Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: 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;374:1134-44. DOI: 10.1056/NEJMoa1507652

Supplementary Appendix for

Coding variation in *ANGPTL4*, *LPL*, and *SVEP1* and risk of coronary disease

Nathan O. Stitziel, M.D., Ph.D.[*,†], Kathleen E. Stirrups, Ph.D.[†], Nicholas G.D. Masca, Ph.D.[†], Jeanette Erdmann, Ph.D.[†], Paola G. Ferrario, Ph.D., Inke R. König, Ph.D., Peter E. Weeke, M.D., Ph.D., Thomas R. Webb, Ph.D., Paul L. Auer, Ph.D., Ursula M. Schick, Ph.D., Yingchang Lu, M.D., Ph.D., He Zhang, Ph.D., Marie-Pierre Dube, Ph.D., Anuj Goel, M.Sc., Martin Farrall, F.R.C.Path., Gina M. Peloso, Ph.D., Hong-Hee Won, Ph.D., Ron Do, Ph.D., Erik van Iperen, M.Sc., Stavroula Kanoni, Ph.D., Jochen Kruppa, Ph.D., Anubha Mahajan, Ph.D., Robert A. Scott, Ph.D., Christina Willenborg, Ph.D., Peter S. Braund, Ph.D., Julian C. van Capelleveen, M.D., Alex S.F. Doney, M.D., Ph.D., Louise A. Donnelly, Ph.D., Rosanna Asselta, Ph.D., Piera A. Merlini, M.D., Stefano Duga, Ph.D., Nicola Marziliano, Ph.D., Josh C. Denny, M.D., M.S., Christian M. Shaffer, B.S., Nour Eddine El-Mokhtari, M.D., Andre Franke, Ph.D., Omri Gottesman, M.D., Stefanie Heilmann, Ph.D., Christian Hengstenberg, M.D., Per Hoffmann, Ph.D., Oddgeir L. Holmen, M.D., Kristian Hveem, M.D., Ph.D., Jan-Håkan Jansson, M.D., Ph.D., Karl-Heinz Jöckel, Ph.D., Thorsten Kessler, M.D., Jennifer Kriebel, Ph.D., Karl L. Laugwitz, M.D., Eirini Marouli, Ph.D., Nicola Martinelli, M.D., Ph.D., Mark I. McCarthy, M.D., Ph.D., Natalie R. Van Zuydam, Ph.D., Christa Meisinger, M.D., M.P.H., Tõnu Esko, Ph.D., Evelin Mihailov, M.Sc., Stefan A. Escher, Ph.D., Maris Alver, M.Sc., Susanne Moebus, Ph.D., Andrew D. Morris, M.D.,Martina Müller-Nurasyid, Ph.D., Majid Nikpay, Ph.D., Oliviero Olivieri, M.D., Louis-Philippe Lemieux Perreault, Ph.D., Alaa AlQarawi, B.Sc., Neil R. Robertson, M.Sc., Karen O. Akinsanya, Ph.D., Dermot F. Reilly, Ph.D., Thomas F. Vogt, Ph.D., Wu Yin, Ph.D., Folkert W. Asselbergs, M.D., Ph.D., Charles Kooperberg, Ph.D., Rebecca D. Jackson, M.D., Eli Stahl, Ph.D., Konstantin Strauch, Ph.D., Tibor V. Varga, M.Sc., Melanie Waldenberger, Ph.D., Lingyao Zeng, M.Sc., Aldi T. Kraja, D.Sc., Ph.D., Chunyu Liu, Ph.D., Georg B. Ehret, M.D., Christopher Newton-Cheh, M.D., M.P.H., Daniel I. Chasman, Ph.D., Rajiv Chowdhury, M.D., Ph.D., Marco Ferrario, M.D., Ian Ford, Ph.D., J. Wouter Jukema, M.D., Ph.D., Frank Kee, M.D., M.Sc., Kari Kuulasmaa, Ph.D., Børge G. Nordestgaard, M.D., D.M.Sc., Markus Perola, M.D., Ph.D., Danish Saleheen, MBBS, Ph.D., Naveed Sattar, FRCP, Ph.D., Praveen Surendran, Ph.D., David Tregouet, Ph.D., Robin Young, Ph.D., Joanna M. M. Howson, Ph.D., Adam S. Butterworth, Ph.D., John Danesh, FRCP, D.Phil., Diego Ardissino, M.D., Erwin P. Bottinger, M.D., Raimund Erbel, M.D., Paul W. Franks, Ph.D., Domenico Girelli, M.D., Ph.D., Alistair S. Hall, M.D., Ph.D., G. Kees Hovingh, M.D., Ph.D., Adnan Kastrati, M.D., Wolfgang Lieb, M.D., M.Sc., Thomas Meitinger, M.D., William E. Kraus, M.D., Svati H. Shah, M.D., M.P.H., Ruth McPherson, M.D., Ph.D., Marju Orho-Melander, Ph.D., Olle Melander, M.D., Ph.D., Andres Metspalu, M.D., Ph.D., Colin N.A. Palmer, Ph.D., Annette Peters, Ph.D., Daniel J. Rader, M.D., Muredach P. Reilly, M.B., B.Ch., MSCE, Ruth J.F. Loos, Ph.D., Alex P. Reiner, M.D., M.Sc., Dan M. Roden, M.D., Jean-Claude Tardif, M.D., John R. Thompson, Ph.D., Nicholas J. Wareham, M.B., B.S., Ph.D., Hugh Watkins, M.D., Ph.D., Cristen J. Willer, Ph.D., Sekar Kathiresan, M.D.[*,†], Panos Deloukas, Ph.D.[*,†], Nilesh J Samani, M.D., FRCP[*,†], Heribert Schunkert, M.D.[*,†]

*Address correspondence to:

Nathan Stitziel, M.D., Ph.D.
Washington University School of Medicine
660 S. Euclid Ave
Campus Box 8086
Saint Louis, MO 63110
nstitziel@wustl.edu

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

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

Nilesh J. Samani, M.D.
Department of Cardiovascular Sciences, University of Leicester
BHF Cardiovascular Research Centre
Glenfield Hospital
Groby Rd.
Leicester, LE3 9QP
United Kingdom
njs@le.ac.uk

Heribert Schunkert, M.D.
Deutsches Herzzentrum München
Technische Universität München
Deutsches Zentrum für Herz-Kreislauf-Forschung (DZHK), Munich Heart Alliance
Lazarettstraß 36
80636 München
Germany
schunkert@dhm.mhn.de

†Contributed equally

Table of Contents

Additional acknowledgements

Additional investigator from the Montreal Heart Institute: Sylvie Provost

Additional investigators of the PROSPER study include: Stella Trompet, Anton De Craen, David Scott, and Brendan Buckley.

Additional investigators of the WOSCOPS study include: Chris Packard, Muriel Caslake, and Sandosh Padmanabhan.

Additional investigators of the CCHS/CGPS/CIHDS studies include: Sune F. Nielsen, Gorm B. Jensen, Anne Tybjaerg-Hansen, and Lia E Bang.

Additional investigators of the PROMIS study include: Asif Rasheed and Philippe Frossard.

Additional investigators of the BRAVE study include: Emanuele di Angelantonio, Dewan S. Alam, Abdullah Al Shafi Majumder, and Ismail Ibrahim Fakir.

Sites and key personnel of contributing MORGAM Centres:

Finland: FINRISK, National Institute for Health and Welfare, Helsinki: V. Salomaa (principal investigator), A. Juolevi, E. Vartiainen, P. Jousilahti; ATBC, National Institute for Health and Welfare, Helsinki: J. Virtamo (principal investigator), H. Kilpeläinen; MORGAM Central Laboratory, National Institute for Health and Welfare, Helsinki: M. Perola (responsible person), P. Laiho; MORGAM Data Centre, National Institute for Health and Welfare, Helsinki: K. Kuulasmaa (responsible person), Z. Cepaitis, A. Haukijärvi, B. Joseph, J. Karvanen, J. Kontto, S. Kulathinal, M. Niemelä, T. Palosaari, O. Saarela; MORGAM Central Laboratory, National Institute for Health and Welfare, Helsinki: M. Perola (responsible person), P. Laiho, M. Sauramo. The ATBC Study was supported by US Public Health Service contracts N01-CN-45165, N01-RC-45035 and N01-RC-37004 from the National Cancer Institute.

France: National Coordinating Centre and PRIME/Lille, Department of Epidemiology and Public Health, INSERM U744-Université Lille Nord de France – Institut Pasteur de Lille: P. Amouyel (principal investigator), M. Montaye, B. Lemaire, S. Beauchant, D. Cottel, C. Graux, N. Marecaux, C. Steclebout, S. Szeremeta Former National Coordinating Centre, National Institute of Health and Medical Research (U258), Paris: P. Ducimetière (national coordinator), A. Bingham; PRIME/Strasbourg, Department of Epidemiology and Public Health, EA 3430, University of Strasbourg, Faculty of Medicine, Strasbourg: D. Arveiler (principal investigator), B. Haas, A. Wagner; PRIME/Toulouse, UMR INSERM 1027; and Department of Epidemiology, Toulouse University School of Medicine, Universite Paul Sabatier, Toulouse: J. Ferrières (principal investigator), J-B. Ruidavets, V. Bongard, D. Deckers, C. Saulet, S. Barrere; MORGAM Laboratory, INSERM UMR_S 1166, Paris: F. Cambien (responsible person), L. Tiret, DA. Tregouet. INSERM and InVS are acknowledged for their support.

Germany: Augsburg, Helmholtz Zentrum München, German Research Centre for Environmental Health, Institute of Epidemiology, Neuherberg, Germany: A. Peters (principal investigator), A. Döring (former principal investigator), E. Wichmann, M. Müller-Nurasyid; MORGAM Biomarker Laboratory, Universitätsklinikum Hamburg-Eppendorf, Hamburg, Germany: S Blankenberg (responsible person), T. Zeller, S. Schnella.

Italy: EPIMED Research Center Department of Clinical and Experimental Medicine. University of Insubria, Varese: M. Ferrario (principal investigator), G. Veronesi. Research Centre on Public Health, University of Milano-Bicocca, Monza, Italy, Giancarlo Cesana. This study was supported by the Health Administration of Regione Lombardia [grant numbers 9783/1986, 41795/1993, 31737/1997,17155/2004 and 10800/2009], for the baseline examinations and the follow-up. Paolo Brambilla and Stefano Signorini, Laboratory Medicine, Hospital of Desio are thanked for their support.

United Kingdom: PRIME/Belfast, Queen's University Belfast, Belfast, Northern Ireland: F. Kee (principal investigator) A. Evans (former principal investigator), J. Yarnell, E. Gardner; Former MORGAM Coordinating Centre, Queen's University Belfast, Belfast, Northern Ireland: A. Evans (MORGAM coordinator), S. Cashman, F Kee. UKCRC are acknowledged for their support.

MORGAM Management Group: K. Kuulasmaa (chair, Helsinki, Finland), S. Blankenberg (coordinator of Biomarker SubStudy, Hamburg, Germany), A. Evans (former chair, Belfast, UK), M. Ferrario (Varese, Italy), F. Kee (Belfast, UK), A. Palotie (Helsinki, Finland), M. Perola (Coordinator of Genetic SubStudy, Helsinki, Finland), A. Peters (Neuherberg, Germany), V. Salomaa (Helsinki, Finland), D. Tregouet (Paris, France), H. Tunstall-Pedoe (Dundee, Scotland); Previous members: K. Asplund (Stockholm, Sweden), F. Cambien/L. Tiret (Paris, France), L. Peltonen (Helsinki, Finland), D. Shields (Dublin, Ireland), B. Stegmayr (Umeå, Sweden) , P.G. Wiklund (Umeå, Sweden).

The Exome Sequencing Project of the U.S. National Heart, Lung, and Blood Institute supported genotyping (RC2HL102925 to S Gabriel and D Altshuler). The study was also supported by the German Federal Ministry of Education and Research (BMBF) in the context of the e:Med program (e:AtheroSysMed) and the FP7 European Union project CVgenes@target (261123). Further grants were received by the Fondation Leducq (CADgenomics: Understanding CAD Genes, 12CVD02). This work has been supported by the “Programma di ricerca Regione-Università, Regione Emilia-Romagna, bando Ricerca Innovativa 2010-2012 to Dr. Diego Ardissino, Cardiovascular genetics: from bench to bedside - Genomic & transcriptomic of ischemic heart disease - CUP E35E09000880002”. The PopGen 2.0 network is supported by a grant from the German Ministry for Education and Research (01EY1103). Recruitment of the BHF-FHS Study was funded by the British Heart Foundation (BHF) with additional support from the Medical Research Council. Genotyping of the BHF-FHS controls was funded by the Wellcome Trust (through the Wellcome Trust Case Control Consortium, WTCCC) and

the cases by the WTCCC, the National Institute for Health Research (NIHR) and the BHF. Data were obtained from Vanderbilt University Medical Center's BioVU which is supported by institutional funding and by the Vanderbilt CTSA grant UL1 TR000445 from NCATS/NIH. This work was also in part supported by NIH grants U19 HL65962, and R01 HL092217. The Verona Heart Study is supported by the CariVerona Foundation. PROCARDIS was supported by the European Community Sixth Framework Program (LSHM-CT- 2007-037273), AstraZeneca, the British Heart Foundation, the Swedish Research Council, the Knut and Alice Wallenberg Foundation, the Swedish Heart-Lung Foundation, the Torsten and Ragnar Söderberg Foundation, the Strategic Cardiovascular Program of Karolinska Institutet and Stockholm County Council, the Foundation for Strategic Research and the Stockholm County Council (560283). EGCUT received financing from European Regional Development Fund, road-map grant no.3.2.0304.11-0312 and grant "Center of Excellence in Genomics" (EXCEGEN). EGCUT studies were covered also by targeted financing from Estonian Government (IUT24-6, IUT20-60) and CTG grant (SP1GVARENG) from Development Fund of the University of Tartu. GoDARTS acknowledges the support of the Health Informatics Centre, University of Dundee for managing and supplying the anonymised data and NHS Tayside, the original data owner. We are grateful to all the participants who enrolled in the GoDARTS study, to the general practitioners, to the Scottish School of Primary Care for their help in recruiting the participants, and to the whole team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists, and nurses. The GoDARTS study is supported by the Wellcome Trust (Awards 072960, 084726 and 104970). The 1958 Birth Cohort sample collection was funded by the Medical Research Council grant G0000934 and the Wellcome Trust grant 068545/Z/02 and genotyping was funded by the Wellcome Trust. Jansson J-H was responsible for the identification of MI cases in the FIA3 study. The FIA3 study was supported in part by a grant from the Swedish Heart-Lund Foundation (grant no. 2020389 to Franks PW). Analysis was in part funded by BHF Programme Grant RG/14/5/30893 to P Deloukas. The KORA research platform (KORA, Cooperative Research in the Region of Augsburg) was initiated and financed by the Helmholtz Zentrum München - German Research Center for Environmental Health, which is funded by the German Federal Ministry of Education and Research and by the State of Bavaria. Furthermore, KORA research was supported within the Munich Center of Health Sciences (MC Health), Ludwig-Maximilians-Universität, as part of LMUinnovativ. We thank the Heinz Nixdorf Foundation (Germany), the Ministerium für Innovation, Wissenschaft und Forschung des Landes Nordrhein-Westfalen and the Faculty of Medicine University Duisburg-Essen for the generous support of the Heinz Nixdorf Recall Study. The BRAVE study genetic epidemiology working group is a collaboration between the Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, UK, the Centre for Control of Chronic Diseases, icddr,b, Dhaka, Bangladesh and the National Institute of Cardiovascular Diseases, Dhaka, Bangladesh. CHD case ascertainment and validation, genotyping, and clinical chemistry assays in EPIC-CVD were principally supported by grants awarded to the University of Cambridge from the EU Framework Programme 7 (HEALTH-F2-2012-279233), the UK Medical Research Council (G0800270) and British Heart Foundation (SP/09/002) , and the European Research Council (268834). We thank all EPIC participants and staff for

their contribution to the study, the laboratory teams at the Medical Research Council Epidemiology Unit for sample management and Cambridge Genomic Services for genotyping, Sarah Spackman for data management, and the team at the EPIC-CVD Coordinating Centre for study coordination and administration. Field-work, genotyping, and standard clinical chemistry assays in PROMIS were principally supported by grants awarded to the University of Cambridge from the British Heart Foundation, UK Medical Research Council, Wellcome Trust, EU Framework 6–funded Bloodomics Integrated Project, Pfizer, Novartis, and Merck. The MORGAM Project received funding during the work from European Union FP 7 projects CHANCES (HEALTH-F3-2010-242244) and BiomarCaRE (278913). This has supported central coordination and part of the activities of the MORGAM Data Centre, at THL in Helsinki, Finland. MORGAM Participating Centres are funded by regional and national governments, research councils, charities, and other local sources. The Ottawa Heart Genomics Study was funded by Canadian Institutes of Health Research # MOP-2380941, #MOP82810, #MOP77682, Canada Foundation for Innovation #11966, Heart & Stroke Foundation of Canada T7268. The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° HEALTH-F2-2009-223004. The WHI program is funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, U.S. Department of Health and Human Services through contracts HHSN268201100046C, HHSN268201100001C, HHSN268201100002C, HHSN268201100003C, HHSN268201100004C, and HHSN271201100004C. The authors thank the WHI investigators and staff for their dedication, and the study participants for making the program possible. A full listing of WHI investigators can be found at: http://www.whi.org/researchers/Documents%20%20Write%20a%20Paper/WHI%20Investigator%20Short%20List.pdf. Exome-chip data and analysis were supported through the Women's Health Initiative Sequencing Project (NHLBI RC2 HL-102924), the Genetics and Epidemiology of Colorectal Cancer Consortium (NCI CA137088), the Genomics and Randomized Trials Network (NHGRI U01-HG005152), and an NCI training grant (R25CA094880). Malmö Diet and Cancer Study: Supported by the Swedish Research Council, the Swedish Heart and Lung Foundation, ERC-Stg-282255, the Novo Nordic Foundation, the Swedish Diabetes Foundation, and the Påhlsson Foundation, and by equipment grants from the Knut and Alice Wallenberg Foundation, the Region Skåne, Skåne University Hospital, and the Linneus Foundation for the Lund University Diabetes Center. The work was funded as part of the DZHK and the eAtheroSysMed project BMBF 1ZX1313C. This work was also supported by grants from the European Union (CVgenes@target), the Leducq Foundation (CADgenomics), and the Bundesministerium für Bildung und Forschung (e:AtheroSysMed) to Jeanette Erdmann, Nilesh Samani, Hugh Watkins and/or Heribert Schunkert.

Supplementary Methods

Genotyping

Samples from the ATVB, Duke, OHS, PAS-AMC, PennCath, PROCARDIS, and VHS studies were genotyped on the Illumina HumanExome BeadChip v1.0 at the Broad Institute according to the manufacturer's recommended protocol. Samples from the MDC study were genotyped on the Illumina OmniExome array according to the manufacturer's recommended protocol. Genotypes were assigned using GenomeStudio v2010.3 module version 1.8.4 along with the custom cluster file StanCtrExChp_CEPH.egt. These were then supplemented with the zCall algorithm to enhance the accuracy of rare variant genotypes[1].

Samples from the BHF-FHS, FIA3, EPIC, and GoDARTS studies were genotyped on the Illumina HumanExome Beadchip v1.0 at the Wellcome Trust Sanger Centre, UK, according to the manufacturer's recommended protocol. Genotypes were assigned using GenCall, the default clustering algorithm within GenomeStudio v2011.1, module version 1.9.4, using the cluster file HumanExome_12v1_A.egt. GenCall data were then subjected to QC, before the post-processing zCall algorithm was used to enhance the accuracy of rare and infrequent SNP genotypes.

Samples from BioVU were genotyped on the Illumina HumanExome BeadChip v1.0 at Vanderbilt University according to the manufacturer's recommended protocol. Genotypes were assigned using GenomeStudio v2010.2 genotyping module version 1.7.4 along with the custom cluster file HumanExome-12v1.egt.

Samples from GerMIFS3 and GerMIFS4, PopGen cases, and Munich-MI were genotyped at the Helmholtz Zentrum München, Germany. Samples from HNR were

genotyped at the Forschungszentrum Life & Brain, Department of Genomics, Bonn, Germany and samples from PopGen controls were genotyped at the Institute of Clinical Molecular Biology (IKMB), Kiel, Germany, respectively. All genotyping was done with the Illumina HumanExome v1.0 array according to the manufacturer's protocol. The analysis was done with the GenomeStudio V2011.1 software and the Genotyping module version 1.9.4 using the original Illumina cluster and manifest files (HumanExome-12v1_A.egt and HumanExome-12v1_A.bpm). The GenCall score cutoff was 0.15 as recommended by Illumina. The Genotypes were exported using the Report Wizard and post-processed with zCall to add rare variant calls that were otherwise missed by GenomeStudio.

Samples from the EGCUT study were genotyped on the Illumina HumanExome BeadChip v1.1 at the Estonian Genome Centre, University of Tartu, Estonia and at the Broad according to the manufacturer's recommended protocol. Genotypes were assigned using first GenomeStudio GenomeStudio v2011.1 module version 1.9.4 and then the zCall algorithm[1].

Samples from the HUNT study were genotyped using the iSelect HumanExome BeadChip V1.0 and the Infinium HD ultra protocol at the Norwegian University of Science and Technology, Norway. Each 96-well plate included both case and control individuals in random order and one sample of reference DNA that was present on every plate. Genotypes were assigned using GenomeStudio V2011.1 followed by zCall version 2.2.

Samples from the Bio*Me* Biobank were genotyped on the Illumina HumanOmniExpressExome array and Illumina HumanExome BeadChip v1.0 at the

Mount Sinai Medical Center according to the manufacturer's recommended protocol. Illumina's Genome Studio was used to call the raw genotyped data, which was subsequently updated with zCall that applies stringent criteria to remove samples based on call rate (< 98%), heterozygosity (>1% or <1%), and gender discordance in addition to markers based on call rate (<95%) and Hardy-Weinberg equilibrium ($P < 10^{-4}$). We first called the genotypes of the 12,726 participants, whose genotype cluster file was used to call the genotypes of the remaining 2,867 participants. A total of 13,710 individuals and 239,035 markers passed these quality control criteria. For the current analysis, 704 CAD cases and 1,729 controls with European American ancestry from the Bio*Me* biobank were analyzed.

Samples from the MHI study were genotyped on the Illumina HumanExome BeadChip v1.1 at the Beaulieu-Saucier Pharmacogenomic Centre according to the manufacturer's recommended protocol. Genotypes were assigned using GenomeStudio v2011.1 module version 1.9.4 along with the custom cluster file HumanExome-12v1.egt.

Samples from the WHI study were genotyped at the Broad Institute or the Translational Genomics Research Institute using the Illumina HumanExome v1.0 SNP array. Genotypes were assigned using GenomeStudio v2010.3. WHI genotypes from both genotyping centers were then merged into a master-file and quality control procedures were performed on this master-file using the PLINK and R47 computing platforms as described below.

Samples from the replication cohorts were genotyped in batches at the Herlev Hospital in Copenhagen (CCHS, CGPS and CIHDS) or Cambridge Genomic Services (BRAVE, EPIC-CVD, MORGAM, PROMIS, PROSPER and WOSCOPS) on

customised versions of the Illumina HumanExome v1.1 SNP array. Genotype calling was performed centrally for all batches at the University of Cambridge using optiCall[2] (0.7.0), followed by zCall for variants with minor allele frequency (MAF) <5%.

Quality control procedures

In the discovery study, various quality control filters were implemented to remove low quality samples and variants. Sample QC was performed on genotypes assigned before the zCall algorithm application. Samples were excluded if they met any of the following criteria: poor concordance with previous genotyping array; missing ≥ 5% genotypes; statistical outliers for heterozygosity; discordance between inferred and reported gender; duplicated samples; unexpected first or second degree related samples; or statistical outliers in principal components analysis. Using a set of common (minor allele frequency > 5%) independent (linkage disequilibrium pruned) markers, we identified samples sharing a high proportion of genotypes identical by descent (PI_HAT > 0.2) who were not known to be related and removed these from the analysis. From the samples that passed quality control, variants were removed if they met any of the following criteria: missing pre-zCall genotypes > 2% of cases or > 2% of controls; missing zCall genotypes in >1% of cases or >1% of controls; Hardy-Weinberg equilibrium (HWE) $P < 1x10^{-5}$ in cases or controls in either pre-zCall or zCall genotypes when available. These procedures were outlined in a centrally developed quality control and analysis protocol.

For the replication study, samples with extreme intensity values, and outlying plates or arrays were removed prior to all genotype calling. Samples with call rates more

than 3 standard deviations below the mean were removed prior to post-processing optiCall calls with zCall. Within each batch, variants were removed if variant call rate < 0.97; HWE $P < 1x10^{-6}$ for common variants or HWE $P < 1x10^{-15}$ for variants with MAF<0.05. Variants within each genotyping batch were aligned to human genome reference sequence plus strand and the standardized files were then used for sample QC. Samples were excluded from each batch/study if sample heterozygosity > ±3 standard deviations from the mean heterogeneity or sample call rate >3 standard deviations from the mean call rate. Duplicates within each batch and ancestral outliers identified by PCA were removed. Samples and variants that failed QC were removed from individual batches. Where studies were analyzed in multiple batches, the batches were combined and any variants out of HWE across the study as a whole were also removed.

Follow-up ANGPTL4 sequencing

We sequenced the seven exons of *ANGPTL4* using next-generation sequencing as previously described[3]. In brief, 4,865 cases with early CAD along with 4,866 CAD-free controls underwent exome sequencing at the Broad Institute. First, for each sample we used 3μg of genomic DNA to perform library construction and in-solution hybrid selection to target 33Mb of genomic sequence. The resulting exome-enriched DNA was sequenced on either Genome Analyzer II using v3 and v4 Sequencing-by-Synthesis Kits, then analyzed using RTA v1.7.48 or on HiSeq 2,000 using HiSeq 2,000 v2 Sequencing-by-Synthesis Kits, then analyzed using RTA v1.10.15. Sequencing was performed using 76 cycle paired-end runs. Sequencing was considered complete when ≥ 80% of targeted bases were covered with ≥ 20 sequencing reads. Raw sequence reads were aligned to the

human reference genome (HG19) using the Burroughs-Wheeler Alignment tool[4] in paired-end mode. Duplicate reads and reads aligned outside of the exome target were removed. The Genome Analysis ToolKit[5] (GATK) was then used to locally realign reads, recalibrate base qualities, identify and genotype single nucleotide variants (SNVs) and short insertion and deletion events (indels), and recalibrate the resulting variant quality scores. SnpEff[6] was used to predict the functional consequences of the identified variants.

We defined null mutations in *ANGPTL4* as single nucleotide variants leading to the introduction of a stop codon (nonsense) or occurring within two base pairs of an exon/intron boundary (splice-site), or insertions and deletions of DNA predicted to alter the open reading frame of the protein and introduce a premature stop codon (frameshift). *ANGPTL4* null mutations were annotated based on the cDNA reference sequence for ANGPTL4 (NM_139314.2) with the ATG initiation codon, encoding methionine, numbered as residue 1 or p.M1.

Statistical analysis

In discovery samples that passed quality control procedures, we performed individual tests for association between QC-passed variants and CAD within each study separately. For variants that were polymorphic in cases and controls, we performed logistic regression with CAD as the dependent variable and genotype (coded as 0, 1, or 2 copies of the effect allele) as the independent variable with the first ten principal components of ancestry as covariates. We combined evidence across individual studies in the discovery phase using an inverse-variance weighted fixed-effects meta-analysis. We annotated the functional effect of each variant using the Genome Variation Server

(http://gvs.gs.washington.edu/GVS138/). We restricted our analysis to autosomal variants with a minor allele frequency of ≥ 0.1% across the 120,575 samples in the discovery study. In loci with previously identified low-frequency CAD variants, we performed conditional association testing using the GCTA v1.24 software as previously described[7], using individual genotypes from 15,011 unrelated individuals from the ATVB, MDC, OHS, PAS-AMC, PennCath, and PROCARDIS studies to estimate linkage disequilibrium patterns between markers on the array.

Outside of known CAD loci, we defined suggestive association with CAD as a meta-analysis *P* value ≤ $1x10^{-4}$. For variants with suggestive association, we performed association analysis in the replication cohorts. In each replication study individually, association with CAD was tested using a linear mixed model using fixed effects of genotype and principal components and a kinship matrix as random effects. Results were combined within ancestry groups and then across all studies using inverse-variance weighted fixed effects meta-analyses. We defined significant novel associations as those nominally significant ($P < 0.05$) in the replication study and with overall (discovery and replication combined) $P < 7.7x10^{-8}$ (a Bonferroni-corrected threshold accounting for 54,003 markers with MAF > 0.01% being tested initially along with 12 replication tests).

To test for association between novel variants and plasma lipids, we first generated score statistics from each cohort listed in Table S4 using raremetalwork or rvtests. We then meta-analyzed the genetic associations centrally using the R-package rareMETALs (version 6.0) to test the association between significantly associated low-frequency variants with LDL cholesterol, high-density lipoprotein (HDL) cholesterol, or the natural logarithm of triglycerides (TG) using covariates of age, gender, and principal

components of ancestry. A linear mixed model was used to test the association between significantly associated low-frequency variants with systolic blood pressure (SBP) and diastolic blood pressure (DBP) in the CHARGE+ consortium (Table S5) using fixed effects of genotype, principal component of ancestry, and study-specific covariates (these included age, age-squared, sex, and body mass index), along with random effects to account for relatedness and ancestry via a kinship matrix. To account for the effect of lipid-lowering and anti-hypertensive medications, we increased the measured value of LDL by 30% and increased the measured values of SBP and DBP by 15 mmHg and 10mmHg, respectively, for those taking such medications.

We used linear regression to test the association between *ANGPTL4* null alleles and plasma lipids in models where the outcome was specified as either low-density lipoprotein cholesterol, high-density lipoprotein cholesterol, or the natural logarithm of triglyceride concentration, the independent variable was the presence or absence of any *ANGPTL4* null allele, and covariates included age, sex, and an indicator variable for study. We accounted for the effect of lipid lowering therapy in 714 individuals known to be taking such medications by increasing the observed LDL value by 30%. We calculated the statistical significance of the association between *ANGPTL4* null alleles and risk for CAD using 100,000 study-stratified permutations of case-control phenotypes.

Coverage and power analysis

To estimate the coverage of the exome array, we evaluated missense variation observed in 7,394 exomes of European ancestry that were sequenced at the Broad Institute as part of a different study and did not contribute to the design of the exome

array. We compared the chromosomal position and alternate allele of variants with MAF between 0.1% and 5% observed in the exome sequences and the content available on the array. We observed that about 82% of non-synonymous variants with a MAF between 0.1% and 5% were present on the exome array (Figure S1).

We used the Genetic Analysis Package library (v1.1-10) in R to estimate statistical power at various combinations of MAF and genotypic effect sizes. In our discovery study we had 80% power to detect alleles with frequency > 0.1% conferring a two-fold increased risk for disease at an alpha level accounting for multiple hypothesis testing (Figure S2). Similarly, we had 80% power to discover 0.5% alleles associated with 50% increased risk (or alternatively 35% decreased risk), 2% alleles associated with 25% increased risk (or alternatively 20% decreased risk), and 5% alleles associated with 15% increased (or decreased) risk of coronary disease (Figure S2).

Figure S1. Estimated coverage of European ancestry variation by the exome array. Coverage estimates were obtained by comparing variation observed in 7,394 European ancestry exome sequences with the content present on the Illumina HumanExome BeadChip v1.0. A locally-weighted polynomial regression was used to calculate a continuous estimate of coverage according to minor allele frequency (blue line) along with 95% confidence intervals (shaded area).

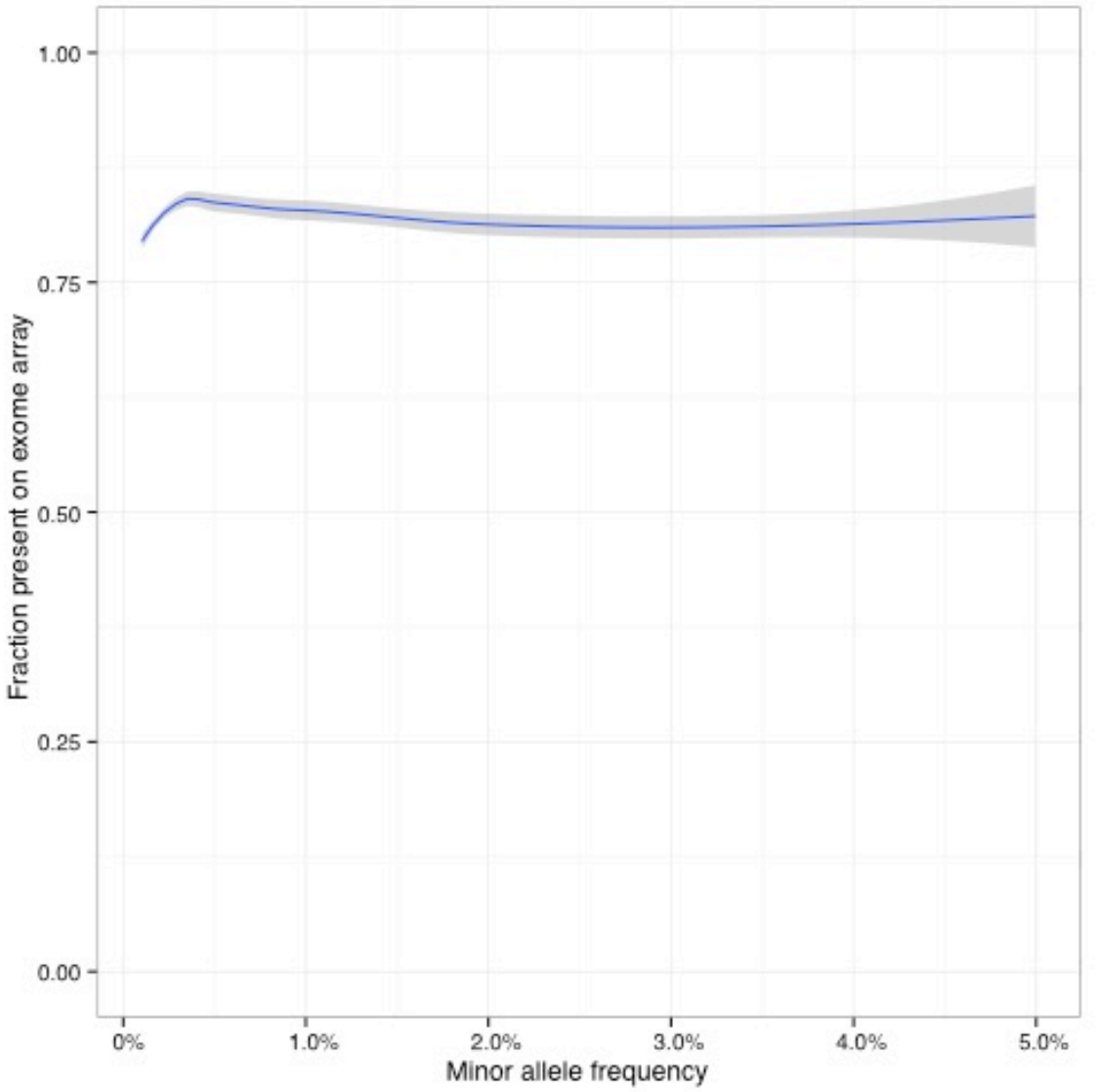

Figure S2. Statistical power for detecting a significant association in the discovery study. Lines corresponding to 80% power for detecting an association at our pre-specified level of significance ($P<8.8\text{x}10^{-7}$) are plotted for combinations of minor allele frequency (x-axis) and genotypic odds ratio (y-axis) assuming an additive genetic model. We used the number of cases and controls genotyped in the discovery study and assumed 5% disease prevalence.

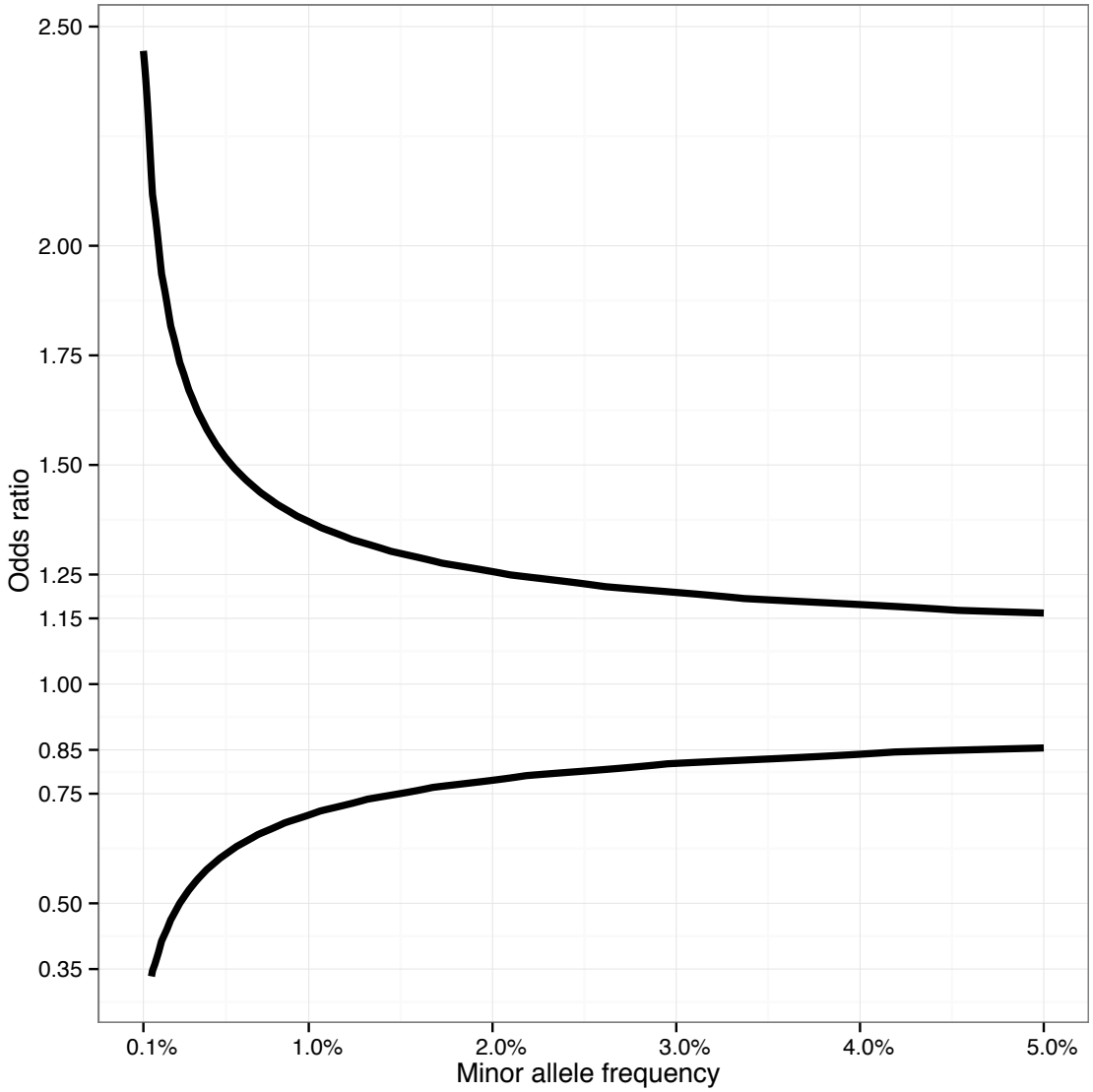

Table S1. Sources of cases and controls in the discovery study

Study	Design	Case definition	Control definition	N Cases	N Controls	Reference
ATVB	Case-control	MI in men or women ≤ 45 years of age	No history of thromboembolic disease	1,428	1,069	[8]
BHF-FHS	Case-control	CAD cases were recruited from the British Heart Foundation Family Heart Study and supplemented by additional cases from WTCCC-CAD2	Controls were selected from the UK 1958 Birth Cohort	2,833	5,912	[9,10]
BioVU	Case-control	Cases with MI or CAD were ascertained from the Vanderbilt University Medical Center Biorepository by searching the electronic medical record for ≥ 2 instances of ICD-9 codes 410.x – 414.x	Controls were individuals from the Vanderbilt University Biorepository who did not have any record of ICD-9 codes 410.x – 414.x	4,587	16,556	[11]
Duke	Case-control	MI or coronary stenosis ≥ 50%	Controls were > 50 years old without coronary stenosis > 30% and without history of MI, coronary artery bypass grafting, percutaneous coronary intervention, or heart transplant	660	515	[12]
EPIC CAD	Nested case-control	The EPIC (European Prospective Study into Cancer and Nutrition) study sub-cohorts from the UK were used. Subjects were collected in collaboration with general practitioners, mainly in Cambridgeshire and Norfolk. Cases were individuals who developed fatal or non-fatal CAD during an average follow-up of 11 years ending June 2006. Participants were identified if they had a hospital admission and/or died with CAD as the underlying cause. CAD was defined as cause of death codes ICD-9 410-414 or ICD-10 I20-I25, and hospital discharge codes ICD-10 I20.0, I21, I22, or I23 according to the International Classification of Diseases, 9th and 10th revisions, respectively.	Controls were study participants who remained free of any cardiovascular disease during follow-up (defined as ICD-9 401-448 and ICD-10 I10-I79)	1,386	7,037	[13]

FIA3	Nested case-control	Cases of MI occurring in participants from Vasterbotten Intervention Program (VIP), WHO's Multinational Monitoring of Trends and Determinants in Cardiovascular Disease (MONICA) study in northern Sweden and the Mammography Screening Project (MSP) in Vasterbotten	Individuals free of MI from VIP and MSP	2,473	2,047	[14,15]
GoDARTS CAD	Case-control	The GoDARTS (Genetics of Diabetes Audit and Research in Tayside Scotland) study is a joint initiative of the Department of Medicine and the Medicines Monitoring Unit (MEMO) at the University of Dundee, the diabetes units at three Tayside healthcare trusts (Ninewells Hospital and Medical School, Dundee; Perth Royal Infirmary; and Stracathro Hospital, Brechin), and a large group of Tayside general practitioners with an interest in diabetes care. Cases were first-ever CAD event, defined as fatal and non-fatal myocardial infarction, unstable angina, or coronary revascularization.	Controls were free of CAD, stroke, and peripheral vascular disease	1,568	2,772	[16]
EGCUT	Nested case-control	CAD or MI cases were ascertained from the Estonian Biobank (Estonian Genome Center at the University of Tartu) using the medical history and current health status that is recorded according to ICD-10 codes (CAD defined with ICD-10 I20-I25).	Controls were selected from the Estonian Biobank (Estonian Genome Center at the University of Tartu) who did not have any record of cardiovascular diseases (ICD-10 I10-I79).	392	777	[17]
German CAD North	Case-control	The German North cohort includes individuals from GerMIFS4, PopGen, and HNR with MI or CAD.	Controls were derived from population-based studies in Germany.	4,464	2,886	[18-20]
German CAD South	Case-control	The German South cohort includes samples from GerMIFS3 and Munich-MI with MI or CAD.	Controls were derived from population-based studies in Germany.	5,255	2,921	[21,22]

HUNT	Case-control	MI Cases were retrospectively identified as HUNT 2 and HUNT 3 participants diagnosed with acute MI (ICD-10 I21 or ICD-9 410) in the medical departments at the two local hospitals in Nord-Trøndelag County from December 1987 to June 2011.	Controls were selected among HUNT 2 and HUNT 3 participants with available DNA (N = 70,300) after excluding individuals with the following hospital diagnosed or self-reported conditions in themselves or known 1st and/or 2nd degree family members: MI, angina, heart failure, stroke, aortic aneurysm, atherosclerosis, intermittent claudication, and registered percutaneous coronary angioplasty procedures or bypass surgery.	2,351	2,348	[23]
Bio*Me* Biobank	Case-control	CAD cases were ascertained from the Bio*Me* Biobank using the electronic health record with ICD9 codes 410.xx to 414.xx and abnormal stress test or abnormal coronary angiography	Controls were individuals from the Bio*Me* Biobank who did not meet the criteria for cases	704	1,729	NIH dbGaP Study Accession phs000388.v1.p1
MDC	Prospective cohort	Prevalent and incident nonfatal or fatal MI	Participants free of CHD at baseline and during follow-up	2,283	4,511	[24]
MHI	Case-control	Cases were ascertained from the Montreal Heart Institute Biobank. CAD was defined as the presence of MI, percutaneous coronary intervention, or coronary artery bypass grafting	Controls were individuals from the Montreal Heart Institute Biobank who were free of history of MI, percutaneous coronary intervention, or coronary artery bypass grafting	3,990	6,585	[25,26]
OHS	Case-control	Cases had angiographically confirmed coronary artery disease (>1 coronary artery with >50% stenosis) and did not have type 2 diabetes; ≤ 50 years old for males and ≤ 50 years old for females	Asymptomatic males > 65, females > 70	1,024	2,267	[27]

PAS-AMC	Case-control	Symptomatic CAD before 51 years of age, defined as MI, coronary revascularization, or evidence of at least 70% stenosis in a major epicardial coronary artery	More than 95% of the controls are from the same region as cases	728	808	[28]
PennCath	Case-control	Cases had angiographically confirmed coronary artery disease (>1 coronary artery with 50% stenosis); ≤ 55 years old for males and ≤ 60 years old for females	Normal coronary angiography in men > 40 years old and women > 45 years old	683	156	[29]
PROCARDIS	Case-control	Symptomatic CAD before age 66. CAD was defined as clinically documented evidence of myocardial infarction, coronary artery bypass grafting, acute coronary syndrome, coronary angioplasty, or stable angina	No personal or sibling history of CAD before age 66	2,490	2,220	[30]
VHS	Case-control	Documented MI, coronary artery bypass grafting, CAD (by angiography) in males ≤ 45 years old and females ≤ 50 years old	Normal coronary angiography in males > 60 years old or females > 65 years old.	176	164	[31]
WHI	Prospective cohort	Cases were individuals from the Women's Health Initiative who had incident MI, coronary revascularization, hospitalized angina or death due to coronary disease	Participants free of CHD on follow-up	2,860	14,960	[32]
Discovery study total				**42,335**	**78,240**	

ATVB: Italian Atherosclerosis, Thrombosis, and Vascular Biology Study; BHF-FHS: British Heart Foundation Family Heart Study; BioVU: Vanderbilt University Medical Center Biorepository; GoDARTS: Genetics of Diabetes Audit and Research Tayside; FIA3: First-time incidence of myocardial infarction in the AC county 3; EGCUT: Estonian Genome Centre, University of Tartu; EPIC: European Prospective Study into Cancer and Nutrition; HUNT: Nord-Trøndelag health study; IPM: Mt. Sinai Institute for Personalized Medicine Biobank; MDC: Malmo Diet and Cancer Study-Cardiovascular Cohort; MHI: Montreal Heart Institute Study; OHS: Ottawa Heart Study; PAS-AMC; Premature Atherosclerosis Study at Academic Medical Center Amsterdam; PennCath: University of Pennsylvania Catheterization Study; PROCARDIS: Precocious Coronary Artery Disease Study; VHS: Verona Heart Study; WHI: Women's Health Initiative. MI: myocardial infarction; CAD: coronary artery disease.

Table S2. **Sources of cases and controls in the replication study**

Study (Ancestry)	Design	Case definition	Control definition	N Cases	N Controls	Reference
BRAVE (SA)	Case-control	First-ever troponin-confirmed acute MI	Hospital controls frequency matched by age and sex	2,971	2,784	N/A
CCHS (EA)	Prospective cohort	Fatal and non-fatal MI and other coronary events according to ICD-10 codes I20-I25	Participants from the CCHS cohort who were free from coronary disease at baseline and after follow-up	2,020	6,087	[33]
CIHDS/ CGPS (EA)	Case-control	Fatal and non-fatal MI and other coronary events according to ICD-10 codes I20-I25	Age- and sex-matched population controls free from coronary disease	8,079	10,367	[33]
EPIC-CVD (EA)	Case-cohort	Fatal and non-fatal MI and other coronary events according to ICD-10 codes I20-I25	A randomly-selected subcohort of participants from the EPIC cohort who were free from coronary disease at baseline and after follow-up	3,873	7,914	[34]
MORGAM (EA)	Case-cohort	Fatal and non-fatal MI and other coronary events according to ICD-10 codes I20-I25	A randomly-selected subcohort of participants from the MORGAM cohorts who were free from coronary disease and stroke at baseline and after follow-up	2,153	2,118	[35,36]
PROMIS (SA)	Case-control	First-ever troponin-confirmed acute MI	Hospital controls frequency matched by age and sex	10,137	11,935	[37]
PROSPER (EA)	Nested case-control	Fatal and non-fatal MI and other coronary events according to ICD-10 codes I20-I25	Age- and sex-matched participants from the PROSPER trial free of coronary disease at baseline and after follow-up	641	638	[38]
WOSCOPS (EA)	Nested case-control	Fatal and non-fatal MI and other coronary events according to ICD-10 codes I20-I25	Age-matched men from the WOSCOPS trial free of coronary disease at baseline and after follow-up	659	687	[39]
Replication study total				**30,533**	**42,530**	

EA: European Ancestry; SA: South Asian Ancestry; BRAVE: Bangladesh Risk of Acute Vascular Events Study; CCHS: Copenhagen City Heart Study; CGPS: Copenhagen General Population Study; CIHDS: Copenhagen Ischaemic Heart Disease Study; EPIC-CVD: European Prospective Investigation into Cancer and Nutrition Study; MORGAM: MOnica Risk, Genetics, Archiving and Monograph project; PROMIS: Pakistan Risk of Myocardial Infarction Study; PROSPER: Prospective Study of Pravastatin in the Elderly at Risk clinical trial; WOSCOPS: West of Scotland Coronary Prevention Study; N/A: None available

Table S3. Sources of cases and controls for *ANGPTL4* sequencing

Study	Cases	Controls	Case definition	Control definition	Ref
ATVB	1,794	1,745	MI in men or women ≤ 45 years of age	Free of MI, coronary revascularization; men ≥ age 50 or women ≥ age 60	[3]
BHF-FHS/ BRICCS/ UKAGS	1,201	1,090	Clinically documented and validated MI in men ≤ 50 years of age, or women ≤ 60 years of age	No history or symptoms of CAD at age 65 yearsd	[9,10]
ESP EOMI	770	860	MI in men or women ≤ age 45	Free of MI, coronary revascularization; men ≥ age 50, women ≥ age 60	[3]
Lubeck MI	858	878	MI in men and women ≤ age 60	Controls without CAD; men and women ≤ age 65	
Munich MI	369	338	MI in men ≤ age 40 or women ≤ age 55	Controls without CAD; men ≥ age 65, women ≥ age 75	[40]
OHS	966	987	Angiographic CAD (>1 coronary artery with >50% stenosis) without history of diabetes at age ≤ 50 for men or ≤ 60 for women	Asymptomatic, men > age 65, women > age 70	[3,27]
PROCARDIS	966	936	Symptomatic CAD before age 66. CAD was defined as clinically documented evidence of myocardial infarction, coronary artery bypass grafting, acute coronary syndrome, coronary angioplasty, or stable angina	No personal or sibling history of CAD before age 66	[3,30]
***ANGPTL4* sequencing totals**	**6,924**	**6,834**			

Study abbreviations as in Table S1. BRICCS: Biomedical Research Informatics Centre for Cardiovascular Science; UKAGS: United Kingdom Aneurysm Growth Study; MI: myocardial infarction; CAD: coronary artery disease.

Table S4. Sources of samples for testing association with lipids

Study	Number of samples	Description of samples	Ref
ATVB	1,010	Controls from ATVB who were free of MI and coronary revascularization; men ≥ 50 years of age or women ≥ 60 years of age	[8]
OHS	2,103	Controls from OHS who were asymptomatic, men > 65, women > 70	[27]
PROCARDIS	2,086	Controls from PROCARDIS who had no personal or sibling history of CAD before 66 years of age	[30]
MDC	4,889	Prospective population-based epidemiologic cohort from Malmö, Sweden	[24]
Total samples	**10,088**		

CAD: Coronary artery disease

Table S5. Sources of samples from the CHARGE+ BP consortium for testing association with blood pressure

Study	Ancestry	Number of samples	Ref
AGES	European	2,973	[41]
ARIC	European	10,865	[42]
BioVU	European	18,875	[11]
CARDIA	European	2,175	[43]
CHS	European	4,132	[44]
FamHS	European	3,723	[45]
FHS	European	7,495	[46]
HABC	European	1,646	[47]
HRS	European	9,625	[48]
MESA	European	2,494	[49]
Mt. Sinai	European	1,337	[50]
RS	European	3,015	[51]
SHIP	European	7,161	[52]
WGHS	European	22,648	[53]
WHI	European	22,309	[54]
ARIC	African American	3,354	[42]
BioVU	African American	2,004	[11]
CARDIA	African American	1,986	[43]
CHS	African American	796	[44]
HABC	African American	1,105	[47]
HRS	African American	2,029	[48]
JHS	African American	2,300	[55]
MESA	African American	1,607	[49]
Mt. Sinai	African American	2,836	[50]
WHI	African American	3,486	[54]
MESA	Hispanic American	1,440	[49]
Mt. Sinai	Hispanic American	3,146	[50]
Total samples		**146,562**	

Table S6. Low-frequency coding variants outside known GWAS loci demonstrating suggestive association with CAD in the discovery study

Locus	rsID	Chromosome: Position	Allele1/ Allele2	Frequency (Allele1)	Functional effect	Stage	OR	*P*
SVEP1	rs111245230	9:113169775	C/T	3.6%	p.D2702G	Discovery	1.14	$1.1x10^{-7}$
						Replication	1.13	$1.0x10^{-3}$
						Combined	1.14	$4.2x10^{-10}$
CHTOP	rs74844193	1:153615820	A/G	1.8%	p.R175H	Discovery	1.18	$4.2x10^{-6}$
						Replication	1.07	0.28
						Combined	1.15	$6.2x10^{-6}$
PLCH2	rs41315664	1:2411245	A/G	1.3%	p.S115N	Discovery	1.29	$1.3x10^{-5}$
						Replication	0.53	0.40
						Combined	1.28	$1.8x10^{-5}$
PRSS53	rs72785539	16:31096495	C/G	0.4%	p.L324V	Discovery	1.53	$1.9x10^{-5}$
						Replication	1.06	0.72
						Combined	1.40	$1.0x10^{-4}$
ABLIM3	rs148615457	5:148596546	G/A	0.1%	p.T232A	Discovery	1.80	$2.1x10^{-5}$
						Replication	0.89	0.50
						Combined	1.34	$5.0x10^{-3}$
APOH	rs1801689	17:64210580	C/A	3.3%	p.C325G	Discovery	1.12	$2.9x10^{-5}$
						Replication	1.02	0.52
						Combined	1.08	$2.4x10^{-4}$
ANGPTL4	rs116843064	19:8429323	A/G	2.0%	p.E40K	Discovery	0.87	$3.0x10^{-5}$
						Replication	0.86	$3.4x10^{-4}$
						Combined	0.86	$4.0x10^{-8}$
OVCH2	rs200352564	11:7716849	G/C	0.1%	p.A412P	Discovery	1.74	$3.7x10^{-5}$
						Replication	0.62	0.04
						Combined	1.33	0.01
OR2J2	rs3129157	6:29141743	A/G	3.4%	p.T111A	Discovery	0.89	$6.4x10^{-5}$
						Replication	1.00	0.95
						Combined	0.91	$2.9x10^{-4}$
TAS2R16	rs34215184	7:122635469	C/A	0.2%	p.L74V	Discovery	3.58	$6.4x10^{-5}$
						Replication	2.31	0.28
						Combined	3.36	$4.0x10^{-5}$
ANKLE1	rs77683348	19:17396344	A/G	2.8%	p.R494Q	Discovery	0.89	$8.1x10^{-5}$
						Replication	1.04	0.77
						Combined	0.90	$1.6x10^{-4}$
TEX15	rs183854485	8:30699807	G/A	0.1%	p.C2243R	Discovery	3.04	$9.7x10^{-5}$
						Replication	0.66	0.43
						Combined	2.17	$2.1x10^{-3}$

GWAS: Genome-wide association study; CAD: coronary artery disease; OR: odds ratio of disease for carriers of Allele 1

Table S7: Association between low-frequency CAD variants outside of known GWAS loci and blood pressure, stratified by ancestry

Variant	Trait	Ancestry	MAF	Effect	*P*
SVEP1	SBP	EA	0.037	0.86	4.4×10^{-6}
rs111245230		AA	0.006	2.57	0.027
		HA	0.028	2.54	0.044
		All	0.032	0.94	3.0×10^{-7}
	DBP	EA	0.037	0.56	1.4×10^{-6}
		AA	0.006	1.45	0.049
		HA	0.028	0.16	0.84
		All	0.032	0.57	4.4×10^{-7}
ANGPTL4	SBP	EA	0.020	-0.18	0.47
rs116843064		AA	0.003	1.66	0.28
		HA	0.023	-1.61	0.24
		All	0.018	-0.18	0.46
	DBP	EA	0.020	-0.13	0.42
		AA	0.003	0.48	0.63
		HA	0.023	-0.69	0.40
		All	0.018	-0.13	0.38

MAF: minor allele frequency; Effect is in units of mm Hg difference for carriers of the minor allele; SBP: systolic blood pressure; DBP: diastolic blood pressure; EA: European ancestry; AA: African ancestry; HA: Hispanic ancestry

Table S8. **Conditional analysis of plasma lipids found to be significantly associated with *ANGPTL4* p.E40K**

Lipid fraction	Adjustment	Effect	*P*
HDL	None	0.29	8.2×10^{-11}
HDL	TG	0.13	0.001
TG	None	-0.33	1.6×10^{-13}
TG	HDL	-0.21	1.8×10^{-7}

HDL: high-density lipoprotein cholesterol; TG: log-transformed triglycerides. Adjustment refers to the additional covariate used in a conditional analysis. Effect refers to units of standard deviation.

Table S9. Null alleles discovered during follow-up sequencing of *ANGPTL4*

Chr	Pos	Ref	Alt	Class	Protein effect
19	8429441	C	-	Frameshift	p.C80Vfs12*
19	8430916	C	T	Nonsense	p.Q133*
19	8431137	C	T	Nonsense	p.R161*
19	8431204	G	A	Splice-site (c.547+1G>A)	N/A
19	8436303	G	-	Frameshift	p.G313Afs84*
19	8438599	G	A	Nonsense	p.W350*
19	8438628	-	CGGC	Frameshift	p.Q362Rfs13*
19	8438638	C	G	Nonsense	p.Y363*
19	8438654	C	T	Nonsense	p.Q369*
19	8438697	G	A	Nonsense	p.W383*

Chr=Chromosome; Pos=Position (HG19); Ref=reference allele; Alt=alternate allele; '-' = no allele (i.e. indicates insertion when '-' is reference and deletion when '-' is alternate); N/A=not applicable

Table S10. **Association between *ANGPTL4* null alleles and plasma lipid concentrations**

	Null allele carriers	Non-carriers	Estimated difference between carriers and non-carriers*	*P* value
LDL	14	6,951	-11.53 mg/dl	0.30
HDL	14	7,202	4.77 mg/dl	0.19
TG	16	8,085	-35%	0.003

*Estimated difference is summary effect estimate for carriers of ANGPTL4 null alleles when compared with non-carriers after adjusting for age, gender, study, and race. LDL: low-density lipoprotein cholesterol; HDL: high-density lipoprotein cholesterol; TG: log-transformed triglycerides

Table S11. **Association between *ANGPTL4* null alleles and risk for CAD**

Study	Null allele carriers with CAD	Total CAD cases	Null allele carriers without CAD	Total CAD controls
ATVB	1	1794	7	1745
BHF-FHS/ BRICCS/ UKAGS	1	1201	1	1090
ESP EOMI	3	770	1	860
Lubeck MI	2	858	4	878
Munich MI	1	369	3	338
OHS	1	966	1	987
PROCARDIS	0	966	2	936
Total	**9**	**6924**	**19**	**6834**
	Odds ratio of disease for carriers = 0.47		***P*=0.041**	

Table S12. Association between *LPL* variation and risk for CAD

rsID	Chromosome : Position	Allele1/ Allele2	Frequency (Allele1)	Functional effect	Stage	OR	P
rs328	8:19819724	G/C	9.94%	p.S447*	Discovery	0.93	5.0×10^{-6}
					Replication	0.95	8.8×10^{-3}
					Combined	0.94	2.5×10^{-7}
rs1801177	8:19805708	A/G	1.9%	p.D36N	Discovery	1.12	1.6×10^{-3}
					Replication	1.16	0.04
					Combined	1.13	2.0×10^{-4}

Supplementary References

1. Goldstein JI, Crenshaw A, Carey J, et al. zCall: a rare variant caller for array-based genotyping: genetics and population analysis. Bioinformatics 2012;28:2543-5.

2. Shah TS, Liu JZ, Floyd JA, et al. optiCall: a robust genotype-calling algorithm for rare, low-frequency and common variants. Bioinformatics 2012;28:1598-603.

3. Do R, Stitziel NO, Won HH, et al. Exome sequencing identifies rare LDLR and APOA5 alleles conferring risk for myocardial infarction. Nature 2015;518:102-6.

4. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009;25:1754-60.

5. DePristo MA, Banks E, Poplin R, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet 2011;43:491-8.

6. Cingolani P, Platts A, Wang le L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly 2012;6:80-92.

7. Yang J, Ferreira T, Morris AP, et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. Nat Genet 2012;44:369-75, S1-3.

8. Atherosclerosis, Thrombosis, and Vascular Biology Italian Study Group. No evidence of association between prothrombotic gene polymorphisms and the development of acute myocardial infarction at a young age. Circulation 2003;107:1117-22.

9. Deloukas P, Kanoni S, Willenborg C, et al. Large-scale association analysis identifies new risk loci for coronary artery disease. Nat Genet 2013;45:25-33.

10. Samani NJ, Erdmann J, Hall AS, et al. Genomewide association analysis of coronary artery disease. N Engl J Med 2007;357:443-53.

11. Weeke P, Denny JC, Basterache L, et al. Examining Rare and Low-Frequency Genetic Variants Previously Associated with Lone or Familial Forms of Atrial Fibrillation in an Electronic Medical Record System: A Cautionary Note. Circ Cardiovasc Genet 2014.

12. 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.

13. Day N, Oakes S, Luben R, et al. EPIC-Norfolk: study design and characteristics of the cohort. European Prospective Investigation of Cancer. British journal of cancer 1999;80 Suppl 1:95-103.

14. Norberg M, Blomstedt Y, Lonnberg G, et al. Community participation and sustainability--evidence over 25 years in the Vasterbotten Intervention Programme. Global health action 2012;5:1-9.

15. Stegmayr B, Lundberg V, Asplund K. The events registration and survey procedures in the Northern Sweden MONICA Project. Scandinavian journal of public health Supplement 2003;61:9-17.

16. Morris AD, Boyle DI, MacAlpine R, et al. The diabetes audit and research in Tayside Scotland (DARTS) study: electronic record linkage to create a diabetes register. DARTS/MEMO Collaboration. Bmj 1997;315:524-8.

17. Leitsalu L, Haller T, Esko T, et al. Cohort Profile: Estonian Biobank of the Estonian Genome Center, University of Tartu. International journal of epidemiology 2014.

18. Erdmann J, Stark K, Esslinger UB, et al. Dysfunctional nitric oxide signalling increases risk of myocardial infarction. Nature 2013;504:432-6.

19. Krawczak M, Nikolaus S, von Eberstein H, Croucher PJ, El Mokhtari NE, Schreiber S. PopGen: population-based recruitment of patients and controls for the analysis of complex genotype-phenotype relationships. Community genetics 2006;9:55-61.

20. Schmermund A, Mohlenkamp S, Stang A, et al. Assessment of clinically silent atherosclerotic disease and established and novel risk factors for predicting myocardial infarction and cardiac death in healthy middle-aged subjects: rationale and design of the Heinz Nixdorf RECALL Study. Risk Factors, Evaluation of Coronary Calcium and Lifestyle. American heart journal 2002;144:212-8.

21. Erdmann J, Willenborg C, Nahrstaedt J, et al. Genome-wide association study identifies a new locus for coronary artery disease on chromosome 10p11.23. Eur Heart J 2011;32:158-68.

22. Koch W, Turk S, Erl A, et al. The chromosome 9p21 region and myocardial infarction in a European population. Atherosclerosis 2011;217:220-6.

23. Krokstad S, Langhammer A, Hveem K, et al. Cohort Profile: the HUNT Study, Norway. International journal of epidemiology 2013;42:968-77.

24. Kathiresan S, Melander O, Anevski D, et al. Polymorphisms associated with cholesterol and risk of cardiovascular events. N Engl J Med 2008;358:1240-9.

25. Auer PL, Teumer A, Schick U, et al. Rare and low-frequency coding variants in CXCR2 and other genes are associated with hematological traits. Nat Genet 2014;46:629-34.

26. Dube MP, Zetler R, Barhdadi A, et al. CKM and LILRB5 Are Associated With Serum Levels of Creatine Kinase. Circ Cardiovasc Genet 2014;7:880-6.

27. McPherson R, Pertsemlidis A, Kavaslar N, et al. A common allele on chromosome 9 associated with coronary heart disease. Science 2007;316:1488-91.

28. Trip MD, Smulders YM, Wegman JJ, et al. Frequent mutation in the ABCC6 gene (R1141X) is associated with a strong increase in the prevalence of coronary artery disease. Circulation 2002;106:773-5.

29. Reilly MP, Li M, He J, et al. Identification of ADAMTS7 as a novel locus for coronary atherosclerosis and association of ABO with myocardial infarction in the presence of coronary atherosclerosis: two genome-wide association studies. Lancet 2011;377:383-92.

30. Clarke R, Peden JF, Hopewell JC, et al. Genetic variants associated with Lp(a) lipoprotein level and coronary disease. N Engl J Med 2009;361:2518-28.

31. Kathiresan S, Voight BF, Purcell S, et al. Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. Nat Genet 2009;41:334-41.

32. Women's Health Initiative Study Group. Design of the Women's Health Initiative clinical trial and observational study. The Women's Health Initiative Study Group. Controlled clinical trials 1998;19:61-109.

33. Nordestgaard BG, Benn M, Schnohr P, Tybjaerg-Hansen A. Nonfasting triglycerides and risk of myocardial infarction, ischemic heart disease, and death in men and women. JAMA 2007;298:299-308.

34. Danesh J, Saracci R, Berglund G, et al. EPIC-Heart: the cardiovascular component of a prospective study of nutritional, lifestyle and biological factors in 520,000 middle-aged participants from 10 European countries. European journal of epidemiology 2007;22:129-41.

35. Evans A, Salomaa V, Kulathinal S, et al. MORGAM (an international pooling of cardiovascular cohorts). International journal of epidemiology 2005;34:21-7.

36. Kulathinal S, Karvanen J, Saarela O, Kuulasmaa K. Case-cohort design in practice - experiences from the MORGAM Project. Epidemiologic perspectives & innovations : EP+I 2007;4:15.

37. Saleheen D, Zaidi M, Rasheed A, et al. The Pakistan Risk of Myocardial Infarction Study: a resource for the study of genetic, lifestyle and other determinants of myocardial infarction in South Asia. European journal of epidemiology 2009;24:329-38.

38. Shepherd J, Blauw GJ, Murphy MB, et al. Pravastatin in elderly individuals at risk of vascular disease (PROSPER): a randomised controlled trial. Lancet 2002;360:1623-30.

39. Shepherd J, Cobbe SM, Ford I, et al. Prevention of coronary heart disease with pravastatin in men with hypercholesterolemia. West of Scotland Coronary Prevention Study Group. N Engl J Med 1995;333:1301-7.

40. Stitziel NO, Won HH, Morrison AC, et al. Inactivating mutations in NPC1L1 and protection from coronary heart disease. N Engl J Med 2014;371:2072-82.

41. Harris TB, Launer LJ, Eiriksdottir G, et al. Age, Gene/Environment Susceptibility-Reykjavik Study: multidisciplinary applied phenomics. Am J Epidemiol 2007;165:1076-87.

42. The ARIC investigators. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. Am J Epidemiol 1989;129:687-702.

43. Friedman GD, Cutter GR, Donahue RP, et al. CARDIA: study design, recruitment, and some characteristics of the examined subjects. Journal of clinical epidemiology 1988;41:1105-16.

44. Fried LP, Borhani NO, Enright P, et al. The Cardiovascular Health Study: design and rationale. Annals of epidemiology 1991;1:263-76.

45. Higgins M, Province M, Heiss G, et al. NHLBI Family Heart Study: objectives and design. Am J Epidemiol 1996;143:1219-28.

46. Dawber TR, Kannel WB, Lyell LP. An approach to longitudinal studies in a community: the Framingham Study. Annals of the New York Academy of Sciences 1963;107:539-56.

47. Cesari M, Penninx BW, Newman AB, et al. Inflammatory markers and cardiovascular disease (The Health, Aging and Body Composition [Health ABC] Study). Am J Cardiol 2003;92:522-8.

48. Juster FT, Suzman R. An overview of the health and retirement study. Journal of Human Resources 1995;30:S7-S56.

49. Bild DE, Bluemke DA, Burke GL, et al. Multi-ethnic study of atherosclerosis: objectives and design. Am J Epidemiol 2002;156:871-81.

50. Tayo BO, Teil M, Tong L, et al. Genetic background of patients from a university medical center in Manhattan: implications for personalized medicine. PLoS One 2011;6:e19166.

51. Hofman A, van Duijn CM, Franco OH, et al. The Rotterdam Study: 2012 objectives and design update. European journal of epidemiology 2011;26:657-86.

52. Volzke H, Alte D, Schmidt CO, et al. Cohort profile: the study of health in Pomerania. International journal of epidemiology 2011;40:294-307.

53. Ridker PM, Chasman DI, Zee RY, et al. Rationale, design, and methodology of the Women's Genome Health Study: a genome-wide association study of more than 25,000 initially healthy american women. Clin Chem 2008;54:249-55.

54. Design of the Women's Health Initiative clinical trial and observational study. The Women's Health Initiative Study Group. Controlled clinical trials 1998;19:61-109.

55. Wyatt SB, Diekelmann N, Henderson F, et al. A community-driven model of research participation: the Jackson Heart Study Participant Recruitment and Retention Study. Ethnicity & disease 2003;13:438-55.