Supplementary Information

Large-scale genome-wide association study of coronary artery disease in genetically diverse populations

Supplementary Methods

Description of the Black and Hispanic cohorts used for the two stage meta-analysis

Atherosclerosis Risk in Communities (ARIC). The ARIC study is a prospective population-based study of atherosclerosis and cardiovascular diseases in 15,792 men and women, including 11,478 non-Hispanic whites and 4,314 African Americans, drawn from 4 U.S. communities (suburban Minneapolis, Minnesota; Washington County, Maryland; Forsyth County, North Carolina, and Jackson, Mississippi)[1]. In the first three communities, the sample reflects the demographic composition of the community. In Jackson, only black residents were enrolled. Participants were between age 45 and 64 years at their baseline examination in 1987-1989 when blood was drawn for DNA extraction and participants consented to genetic testing. In ARIC, coronary heart diease (CHD) events were ascertained from annual follow-up interviews and morbidity and mortality surveillance of communities including hospitalizations and deaths. Events were reviewed by two physicians and differences adjudicated. CHD events were defined as acute hospitalized myocardial infarction (MI) (definitive or probable), definite fatal CHD, electrocardiography (ECG) diagnosis of MI, or revascularization. Acute MI was defined based on criteria that included cardiac pain, cardiac markers and ECG readings. For this analysis, subjects were considered cases if they either had evidence of any of these manifestation of CHD at baseline (prevalent) or during follow up (incident).
Genotyping/Imputation: A total of 2,822 African American participants were genotyped with the Affymetrix Genome-Wide Human SNP Array 6.0. Quality control (QC) metrics included: a SNP call rate of > 90%, sample call rate >95%, and HWE p-value $<10^{-6}$, samples with <18 fingerprinting assays working, samples with>3 discordant fingerprinting assays, duplicates. samples creating large number of haploid heterozygous calls, extreme heterozygosity value, low-level IBD/IBS sharing with large number of samples, nearest neighbor analysis outlier and outliers from clustering based on missingness. Imputation was conducted to the 1000 Genome Phase 3 panel using IMPUTE (version 2.3.2) and SNPs with info score less than 0.4 were excluded.

Biobank of Institute for Personalized Medicine at Mount Sinai (BioMe): The BioMe Biobank, founded in September 2007, is an ongoing, consented electronic medical record (EMR)-linked bio- and data repository that enrolls participants non-selectively from the Mount Sinai Medical Center patient population. The BioMe Biobank currently (as of Jan 2020) comprises >52,000 participants from diverse ancestries characterized by a broad spectrum of (longitudinal) biomedical traits (https://icahn.mssm.edu/research/ipm/programs/biome-biobank)[2,3]. BioMe participants represent the broad ancestral, ethnic and socioeconomic diversity with a distinct and population-specific disease burden, characteristic of Northern Manhattan communities served by Mount Sinai Hospital. Enrolled participants consent to be followed throughout their clinical care (past, present, and future) at Mount Sinai in real-time, integrating their genomic information with their electronic health record for discovery research and clinical care implementation. BioMe participants are predominantly of African, Hispanic/Latino, and European ancestry. Participants who self-identify as Hispanic/Latino further report to be of Puerto Rican (39%), Dominican (23%), Central/South American (17%), Mexican (5%) or other Hispanic (16%) ancestry. More than 40% of European ancestry participants are genetically determined to be of Ashkenazi Jewish ancestry. The IRB-approved BioMe Biobank consent permits use of samples and de-identified linkable past, present and future clinical information from EMRs; re-contacting participants for enrollment in future research; unlimited duration of storage, and access to clinical information from the entire medical records, as well as local and external sharing of specimens and data. The BioMe Biobank has a longitudinal design as participants consent to make any EMR data from past (dating back as far as 2003), present and future inpatient or outpatient encounters available for research. As a result, participants have a high median number of encounters per patient reflecting predominant enrollment of participants with common chronic conditions from primary care facilities. Mount Sinai s system-wide Epic EMR implementation captures a full spectrum of biomedical phenotypes, including clinical outcomes, covariate and exposure data. This clinical information is complemented by detailed information on ancestry, residence history, familial medical

history, education, socio-economic status, physical activity, smoking, alcohol use, and weight history being collected in a systematic manner by interview-based questionnaire at time of enrollment. Phenotype harmonization and validation is critical to facilitate consortium-wide analyses. By applying advanced medical informatics and data mining tools, high-quality and validated phenotype data can be culled from Mount Sinai s Epic EMR. Fully-implemented phenotype algorithms include; T2D, CKD, CAD, lipid disorders, peripheral artery disease, resistant hypertension, blood cell traits, abdominal aortic aneurism, venous thromboembolism among others (see also Phenotype KnowledgeBase (PheKB) of the eMERGE Network. A total of 14,017 participants have been genotyped for both GWAS (11,150 Illumina OmniExpress BeadChip, 2,867 Affymetrix Human SNP Array 6.0) and ExomeChip (Illumina HumanExome v1.0 BeadChip) arrays funded by institutional sources. An additional ~16,000 BioMe participants were genotyped using the Illumina MEGA Chip, funded by NHGRI through PAGE II (U01HG007417) (n~12,500) and through institutional funds (n~3,500). For PAGE, a total of 1,705,969 variants were genotyped on MEGA[4]. Variant-level quality control (QC) was completed by were filtered through various criteria, including the exclusion of (1) Center for Inherited Disease Research (CIDR) technical filters, (2) variants with missing call rate >= 2%, (3) variants with more than 6 discordant calls in 988 study duplicates, (4) SNPs with greater than 1 Mendelian errors in 282 trios and 1,439 duos, (5) variants with a Hardy-Weinberg p-value less than 10-4, (6) variants with sex difference in allele frequency >= 0.2 for 373 autosomes/XY, and (7) variants with sex difference in heterozygosity > 0.3 for autosomes/XY. After variant QC, a total of 1,438,399 variants remained. Imputation was conducted at the University of Washington Genetic Coordinating Center. Sites were further restricted to variants with (1) known chromosome and position; (2) located on chromosomes 1-22, X, or XY (pseudo-autosomal); (3) with unique positions, which involved removing redundant and duplicate sites; and (4) sites with available Supplementary Information for strand annotation. After these restrictions, a total of 1,402,653 sites remained. The study samples were phased with SHAPEIT2 and imputed with IMPUTE2 to the 1000 Genomes Project Phase 3 data release. Reference panel variants were filtered to exclude all monomorphs and singletons (i.e. restricting to minor allele count (MAC) >= 2 across all 1000 Genomes Phase 2 samples). Imputed variants were excluded if the IMPUTE2 info score was less than 0.4.

The **Multiethnic Cohort study** (MEC) is a large prospective cohort study that was established between 1993 and 1996. The MEC includes primarily African Americans, Japanese American, Native Hawaiians, Latinos and European Americans living in Hawaii and California[5,6]. Cohort members were recruited through Department of Motor Vehicle license files and supplemented by voter registration and Health Care Financing Administration (Medicare) files. Participating individuals were between 45 and 75 years of age, and completed a 26-page self-administered, detailed questionnaire at cohort entry (baseline data, 1993–1996). The questionnaire included basic demographic factors (including race/ethnicity and education), lifestyle factors (e.g., diet, medication use and smoking history), and chronic medical conditions. Follow-up questionnaires are administered every five years which contain updates on participant s CHD status and lifestyle factors. Several nested case-control studies have been assembled in the MEC for GWAS of a number of cancer and non-cancer traits including breast cancer, prostate cancer, and type-2 diabetes, as well as a cohort of smokers, mainly in populations of non-European ancestry[7-11]. In the current study, we identified CHD cases and non-cases within these nested studies for the genetic analysis of CHD risk SNPs following a previously published protocol[12]. The MEC study obtained written informed consent from study participants for genetic analysis, approval from the Health Science Review Board (HSIRB) at the University of Southern California, and IRB certification permitting data sharing in accordance with the NIH Policy for Sharing of Data Obtained in NIH Supported or Conducted Genome-Wide Association Studies (GWAS). Genetic data for the MEC is available on dbGAP (phs000517.v3.p1, phs000851.v1.p1, phs000356.v2.p1, phs000306.v4.p1, phs000683.v1.p1). CHD cases were identified through linkage of the MEC to the California Hospital Discharge Data (1990–2012) (CHDD) and the Centers for Medicare and Medicaid Services (CMS) claim files (MedPAR, outpatient) (1999–2011). Analysis for this study were restricted to Black and Hispanic/Latin participants. A CHD case was defined as having ischemic heart disease under ICD-9 codes (DX 410–414), by the principal or first diagnosis code and the principal or first procedure code.

We also included cases with a primary cause of death due to myocardial infarction (ICD-9 DX410, ICD-10 I21), or other CHD conditions (ICD-9 DX411–414, ICD-10 I20, I22–25). Both prevalent (~20%) and incident (~80%) CHD cases were included in this study. Controls in this study were subjects with no history of heart attack or angina based on the baseline questionnaire or all subsequent follow-up questionnaires. Those taking nitrates at blood draw in subsequent examinations were also excluded. Individuals with non-primary CHD diagnosis codes (i.e., 2–24) from the CHDD and Medicare data were excluded from being either a case or control. Genotyping and Quality Control: We utilized genetic data generated from case-control studies in the MEC of breast cancer, prostate cancer, type 2 diabetes, and a cohort of smokers from African Americans and Latinos. Genotyping was conducted using the Illumina platform with different arrays, including the Human 1M-Duo v3.0 BeadChip, HumanOmni2.5-Quad BeadChip, Human 660W-Quad BeadChip, and the PAGE-MEGA chip (see BioME for genotyping/imputation QC details)[4].

WHI (Women s Health Initiative): WHI is a long-term national health study that focuses on strategies for preventing common diseases such as heart disease, cancer and fracture in postmenopausal women. A total of 161,838 women aged 50–79 years old were recruited from 40 clinical centers in the US between 1993 and 1998. WHI consists of an observational study, two clinical trials of postmenopausal hormone therapy (HT, estrogen alone or estrogen plus progestin), a calcium and vitamin D supplement trial, and a dietary modification trial. Study recruitment and exclusion criteria have been described previously[13-16]. CHD related outcomes were collected at baseline through self-report questionnaires and adjudicated by a panel of physicians after chart review during follow up[15]. For this analysis, cases were defined as women who either self-reported a myocardial infarction or a revascularization procedure at baseline (i.e. prevalent CHD), and/or either developed a definitive myocardial infarction, underwent a revascularization procedure, or definitively died from CHD during follow up. Genotyping was performed in subsets of participants using case-control sampling. GWAS studies contributing to the analysis here include the WHI Genomics and Randomized Trials Network (GARNET), which included all coronary heart disease, stroke, venous thromboembolic events and selected diabetes cases that happened during the active intervention phase in the WHI HT clinical trials and aged matched controls), the WHI SNP Health Association Resource (WHI SHARe), a randomly selected sample of 8,515 African American and 3,642 Hispanic women from WHI, and PAGE. GWAS was performed using Affymetrix 6.0 (WHI-SHARe), HumanOmniExpressExome-8v1_B (WHIMS), Illumina HumanOmni1-Quad v1-0 B (GARNET), and the Illumina Multi-Ethnic Genotyping Array (WHI-PAGE). For GARNET and SHARE genotyping, extensive quality control (QC) of the GWAS data included alignment (“flipping”) to the same reference panel, imputation to the 1000G data (using the recent reference panel - v3.20101123), identification of genetically related individuals, and computations of principal components (PCs) using methods developed by Price et al. (using EIGENSOFT software 53), and finally the comparison with self-reported ethnicity. After QC and exclusions from analysis protocol, the number of women included in analysis is 4,423 whites for GARNET, 7,919 for SHARe African American and 3,377 for SHARe Hispanics. For PAGE, (see BioME for genotyping/imputation QC details)[4]. There was substantial overlap between WHI-SHARE and WHI-MEGA. When both sets of genotypes were available for analyses, MEGA genotypes/imputations were prioritized.

BioVu is Vanderbilt s biorepository of DNA extracted from discarded blood collected during routine clinical testing and linked to de-identified medical records in the Synthetic Derivative[17]. Planning for BioVU began in mid-2004 and the first samples were collected in February 2007. Prior to collecting DNA samples, all aspects of the BioVU project were extensively tested. BioVU now accrues 500-1000 samples per week, totaling more than 225,000 DNA samples. Vanderbilt clinic patients may sign the BioVU Consent Form if they wish to donate their excess blood samples, or not sign the form if they do not wish to participate. Samples are scanned via a custom-developed sample acceptance program that includes automated exclusion based on specific criteria. Manual exclusions include poor quality of the blood sample, insufficient volume of blood and/or an unreadable label on the sample tube. Automated exclusions include opt-out, no signed form documenting notification of the program, duplicate samples not targeted for replenishment and random exclusion. Once a sample passes the necessary criteria, it is accepted by the program. With the help of the

bioinformatics expertise at Vanderbilt a "mirror image" of the EMR, the Synthetic Derivative, was created. It contains over 2 million individual patients with all clinical information available in a searchable form for more than the past ten years. The Synthetic Derivative is scrubbed of HIPAA identifiers with an error rate of ~0.01%. New clinical data are added to the database as they are created. For this study, analysis was restricted to African American participants with genotyping data. CAD was defined as present if patients had a) ICD9 codes for by pass (36.10 to 316.14), percutaneous transluminal coronary angioplasty (36.00 to 36.09) or text descriptions for CAD procedures, or b) ICD9 codes for ischemic heart disease (410* to 414*) with or without abnormal cardiac catheterization, or c) abnormal stress test with abnormal cardiac catheterization. Controls needed to have absence of criteria a, b and c. Samples were genotyped with the Illumina 1M array and standard quality control measures were taken. Analyses were limited to African Americans based on self-report. Briefly, SNPs with low genotyping quality (<98%), minor allele frequency (< 0.01), non-autosomal SNPs, SNPs which did not map to a chromosomal position and SNPs deviating from Hardy-Weinberg equilibrium were removed. Individuals with low genotyping quality (98%), those with inconsistent reported versus genetically determined sex and cryptic relatedness were also excluded. Principal components were generated using EIGENSTRAT. Phasing was performed using SHAPEIT2 and imputed with IMPUTE2 using the 1000 genomes reference panels (phase 1 version 3). Multiple logistic regression was used to model with CAD as the dependent variable with SNPs (additive model) while adjusting for the top two principal components were performed using SNPTESTv2.5.

CHS (Cardiovascular Health Study): CHS is a population-based cohort study of risk factors for cardiovascular disease in adults 65 years of age or older conducted across four field centers[18]. The original predominantly European ancestry cohort of 5,201 persons was recruited in 1989-1990 from random samples of the Medicare eligibility lists and an additional predominately African-American cohort of 687 persons was enrolled in 1992-93. Blood samples were drawn from all participants at their baseline examination and DNA was subsequently extracted from available samples. CHS was approved by institutional review committees at each site and individuals in the present analysis gave informed consent including consent to use of genetic information for the study of cardiovascular disease. For this analysis, European ancestry participants were excluded. Both prevalent and incident CHD events were used to form the case sets using previously published protocols[19,20]. Genotyping/Imputation: Blood samples were drawn from all participants at their baseline examination and DNA was subsequently extracted from available samples. Genotyping was performed at the General Clinical Research Center's Phenotyping/Genotyping Laboratory at Cedars-Sinai among CHS participants who consented to genetic testing and had DNA available using the Illumina HumanOmni1-Quad_v1 BeadChip system. Participant-level exclusions: Beyond laboratory genotyping failures, participants were excluded if they had a call rate ≤ 95% or if their genotype was discordant with known sex or prior genotyping. After QC, genotyping was successful for 823 African-American participants. SNP exclusions: In CHS, the following exclusions were applied: call rate < 97%, HWE $P < 10^{-5}$, > 2 duplicate errors or Mendelian inconsistencies (for reference trios), heterozygote frequency = 0, SNP not found in HapMap. TOPMed Imputation: Imputation to the TOPMed Freeze5 panel was performed on the Michigan imputation server using minimac 4. Prephasing was performed with Eagle v2.4. SNPs with variance on the allele dosage ≤0.01 were excluded from analyses. Consent: CHS was approved by institutional review committees at each site, the subjects gave informed consent, and those included in the present analysis consented to the use of their genetic information for the study of cardiovascular disease.

Health, Aging, and Body Composition (Health ABC) Study. The Health ABC study is a prospective cohort study investigating the associations between body composition, weight-related health conditions, and incident functional limitation in older adults. Health ABC enrolled well-functioning, community-dwelling black (n=1281) and white (n=1794) men and women aged 70-79 years between April 1997 and June 1998. Participants were recruited from a random sample of white and all black Medicare eligible residents in the Pittsburgh, PA, and Memphis, TN, metropolitan areas. Participants have undergone annual exams and semi-annual phone interviews. Follow-up for exam measures continued for 16 years, ending in 2013 (December 31st, 2012). The current study sample is restricted to blacks with available genotyping data. Individuals with

both prevalent and incident CHD were assigned as cases using established protocols described in more detail at https://healthabc.nia.nih.gov/dataset-documentation/outcomes-datasets. Briefly, individuals with prevalent CHD were defined as MI, angina, bypass surgery, or percutaneous transluminal coronary angioplasty (PTCA) procedures or diagnoses reported to the Centers for Medicare and Medicaid Services in the five years prior to baseline. Incident CHD was defined as adjudicated event of MI, hospitalization for angina, or adjudicated CHD death in participants who did not have prevalent CHD. The sample included 1105 participants and 322 events. With follow through December 31, 2012. Genotyping was performed by the Center for Inherited Disease Research (CIDR) using the Illumina Human1M-Duo BeadChip system. Samples were excluded from the dataset for the reasons of sample failure, genotypic sex mismatch, and first-degree relative of an included individual based on genotype data. Analysis was restricted to SNPs with minor allele frequency ≥ 1%, call rate ≥97% and HWE p≥10-6. Genotypes were available on 914,263 high quality SNPs for imputation based on the HapMap CEU (release 22, build 36) using the MACH software (version 1.0.16).

Healthy Aging in Neighborhoods of Diversity across the Life Span (HANDLS): HANDLS is a community-based, longitudinal epidemiologic study examining the influences of race and socioeconomic status (SES) on the development of age-related health disparities among a sample of socioeconomically diverse African Americans and whites. This unique study will assess over a 20-year period physical parameters and also evaluate genetic, biologic, demographic, and psychosocial, parameters of African American and white participants in higher and lower SES to understand the driving factors behind persistent black-white health disparities in overall longevity, cardiovascular disease, and cognitive decline. The study recruited 3,722 participants from Baltimore, MD with a mean age of 47.7 years, 2,200 African Americans and 1,522 whites, with 41% reporting household incomes below the 125% poverty delimiter. Analysis for this study was restricted to African American participants. CHD was defined as self-reported hospital-diagnosis of prevalent or incident coronary artery disease. Genotyping was done on a subset of self-reporting African American participants by the Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health (NIH). The Illumina Human 1M SNP array was used for genotyping. The samples were excluded if sample call rate < 95%, heterozygosity outliers (mean±3×s.d.) or gender mismatch. The variants were excluded if the genotyping call rate < 98%, Hardy-Weinberg Equilibrium (HWE) P-value < 1×10-6, and minor allele frequency < 1%. The imputing of genotype was based on 1000 Genome reference panel. A larger genotyping effort included a small subset of self-reporting European ancestry samples. This research was supported by the Intramural Research Program of the NIH, NIA and the National Center on Minority Health and Health Disparities.

Jackson Heart Study (JHS). The JHS is a prospective population-based study to seek the causes of the high prevalence of common complex diseases among African Americans in the Jackson, Mississippi metropolitan area, including cardiovascular disease, type-2 diabetes, obesity, chronic kidney disease, and stroke[21-23]. During the baseline examination period (2000-2004) 5,301 self-identified African Americans were recruited from four sources, including (1) randomly sampled households from a commercial listing; (2) ARIC participants; (3) a structured volunteer sample that was designed to mirror the eligible population; and (4) a nested family cohort. Unrelated participants were between 35 and 84 years old, and members of the family cohort were ≥ 21 years old when consent for genetic testing was obtained and blood was drawn for DNA extraction. Based on DNA availability, appropriate informed consent, and genotyping results that met quality control procedures, genotype data were available for 3,030 individuals, including 885 who are also ARIC participants. In the current study, JHS participants who were also enrolled in the ARIC study were analyzed with the ARIC dataset – for this reason, the JHS dataset analyzed here had 2,145 individuals. Many key aspects of the JHS were modeled on, and are essentially identical to, the methods used in the ARIC study (see above), including phlebotomy procedures, blood pressure measurement, laboratory methods for lipids, cholesterol, and glucose analysis, definitions of prevalent diabetes mellitus and hypertension, and survey methods and definitions related to medical history, cigarette smoking, and CHD. The JHS Affy6.0 GWAS data were imputed to the 1000G Phase 3 version 5 reference panel. Imputation was completed using Minimac3

on the Michigan Imputation Server[24]. The reference panel includes 5,008 haplotypes from 26 populations across the world (http://www.internationalgenome.org). Prior to imputation, SNPs were filtered for minor allele frequency ≥1%, call rate ≥ 90%, HWE p-value > 10-6, as well as exclusion of sites with invalid or mismatched alleles for the reference panel. The total number of SNPs imputed was 49,143,605 (prior to filters for imputation quality or minor allele count).

eMERGE Network: The eMERGE network is a US-based consortium of cohorts in which DNA samples are linked to EHR data to enable large-scale, high-throughput genomic studies. At the time of these analysis, the network, in its third phase, included 105,108 genotyped participants from 12 geographically distinct healthcare institutions located across the US, each with its own biorepository and site-specific eligibility criteria for enrollment. The majority of the eMERGE sites recruited primarily from the outpatient setting. The Mayo Clinic cohort partially included individuals referred for noninvasive vascular evaluation or cardiac stress testing (n= 3,640, 35% of the Mayo Clinic genotyped cohort). Details of the enrollment process of each biorepository contributing to the eMERGE network have been previously published[25,26]. Each member site obtained approval from its respective institutional review board. Genotype Data: High-density genotype data were available for 105,108 participants from the eMERGE network. To harmonize the genotype data from 12 member sites, we imputed each of the 80 Illumina and Affymetrix genotype batches via the Michigan Imputation Server by using the minimac3 algorithm and the genotype reference panel from the Haplotype Reference Consortium[24,27]. A detailed description of quality-control procedures has been previously published[28]. We restricted the study cohort to adult participants (≥18 years of age) of two mutually exclusive racial/ethnic groups, namely, non-Hispanic African Ancestry [AA] (matching genetic and self-reported ancestry, self-identified to be "non-Hispanic") and Hispanic ethnicity [HE] (self-identified as "Hispanic"). As part of phenotype data quality control, we only kept individuals with at least two EHR records and ≥1-year of EHR history. In accordance with the MVP phenotype description, we defined coronary heart disease (CHD) cases as individuals with a history of myocardial infarction or coronary revascularization (minimum of 1 diagnostic/procedure code). Cases of soft CHD were ascertained using additional diagnostic codes provided by MVP (minimum of 2 diagnostic codes on different dates). Controls were defined as individuals free of any diagnostic/procedure codes for CHD. Participants who do not fit into any of these designations were excluded. We supplemented the provided list of diagnostic/procedure codes for CHD with our previously validated list of codes in the eMERGE SEARCH algorithm[29]. Separately in AA and HE cohorts, we used KING pipeline[30] (version 2.2.4) to extract lists of participants that contain no pairs of individuals with a first-degree relationship. Following implementation of these quality control measures and phenotype definitions, our replication cohort included 9,106 AA and 3,969 HE unrelated participants.

Penn Medicine Biobank (PMBB): All individuals who were recruited for the PMBB are patients of clinical practice sites of the University of Pennsylvania Health System. Appropriate consent was obtained from each participant regarding storage of biological specimens, genetic sequencing, access to all available EHR data and permission to recontact for future studies. The study was approved by the Institutional Review Board of the University of Pennsylvania and complied with the principles set out in the Declaration of Helsinki. Currently >60,000 participants have actively consented for linkage of biospecimens with electronic health record data for broad health related research. CAD was defined by the presence of 2 or more of the following ICD 9 and 10 codes as an outpatient or 1 of the following as an inpatient: 410, 410.0, 410.00-02, 410.1, 410.10-12, 410.2, 410.20-22, 410.3, 410.30-33, 410.4, 410.40-42, 410.5, 410.50-52, 410.6, 410.60-62, 410.7, 410.70-72, 410.8, 410.80-82, 410.9, 410.90-92, 411.0, 411.1, 411.81, 411.89, 412, 414.00, 414.01-05, 414.2-4, 414.8, 414.9, V45.81-82, I21, I21.0, I21.01-02, I21.09, I21.1, I21.11, I21.19, I21.2, I21.21, I21.29, I21.3 4, I21.9, I21.A, I21.A1, I21.A9, I22, I22.0-2, I22.8, I22.9, I23, I23.0, I23.1-8, I24, I24.0-1, I24.8-9, I25.1-2, I25.5-6, I25.70-73, I25.79, I25.810, I25.82-84, I25.89, I25.9, Z95.1, Z98.61, I20.0 . Participants have been genotyped on either the Illumina QuadOmni platform at the Regeneron Genetics Center, or either the Illumina GSA V1 or GSA V2 platforms at the Children's Hospital of Philadelphia Center for Applied Genomics. Standard quality control pipelines were employed and the datasets were combined using an imputation based approach[31] The data was imputed to 1000G Phase3 v5 reference panel using the Michigan

Imputation Server[24] Genetic ancestry was inferred from principal components derived from common, high-quality variants using SMARTPCA. Related individuals were removed from the dataset based on a kinship coefficient of 0.25 or greater. Analyses were restricted to subjects of African American ancestry.

UK-Biobank. UK Biobank is a large-scale biomedical database and research resource, containing in-depth genetic and health information from half a million UK participants aged 37-73 years recruited between the years of 2006 and 2010. The database is regularly augmented with additional data and is globally accessible to approved researchers undertaking vital research into the most common and life-threatening diseases. Ethical approval for the use of UKB clinical data along with consent from participants was obtained by the National Health Service National Research Ethics Service (ref: 11/NW/0382) and data use approved under applications Data was accessed through UK Biobank Resource under Application Numbers 19416 and 13721. Approach to case-definition of CHD and analyses mirrored previous work in the UK Biobank for whites[32] but analyses were restricted to participants of African ancestry. Both prevalent and incident cases were included in the case set. Participants were assayed two very similar genotyping arrays, the UK Biobank Lung Exome Variant Evaluation (UK BiLEVE) study were genotyped at 807,411 markers using the Applied Biosystems UK BiLEVE Axiom Array by Affymetrix (now part of Thermo Fisher Scientific)[33] (n = 49,950) and the closely related Applied Biosystems UK Biobank Axiom Array (825,927 markers) that shares 95% of marker content with the UK BiLEVE Axiom Array (n = 438,427). The quality control pipeline for the genotyping, the algorithms used for the identification of ancestral diversity and cryptic relatedness, and the haplotype estimation and genotype imputation algorithm have been described extensively elsewhere[34].

Supplementary results

Druggability of all implicated genes among 106 novel CAD susceptibility loci – summary of findings from the Therapeutic Target Database

INFLAMMATION/IMMUNE/INFECTION RELATED
Diethylcarbamazine and Zileuton are both approved drugs that target ALOX5 (leukotriene synthesis inhibitors) and used for the treatment of filiariasis and asthma. The anti-inflammatory effects of these drugs may be relevant to the treatment of CAD given the recent positive trials of other anti-inflammation drugs. However, there is no observational evidence that leukotriene synthesis inhibitors protect against CAD.
IL-3 approval for aplastic anemia targets candidate gene CSF3.

Risankizumab, tidrakizumab, ustekinumab, and CNTO-1959 are all approved for psoriasis and target IL23.
Tofacitinib targets Janus kinase 3 (JAK-3) and is approved for the treatment of Rheumatoid arthritis.
Meprobamate is approved for malaria and targets Protein kinase C epsilon (PRKCE)
Acitretin is approved for psoriasis and targets Signal transducer and activator of transcription 3 (STAT3)
Dihydroxyaluminium and levosimendan are approved for treatment of skin inflammation and CHF, respectively.
Glassia and zemaira are approved for emphysema and target Alpha-1-antitrypsin (SERPINA1)
Isoxyl is approved to treat bacterial infection and is a target for Acyl-CoA desaturase (SCD).

CHEMOTHERAPY/CANCER DRUGS
KU-0058948, nicotinamide, and niraparib tosylate are approved to treat ovarian cancer, inflammatory skin condition and peritoneal cancer.
Dacomitinib, lapatinib, margetuximab, masoprocol, merimepodib, pertuzamam, trastuzumab, tucatinib are all approved drugs for the treatment of lung, breast, prostate cancers and target Erbb2 tyrosine kinase receptor (HER2)
Topotecan and irinotecan and belotecan are all approved for small cell lung, colorectal and ovarian CA and target TOP1 (or topoisomerase I).
Afibercept, bevacizumab, brolucizumab, and ranibizumab are approved for treatment of metastatic colorectal cancer, macular regeneration, and Hereditary hemorrhagic telangiectasia and all target VEGFA

CHOLESTEROL/LIPIDS METABOLISM RELATED

Probucol is an approved drug and the target for ABCA1 but the main indication of this drug is already to lower cholesterol and risk of CAD.

Mipomersen is an approved drug for Familial hypercholesterolemia and it targets APOB messenger RNA.

Aminoglutethimide is approved for Cushings and targets Cholesterol desmolase (CYP11A1) which catalyzes the side-chain cleavage reaction of cholesterol to pregnenolone, the precursor of most steroid hormones.

SBC-102 is approved to treat enzyme deficiency for Lysosomal acid lipase (LIPA)

Clofibrate and gemfibrozil target lipoprotein lipase (LPL) and are approved to treat hyperlipidemia.

CARDIAC/VESSEL WALL

Ambrisentan, bosentan, LU302146 are approved for pulmonary hypertension which target Endothelian A receptor (EDNRA).

Benzbromarone (Gout) and Crofelemer (HIV associated diarrhea) target ANO, a Calcium-dependent chloride channel anoctamin (ANO). This gene is required for pacemaker activity in the GI smooth muscle cells.

Erenumab, Framanezumab, and Galcanezumab are all approved drugs for hyperprolactonemia. They target Calcitonin gene-related peptide receptor (CGRPR). Rimegepant and ubrogepant are approved for migraine and also target same gene.

Pegaptanib is approved to treat Neovascular age-related macular degeneration and is a target of Neuropilin-1 (NRP1).

Dronedarone is approved to treat angina and is a target for Voltage-gated potassium channel Kv1.5 (KCNA5).

VESNARINONE targets Voltage-gated potassium channel Kv11.1 (KCNH2) and is approved for the treatment of cardiac failure.

Crizanlizumab is approved for vaso-occlusive disease and target p-selectin.

7 drugs (avanafil, dipyridamole, papaverine, rhucin, tadalafil, udenafil, vardenafil) are approved for erectile dysfunction, hypertension, spasm, hereditary angioedema, ED.

BONE MARROW STIMULANTS

Romiplostim an approved treatment for thrombocytopenia is a target for AHR and some additional genes related to the vascular wall including Fibroblast growth factor receptor 1 (FGFR1), Platelet-derived growth factor receptor alpha and beta (PDGFRA, PDGFRB), Vascular endothelial growth factor receptor 2 (KDR), and VEGFR1 messenger RNA (VEGFR1 mRNA).

GDC-0199, MCI-186, and Taxol are approved drugs for CLL, ALS, and Solid Tumour/cancer. They target BCL2.

AS-1670542, Eltrombopag, Lusutrombopag, and revolade/promacta, and romiplostim are approved to treat thrombocytopenia and target Thrombopoietin receptor (MPL)

OTHER

27 approved drugs (Acetaminophen, Amfepramone, Amitriptyline, Amoxapine, Atomoxetine, Bupropion, Cocaine, Dasotraline, Desipramine, Desvenalfaxine succinate, Duloxetine, Imipramine, Iobenguane I-123, Levomilnacipran, Maprotiline, Mazindol, Milnacipran, Netarsudil, Phenmetrazine, Phentermine, Protriptyline, Reboxetine, Sibutramine, Tapentadol Hcl, Trimipramine, Venlafaxine, SPN-812) target Norepinephrine transporter (NET) of the gene SLC6A2 and these drugs are used to treat pain, obesity, depression, ADHD, smoking dependence, aneasthesia, fibromyalgia, neuroendocrine cancer, glaucoma

26 approved drugs (acetaminophen, amfepramone, amitriptaline, bupropion, Chlorphentermine Hydrochloride, citalopram, clomipramine, cocaine, dasotraline, desvenlafaxine, dextromethorphan, duloxetine, escitalopram, fluoxetine, fluvoxamine, levomilancipran, luvox, nortriptyline, paroxetine, sertraline, sibutramine, tianeptine, trazodone, venlafaxine, vilazodone, vortioxetine) that target Serotonin transporter (SERT) of the gene SLC6A4

Regeneron Genetics Center consortium author list

Affiliation address: Regeneron Genetics Center (RGC), 777 Old Saw Mill River Road, Tarrytown, NY 10591 +1 914-847-7000

RGC Management and Leadership Team

Goncalo Abecasis, Ph.D., Aris Baras, M.D., Michael Cantor, M.D., Giovanni Coppola, M.D., Aris Economides, Ph.D., Luca A. Lotta, M.D., Ph.D., John D. Overton, Ph.D., Jeffrey G. Reid, Ph.D., Alan Shuldiner, M.D.

Sequencing and Lab Operations

Christina Beechert, Caitlin Forsythe, M.S., Erin D. Fuller, Zhenhua Gu, M.S., Michael Lattari, Alexander Lopez, M.S., John D. Overton, Ph.D., Thomas D. Schleicher, M.S., Maria Sotiropoulos Padilla, M.S., Louis Widom, Sarah E. Wolf, M.S., Manasi Pradhan, M.S., Kia Manoochehri, Ricardo H. Ulloa.

Genome Informatics

Xiaodong Bai, Ph.D., Suganthi Balasubramanian, Ph.D., Andrew Blumenfeld, Boris Boutkov, Ph.D., Gisu Eom, Lukas Habegger, Ph.D., Alicia Hawes, B.S., Shareef Khalid, Olga Krasheninina, M.S., Rouel Lanche, Adam J. Mansfield, B.A., Evan K. Maxwell, Ph.D., Mrunali Nafde, Sean O'Keeffe, M.S., Max Orelus, Razvan Panea, Ph.D., Tommy Polanco, B.A., Ayesha Rasool, M.S., Jeffrey G. Reid, Ph.D., William Salerno, Ph.D., Jeffrey C. Staples, Ph.D.

Research Program Management

Marcus B. Jones, Ph.D., Jason Mighty, Ph.D., Lyndon J. Mitnaul, Ph.D.

BioBank Japan Project consortium author list

Koichi Matsuda. Laboratory of Genome Technology, Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo, Japan. Laboratory of Clinical Genome Sequencing, Graduate School of Frontier Sciences, The University of Tokyo, Tokyo, Japan.

Yuji Yamanashi. Division of Genetics, The Institute of Medical Science, The University of Tokyo, Tokyo, Japan.

Yoichi Furukawa. Division of Clinical Genome Research, Institute of Medical Science, The University of Tokyo, Tokyo, Japan.

Takayuki Morisaki. Division of Molecular Pathology IMSUT Hospital, Department of Internal Medicine Project Division of Genomic Medicine and Disease Prevention The Institute of Medical Science The University of Tokyo, Tokyo, Japan.

Yoshinori Murakam. Department of Cancer Biology, Institute of Medical Science, The University of Tokyo, Tokyo, Japan.

Yoichiro Kamatani. Laboratory of Complex Trait Genomics, Graduate School of Frontier Sciences, The University of Tokyo, Tokyo, Japan. Laboratory of Clinical Genome Sequencing, Graduate School of Frontier Sciences, The University of Tokyo, Tokyo, Japan.

Kaori Muto. Department of Public Policy, Institute of Medical Science, The University of Tokyo, Tokyo, Japan.

Akiko Nagai. Department of Public Policy, Institute of Medical Science, The University of Tokyo, Tokyo, Japan.

Wataru Obara. Department of Urology, Iwate Medical University, Iwate, Japan.

Ken Yamaji. Department of Internal Medicine and Rheumatology, Juntendo University Graduate School of Medicine, Tokyo, Japan.

Kazuhisa Takahashi. Department of Respiratory Medicine, Juntendo University Graduate School of Medicine, Tokyo, Japan.

Satoshi Asai. Division of Pharmacology, Department of Biomedical Science, Nihon University School of Medicine, Tokyo, Japan.
Division of Genomic Epidemiology and Clinical Trials, Clinical Trials Research Center, Nihon University. School of Medicine, Tokyo, Japan.

Yasuo Takahashi. Division of Genomic Epidemiology and Clinical Trials, Clinical Trials Research Center, Nihon University School of Medicine, Tokyo, Japan.

Takao Suzuki. Tokushukai Group, Tokyo, Japan.

Nobuaki Sinozaki. Tokushukai Group, Tokyo, Japan.

Hiroki Yamaguchi. Departmentof Hematology, Nippon Medical School, Tokyo, Japan.

Shiro Minami. Department of Bioregulation, Nippon Medical School, Kawasaki, Japan.

Shigeo Murayama. Tokyo Metropolitan Geriatric Hospital and Institute of Gerontology, Tokyo, Japan.

Kozo Yoshimori. Fukujuji Hospital, Japan Anti-Tuberculosis Association, Tokyo, Japan.

Satoshi Nagayama. The Cancer Institute Hospital of the Japanese Foundation for Cancer Research, Tokyo, Japan.

Daisuke Obata. Center for Clinical Research and Advanced Medicine, Shiga University of Medical Science, Shiga, Japan.

Masahiko Higashiyama. Department of General Thoracic Surgery, Osaka International Cancer Institute, Osaka, Japan.

Akihide Masumoto. IIZUKA HOSPITAL, Fukuoka,Japan.

Yukihiro Koretsune. National Hospital Organization Osaka National Hospital, Osaka, Japan.

Kazuyoshi Ishigaki. Laboratory for Statistical and Translational Genetics, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan.

Satoshi Koyama. Laboratory for Cardiovascular Genomics and Informatics, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan.

Kaoru Ito. Laboratory for Cardiovascular Genomics and Informatics, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan.

CARDIoGRAMplusC4D Consortium Steering Committee
Affiliation that follow are at time of contribution to CARDIoGRAMplusC4D
(last updated December 6, 2021)

Krishna G Aragam[1], David O Arnar[2], Themistocles L Assimes[3], Bjørn O Åsvold[4], Deepak S Atri[5], Aris Baras[6], Johan LM Björkegren[7], Stefan Blankenberg[8], Bernhard O Boehm[9], Eric Boerwinkle[10], Michael R Brown[11], Ben Brumpton[12], Noël P Burtt[13], Adam S Butterworth[14], John C Chambers[15], Robert Clarke[16], Rory Collins[16], Maria C Costanzo[13], John Danesh[17], Paul S de Vries[18], George Dedoussis[19], Panos Deloukas[20], Patrick T Ellinor[21], Jeanette Erdmann[22], Martin Farrall[23], Hilary K Finucane[24], Jason Flannick[25], Paul W Franks[26], Olga Giannakopoulou[27], Panagiota Giardoglou[19], Anuj Goel[28], Christopher Grace[23], Dongfeng Gu[29], Daniel F Gudbjartsson[30], Ulrich Güldener[31], Rajat M Gupta[32], Syed M. Ijlal Haider[33], Anders Hamsten[34], Anna Helgadottir[35], George Hindy[36], Hilma Holm[37], Jemma C Hopewell[16], G Kees Hovingh[38], Kristian Hveem[12], Maysson Ibrahim[16], Erik Ingelsson[39], Kazuyoshi Ishigaki[40], Kaoru Ito[41], Dong-Keun Jang[13], Tao Jiang[42], Frederick K Kamanu[43], Yoichiro Kamatani[44], Stavroula Kanoni[27], Adnan Kastrati[45], Sekar Kathiresan[46], Thorsten Kessler[45], Amit V Khera[47], Bong-Jo Kim[48], Issei Komuro[49], Tomasz Konopka[27], Jaspal S Kooner[50], Satoshi Koyama[41], Iftikhar J Kullo[51], Theodosios Kyriakou[52], Terho Lehtimäki[53], Ling Li[31], Luca A Lotta[6], Lijiang Ma[54], Nicholas A Marston[43], Winfried März[55], Ruth McPherson[56], Thomas Meitinger[57], Olle Melander[58], Andres Metspalu[59], Alanna C Morrison[11], Sören Mucha[33], Loreto Muñoz Venegas[60], Matthias Munz[33], Federico Murgia[16], Christopher P Nelson[61], Jonas B Nielsen[62], Markku S Nieminen[63], Markus M Nöthen[64], Christopher O'Donnell[65], Marju Orho-Melander[58], Colin N A Palmer[66], Shichao Pang[31], Markus Perola[67], Loukianos S Rallidis[68], Muredach P Reilly[69], Tobias Reinberger[33], Samuli Ripatti[70], Robert Roberts[71], Carolina Roselli[72], Christian T Ruff[43], Arno Ruusalepp[73], Marc S Sabatine[43], Danish Saleheen[74], Nilesh J Samani[75], Manjinder S Sandhu[76], Gavin Schnitzler[77], Stefan Schreiber[78], Heribert Schunkert[79], Paul Sherliker[80], Agneta Siegbahn[81], Damian Smedley[27], Kari Stefansson[82], Ida Surakka[83], Gudmundur Thorgeirsson[84], Gudmar Thorleifsson[37], Unnur Thorsteinsdottir[85], Jacob K Ulirsch[86], Moritz von Scheidt[87], Minxian Wang[88], Hugh Watkins[23], Thomas R Webb[75], Elle M Weeks[89], Cristen J Willer[90], Brooke N Wolford[91], Pierre A Zalloua[92], Wei Zhou[93]

[1]Cardiovascular Research Center, Massachusetts General Hospital; Center for Genomic Medicine, Massachusetts General Hospital, Boston MA USA; Cardiovascular Disease Initiative, Broad Institute of MIT and Harvard; Program in Medical and Population Genetics Broad Institute of MIT and Harvard Cambridge MA USA, [2]deCODE genetics/Amgen, Inc. Reykjavik, Iceland; Faculty of Medicine, University of Iceland, Reykjavik, Iceland, Department of Internal Medicine, Division of Cardiology, Landspitali – National University Hospital of Iceland, Reykjavik, Iceland, [3]Department of Medicine, Stanford University School of Medicine, Stanford, California, USA, [4]K.G. Jebsen Center for Genetic Epidemiology, Department of Public Health and Nursing, Norwegian University of Science and Technology, NTNU Trondheim, Norway; HUNT Research Centre, Norwegian University of Science and Technology, Levanger Norway; Department of Endocrinology, Clinic of Medicine, St. Olavs Hospital, Trondheim, Norway, [5]Divisions of Cardiovascular Medicine and Genetics, Brigham and Women's Hospital, Harvard Medical School, Broad Institute of MIT and Harvard, Cardiovascular Disease Initiative, Broad Institute of MIT and Harvard, Cambridge, MA, USA, [6]Regeneron Genetics Center, Regeneron Pharmaceuticals, Tarrytown, NY, USA, [7]Clinical Gene Networks AB, Stockholm, Sweden; Department of Genetics & Genomic Sciences, Institute of Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York, USA; Integrated Cardio Metabolic Centre, Karolinska Institutet, Karolinska Universitetssjukhuset, Huddinge, Sweden, [8]University Heart Center Hamburg, Clinic for general and interventional Cardiology, Hamburg, Germany, [9]Division of Endocrinology and Diabetes, Department of Internal Medicine, Ulm University Medical Centre, Ulm, Germany, [10]Human Genetics Center, Department of Epidemiology, Human Genetics, and Environmental Sciences, School of Public Health, The University of Texas Health Science Center at Houston, Houston, TX, USA; Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX, USA, [11]Human Genetics Center, Department of Epidemiology, Human Genetics, and Environmental Sciences, School of Public Health, The University of Texas Health Science Center at Houston, Houston, TX, USA, [12]K.G. Jebsen Center for Genetic

Epidemiology, Department of Public Health and Nursing, Norwegian University of Science and Technology, NTNU, Trondheim, Norway; HUNT Research Centre, Norwegian University of Science and Technology, Levanger, Norway, [13]Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA, USA, [14]BHF Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK; Health Data Research UK Cambridge, Wellcome Genome Campus and University of Cambridge, Cambridge, UK; British Heart Foundation Centre of Excellence, Division of Cardiovascular Medicine, Addenbrooke's Hospital, Cambridge, UK; National Institute for Health Research Cambridge Biomedical Research Centre, Cambridge University Hospitals, Cambridge, UK; The National Institute for Health Research Blood and Transplant Unit (NIHR BTRU) in Donor Health and Genomics at the University of Cambridge, Cambridge, UK, [15]Epidemiology and Biostatistics, Imperial College London, London, UK, [16]CTSU - Nuffield Department of Population Health, Medical Sciences Division, University of Oxford, Oxford, UK, [17]BHF Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge; Health Data Research UK, Wellcome Genome Campus and University of Cambridge; The National Institute for Health Research Blood and Transplant Unit (NIHR BTRU) in Donor Health and Genomics at the University of Cambridge, Cambridge, UK; Human Genetics, Wellcome Sanger Institute, Saffron Walden, UK; National Institute for Health Research Cambridge Biomedical Research Centre, Cambridge University Hospitals, Cambridge, UK; British Heart Foundation Centre of Excellence, Division of Cardiovascular Medicine, Addenbrooke's Hospital, Cambridge, UK, [18]Human Genetics Center, Department of Epidemiology, Human Genetics, and Environmental SciencesSchool of Public Health, The University of Texas Health Science Center at Houston, Houston, TX, USA, [19]School of Health Science & Education, Department of Nutrition-Dietetics, Harokopio University, Athens, Greece, [20]William Harvey Research Institute, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, U; KPrincess Al-Jawhara Al-Brahim Centre of Excellence in Research of Hereditary Disorders (PACER-HD) King Abdulaziz University, Jeddah, Saudi Arabia; Princess Al-Jawhara Al-Brahim Centre of Excellence in Research of Hereditary Disorders (PACER-HD) King Abdulaziz University, Jeddah, Saudi Arabia, [21]Cardiovascular Research Center, Massachusetts General Hospital, Boston, MA, USA; Cardiovascular Disease Initiative, Broad Institute of MIT and Harvard, Cambridge, MA, USA, [22]Institute for Cardiogenetics University of Lübeck, Lübeck, Germany; DZHK (German Research Center for Cardiovascular Research), partner site Hamburg-Lübeck-Kiel, Germany, [23]Radcliffe Department of Medicine, Division of Cardiovascular Medicine, University of Oxford; Wellcome Centre for Human Genetics, University of Oxford, Oxford, UK, [24]Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA, USA; Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, Massachusetts, USA; Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA, USA, [25]Division of Genetics and Genomics, Boston Children's Hospital, Boston, MA, USA, [26]Department of Clinical Sciences, Genetic & Molecular Epidemiology Unit, Lund University Diabetes Center, Skåne University Hosptial, Malmö, Sweden Department of Public Health & Clinical Medicine, Genetic Epidemiology & Clinical Research Group, Section for Medicine, Umeå University, Umeå, Sweden; Department of Nutrition, Harvard School of Public Health, Boston, MA USA, [27]William Harvey Research Institute, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK, [28]Radcliffe Department of MedicineDivision of Cardiovascular MedicineUniversity of Oxford; Wellcome Centre for Human Genetics, University of Oxford, Oxford, UK, [29]Department of Population Genetics and Prevention, Fu Wai Hospital & Cardiovascular Institute, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China, [30]deCODE genetics/Amgen, Inc. Reykjavik Iceland; School of Engineering and Natural Sciences, University of Iceland, Reykjavik, Iceland, [31]German Heart Centre Munich, Department of Cardiology, Technical University of Munich, Munich, Germany, [32]Divisions of Cardiovascular Medicine and Genetics, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA; Cardiovascular Disease Initiative, Broad Institute of MIT and Harvard, Cambridge, MA, USA; Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA, USA, [33]Institute for Cardiogenetics, University of Lübeck, Lübeck, Germany, [34]Atherosclerosis Research Unit, Department of Medicine Solna, Karolinska Institutet, Stockholm, Sweden Center for Molecular Medicine and Department of Cardiology,

Karolinska University Hospital, Stockholm, Sweden, [35]deCODE genetics/Amgen, Inc.Reykjavik, Iceland, [36]Department of Population Medicine, Qatar University College of Medicine, Doha, Qatar, [37]deCODE genetics/Amgen, Inc. Reykjavik, Iceland, [38]Department of Vascular Medicine, Academic Medical Center, Amsterdam, The Netherlands, [39]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden, [40]Laboratory for Statistical and Translational Genetics, RIKEN Center for Integrative Medical Sciences,Yokohama Kanagawa, Japan, [41]Laboratory for Cardiovascular Genomics and Informatics, RIKEN Center for Integrative Medical Sciences, Yokohama Kanagawa, Japan, [42]BHF Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge UK, [43]TIMI Study Group, Division of Cardiovascular Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, USA, [44]Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Tokyo, Japan, [45]German Heart Centre Munich, Department of Cardiology, Technical University of Munich; German Research Center for Cardiovascular Research (DZHK e.V.), partner site Munich Heart Alliance,Munich Germany, [46]Verve Therapeutics, Cambridge, MA, USA; Cardiovascular Research Center and Center for Human Genetic Research, Massachusetts General Hospital, Boston, Massachusetts, USA; Broad Institute, Cambridge, Massachusetts, USA; Department of Medicine, Harvard Medical School, Boston, Massachusetts, USA, [47]Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA, USA; Cardiovascular Disease Initiative, Broad Institute of MIT and Harvard, Cambridge, MA, USA; Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge MA, USA, [48]Division of Structural and Functional Genomics Center for Genome Science, Korea National Institute of Health, Korea, [49]Department of Cardiovascular Medicine, The University of Tokyo, Tokyo, Japan, [50]National Heart and Lung Institute (NHLI), Imperial College London, Hammersmith Hospital, London, UK., [51]Department of Cardiovascular Medicine, Mayo Clinic, MN, USA, [52]Wellcome Centre for Human Genetics, University of Oxford, Oxford, UK, [53]Department of Clinical Chemistry, Fimlab Laboratories, Tampere University Hospital, Tampere, Finland; Department of Clinical Chemistry, University of Tampere School of Medicine, Tampere, Finland, [54]Department of Genetics and Genomic Science, Icahn Institute for Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai; The Zena and Michael A. Wiener Cardiovascular Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA, [55]Synlab Academy, Mannheim, Germany; Medical Clinic V (Nephrology, Hypertensiology, Endocrinology, Diabetology, Rheumatology), Medical Faculty of Mannheim, University of Heidelberg, Mannheim, Germany, [56]Department of Medicine, Division of Cardiology, University of Ottawa Heart Institute, Ottawa, Canada, [57]Institute of Human Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany; Klinikum rechts der IsarInstitute of Human Genetics, Technical University of Munich, Munich, Germany, German Research Center for Cardiovascular Research (DZHK e.V.), partner site Munich Heart Alliance, Munich, Germany, [58]Department of Clinical Sciences in Malmö, Lund University, Malmo, Sweden, [59]Estonian Genome Center, University of Tartu, Tartu, Estonia; Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia, [60]Institute for Cardiogenetics, University of Lübeck; DZHK (German Research Center for Cardiovascular Research), partner site Hamburg-Lübeck-Kiel, Lübeck, Germany, [61]Department of Cardiovascular Sciences and NIHR Leicester Biomedical Research Centre, University of Leicester, Leicester UK, [62]Department of Internal Medicine, CardiologyUniversity of Michigan, Ann Arbor, Michigan, US; K.G. Jebsen Center for Genetic Epidemiology, Department of Public Health and Nursing, Norwegian University of Science and Technology, NTNU, Trondheim, Norway, [63]Division of Cardiology, Department of Medicine, Helsinki University Central Hospital (HUCH), Helsinki, Finland, [64]School of Medicine & University Hospital Bonn, Institute of Human Genetics, University of Bonn, Bonn, Germany, [65]Framingham Heart Study, National Heart, Lung and Blood Institute, Framingham, MA 01702, USA, [66]Medical Research Institute, University of Dundee, Ninewells Hospital and Medical School, Dundee, UK, [67]Department of Chronic Disease Prevention, National Institute for Health and Welfare, Helsinki, Finland, [68]Second Department of Cardiology, Medical School, National and Kapodistrian University of Athens, University General Hospital Attikon, Athens, Greece, [69]Cardiovascular Institute, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, USA, [70]Department of Chronic Disease Prevention, National Institute for Health and Welfare, Helsinki FIN-

00271, Finland; Wellcome Trust Sanger Institute, Hinxton, UK, [71]Cardiovascular Genomics & Genetics, University of Arizona, College of Medicine, Phoenix, Arizona, USA; University of Ottawa Heart Institute, Cardiovascular Research Methods Centre and Ruddy Canadian Cardiovascular Genetics Centre, Ontario, Canada, [72]Cardiovascular Disease Initiative, Broad Institute of MIT and Harvard, Cambridge, MA, USA, [73]Department of Cardiac Surgery and The Heart Clinic, Tartu University Hospital, Tartu, Estonia, [74]Center for Non-Communicable Diseases, Karachi, Pakistan; Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK; Department of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA, [75]Department of Cardiovascular Sciences and NIHR Leicester Biomedical Research Centre, University of Leicester, Leicester, UK, [76]MRC Epidemiology Unit, Institute of Metabolic Science, Addenbrooke's Hospital, Cambridge, UK; Wellcome Trust Sanger Institute, Hinxton, UK, [77]Cardiovascular Disease Initiative, Broad Institute, [78]Institut für Klinische Molekularbiologie, Christian-Albrechts Universität, Kiel, Germany, [79]German Heart Centre Munich, Department of Cardiology, Technical University of Munich, Munich, Germany; German Research Center for Cardiovascular Research (DZHK e.V.), partner site Munich Heart Alliance, Munich, Germany, [80]Medical Research Council Population Health Research Unit, CTSU - Nuffield Department of Population Health, Medical Sciences Division, University of Oxford, Oxford, UK, [81]Uppsala Clinical Research Center, Uppsala University, Uppsala, Sweden, [82]deCODE genetics/Amgen, Inc. Reykjavik, Iceland; Faculty of Medicine, University of Iceland, Reykjavik, Iceland, [83]Department of Internal Medicine, Cardiology, University of Michigan, Ann Arbor, Michigan USA, [84]deCODE genetics/Amgen, Inc., Faculty of Medicine, University of Iceland; Department of Internal Medicine, Division of Cardiology, Landspitali – National University Hospital of Iceland, Reykjavik, Iceland, [85]deCODE genetics/Amgen, Inc.; Faculty of Medicine, University of Iceland, Reykjavik, Iceland, [86]Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA, USA; Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, Massachusetts, USA; Program in Biological and Biomedical Sciences, Harvard Medical School, Boston, Massachusetts, USA;, [87]German Heart Centre Munich, Department of Cardiology, Technical University of Munich; German Research Center for Cardiovascular Research (DZHK e.V.), partner site Munich Heart Alliance, Munich, Germany, [88]Cardiovascular Disease Initiative, Broad Institute of MIT and Harvard; Program in Medical and Population Genetics Broad Institute of MIT and Harvard, Cambridge MA USA, [89]Program in Medical and Population Genetics, Broad Institute of MIT and Harvard,Cambridge MA USA, [90]Department of Internal Medicine, Cardiology, University of Michigan, Ann Arbor, Michigan, USA; Department of Computational Medicine & Bioinformatics, University of Michigan, Ann Arbor, Michigan, USA; Department of Human Genetics, University of Michigan, Ann Arbor, Michigan, USA, [91]Department of Computational Medicine & Bioinformatics, University of Michigan, Ann Arbor Michigan USA, [92]University of Balamand, East Med Res Institute, School of Medicine, Amioun, Lebanon; Lebanese American University, Chouran, Beirut, Lebanon; Harvard T.H.Chan School of Public Health, Boston, MA, USA, [93]Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge MA USA; Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, Massachusetts, USA; Stanley Center for Psychiatric Research Broad Institute of MIT and Harvard, Cambridge, MA, USA; Department of Computational Medicine & Bioinformatics, University of Michigan, Ann Arbor, Michigan, USA

VA Million Veteran Program Core Acknowledgement for Publications
Last updated August 3, 2021

MVP Executive Committee

- Co-Chair: J. Michael Gaziano, M.D., M.P.H.
  VA Boston Healthcare System, 150 S. Huntington Avenue, Boston, MA 02130
- Co-Chair: Sumitra Muralidhar, Ph.D.
  US Department of Veterans Affairs, 810 Vermont Avenue NW, Washington, DC 20420
- Rachel Ramoni, D.M.D., Sc.D., Chief VA Research and Development Officer
  US Department of Veterans Affairs, 810 Vermont Avenue NW, Washington, DC 20420
- Jean Beckham, Ph.D.
  Durham VA Medical Center, 508 Fulton Street, Durham, NC 27705
- Kyong-Mi Chang, M.D.
  Philadelphia VA Medical Center, 3900 Woodland Avenue, Philadelphia, PA 19104
- Philip S. Tsao, Ph.D.
  VA Palo Alto Health Care System, 3801 Miranda Avenue, Palo Alto, CA 94304
- James Breeling, M.D., Ex-Officio
  US Department of Veterans Affairs, 810 Vermont Avenue NW, Washington, DC 20420
- Grant Huang, Ph.D., Ex-Officio
  US Department of Veterans Affairs, 810 Vermont Avenue NW, Washington, DC 20420
- Juan P. Casas, M.D., Ph.D., Ex-Officio
  VA Boston Healthcare System, 150 S. Huntington Avenue, Boston, MA 02130

MVP Program Office

- Sumitra Muralidhar, Ph.D.
  US Department of Veterans Affairs, 810 Vermont Avenue NW, Washington, DC 20420
- Jennifer Moser, Ph.D.
  US Department of Veterans Affairs, 810 Vermont Avenue NW, Washington, DC 20420

MVP Recruitment/Enrollment

- MVP Cohort Management Director/Recruitment/Enrollment Director, Boston – Stacey B. Whitbourne, Ph.D.; Jessica V. Brewer, M.P.H.
  VA Boston Healthcare System, 150 S. Huntington Avenue, Boston, MA 02130
- VA Central Biorepository, Boston – Mary T. Brophy M.D., M.P.H.; Donald E. Humphries, Ph.D.; Luis E. Selva, Ph.D.
  VA Boston Healthcare System, 150 S. Huntington Avenue, Boston, MA 02130
- MVP Informatics, Boston – Nhan Do, M.D.; Shahpoor (Alex) Shayan, M.S.
  VA Boston Healthcare System, 150 S. Huntington Avenue, Boston, MA 02130
- MVP Data Operations/Analytics, Boston – Kelly Cho, M.P.H., Ph.D.
  VA Boston Healthcare System, 150 S. Huntington Avenue, Boston, MA 02130
- Director of Regulatory Affairs – Lori Churby, B.S.
  VA Palo Alto Health Care System, 3801 Miranda Avenue, Palo Alto, CA 94304
- MVP Coordinating Centers
  - Cooperative Studies Program Clinical Research Pharmacy Coordinating Center, Albuquerque – Todd Connor, Pharm.D.; Dean P. Argyres, B.S., M.S.

New Mexico VA Health Care System, 1501 San Pedro Drive SE, Albuquerque, NM 87108

- Genomics Coordinating Center, Palo Alto – Philip S. Tsao, Ph.D.
  VA Palo Alto Health Care System, 3801 Miranda Avenue, Palo Alto, CA 94304
- MVP Boston Coordinating Center, Boston - J. Michael Gaziano, M.D., M.P.H.
  VA Boston Healthcare System, 150 S. Huntington Avenue, Boston, MA 02130
- MVP Information Center, Canandaigua – Brady Stephens, M.S.
  Canandaigua VA Medical Center, 400 Fort Hill Avenue, Canandaigua, NY 14424

MVP Science

Saiju Pyarajan Ph.D.
VA Boston Healthcare System, 150 S. Huntington Avenue, Boston, MA 02130
Philip S. Tsao, Ph.D.
VA Palo Alto Health Care System, 3801 Miranda Avenue, Palo Alto, CA 94304

- Data Core - Kelly Cho, M.P.H, Ph.D.
  VA Boston Healthcare System, 150 S. Huntington Avenue, Boston, MA 02130
- VA Informatics and Computing Infrastructure (VINCI) – Scott L. DuVall, Ph.D.
  VA Salt Lake City Health Care System, 500 Foothill Drive, Salt Lake City, UT 84148
- Data and Computational Sciences – Saiju Pyarajan, Ph.D.
  VA Boston Healthcare System, 150 S. Huntington Avenue, Boston, MA 02130
- Statistical Genetics – Elizabeth Hauser, Ph.D.
  Durham VA Medical Center, 508 Fulton Street, Durham, NC 27705
  Yan Sun, Ph.D.
  Atlanta VA Medical Center, 1670 Clairmont Road, Decatur, GA 30033
  Hongyu Zhao, Ph.D.
  West Haven VA Medical Center, 950 Campbell Avenue, West Haven, CT 06516

Current MVP Local Site Investigators

- Atlanta VA Medical Center (Peter Wilson, M.D.)
  1670 Clairmont Road, Decatur, GA 30033
- Bay Pines VA Healthcare System (Rachel McArdle, Ph.D.)
  10,000 Bay Pines Blvd Bay Pines, FL 33744
- Birmingham VA Medical Center (Louis Dellitalia, M.D.)
  700 S. 19th Street, Birmingham AL 35233
- Central Western Massachusetts Healthcare System (Kristin Mattocks, Ph.D., M.P.H.)
  421 North Main Street, Leeds, MA 01053
- Cincinnati VA Medical Center (John Harley, M.D., Ph.D.)
  3200 Vine Street, Cincinnati, OH 45220
- Clement J. Zablocki VA Medical Center (Jeffrey Whittle, M.D., M.P.H.)
  5000 West National Avenue, Milwaukee, WI 53295
- VA Northeast Ohio Healthcare System (Frank Jacono, M.D.)
  10701 East Boulevard, Cleveland, OH 44106
- Durham VA Medical Center (Jean Beckham, Ph.D.)
  508 Fulton Street, Durham, NC 27705
- Edith Nourse Rogers Memorial Veterans Hospital (John Wells., Ph.D.)
  200 Springs Road, Bedford, MA 01730

- Edward Hines, Jr. VA Medical Center (Salvador Gutierrez, M.D.)
  5000 South 5th Avenue, Hines, IL 60141
- Veterans Health Care System of the Ozarks (Kathrina Alexander, M.D.)
  1100 North College Avenue, Fayetteville, AR 72703
- Fargo VA Health Care System (Kimberly Hammer, Ph.D.)
  2101 N. Elm, Fargo, ND 58102
- VA Health Care Upstate New York (James Norton, Ph.D.)
  113 Holland Avenue, Albany, NY 12208
- New Mexico VA Health Care System (Gerardo Villareal, M.D.)
  1501 San Pedro Drive, S.E. Albuquerque, NM 87108
- VA Boston Healthcare System (Scott Kinlay, M.B.B.S., Ph.D.)
  150 S. Huntington Avenue, Boston, MA 02130
- VA Western New York Healthcare System (Junzhe Xu, M.D.)
  3495 Bailey Avenue, Buffalo, NY 14215-1199
- Ralph H. Johnson VA Medical Center (Mark Hamner, M.D.)
  109 Bee Street, Mental Health Research, Charleston, SC 29401
- Columbia VA Health Care System (Roy Mathew, M.D.)
  6439 Garners Ferry Road, Columbia, SC 29209
- VA North Texas Health Care System (Sujata Bhushan, M.D.)
  4500 S. Lancaster Road, Dallas, TX 75216
- Hampton VA Medical Center (Pran Iruvanti, D.O., Ph.D.)
  100 Emancipation Drive, Hampton, VA 23667
- Richmond VA Medical Center (Michael Godschalk, M.D.)
  1201 Broad Rock Blvd., Richmond, VA 23249
- Iowa City VA Health Care System (Zuhair Ballas, M.D.)
  601 Highway 6 West, Iowa City, IA 52246-2208
- Eastern Oklahoma VA Health Care System (River Smith, Ph.D.)
  1011 Honor Heights Drive, Muskogee, OK 74401
- James A. Haley Veterans' Hospital (Stephen Mastorides, M.D.)
  13000 Bruce B. Downs Blvd, Tampa, FL 33612
- James H. Quillen VA Medical Center (Jonathan Moorman, M.D., Ph.D.)
  Corner of Lamont & Veterans Way, Mountain Home, TN 37684
- John D. Dingell VA Medical Center (Saib Gappy, M.D.)
  4646 John R Street, Detroit, MI 48201
- Louisville VA Medical Center (Jon Klein, M.D., Ph.D.)
  800 Zorn Avenue, Louisville, KY 40206
- Manchester VA Medical Center (Nora Ratcliffe, M.D.)
  718 Smyth Road, Manchester, NH 03104
- Miami VA Health Care System (Ana Palacio, M.D., M.P.H.)
  1201 NW 16th Street, 11 GRC, Miami FL 33125
- Michael E. DeBakey VA Medical Center (Olaoluwa Okusaga, M.D.)
  2002 Holcombe Blvd, Houston, TX 77030
- Minneapolis VA Health Care System (Maureen Murdoch, M.D., M.P.H.)
  One Veterans Drive, Minneapolis, MN 55417

- N. FL/S. GA Veterans Health System (Peruvemba Sriram, M.D.)
  1601 SW Archer Road, Gainesville, FL 32608
- Northport VA Medical Center (Shing Shing Yeh, Ph.D., M.D.)
  79 Middleville Road, Northport, NY 11768
- Overton Brooks VA Medical Center (Neeraj Tandon, M.D.)
  510 East Stoner Ave, Shreveport, LA 71101
- Philadelphia VA Medical Center (Darshana Jhala, M.D.)
  3900 Woodland Avenue, Philadelphia, PA 19104
- Phoenix VA Health Care System (Samuel Aguayo, M.D.)
  650 E. Indian School Road, Phoenix, AZ 85012
- Portland VA Medical Center (David Cohen, M.D.)
  3710 SW U.S. Veterans Hospital Road, Portland, OR 97239
- Providence VA Medical Center (Satish Sharma, M.D.)
  830 Chalkstone Avenue, Providence, RI 02908
- Richard Roudebush VA Medical Center (Suthat Liangpunsakul, M.D., M.P.H.)
  1481 West 10th Street, Indianapolis, IN 46202
- Salem VA Medical Center (Kris Ann Oursler, M.D.)
  1970 Roanoke Blvd, Salem, VA 24153
- San Francisco VA Health Care System (Mary Whooley, M.D.)
  4150 Clement Street, San Francisco, CA 94121
- South Texas Veterans Health Care System (Sunil Ahuja, M.D.)
  7400 Merton Minter Boulevard, San Antonio, TX 78229
- Southeast Louisiana Veterans Health Care System (Joseph Constans, Ph.D.)
  2400 Canal Street, New Orleans, LA 70119
- Southern Arizona VA Health Care System (Paul Meyer, M.D., Ph.D.)
  3601 S 6th Avenue, Tucson, AZ 85723
- Sioux Falls VA Health Care System (Jennifer Greco, M.D.)
  2501 W 22nd Street, Sioux Falls, SD 57105
- St. Louis VA Health Care System (Michael Rauchman, M.D.)
  915 North Grand Blvd, St. Louis, MO 63106
- Syracuse VA Medical Center (Richard Servatius, Ph.D.)
  800 Irving Avenue, Syracuse, NY 13210
- VA Eastern Kansas Health Care System (Melinda Gaddy, Ph.D.)
  4101 S 4th Street Trafficway, Leavenworth, KS 66048
- VA Greater Los Angeles Health Care System (Agnes Wallbom, M.D., M.S.)
  11301 Wilshire Blvd, Los Angeles, CA 90073
- VA Long Beach Healthcare System (Timothy Morgan, M.D.)
  5901 East 7th Street Long Beach, CA 90822
- VA Maine Healthcare System (Todd Stapley, D.O.)
  1 VA Center, Augusta, ME 04330
- VA New York Harbor Healthcare System (Peter Liang, M.D., M.P.H.)
  423 East 23rd Street, New York, NY 10010
- VA Pacific Islands Health Care System (Daryl Fujii, Ph.D.)
  459 Patterson Rd, Honolulu, HI 96819

- VA Palo Alto Health Care System (Philip Tsao, Ph.D.)
  3801 Miranda Avenue, Palo Alto, CA 94304-1290
- VA Pittsburgh Health Care System (Patrick Strollo, Jr., M.D.)
  University Drive, Pittsburgh, PA 15240
- VA Puget Sound Health Care System (Edward Boyko, M.D.)
  1660 S. Columbian Way, Seattle, WA 98108-1597
- VA Salt Lake City Health Care System (Jessica Walsh, M.D.)
  500 Foothill Drive, Salt Lake City, UT 84148
- VA San Diego Healthcare System (Samir Gupta, M.D., M.S.C.S.)
  3350 La Jolla Village Drive, San Diego, CA 92161
- VA Sierra Nevada Health Care System (Mostaqul Huq, Pharm.D., Ph.D.)
  975 Kirman Avenue, Reno, NV 89502
- VA Southern Nevada Healthcare System (Joseph Fayad, M.D.)
  6900 North Pecos Road, North Las Vegas, NV 89086
- VA Tennessee Valley Healthcare System (Adriana Hung, M.D., M.P.H.)
  1310 24th Avenue, South Nashville, TN 37212
- Washington DC VA Medical Center (Jack Lichy, M.D., Ph.D.)
  50 Irving St, Washington, D. C. 20422
- W.G. (Bill) Hefner VA Medical Center (Robin Hurley, M.D.)
  1601 Brenner Ave, Salisbury, NC 28144
- White River Junction VA Medical Center (Brooks Robey, M.D.)
  163 Veterans Drive, White River Junction, VT 05009
- William S. Middleton Memorial Veterans Hospital (Prakash Balasubramanian, M.D.)
  2500 Overlook Terrace, Madison, WI 53705

References

1. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. *Am J Epidemiol* **129**, 687-702 (1989).
2. Tayo, B.O.*, et al.* Genetic Background of Patients from a University Medical Center in Manhattan: Implications for Personalized Medicine. *PLoS ONE* **6**, e19166 (2011).
3. Damrauer, S.M.*, et al.* Association of the V122I Hereditary Transthyretin Amyloidosis Genetic Variant With Heart Failure Among Individuals of African or Hispanic/Latino Ancestry. *JAMA* **322**, 2191-2202 (2019).
4. Wojcik, G.L.*, et al.* Genetic analyses of diverse populations improves discovery for complex traits. *Nature* **570**, 514-518 (2019).
5. Kolonel, L.N.*, et al.* A multiethnic cohort in Hawaii and Los Angeles: baseline characteristics. *Am J Epidemiol* **151**, 346-357 (2000).
6. Kolonel, L.N., Altshuler, D. & Henderson, B.E. The multiethnic cohort study: exploring genes, lifestyle and cancer risk. *Nature Reviews Cancer* **4**, 519-527 (2004).
7. Han, Y.*, et al.* Integration of multiethnic fine-mapping and genomic annotation to prioritize candidate functional SNPs at prostate cancer susceptibility regions. *Hum Mol Genet* **24**, 5603-5618 (2015).
8. Feng, Y.*, et al.* A comprehensive examination of breast cancer risk loci in African American women. *Hum Mol Genet* **23**, 5518-5526 (2014).
9. Han, Y.*, et al.* Generalizability of established prostate cancer risk variants in men of African ancestry. *Int J Cancer* **136**, 1210-1217 (2015).
10. Stram, D.O.*, et al.* Racial/Ethnic Differences in Lung Cancer Incidence in the Multiethnic Cohort Study: An Update. *J Natl Cancer Inst* **111**, 811-819 (2019).
11. Murphy, S.E.*, et al.* Nicotine N-glucuronidation relative to N-oxidation and C-oxidation and UGT2B10 genotype in five ethnic/racial groups. *Carcinogenesis* **35**, 2526-2533 (2014).
12. Ke, W.*, et al.* Evaluation of 71 Coronary Artery Disease Risk Variants in a Multiethnic Cohort. *Front Cardiovasc Med* **5**, 19 (2018).
13. Design of the Women's Health Initiative clinical trial and observational study. The Women's Health Initiative Study Group. *Control Clin Trials* **19**, 61-109 (1998).
14. Anderson, G.L.*, et al.* Implementation of the Women's Health Initiative study design. *Ann Epidemiol* **13**, S5-17 (2003).
15. Curb, J.D.*, et al.* Outcomes ascertainment and adjudication methods in the Women's Health Initiative. *Ann Epidemiol* **13**, S122-128 (2003).
16. Hays, J.*, et al.* The Women's Health Initiative recruitment methods and results. *Ann Epidemiol* **13**, S18-77 (2003).
17. Roden, D.M.*, et al.* Development of a large-scale de-identified DNA biobank to enable personalized medicine. *Clin Pharmacol Ther* **84**, 362-369 (2008).
18. Fried, L.P.*, et al.* The Cardiovascular Health Study: design and rationale. *Ann Epidemiol* **1**, 263-276 (1991).
19. Ives, D.G.*, et al.* Surveillance and ascertainment of cardiovascular events. *Annals of Epidemiology* **5**, 278-285 (1995).
20. Psaty, B.M.*, et al.* Methods of assessing prevalent cardiovascular disease in the Cardiovascular Health Study. *Annals of Epidemiology* **5**, 270-277 (1995).
21. Sempos, C.T., Bild, D.E. & Manolio, T.A. Overview of the Jackson Heart Study: a study of cardiovascular diseases in African American men and women. *Am J Med Sci* **317**, 142-146 (1999).
22. Taylor, H.A., Jr. The Jackson Heart Study: an overview. *Ethn Dis* **15**, S6-1-3 (2005).
23. Taylor, H.A., Jr.*, et al.* Toward resolution of cardiovascular health disparities in African Americans: design and methods of the Jackson Heart Study. *Ethn Dis* **15**, S6-4-17 (2005).

24. Das, S.*, et al.* Next-generation genotype imputation service and methods. *Nat Genet* **48**, 1284-1287 (2016).
25. Gottesman, O.*, et al.* The Electronic Medical Records and Genomics (eMERGE) Network: past, present, and future. *Genet Med* **15**, 761-771 (2013).
26. McCarty, C.A.*, et al.* The eMERGE Network: a consortium of biorepositories linked to electronic medical records data for conducting genomic studies. *BMC Med Genomics* **4**, 13 (2011).
27. McCarthy, S.*, et al.* A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet* (2016).
28. Stanaway, I.B.*, et al.* The eMERGE genotype set of 83,717 subjects imputed to ~40 million variants genome wide and association with the herpes zoster medical record phenotype. *Genet Epidemiol* **43**, 63-81 (2019).
29. Safarova, M.S., Liu, H. & Kullo, I.J. Rapid identification of familial hypercholesterolemia from electronic health records: The SEARCH study. *J Clin Lipidol* **10**, 1230-1239 (2016).
30. Manichaikul, A.*, et al.* Robust relationship inference in genome-wide association studies. *Bioinformatics* **26**, 2867-2873 (2010).
31. Verma, S.S.*, et al.* Imputation and quality control steps for combining multiple genome-wide datasets. *Front Genet* **5**, 370 (2014).
32. van der Harst, P. & Verweij, N. Identification of 64 Novel Genetic Loci Provides an Expanded View on the Genetic Architecture of Coronary Artery Disease. *Circ Res* **122**, 433-443 (2018).
33. Wain, L.V.*, et al.* Novel insights into the genetics of smoking behaviour, lung function, and chronic obstructive pulmonary disease (UK BiLEVE): a genetic association study in UK Biobank. *Lancet Respir Med* **3**, 769-781 (2015).
34. Bycroft, C.*, et al.* The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203-209 (2018).