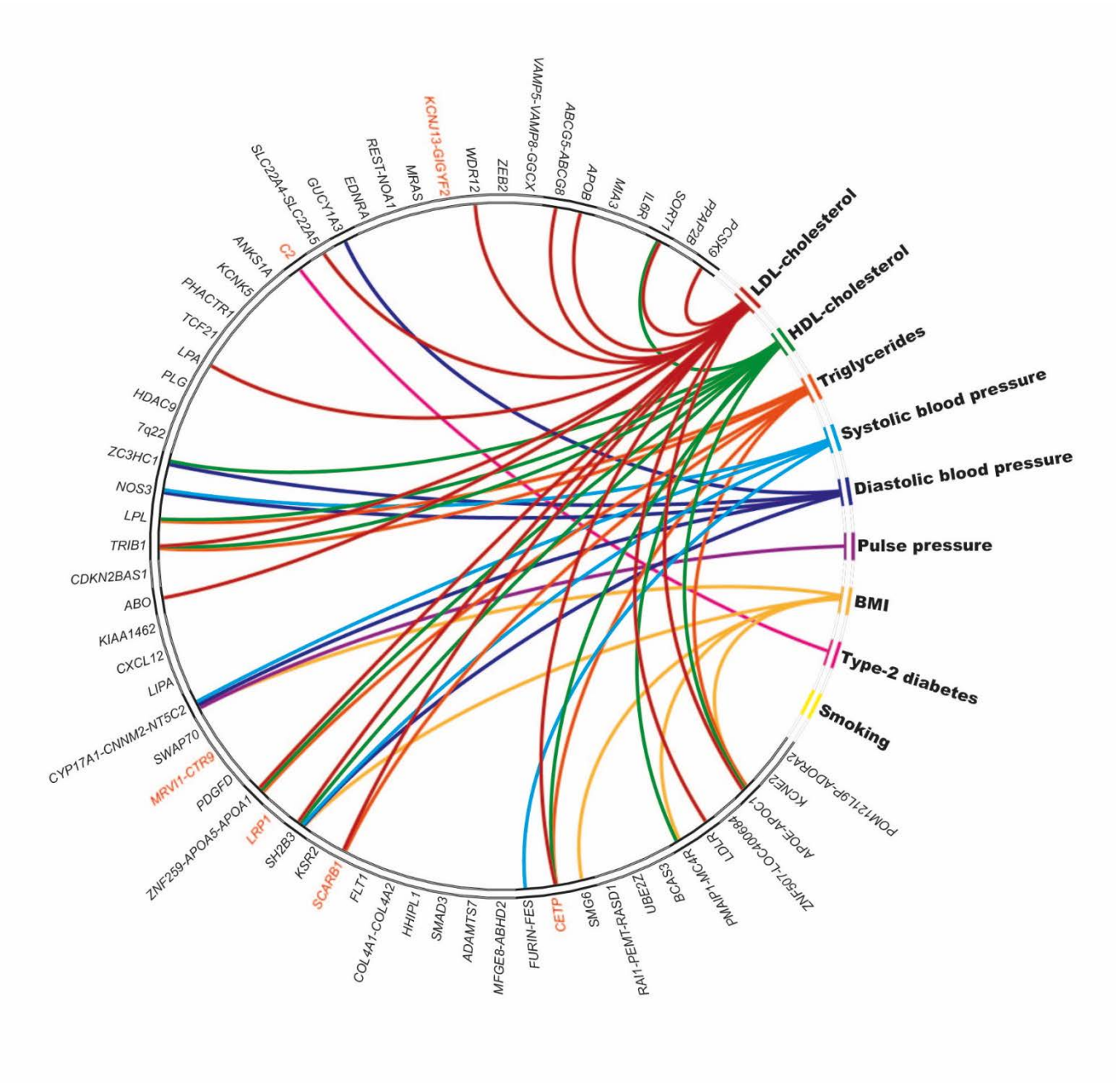
Table 3. Association of CAD loci with other diseases and traits

Locus	Locus Name	Disease or trait	Disease or trait Lead SNP	p value	Direction	r2 (CAD Lead and disease or trait Lead)
New Loci						
6p21	*C2*	Systemic lupus erythematosus (SLE)	rs3130342	$9.3x10^{-7}$	+	0.87
		Primary biliary cirrhosis	rs3134954	$1x10^{-5}$	+	0.86
		Multiple sclerosis	rs3134954	$3.2x10^{-9}$	-	0.86
11p15	*MRVI1-CTR9*	Bipolar disorder and schizophrenia	rs2018368	$1x10^{-6}$	+	0.96
12q13	*LRP1*	Migraine	rs11172113	$4.3x10^{-9}$	-	Same SNP
		Lung function (FEV1/FVC)	rs11172113	$1.2x10^{-8}$	+	Same SNP
		Cervical artery dissection	rs11172113	$3x10^{-7}$	-	Same SNP
12q24	*SCARB1*	Lp-PLA2 activity	rs11057841	$6.1x10^{-14}$	+	0.92
		Circulating vitamin E levels	rs11057830	$8.2x10^{-9}$	+	Same SNP
		Lp-PLA2 mass	rs10846744	$6.8x10^{-6}$	+	0.94
Known Loci						
1p32	*SORT1*	Lp-PLA2 activity	rs7528419	$1.3x10^{-17}$	+	0.9
		Metabolic syndrome domains (Atherogenic Dyslipidemia - PC1)	rs12740374	$8x10^{-16}$	+	Same SNP
		Lp-PLA2 mass	rs7528419	$7.1x10^{-5}$	+	0.9
1q21	*IL6R*	C-reactive protein	rs4845625	$4.2x10^{-7}$	+	Same SNP
2p21	*ABCG5-ABCG8*	Serum Phytosterol	rs4245791	$2.2x10^{-70}$	+	1
2p11	*VAMP5-VAMP8-GGCX*	Prostate cancer	rs10187424	$2.7x10^{-15}$	-	0.87
2q33	*WDR12*	Cerebral white matter hyperintensities burden	rs6705330	$5.7x10^{-5}$	+	1
3q22	*MRAS*	Coronary artery calcification	rs2306374	$2.7x10^{-5}$	+	1
4q12	*REST-NOA1*	Height	rs17081935	$6.7x10^{-17}$	+	0.95
4q31	*EDNRA*	Carotid intima media thickness	rs1878406	$7x10^{-12}$	+	Same SNP
		Intracranial aneurysm	rs6842241	$2.4x10^{-9}$	+	0.94
6p24	*PHACTR1*	Coronary artery calcification	rs9349379	$4x10^{-22}$	+	Same SNP
		Cervical artery dissection	rs9349379	$1x10^{-11}$	-	Same SNP
		Migraine	rs9349379	$5 x10^{-8}$	-	Same SNP
		Pulse wave velocity	rs7750679	$5.4x10^{-5}$	+	1
6q25	*LPA*	Lipoprotein A (Lp(a))	rs3798220	$1.6x10^{-49}$	+	Same SNP
		Colorectal cancer	rs7758229	$5.6x10^{-9}$	-	0.85
7p21	*HDAC9*	Stroke (large vessel stroke)	rs11984041	$1.9x10^{-11}$	+	1
7q22	7q22	Endometriosis	rs10953541	$3.2x10^{-5}$	+	Same SNP
8q24	*TRIB1*	Metabolic syndrome domains (Atherogenic Dyslipidemia - PC1)	rs2954021	$1.2x10^{-11}$	+	Same SNP
		Adiponectin levels	rs2954021	$1.8x10^{-5}$	+	Same SNP
		Serum creatinine	rs2954021	$2.3x10^{-5}$	+	Same SNP
9p21	*CDKN2BAS1*	Coronary artery calcification	rs1333049	$3.3x10^{-24}$	+	0.97
		Abdominal aortic aneurysm	rs2383207	$1.9x10^{-8}$	+	0.91
		Ankle brachial index	rs10757269	$2.7x10^{-9}$	+	0.9
		Stroke (large vessel stroke)	rs2383207	$2.4x10^{-6}$	+	0.91
		Intracranial aneurysm	rs10733376	$4x10^{-12}$	+	0.94
9q34	*ABO*	Alkaline phosphatase in plasma	rs579459	$3x10^{-123}$	+	Same SNP
		Activated partial thromboplastin time	rs579459	$1.7x10^{-74}$	+	Same SNP
		Soluble P-selectin	rs579459	$1.9x10^{-41}$	+	Same SNP
		Soluble E-selectin	rs579459	$1.3x10^{-29}$	+	Same SNP
		Plasma carcinoembryonic levels	rs579459	$3x10^{-21}$	+	Same SNP
		Red blood cell count	rs579459	$9.3x10^{-18}$	+	Same SNP
		Hemoglobin	rs579459	$1.4x10^{-15}$	+	Same SNP
		Hematocrit	rs579459	$7.6x10^{-14}$	+	Same SNP
		IL-6 levels	rs579459	$3.6x10^{-13}$	+	Same SNP
		Circulating galectin-3 levels	rs579459	$1.9x10^{-10}$	+	Same SNP
		Factor XIII antigen	rs579459	$2.3x10^{-8}$	+	Same SNP
		von Willebrand factor	rs651007	$1x10^{-161}$	+	1
		Venous thromboembolism	rs495828	$2x10^{-17}$	+	1
		Serum alkaline phosphatase levels	rs651007	$1x10^{-56}$	+	1
		Ferritin levels	rs651007	$1x10^{-8}$	+	1
		Soluble ICAM-1	rs507666	$3x10^{-91}$	+	0.83
10q24	*CYP17A1-CNNM2-NT5C2*	Intracranial aneurysm	rs12413409	$1.2x10^{-9}$	+	Same SNP
		Schizophrenia	rs11191580	$1.7x10^{-9}$	+	1
		Autism spectrum disorder	rs11191454	$1.4x10^{-8}$	+	1
		Parkinson's disease	rs17115100	$7.4x10^{-8}$	+	0.9
11q23	*ZNF259-APOA5-APOA1*	Metabolic syndrome domains (Atherogenic Dyslipidemia - PC2)	rs964184	$1.8x10^{-12}$	+	Same SNP
		Vitamin E levels	rs964184	$7.8x10^{-12}$	+	Same SNP
		Lp-PLA2 activity	rs964184	$8.4x10^{-11}$	+	Same SNP
		Metabolic syndrome domains (Atherogenic Dyslipidemia - PC1)	rs964184	$1.2x10^{-10}$	+	Same SNP
12q24	*SH2B3*	Selective immunoglobulin A deficiency	rs3184504	$5.6x10^{-31}$	+	Same SNP
		Type 1 diabetes	rs3184504	$2.8x10^{-27}$	+	Same SNP
		Celiac disease	rs3184504	$5.4x10^{-21}$	+	Same SNP
		Hemoglobin	rs3184504	$4.3x10^{-19}$	+	Same SNP
		Celiac disease	rs3184504	$5.4x10^{-21}$	+	Same SNP
		Blood eosinophil count	rs3184504	$6.5x10^{-19}$	+	Same SNP
		Generalized vitiligo	rs3184504	$2.5x10^{-17}$	+	0.97
		Soluble ICAM-1	rs3184504	$2.9x10^{-17}$	+	Same SNP
		Hematocrit	rs3184504	$7.9x10^{-16}$	+	Same SNP
		Hypothyroidism	rs3184504	$2.6x10^{-12}$	+	Same SNP
		Rheumatoid arthritis and celiac disease	rs3184504	$1.4x10^{-11}$	+	Same SNP
		Primary sclerosing cholangitis	rs3184504	$5.9x10^{-11}$	+	Same SNP
		Serum urate	rs3184504	$2.6x10^{-10}$	+	Same SNP
		Juvenile idiopathic arthritis	rs3184504	$2.6x10^{-9}$	+	0.9
		Lymphocyte count	rs3184504	$1.1x10^{-8}$	+	Same SNP
		Plasma Beta-2 microglobulin levels	rs3184504	$3.1x10^{-8}$	+	Same SNP
		White blood cell count	rs3184504	$6.3x10^{-6}$	+	Same SNP
		Rheumatoid arthritis	rs3184504	$6x10^{-6}$	+	Same SNP
		Tetrology of fallot	rs11065987	$4.6x10^{-8}$	+	0.91
13q34	*COL4A1-COL4A2*	Coronary artery calcification	rs3809346	$8.6x10^{-7}$	+	0.97

15q22	*SMAD3*	Crohn's disease	rs17293632	$2.7x10^{-19}$	-	0.9
		Inflammatory bowel disease	rs17293632	$6x10^{-16}$	-	0.9
		Ulcerative colitis	rs17293632	$9.5x10^{-6}$	-	0.9
		Self-reported allergy	rs17228058	$1.2x10^{-8}$	-	0.9
15q25	*ADAMTS7*	Coronary artery calcification	rs3825807	$6.5x10^{-6}$	+	Same SNP
17p13	*SMG6*	Aortic root size	rs10852932	$2.3x10^{-11}$	+	0.96
17q21	*UBE2Z*	Height	rs318095	$1.5x10^{-16}$	-	1
		Breast size (bra cup size in women)	rs12603969	$3x10^{-5}$	-	1
18q21	*PMAIP1-MC4R*	Obesity	rs17782313	$4.8x10^{-15}$	+	0.86
		Height	rs11152213	$6.9x10^{-13}$	+	0.91
		Antipsychotic drug induced weight gain	rs12967878	$3.6x10^{-7}$	+	0.9
19q13	*APOE-APOC1*	Common carotid artery intima media thickness (IMT)	rs445925	$1.7x10^{-8}$	+	Same SNP
		Lp-PLA2 activity	rs445925	$3.3x10^{-10}$	+	Same SNP
		Metabolic syndrome domains (Atherogenic Dyslipidemia - PC1)	rs445925	$1.3x10^{-35}$	+	Same SNP
		Alzheimer's disease	rs2075650	$1x10^{-295}$	+	Same SNP
		Longevity	rs2075650	$3.4x10^{-17}$	-	Same SNP
		Lp-PLA2 activity	rs2075650	$8.1x10^{-15}$	+	Same SNP
		Age-related macular degeneration	rs2075650	$8.4x10^{-8}$	-	Same SNP
		C-reactive protein	rs2075650	$4.2x10^{-8}$	+	Same SNP
		Cognitive decline	rs2075650	$2x10^{-8}$	+	Same SNP

CAD; coronary artery disease, Disease or trait lead SNP, SNP; single nucleotide polymorphism, Direction, + indicates increase in disease or trait with the CAD risk allele; NA, not available.

Figure 1. Chord diagram showing significant associations of CAD loci with selected cardiovascular risk factors



Legend to Figure 1: Only associations that passed Bonferroni correction are shown (see Table 2.) The new CAD loci are highlighted in red. Connections indicate that SNPs at respective loci associate with both CAD and the respective risk factor; they do not imply, that the risk factor causally explains the association with CAD.

PERSPECTIVES

Competencies in Medical Knowledge: Novel genetic loci affecting risk of coronary artery disease continue to be discovered. Only a third of such loci associate with traditional cardiovascular risk factors. At least half of the loci also associate with other diseases or traits (pleiotropy).

Translational outlook: Leveraging this information may identify additional mechanisms that contribute to CAD which in turn may lead to novel therapeutic targets.