

Supplemental information

Rare coding variants in 35 genes associate with circulating lipid levels—A multi-ancestry analysis of 170,000 exomes

George Hindy, Peter Dornbos, Mark D. Chaffin, Dajiang J. Liu, Minxian Wang, Margaret Sunitha Selvaraj, David Zhang, Joseph Park, Carlos A. Aguilar-Salinas, Lucinda Antonacci-Fulton, Diego Ardisino, Donna K. Arnett, Stella Aslibekyan, Gil Atzmon, Christie M. Ballantyne, Francisco Barajas-Olmos, Nir Barzilai, Lewis C. Becker, Lawrence F. Bielak, Joshua C. Bis, John Blangero, Eric Boerwinkle, Lori L. Bonnycastle, Erwin Bottinger, Donald W. Bowden, Matthew J. Bown, Jennifer A. Brody, Jai G. Broome, Noël P. Burt, Brian E. Cade, Federico Centeno-Cruz, Edmund Chan, Yi-Cheng Chang, Yii-Der I. Chen, Ching-Yu Cheng, Won Jung Choi, Rajiv Chowdhury, Cecilia Contreras-Cubas, Emilio J. Córdova, Adolfo Correa, L. Adrienne Cupples, Joanne E. Curran, John Danesh, Paul S. de Vries, Ralph A. DeFronzo, Harsha Doddapaneni, Ravindranath Duggirala, Susan K. Dutcher, Patrick T. Ellinor, Leslie S. Emery, Jose C. Florez, Myriam Fornage, Barry I. Freedman, Valentin Fuster, Ma. Eugenia Garay-Sevilla, Humberto García-Ortiz, Soren Germer, Richard A. Gibbs, Christian Gieger, Benjamin Glaser, Clicerio Gonzalez, Maria Elena Gonzalez-Villalpando, Mariaelisa Graff, Sarah E. Graham, Niels Grarup, Leif C. Groop, Xiuqing Guo, Namrata Gupta, Sohee Han, Craig L. Hanis, Torben Hansen, Jiang He, Nancy L. Heard-Costa, Yi-Jen Hung, Mi Yeong Hwang, Marguerite R. Irvin, Sergio Islas-Andrade, Gail P. Jarvik, Hyun Min Kang, Sharon L.R. Kardia, Tanika Kelly, Eimear E. Kenny, Alyna T. Khan, Bong-Jo Kim, Ryan W. Kim, Young Jin Kim, Heikki A. Koistinen, Charles Kooperberg, Johanna Kuusisto, Soo Heon Kwak, Markku Laakso, Leslie A. Lange, Jiwon Lee, Juyoung Lee, Seonwook Lee, Donna M. Lehman, Rozenn N. Lemaitre, Allan Linneberg, Jianjun Liu, Ruth J.F. Loos, Steven A. Lubitz, Valeriya Lyssenko, Ronald C.W. Ma, Lisa Warsinger Martin, Angélica Martínez-Hernández, Rasika A. Mathias, Stephen T. McGarvey, Ruth McPherson, James B. Meigs, Thomas Meitinger, Olle Melander, Elvia Mendoza-Caamal, Ginger A. Metcalf, Xuenan Mi, Karen L. Mohlke, May E. Montasser, Jee-Young Moon, Hortensia Moreno-Macías, Alanna C. Morrison, Donna M. Muzny, Sarah C. Nelson, Peter M. Nilsson, Jeffrey R. O'Connell, Marju Orho-Melander, Lorena Orozco, Colin N.A. Palmer, Nicholette D. Palmer, Cheol Joo Park, Kyong Soo Park, Oluf Pedersen, Juan M. Peralta, Patricia A. Peyser, Wendy S. Post, Michael Preuss, Bruce M. Psaty, Qibin Qi, D.C. Rao, Susan Redline, Alexander P. Reiner, Cristina Rella-Monsalve, Stephen S. Rich, Nilesh Samani, Heribert Schunkert, Claudia Schurmann, Daekwan Seo, Jeong-Sun Seo, Xueling Sim, Rob Sladek, Kerrin S. Small, Wing Yee So, Adrienne M. Stilp, E. Shyong Tai, Claudia H.T. Tam, Kent D. Taylor, Yik Ying Teo, Farook Thameem, Brian Tomlinson, Michael Y. Tsai, Tiinamaija Tuomi, Jaakko Tuomilehto, Teresa Tusié-Luna, Miriam S. Udler, Rob M. van Dam, Ramachandran S. Vasan, Karine A. Viaud Martinez, Fei Fei Wang, Xuzhi Wang, Hugh Watkins, Daniel E. Weeks, James G.

The American Journal of Human Genetics, Volume 109

Supplemental information

**Rare coding variants in 35 genes associate
with circulating lipid levels—A multi-ancestry
analysis of 170,000 exomes**

Hindy, et al.

Figure S1. Comparison of effect sizes and p-value in UK Biobank including and excluding individuals on statin treatment.

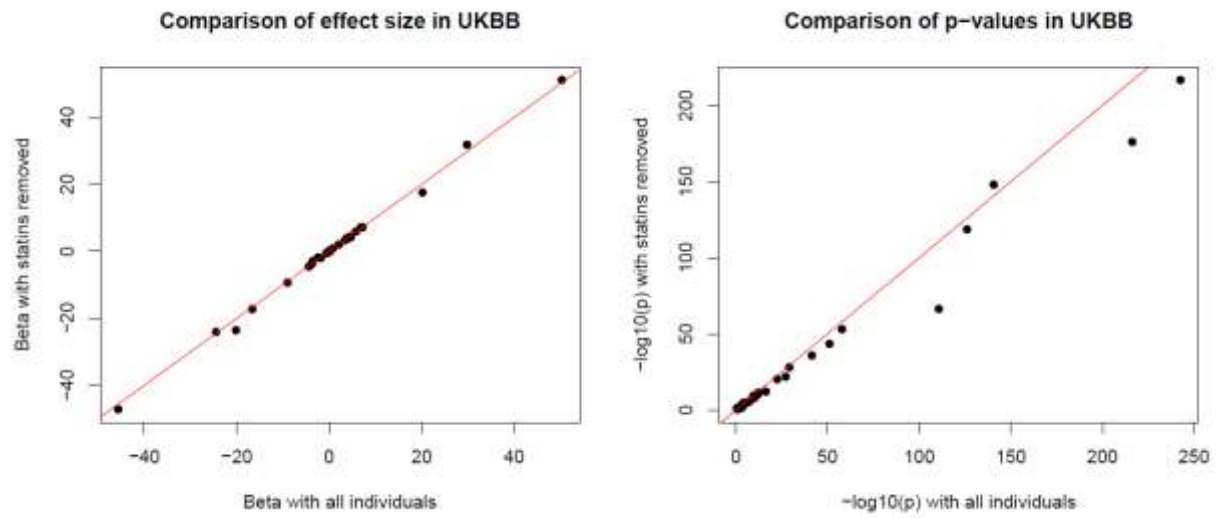


Figure S2. Descriptive variant characteristics by type, ancestry, and minor allele count.



Figure S2: A) Our study included 15,599,513 genetic variants. Variants were annotated as high confidence loss-of-function by LOFTEE ($n=340,214$), splice site altering variants using a deep neural network prediction (SPICE AI) ($n=238,646$), damaging missense variants according to the MetaSVM algorithm ($n=729,098$) and damaging missing in 5 out of 5 prediction algorithms ($n=1,106,309$). Most of the variants had a minor allele count of less than 5 in all ($n=1,171,5189$) and within each of the four different annotations. **B)** The proportion of specific annotations out of the total number of variants that were annotated as coding ($n=5,085,712$). Each of the four annotations demonstrated the highest enrichment among the variants with the lowest frequency. ALL=multi-ancestry, AFR=African ancestry, EAS=East Asian ancestry, EUR=European ancestry, HIS=Hispanic ancestry, SAS=South Asian ancestry.

Figure S3. Overlap among different ancestries for all variants contributing to significant gene-based associations with HDL-C, TG and LDL-C

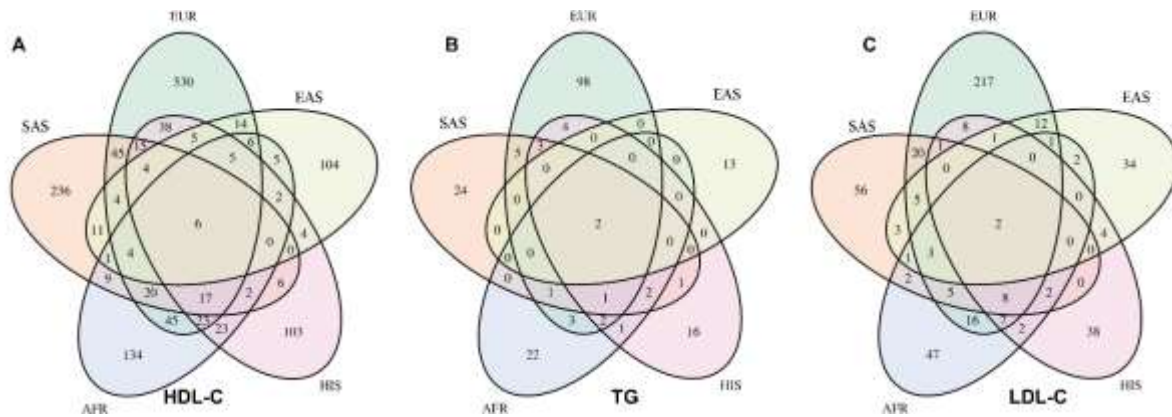


Figure S3. Venn Diagram for the overlap of all variants included in the significant gene-based association analysis among different ancestries. **A)** A total of 4 genes (*CETP*, *ABCA1*, *CD36*, and *LCAT*) showed significant gene-based associations in multi-ancestry and each of the 5 different ancestries analyses with $P < 0.05$ with HDL cholesterol (HDL-C). Ancestry-specific single-variant contributions included a total of 781 from European-, 380 from South Asian-, 302 African-, 253 Hispanic- and 175 East Asian ancestries. **B)** A total of 3 genes (*APOC3*, *ANGPTL3*, and *APOB*) showed significant gene-based associations in multi-ancestry and each of the 5 different ancestries analyses with $P < 0.05$ with triglycerides (TG). Ancestry-specific single-variant contributions included a total of 119 from European-, 39 from South Asian-, 34 African-, 32 Hispanic- and 15 East Asian ancestries. **C)** A total of 3 genes (*LDLR*, *PCSK9*, and *APOB*) showed significant gene-based associations in multi-ancestry and each of the 5 different ancestries analyses with $P < 0.05$ with LDL cholesterol (LDL-C). Ancestry-specific single-variant contributions included a total of 306 from European-, 108 from South Asian-, 98 African-, 73 Hispanic- and 68 East Asian ancestries.

Figure S4. Overlap among different ancestries for variants contributing to significant gene-based associations with HDL-C, TG and LDL-C with P value less than 0.05

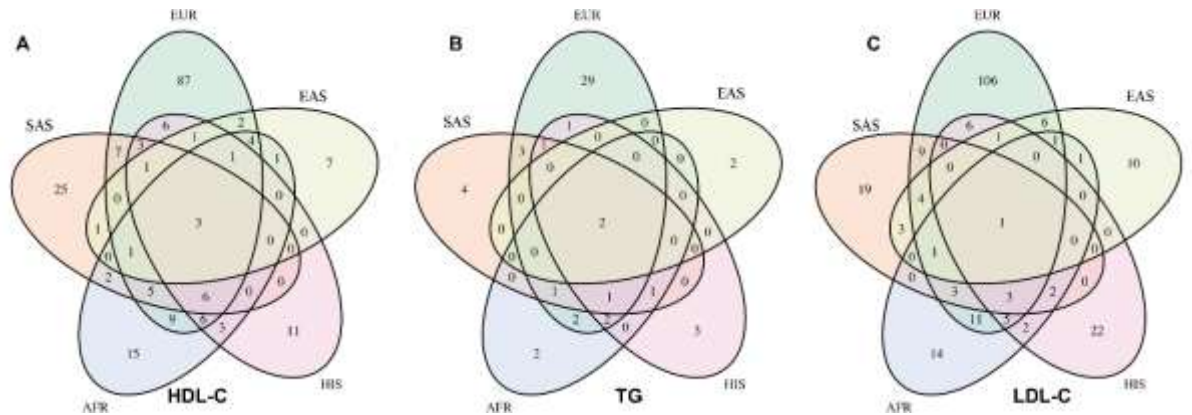


Figure S4. Venn Diagram for the overlap of variants with $P < 0.05$ included in the significant gene-based association analysis among different ancestries. **A)** A total of 4 genes (*CETP*, *ABCA1*, *CD36*, and *LCAT*) showed significant gene-based associations in multi-ancestry and each of the 5 different ancestries analyses with $P < 0.05$ with HDL cholesterol (HDL-C). Ancestry-specific single-variant contributions included a total of 142 from European-, 54 from South Asian-, 56 African-, 41 Hispanic- and 22 East Asian ancestries. **B)** A total of 3 genes (*APOC3*, *ANGPTL3*, and *APOB*) showed significant gene-based associations in multi-ancestry analyses with triglycerides (TG). Ancestry-specific single-variant contributions included a total of 42 from European-, 13 from South Asian-, 11 African-, 11 Hispanic- and 4 East Asian ancestries. **C)** A total of 3 genes (*LDLR*, *PCSK9*, and *APOB*) showed significant gene-based associations in multi-ancestry and each of the 5 different ancestries analyses with $P < 0.05$ with LDL cholesterol (LDL-C). Ancestry-specific single-variant contributions included a total of 157 from European-, 45 from South Asian-, 44 African-, 42 Hispanic- and 28 East Asian ancestries.

Figure S5. Overlap among different ancestries for the top variant contributing to significant gene-based associations with HDL-C, TG and LDL-C

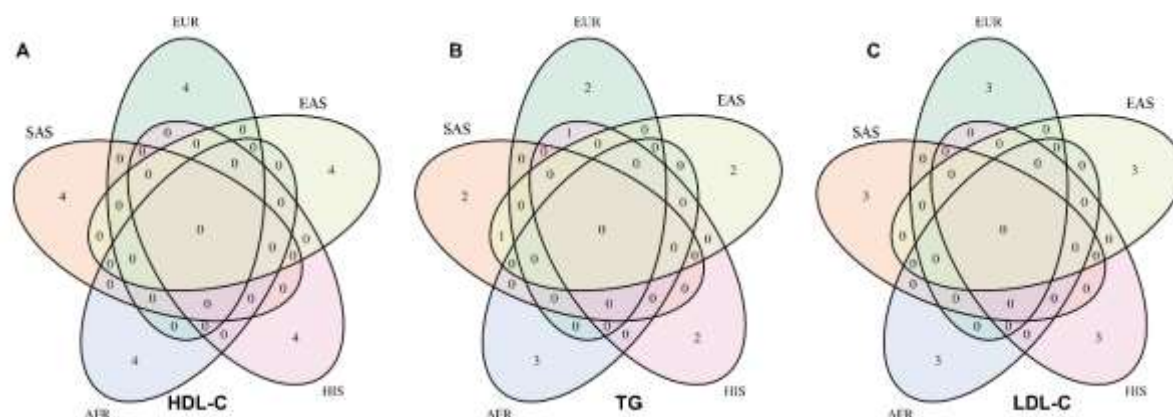


Figure S5. Venn Diagram for the overlap of the top variant included in each of the significant gene-based association analysis among different ancestries. **A)** 4 genes (*CETP*, *ABCA1*, *CD36*, and *LCAT*) showed significant gene-based associations in multi-ancestry and each of the 5 different ancestries analyses with $P < 0.05$ with HDL cholesterol (HDL-C). **B)** A total of 3 genes (*APOC3*, *ANGPTL3*, and *APOB*) showed significant gene-based associations in multi-ancestry analyses with triglycerides (TG). **C)** A total of 3 genes (*LDLR*, *PCSK9*, and *APOB*) showed significant gene-based associations in multi-ancestry and each of the 5 different ancestries analyses with $P < 0.05$ with LDL cholesterol (LDL-C).

Figure S6. Cumulative loss-of-function minor allele count and effect size on LDL cholesterol by ancestry

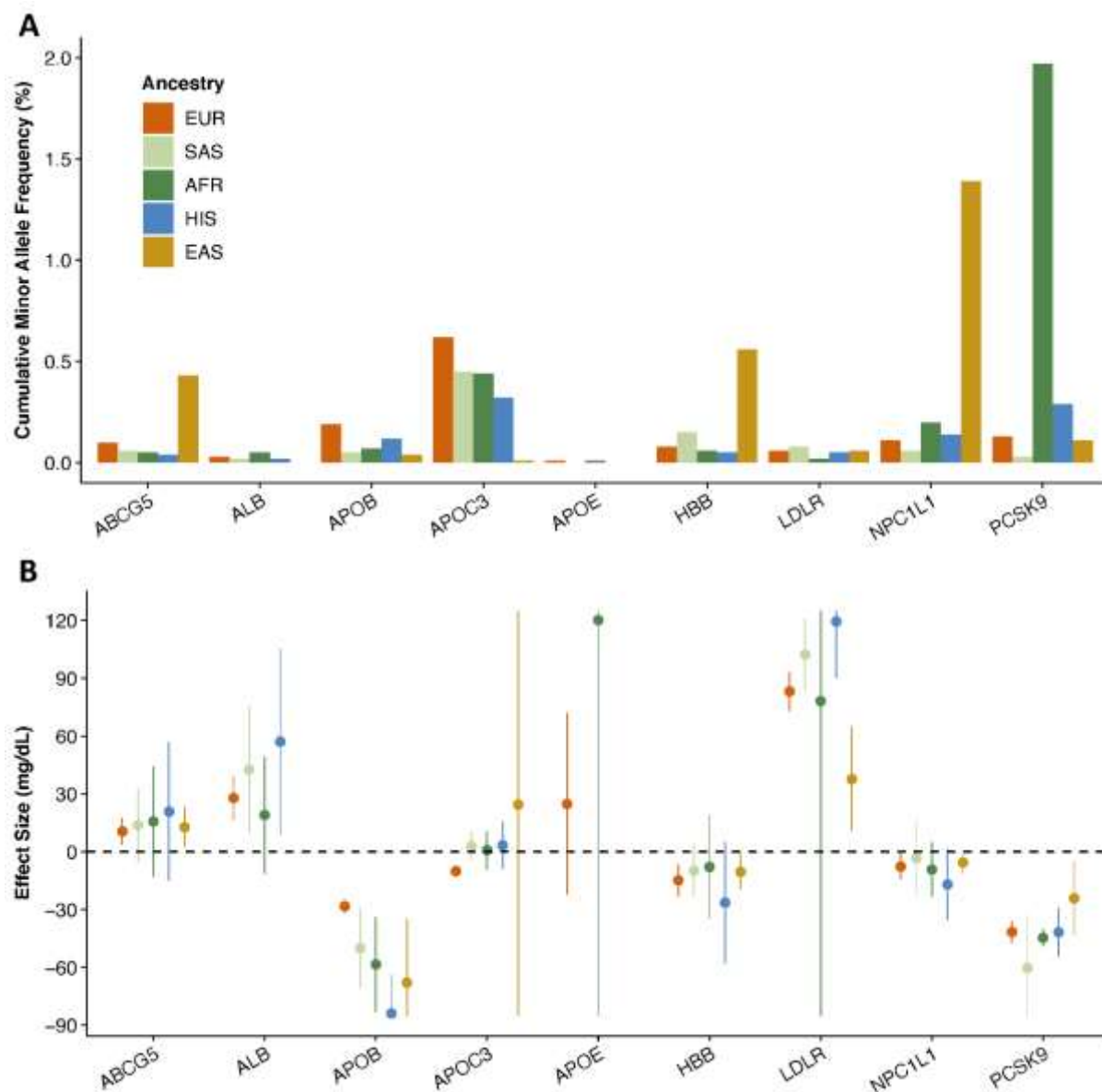


Figure S6: A) Cumulative minor allele frequencies and **B)** burden test effect sizes on LDL cholesterol levels for exome-wide significant genes ($P < 4.3 \times 10^{-7}$) within each of the five major ancestries using variants from the high confidence loss-of-function grouping (LOFTEE). AFR=African ancestry, EAS=East Asian ancestry, EUR=European ancestry, HIS=Hispanic ancestry, SAS=South Asian ancestry.

Figure S7. Cumulative loss-of-function minor allele count and effect size on triglycerides by ancestry

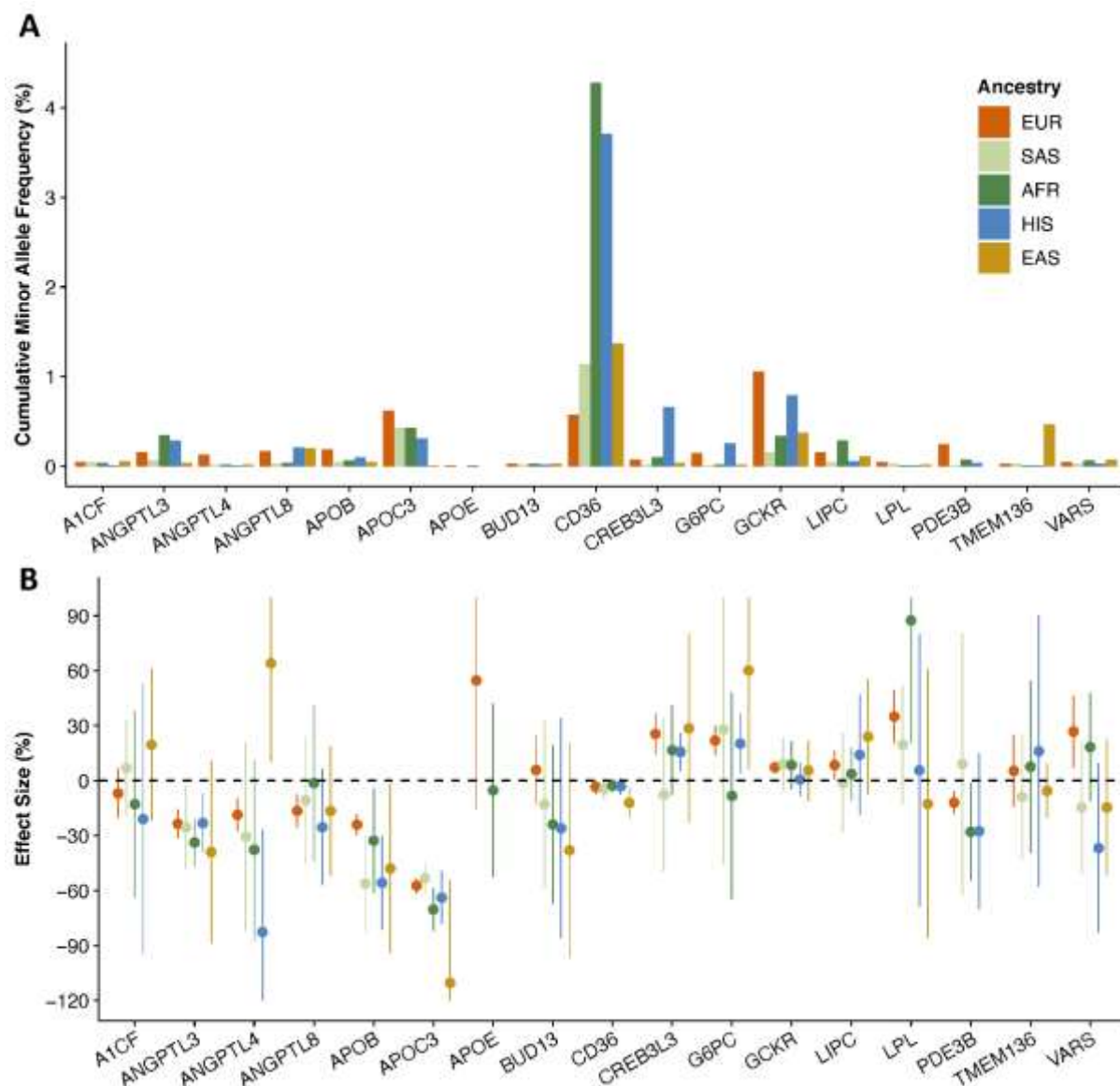


Figure S7: A) Cumulative minor allele frequencies and **B)** burden test effect sizes on triglyceride levels for exome-wide significant genes ($P < 4.3 \times 10^{-7}$) within each of the five major ancestries using variants from the high confidence loss-of-function grouping (LOFTEE). AFR=African ancestry, EAS=East Asian ancestry, EUR=European ancestry, HIS=Hispanic ancestry, SAS=South Asian ancestry.

Figure S8. Cumulative loss-of-function minor allele count and effect size on HDL cholesterol by ancestry

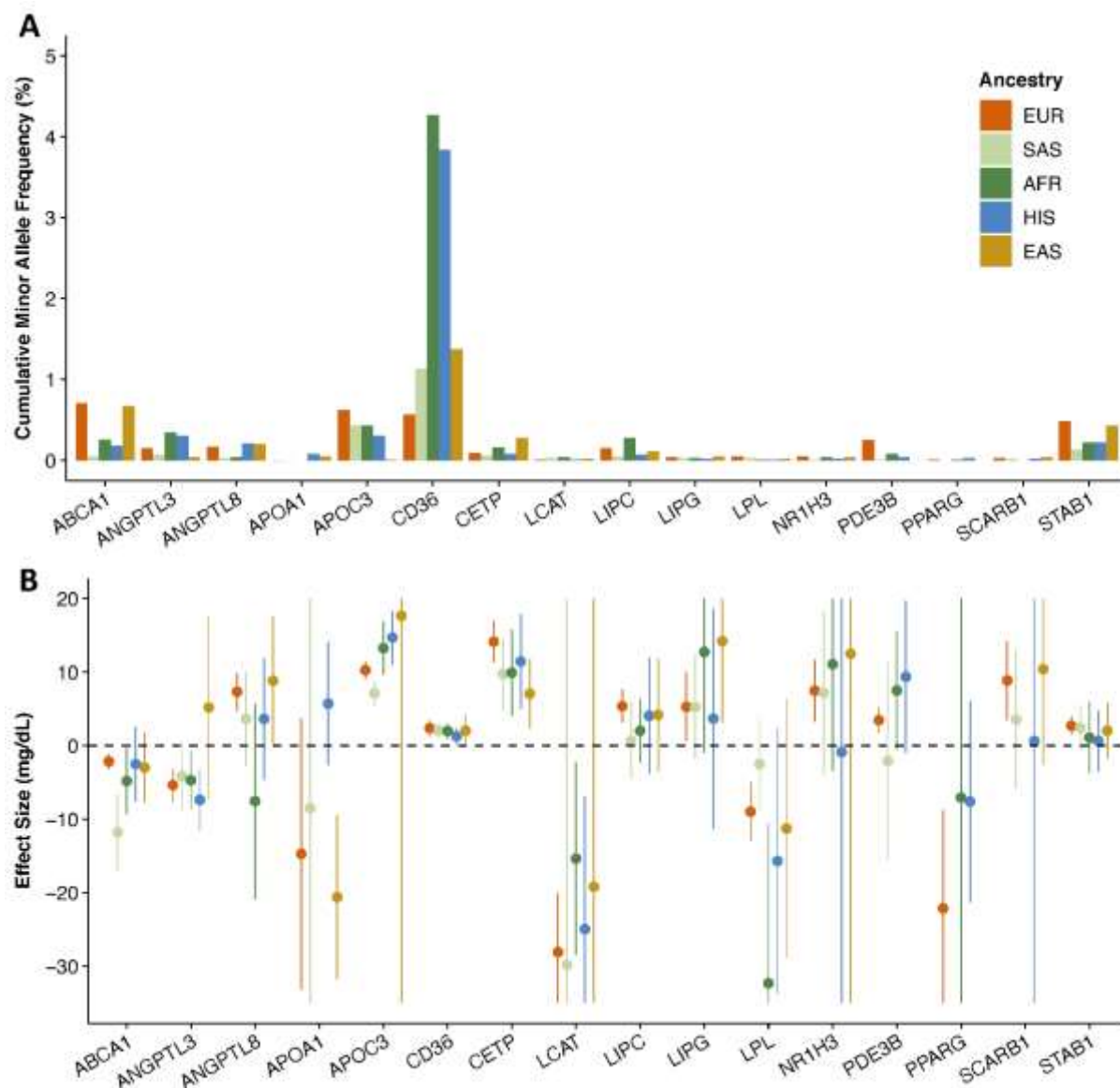


Figure S8: A) Cumulative minor allele frequencies and **B)** burden test effect sizes on HDL cholesterol levels for exome-wide significant genes ($P < 4.3 \times 10^{-7}$) within each of the five major ancestries using variants from the high confidence loss-of-function grouping (LOFTEE). AFR=African ancestry, EAS=East Asian ancestry, EUR=European ancestry, HIS=Hispanic ancestry, SAS=South Asian ancestry.

Supplemental Methods

Sequencing and Quality Control

Myocardial Infarction Genetics Consortium (MIGen)

A set of common variants was extracted for sample quality control including relative inference, principal component analysis, and estimation of heterozygosity. SNPs on autosomes and not in low complexity regions or segmental duplications were extracted. SNPs with quality of depth (QD) > 2 , call rate $>98\%$, self-reported-race-specific Hardy-Weinberg equilibrium (HWE) p-value $>1\times10^{-8}$, Variant Quality Score Recalibration (VQSR) of PASS and MAF $>1\%$ were retained. Sample relatedness was estimated with KING and duplicate samples removed. Genetically inferred ancestry was assigned to each individual by calculating principal components jointly with 1000 Genomes phase 3 version 5 and building a 5-Nearest Neighbor classifier¹ using the top 6 principal components. Heterozygosity was estimated within each genetic ancestry group and samples with F statistic above 0.3 were removed. Genetic sex was inferred based on high quality X-chromosome variation including variants with call rate >0.95 , MAF $>2\%$, a PASS VQSR, QD >3 if the variant is an insertion or deletion and QD >2 if it is SNP. Samples with discordant phenotypic sex and genetic sex were removed. Finally, sample quality control metrics were calculated using Hail and samples with call rate <0.9 , a mean depth (DP) <30 and mean genotype quality (GQ) <0.8 were excluded. A total of 44,240 samples with lipid data measurements were included after further excluding duplicates and relatives with other data sources (**Table S1**).

Variant quality control was performed amongst remaining samples and a total of 8,716,575 autosomal variants were included after removing those that fail HWE as

calculated by genetic ancestry group ($p\text{-value} < 1 \times 10^{-8}$), lie in low complexity regions or segmental duplications, with inbreeding coefficient < -0.3 , are insertions or deletions with $QD \leq 3$ or SNPs with $QD \leq 2$ or variants where VQSR does not PASS with the exception of singletons where variants with VQSRTrancheSNP99.60to99.80 were retained.

Trans-Omics for Precision Medicine (TOPMed)

Whole genome sequencing at 30X mean depth was performed at one of six sequencing centers: Broad Institute of MIT and Harvard, Northwest Genomics Center, New York Genome Center, Illumina Laboratory Services, Psomagen, Inc. (formerly Macrogen USA), Baylor College of Medicine Human Genome Sequencing Center. For most studies, all individuals in the study were sequenced at the same center. Sequence reads were aligned to human genome build GRCh37 or GRCh38 at each center using similar, but not identical, processing pipelines. The resulting sequence data files were transferred from all centers to the TOPMed Informatics Research Center (IRC), where they were re-aligned to build GRCh38, using a common pipeline to produce a set of 'harmonized' .cram files. Processing was coordinated and managed by the 'GotCloud' processing pipeline. The IRC performed joint genotype calling on all samples. Quality control was performed at each stage of the process by the Sequencing Centers, the IRC, and the TOPMed Data Coordinating Center (DCC). Only samples that passed QC were included in the call set.

The two sequence quality criteria that were used to pass sequence data on for joint variant discovery and genotyping are: estimated DNA sample contamination below

3%, and fraction of the genome covered at least 10x 95% or above. DNA sample contamination was estimated from the sequencing center read mapping using software verifyBamId.²

The genotype used for analysis are from “freeze 6a” of the variant calling pipeline performed by the TOPMed Informatics Research Center (Center for Statistical Genetics, University of Michigan, Hyun Min Kang, Tom Blackwell and Gonçalo Abecasis). Variant detection (SNPs and indels) from each sequenced (and aligned) genome was performed by the vt discover2 software tool. The variant calling software tools are under active development; updated versions can be accessed at <http://github.com/atks/vt>, <http://github.com/hyunminkang/apigenome>, and https://github.com/statgen/topmed_variant_calling.

One individual from duplicate pairs identified by the DCC was removed, retaining the individual with lipid levels available when one did not have lipid levels. If both individuals had lipid levels, one individual was randomly selected. Individuals were excluded when their genotype determined sex did not match phenotype reported sex (n=6) and individuals <18 years old were excluded (n=865). Ancestry was defined as reported ancestry, which showed, generally, good concordance with PCs

AMP-T2D-GENES

Sequencing and quality control of the AMP-T2D-GENES study has been previously described.³ Sequencing reads were processed and aligned to the human genome (build hg19) using the Picard (broadinstitute.github.io/picard/), BWA, and GATK software packages, following best-practice pipelines. Single nucleotide and

short indel variants were then called using a series of GATK commands (version nightly-2015-07-31-g3c929b0): ApplyRecalibration, CombineGVCFs, CombineVariants, GenotypeGVCFs, HaplotypeCaller, SelectVariants, and VariantFiltration.

Variants were called within 50bp of any region targeted for capture in any sequenced cohort. Following variant calling, all sites were then lifted over to build GRCh38 using CrossMap.

To perform data quality control, we first calculated a range of metrics measuring sample sequencing quality. We then stratified samples by ancestry and sequence capture technology and excluded from further analysis samples that were outliers according to any metric, based on visual inspection by comparison to other samples within the same stratum. After exclusion of samples, we calculated an additional set of variant metrics and excluded any variant with overall call rate <0.3 , heterozygosity of 1, or heterozygote allele balance of 0 or 1 (i.e. 100% or 0% of reads called non-reference for heterozygous genotypes). After these initial quality control steps, 49,484 samples and 7.02M variants remained in our dataset.

Following initial sample and variant quality control, we performed additional exclusions of samples from association analysis. First, we computed a set of “transethnic” SNPs for use in identity-by-descent (IBD) and principal component (PC) analysis. We began this analysis with variants in the clean dataset (a) with genotype call rate $>95\%$, (b) with minor allele frequency (MAF) $>1\%$ in each ancestry, and (c) further than 250Kb from the HLA region or an established T2D association signal. We LD-pruned variants using PLINK based on maximum $r^2 = 0.2$ (parameters – indep-pairwise 50 5 0.2). We used the remaining 171K variants to estimate pairwise

individual IBD using PLINK, and the top 10 PCs of genetic ancestry using EIGENSTRAT. For each pair of individuals with $IBD > 0.9$, we excluded the individual with the lower call rate (337 duplicate exclusions). We then excluded, for each of the five ancestries, any individual who appeared, based on visual inspection of the first two transethnic PCs, to lie outside of the main PC cluster corresponding to that ancestry (133 ethnic outliers). Finally, we used the subset of transethnic ancestry SNPs on the X chromosome to compare genetic sex to reported sex, using PLINK, and excluded all discordant individuals (273 sex discordances). Exclusion of the samples failing quality control, and variants that became monomorphic as a result of these sample exclusions, yielded a dataset of 45,231 individuals and 6.33M variants.

After these three rounds of sample exclusions, we identified five sets of ancestry-specific “ancestry” SNPs. We used the same procedure as for the transethnic SNPs (described above), except that we applied the MAF threshold only within the appropriate ancestry. We used these ancestry SNPs to estimate, for each ancestry, pairwise IBD values, genetic relatedness matrices (GRMs), and PCs for use in downstream association analysis. Additionally, from the IBD values, we generated a list of unrelated individuals within each ancestry by excluding the individual with the lower call rate in any pair of individuals with $IBD > 0.3$ (leading to 2,157 excluded individuals). The final “unrelated analysis” set consisted of 43,090 individuals and yielded 6.29M non-monomorphic variants.

UK Biobank

We used two UKB datasets with exome sequence data. The first is a CAD case control study with 12,938 individuals. 29 samples were removed as they had

discordant genotypes with genotyping array data, 17 showed mismatch between the reported and genetically inferred sex, 4 had excess heterozygosity and 6 had a call rate <95%. To perform the sex-mismatch analyses, variants on the X-chromosome were selected after filtering out low quality genotypes, call rate<95%, MAF<2%, low QD score (3 for INDELs and 2 for SNPs), low confidence regions and segmental duplications and those that do not have PASS VQSR. A set of high quality common autosomal variants were extracted for relative inference, principal component analysis, and estimation of heterozygosity after removing low confidence regions and segmental duplications, low quality genotypes, QD<2, call rate<98%, self-reported ancestry-specific HWE $p > 1 \times 10^{-6}$ among controls, MAF<1% and do not have PASS VQSR. Heterozygosity was estimated within each ancestry and samples with F statistic>2 were removed. Genetically inferred ancestry was obtained using the 1000 Genomes as reference. Sample QC metrics were then calculated in HAIL using autosomal variants after filtering out low-quality genotypes, variants with ancestry-specific HWE $p < 1 \times 10^{-6}$, low confidence regions and segmental duplications, low QD score (3 for INDELs and 2 for SNPs) and those that do not have PASS VQSR. Samples with call rate below 95%, mean DP below 30 and mean GQ below 80 were removed. Variant QC was done through filtering out monomorphic variants, call rate below 95%, those with HWE ($p < 1 \times 10^{-6}$), lie in low confidence regions or segmental duplications, are insertions or deletions with QD ≤ 3 or SNPs with QD ≤ 2 or variants where VQSR does not PASS unless singleton in which case retain those with VQSRTancheSNP99.60to99.80. A total of 11,216 PC-identified European ancestry participants were included after additional removal of duplicates and relatives across data sources. A total of 2,734,519 variants were included.

The second UKB data set is a population-based dataset. Samples were filtered out if they showed mismatch between genetically determined and reported sex, high rates of heterozygosity or contamination ($D\text{-stat} > 0.4$), low sequence coverage ($<85\%$ of targeted bases achieving $>20X$ coverage), duplicates, and exome sequence variants discordant with genotyping chip. More details are described elsewhere.⁴ The "Functionally Equivalent" (FE) call set was used.⁵ A total of 43,243 PC-identified European ancestry individuals were included after additional removal of duplicates and relatives across data sources.

Replication of gene-based associations

We performed replication of our top gene-based associations with blood lipid levels in the Penn Medicine BioBank (PMBB) and UK Biobank samples that did not contribute to the discovery analysis.

The PMBB is a repository of genotype and phenotype data for 43,731 patients at the University of Pennsylvania Perelman School of Medicine. All individuals recruited for 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, and access to all available EHR data. 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. The six lipid phenotypes studied were HDL-C ($n=21,247$), LDL-C ($n=21,040$), non-HDL-C ($n=21,087$), TC ($n=21,153$), TG ($n=21,418$), and TG:HDL ($n=21,213$). All available lipid trait measurements up to July 2020 were included. HDL-C, LDL-C, TC, and TG levels were measured directly and accessible via PMBB. Non-HDL-C levels

were obtained by subtracting HDL-C from TC levels. TG and TG:HDL levels were logarithmically transformed to normalize their distribution for association testing. Due to the clinical nature of the biobank, samples often had multiple phenotype values corresponding to a patient's various clinical appointments. Gene-based associations were performed on the minimum, median, and maximum phenotype values to account for both potentially protective and pathogenic effects. Conceptually, the idea behind using minimum, median, and maximum phenotype values is to better capture the full range of phenotypes, given that lipid levels can vary over time, including the effects of lipid-lowering medications. For example, in the common situation in which a patient has initiated statin therapy during the course of their EHR record, the maximum LDL-C is more likely to reflect the untreated 'basal' level than the median or the minimum LDL-C. Genetic variants that elevate a specific lipid phenotype are likely to be stronger for maximum values, while genetic variants that reduce a specific lipid phenotype are likely to be stronger for minimum values. For the gene-based association analysis, 10 different variant groupings were used to determine the set of damaging variants within each gene including the six groupings used in the initial study. The additional four groupings used predicted loss-of-function (pLOF) variants that included frameshift, stop gain, and splicing variants as annotated by RefGene. Missense variants were annotated using Rare Exome Variant Ensemble Learner (REVEL) and filtered for those with a pathogenicity score >0.5 . The four additional groupings consisted of, 1) pLOF, $MAF \leq 0.1\%$, 2) pLOF, $MAF \leq 0.1\%$, REVEL missense, 3) pLOF, $MAF \leq 1\%$, and 4) pLOF, $MAF \leq 1\%$, and REVEL missense. Each of the 10 groupings were used in a gene-based association test with the minimum, median, and maximum values of the 6 lipid phenotypes. Furthermore, ancestry-specific associations were also performed to elucidate any potential ancestry-specific

effects. This included associations among African and European ancestries separately, and then the two populations meta-analyzed. All associations were adjusted for sex, age, and principal components. The number of PCs chosen for each ancestry were determined according to ancestry-specific scree plots. The first 5 principal components were used for African ancestry associations, and the first 10 principal components were used for European ancestry associations.

In UK Biobank, we analyzed the association of rare variant aggregates from the 10 genes against four lipid phenotypes in the UK biobank whole exome sequencing (WES) data. Variant aggregates were obtained for the following four categories 1) LOFTEE – HC 2) LOFTEE - HC & predicted splice site altering 3) LOFTEE - HC & deleterious-METAsvm 4) LOFTEE - HC & deleterious-METAsvm & predicted splice site altering. We removed UK Biobank individuals used in the discovery analysis, resulting in 150,694 individuals for replication. The phenotypes were adjusted for lipid lowering medications, where total cholesterol was adjusted by dividing by 0.8 and LDL-C by dividing by 0.7. Triglycerides were natural log transformed for analysis. The phenotypes were inverse rank normalized and scaled by the standard deviation of the trait and adjusted for covariates (sex, age, age², PC1-PC10, if British ancestry). Rare variant aggregate test was conducted using STAAR⁶ with a MAF of 0.01 for the four lipids. Effect estimates were calculated using glmm.wald burden test.

References for Supplemental Methods:

1. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al. (2011). Scikit-learn: Machine Learning in Python. *J Mach Learn Res* 12, 2825–2830.
2. Jun, G., Flickinger, M., Hetrick, K.N., Romm, J.M., Doheny, K.F., Abecasis, G.R., Boehnke, M., and Kang, H.M. (2012). Detecting and estimating contamination

of human DNA samples in sequencing and array-based genotype data. *Am J Hum Genet* 91, 839-848.

3. Flannick, J., Mercader, J.M., Fuchsberger, C., Udler, M.S., Mahajan, A., Wessel, J., Teslovich, T.M., Caulkins, L., Koesterer, R., Barajas-Olmos, F., et al. (2019). Exome sequencing of 20,791 cases of type 2 diabetes and 24,440 controls. *Nature* 570, 71-76.
4. Van Hout, C.V., Tachmazidou, I., Backman, J.D., Hoffman, J.X., Ye, B., Pandey, A.K., Gonzaga-Jauregui, C., Khalid, S., Liu, D., Banerjee, N., et al. (2019). Whole exome sequencing and characterization of coding variation in 49,960 individuals in the UK Biobank. *bioRxiv*, 572347.
5. Regier, A.A., Farjoun, Y., Larson, D.E., Krasheninina, O., Kang, H.M., Howrigan, D.P., Chen, B.J., Kher, M., Banks, E., Ames, D.C., et al. (2018). Functional equivalence of genome sequencing analysis pipelines enables harmonized variant calling across human genetics projects. *Nature communications* 9, 4038.
6. Li, X., Li, Z., Zhou, H., Gaynor, S.M., Liu, Y., Chen, H., Sun, R., Dey, R., Arnett, D.K., Aslibekyan, S., et al. (2020). Dynamic incorporation of multiple in silico functional annotations empowers rare variant association analysis of large whole-genome sequencing studies at scale. *Nat Genet* 52, 969-983.

Study Participant Descriptions

Myocardial Infarction Genetics Consortium (MIGen) study participants

MIGen studies included the Atherosclerosis Risk in Communities study (ARIC), Italian Atherosclerosis Thrombosis and Vascular Biology (ATVB) study,¹ Bangladesh Risk of Acute Vascular Events study (BRAVE),² the Exome Sequencing Project Early-Onset Myocardial Infarction (ESP-EOMI) study,³ a nested case-control cohort of the Jackson Heart Study (JHS),⁴ the South German Myocardial Infarction study,⁵ the Ottawa Heart Study (OHS),⁶ the Precocious Coronary Artery Disease Study (PROCARDIS),⁷ the Pakistan Risk of Myocardial Infarction Study (PROMIS),⁸ the Registre Gironi del COR (Gerona Heart Registry or REGICOR) study,⁹ the Leicester Myocardial Infarction study,¹⁰ and the North German Myocardial Infarction study¹¹ (**Supplemental Table 37**). Clinical data were assessed in each study.

All participants in the study provided written informed consent for genetic studies. The institutional review boards at the Broad Institute and each participating institution approved the study protocol.

In order to minimize the possibility of unintentionally sharing information that can be used to re-identify private information, a subset of the data generated for this study are available at dbGaP and can be accessed at through dbGaP Study Accessions: phs000090.v1.p1 (ARIC), phs000814.v1.p1 (ATVB), phs001398.v1.p1 (BRAVE), phs000279.v2.p1 (EOMI), phs001098.v1.p1 (JHS), phs001000.v1.p1 (Leicester), phs000990.v1.p1 (NorthGermanMI), phs000916.v1.p1 (SouthGermanMI), phs000806.v1.p1 (OHS), phs000883.v1.p1 (PROCARDIS), phs000917.v1.p1 (PROMIS), phs000902.v1.p1 (Regicor).

TOPMed program study participants

Atherosclerosis Risk in Communities study (ARIC, 2868)

TOPMed dbGaP accession#: phs001211, Parent dbGaP accession#: phs000280

ARIC is a large population-based prospective longitudinal cohort study (began 1987) from four U.S. communities: Forsyth County, NC; Jackson, MS; the northwest suburbs of Minneapolis, MN; and Washington County, MD. ARIC was designed to investigate the etiology and natural history of atherosclerosis, its consequences, and related medical care by race, gender, location, and time as previously described.¹² A total of 15,792 participants (55% female and 27% African American) aged 45-64 years were recruited between 1987 and 1989 and received extensive examination, including medical, social and demographic data. The baseline visit was conducted between 1987 and 1989, the second visit in 1990-1992, the third visit in 1993-1995, the fourth visit in 1996-1998, the fifth visit in 2011-2013, the sixth visit in 2016-2017, and the seventh visit in 2018-2019. Follow-up is also conducted semi-annually since 2012 (annually prior to that) by telephone to maintain contact with participants and to assess the health status of the cohort.

The Atherosclerosis Risk in Communities study has been funded in whole or in part with Federal funds from the National Heart, Lung, and Blood Institute, National Institutes of Health, Department of Health and Human Services (contract numbers HHSN268201700001I, HHSN268201700002I, HHSN268201700003I, HHSN268201700004I and HHSN268201700005I). The authors thank the staff and participants of the ARIC study for their important contributions.

Whole genome sequencing (WGS) for the Trans-Omics in Precision Medicine (TOPMed) program was supported by the National Heart, Lung and Blood Institute (NHLBI). WGS for “NHLBI TOPMed: Atherosclerosis Risk in Communities (ARIC)” (phs001211) was performed at the Baylor College of Medicine Human Genome Sequencing Center (HHSN268201500015C and 3U54HG003273-12S2) and the Broad Institute for MIT and Harvard (3R01HL092577-06S1). Centralized read mapping and genotype calling, along with variant quality metrics and filtering were provided by the TOPMed Informatics Research Center (3R01HL-117626-02S1). Phenotype harmonization, data management, sample-identity QC, and general study coordination, were provided by the TOPMed Data Coordinating Center (3R01HL-120393-02S1). We gratefully acknowledge the studies and participants who provided biological samples and data for TOPMed.

The Genome Sequencing Program (GSP) was funded by the National Human Genome Research Institute (NHGRI), the National Heart, Lung, and Blood Institute (NHLBI), and the National Eye Institute (NEI). The GSP Coordinating Center (U24 HG008956) contributed to cross-program scientific initiatives and provided logistical and general study coordination. The Centers for Common Disease Genomics (CCDG) program was supported by NHGRI and NHLBI, and whole genome sequencing was performed at the Baylor College of Medicine Human Genome Sequencing Center (UM1 HG008898 and R01HL059367).

Old Order Amish (Amish, 1,083)

TOPMed dbGaP accession#: phs000956, Parent dbGaP accession#: phs000391

The Amish Complex Disease Research Program includes a set of large community-based studies focused largely on cardiometabolic health carried out in the Old Order Amish (OOA) community of Lancaster, Pennsylvania.¹³ The OOA population of Lancaster County, PA immigrated to the Colonies from Western Europe in the early 1700's. There are now over 30,000 OOA individuals in the Lancaster area, nearly all of whom can trace their ancestry back 12-14 generations to approximately 700 founders. Investigators at the University of Maryland School of Medicine have been studying the genetic determinants of cardiometabolic health in this population since 1993. To date, over 7,000 Amish adults have participated in one or more of our studies.

The Amish studies upon which these data are based were supported by NIH grants R01 AG18728, U01 HL072515, R01 HL088119, R01 HL121007, U01 HL137181, and P30 DK072488, American Heart Association grant AHA 17GRNT33661168 WGS for “NHLBI TOPMed: Genetics of Cardiometabolic Health in the Amish” (phs000956) was performed at the Broad Institute of MIT and Harvard (3R01HL121007-01S1).

Mt Sinai BioMe Biobank (BioMe, 3257)

TOPMed dbGaP accession#: phs001644, Parent dbGaP accession#: phs000925

The Mount Sinai Institute for Personalized Medicine BioMe Biobank is a consented, EMR-linked medical care setting biorepository of the Mount Sinai Medical Center drawing from a population of over 70,000 inpatients and 800,000 outpatient visits annually.¹⁴ The Mount Sinai Medical Center services diverse local communities of upper Manhattan, including Central Harlem (86% African American), East Harlem

(88% Hispanic Latino), and Upper East Side (88% European ancestry/white) with broad health disparities. Biobank operations are fully integrated in clinical care processes, including direct recruitment from clinical sites waiting areas and phlebotomy stations by dedicated Biobank recruiters independent of clinical care providers, prior to or following a clinician standard of care visit. Recruitment currently occurs at a broad spectrum of over 30 clinical care sites.

The Mount Sinai BioMe Biobank has been supported by The Andrea and Charles Bronfman Philanthropies and in part by Federal funds from the NHLBI and NHGRI (U01HG00638001; U01HG007417; X01HL134588). WGS for “NHLBI TOPMed: Mount Sinai BioMe Biobank” (phs001644) was performed at the Baylor College of Medicine Human Genome Sequencing Center (HHSN268201600033I). We thank all participants in the Mount Sinai Biobank. We also thank all our recruiters who have assisted and continue to assist in data collection and management and are grateful for the computational resources and staff expertise provided by Scientific Computing at the Icahn School of Medicine at Mount Sinai.

Coronary Artery Risk Development in Young Adults (CARDIA, 2724)

TOPMed dbGaP accession#: phs001612, Parent dbGaP accession#: phs000285

The Coronary Artery Risk Development in Young Adults Study (CARDIA) is a study examining the etiology and natural history of cardiovascular disease beginning in young adulthood.¹⁵ In 1985-1986, a cohort of 5115 healthy black and white men and women aged 18-30 years were selected to have approximately the same number of people in subgroups of age (18-24 and 25-30), sex, race, and education (high school or less and more than high school) within each of four US Field Centers. These same participants were asked to participate in follow-up examinations during 1987-1988 (Year 2), 1990-1991 (Year 5), 1992-1993 (Year 7), 1995-1996 (Year 10), 2000-2001 (Year 15), 2005-2006 (Year 20), 2010-2011 (Year 25); and 2015-2016 (Year 30). A majority of the group has been examined at each of the follow-up examinations (91%, 86%, 81%, 79%, 74%, 72%, 72%, and 71%, respectively). In addition to the follow-up examinations, participants are contacted regularly for the ascertainment of information on out-patient procedures and hospitalizations experienced between contacts.

The Coronary Artery Risk Development in Young Adults Study (CARDIA) is conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with the University of Alabama at Birmingham (HHSN268201800005I & HHSN268201800007I), Northwestern University (HHSN268201800003I), University of Minnesota (HHSN268201800006I), and Kaiser Foundation Research Institute (HHSN268201800004I). CARDIA was also partially supported by the Intramural Research Program of the National Institute on Aging (NIA) and an intra-agency agreement between NIA and NHLBI (AG0005). WGS for “NHLBI TOPMed: Coronary Artery Risk Development in Young Adults” (phs001612) was performed at the Baylor College of Medicine Human Genome Sequencing Center (HHSN268201600033I).

Cleveland Family Study (CFS, 532)

TOPMed dbGaP accession#: phs000954, Parent dbGaP accession#: phs000284

The Cleveland Family Study (CFS) is a family-based study of sleep apnea, comprising of 2,284 individuals (46% African American) from 361 families studied up to 4 occasions over 16 years, 1990-2006.¹⁶⁻¹⁹ Index probands (n=275) were recruited from 3 area hospital sleep labs if they had a confirmed diagnosis of sleep apnea and at least 2 first-degree relatives available to be studied. In the first 5 years of the study, neighborhood control probands (n=87) with at least 2 living relatives available for study were selected at random from a list provided by the index family and also studied. All available first-degree relatives and spouses of the case and control probands also were recruited. Second-degree relatives, including half-sibs, aunts, uncles and grandparents, were also included if they lived near the first-degree relatives (cases or controls), or if the family had been found to have two or more relatives with sleep apnea. Blood was sampled and DNA isolated for participants seen in the last two exam cycles (n=1,447).

CFS is supported by grants from the NHLBI (HL046389, HL113338, and 1R35HL135818). WGS for “NHLBI TOPMed: Cleveland Family Study - WGS Collaboration” (phs000954) was performed at the University of Washington Northwest Genomics Center (3R01HL098433-05S1 and HHSN268201600032I).

Cardiovascular Health Study (CHS, 2070)

TOPMed dbGaP accession#: phs001368, Parent dbGaP accession#: phs000287

The Cardiovascular Health Study (CHS) originated in 1988 and is a study of risk factors for development and progression of coronary heart disease and stroke in people aged 65 years and older.²⁰⁻²² The 5,888 study participants were recruited from four U.S. communities and have undergone extensive clinic examinations for evaluation of markers of subclinical cardiovascular disease. The original cohort totaled 5,201 participants. A new cohort was recruited in 1992. The 687 participants in the new cohort are predominately African-American and were recruited at three of the four field centers. Starting in 1989, and continuing through 1999, participants underwent annual extensive clinical examinations. Measurements included traditional risk factors such as blood pressure and lipids as well as measures of subclinical disease, including echocardiography of the heart, carotid ultrasound, and cranial magnetic-resonance imaging (MRI). At six-month intervals between clinic visits, and once clinic visits ended, participants were contacted by phone to ascertain hospitalizations and health status. The main outcomes are coronary heart disease (CHD), angina, heart failure (HF), stroke, transient ischemic attack (TIA), claudication, and mortality. Participants continue to be contacted by phone every 6 months.

This CHS research was supported by NHLBI contracts HHSN268201200036C, HHSN268200800007C, HHSN268201800001C, N01HC55222, N01HC85079, N01HC85080, N01HC85081, N01HC85082, N01HC85083, N01HC85086; and NHLBI grants U01HL080295, R01HL087652, R01HL105756, R01HL103612, R01HL120393, R01HL130114, and R01 HL059367, with additional contribution from the National Institute of Neurological Disorders and Stroke (NINDS). Additional support was provided through R01AG023629 from the National Institute on Aging (NIA). A full list of principal CHS investigators and institutions can be found at CHS-NHLBI.org. WGS for “NHLBI TOPMed: Cardiovascular Health Study” (phs001368) was performed at the Baylor College of Medicine Human Genome Sequencing Center (3U54HG003273-12S2,

HHSN268201500015C, and HHSN268201600033I). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Diabetes Heart Study (DHS, 345)

TOPMed dbGaP accession#: phs001412, Parent dbGaP accession#: phs001012

The Diabetes Heart Study (DHS) is a family-based study enriched for type 2 diabetes (T2D).²³ The cohort included 1443 European American and African American participants from 564 families with multiple cases of type 2 diabetes. The cohort was recruited between 1998 and 2006. Participants were extensively phenotyped for measures of subclinical CVD and other known CVD risk factors. Primary outcomes were quantified burden of vascular calcified plaque in the coronary artery, carotid artery, and abdominal aorta all determined from non-contrast computed tomography scans.

This work was supported by R01 HL92301, R01 HL67348, R01 NS058700, R01 AR48797, R01 DK071891, R01 AG058921, the General Clinical Research Center of the Wake Forest University School of Medicine (M01 RR07122, F32 HL085989), the American Diabetes Association, and a pilot grant from the Claude Pepper Older Americans Independence Center of Wake Forest University Health Sciences (P60 AG10484). WGS for “NHLBI TOPMed: Diabetes Heart Study” (phs001412) was performed at the Broad Institute of MIT and Harvard (HHSN268201500014C).

Framingham Heart Study (FHS, 3,961)

TOPMed dbGaP accession#: phs000974, Parent dbGaP accession#: phs000007

The Framingham Heart Study (FHS) is a prospective cohort study of 3 generations of subjects who have been followed up to 65 years to evaluate risk factors for cardiovascular disease.²⁴⁻²⁷ Its large sample of ~15,000 men and women who have been extensively phenotyped with repeated examinations make it ideal for the study of genetic associations with cardiovascular disease risk factors and outcomes. DNA samples have been collected and immortalized since the mid-1990s and are available on ~8000 study participants in 1037 families. These samples have been used for collection of GWAS array data and exome chip data in nearly all with DNA samples, and for targeted sequencing, deep exome sequencing and light coverage whole genome sequencing in limited numbers. Additionally, mRNA and miRNA expression data, DNA methylation data, metabolomics and other 'omics data are available on a sizable portion of study participants. This project will focus on deep whole genome sequencing (mean 30X coverage) in ~4100 subjects and imputed to all with GWAS array data to more fully understand the genetic contributions to cardiovascular, lung, blood and sleep disorders.

FHS acknowledges the support of contracts NO1-HC-25195 and HHSN268201500001I from the National Heart, Lung and Blood Institute and grant supplement R01 HL092577-06S1 for this research. WGS for “NHLBI TOPMed: Whole Genome Sequencing and Related Phenotypes in the Framingham Heart Study” (phs000974) was performed at the Broad Institute of MIT and Harvard (HHSN268201500014C, 3R01HL092577-06S1, and 3U54HG003067-12S2). We also

acknowledge the dedication of the FHS study participants without whom this research would not be possible.

Genetic Epidemiology Network of Arteriopathy (GENOA, 391)

TOPMed dbGaP accession#: phs001345, Parent dbGaP accession#: phs001238

The Genetic Epidemiology Network of Arteriopathy (GENOA) is one of four networks in the NHLBI Family-Blood Pressure Program (FBPP).²⁸ GENOA's long-term objective is to elucidate the genetics of target organ complications of hypertension, including both atherosclerotic and arteriolosclerotic complications involving the heart, brain, kidneys, and peripheral arteries.²⁹ The longitudinal GENOA Study recruited European-American and African-American sibships with at least 2 individuals with clinically diagnosed essential hypertension before age 60 years. All other members of the sibship were invited to participate regardless of their hypertension status. Participants were diagnosed with hypertension if they had either 1) a previous clinical diagnosis of hypertension by a physician with current anti-hypertensive treatment, or 2) an average systolic blood pressure ≥ 140 mm Hg or diastolic blood pressure ≥ 90 mm Hg based on the second and third readings at the time of their clinic visit. Only participants of the African-American Cohort were sequenced through TOPMed.

Support for GENOA was provided by the National Heart, Lung and Blood Institute (HL054457, HL054464, HL054481, and HL087660) of the National Institutes of Health. WGS for "NHLBI TOPMed: Genetic Epidemiology Network of Arteriopathy" (phs001345) was performed at the Broad Institute of MIT and Harvard (HHSN268201500014C) and the University of Washington Northwest Genomics Center (3R01HL055673-18S1).

Genetics of Lipid-Lowering Drugs and Diet Network (GOLDN, 594)

TOPMed dbGaP accession#: phs001359, Parent dbGaP accession#: phs000741

The Genetics of Lipid-Lowering Drugs and Diet Network (GOLDN) study was initiated to assess how genetic factors interact with environmental (diet and drug) interventions to influence blood levels of triglycerides and other atherogenic lipid species and inflammation markers (registered at clinicaltrials.gov, number NCT00083369).³⁰ The study recruited participants of European ancestry primarily from three-generational pedigrees from two NHLBI Family Heart Study (FHS) field centers (Minneapolis, MN and Salt Lake City, UT).³¹ Only families with at least two siblings were recruited and only participants who did not take lipid-lowering agents (pharmaceuticals or nutraceuticals) for at least 4 weeks prior to the initial visit were included. The diet intervention followed the protocol of Patsch et al.³² The whipping cream (83% fat) meal had 700 Calories/m² body surface area (2.93 mJ/m² body surface area): 3% of calories were derived from protein (instant nonfat dry milk) and 14% from carbohydrate (sugar). The ratio of polyunsaturated to saturated fat was 0.06 and the cholesterol content of the average meal was 240 mg. The mixture was blended with ice and flavorings. Blood samples were drawn immediately before (fasting) and at 3.5 and 6 hours after consuming the high-fat meal. The diet intervention was administered at baseline as well as after a 3-week treatment with 160 mg micronized fenofibrate. Participants were given the option to complete one or

both (diet and drug) interventions. Of all participants, 1079 had phenotypic data and provided appropriate consent, and underwent whole genome sequencing through the TOPMed program.

GOLDN biospecimens, baseline phenotype data, and intervention phenotype data were collected with funding from National Heart, Lung and Blood Institute (NHLBI) grant U01 HL072524. WGS for “NHLBI TOPMed: Genetics of Lipid Lowering Drugs and Diet Network” (phs001359) was performed at the University of Washington Northwest Genomics Center (3R01HL104135-04S1 and R01 HL104135).

Genetic Epidemiology Network of Salt Sensitivity (GenSalt, 1,749)

TOPMed dbGaP accession#: phs001217, Parent dbGaP accession#: phs000784

The Genetic Epidemiology Network of Salt-Sensitivity (GenSalt) study, using a family feeding-study design, aims to identify genes which interact with dietary sodium and potassium intake to influence blood pressure in Han Chinese participants from rural north China.³³ The dietary intervention included a 7-day low-sodium feeding (51.3 mmol/day), a 7-day high-sodium feeding (307.8 mmol/day) and a 7-day high-sodium feeding with an oral potassium supplementation (60 mmol/day). Microsatellite markers for genome-wide linkage scan and single nucleotide polymorphism (SNP) markers in candidate genes will be genotyped. Overall, 3153 participants from 658 families were recruited for GenSalt. Whole genome sequencing has been conducted for 1,860 participants as a part of TOPMed.

GenSalt was supported by research grants (U01HL072507, R01HL087263, and R01HL090682) from the National Heart, Lung and Blood Institute, National Institutes of Health, Bethesda, MD. WGS for “NHLBI TOPMed: Genetic Epidemiology Network of Salt Sensitivity” (phs001217) was performed at the Baylor College of Medicine Human Genome Sequencing Center (HHSN268201500015C).

Genetic Studies of Atherosclerosis Risk (GeneSTAR, 1,749)

TOPMed dbGaP accession#: phs001218, Parent dbGaP accession#: phs000375

GeneSTAR began in 1982 as the Johns Hopkins Sibling and Family Heart Study, a prospective longitudinal family-based study conducted originally in healthy adult siblings of people with documented early onset coronary disease under 60 years of age.^{34,35} Commencing in 2003, the siblings, their offspring, and the coparent of the offspring participated in a 2 week trial of aspirin 81 mg/day with pre and post ex vivo platelet function assessed using multiple agonists in whole blood and platelet rich plasma. Extensive additional cardiovascular testing and risk assessment was done at baseline and serially. Follow-up was carried out to determine incident cardiovascular disease, stroke, peripheral arterial disease, diabetes, cancer, and related comorbidities, from 5 to 30 years after study entry. The goal of several additional phenotyping and interventional substudies has been to discover and amplify understanding of the mechanisms of atherogenic vascular diseases and attendant comorbidities.

GeneSTAR was supported by grants from the National Institutes of Health/National Heart, Lung, and Blood Institute (U01 HL72518, HL087698, HL49762, HL58625, HL071025, HL112064), the

National Institutes of Health/National Institute of Nursing Research (NR0224103), and by a grant from the National Institutes of Health/National Center for Research Resources (M01-RR000052) to the Johns Hopkins General Clinical Research Center. WGS for “NHLBI TOPMed: Genetic Studies of Atherosclerosis Risk” (phs001218) was performed at the Broad Institute of MIT and Harvard (HHSN268201500014C), the Macrogen Corp. (3R01HL112064-04S1), and Illumina (R01HL112064).

Hispanic Community Health Study - Study of Latinos (HCHS/SOL, 2540)

TOPMed dbGaP accession#: phs001395, Parent dbGaP accession#: phs000810

The Hispanic Community Health Study / Study of Latinos (HCHS/SOL) is a multi-center epidemiologic study in Hispanic/Latino populations to determine the role of acculturation in the prevalence and development of disease, and to identify risk factors playing a protective or harmful role in Hispanics/Latinos.³⁶ The goals of the HCHS/SOL include studying the prevalence and development of disease in Hispanics/Latinos, including the role of acculturation, and identifying disease risk factors that play protective or harmful roles in Hispanics/Latinos. A total of 16,415 persons of Cuban, Dominican, Mexican, Puerto Rican, Central American, and South American backgrounds were recruited through four Field Centers affiliated with San Diego State University, Northwestern University in Chicago, Albert Einstein College of Medicine in the Bronx area of New York, and the University of Miami. Seven additional academic centers serve as scientific and logistical support centers. Study participants aged 18-74 years took part in an extensive clinic exam and assessments to ascertain socio-demographic, cultural, environmental and biomedical characteristics. Annual follow-up interviews are conducted to determine a range of health outcomes.

The Hispanic Community Health Study/Study of Latinos was carried out as a collaborative study supported by contracts from the National Heart, Lung, and Blood Institute (NHLBI) to the University of North Carolina (N01-HC65233), University of Miami (N01-HC65234), Albert Einstein College of Medicine (N01-HC65235), Northwestern University (N01-HC65236), and San Diego State University (N01-HC65237). The following Institutes/Centers/Offices contribute to the HCHS/SOL through a transfer of funds to the NHLBI: National Center on Minority Health and Health Disparities, the National Institute of Deafness and Other Communications Disorders, the National Institute of Dental and Craniofacial Research, the National Institute of Diabetes and Digestive and Kidney Diseases, the National Institute of Neurological Disorders and Stroke, and the Office of Dietary Supplements. WGS for “NHLBI TOPMed: Hispanic Community Health Study - Study of Latinos” (phs001395) was performed at the Baylor College of Medicine Human Genome Sequencing Center (HHSN268201600033I).

Hypertension Genetic Epidemiology Network and Genetic Epidemiology Network of Arteriopathy (HyperGEN, 1,797)

TOPMed dbGaP accession#: phs001293, Parent dbGaP accession#: phs001293

The Hypertension Genetic Epidemiology Network Study (HyperGEN) - Genetics of Left Ventricular (LV) Hypertrophy is a familial study aimed to understand genetic risk factors for LV hypertrophy by conducting genetic studies of continuous traits from echocardiography exams.³⁷ The originating HyperGEN study aimed to understand genetic risk factors for hypertension.³⁸ HyperGEN recruited 470 multiply-affected population-based hypertensive AA sibships (N=1224 siblings) from 1996-1999. HyperGEN probands were ascertained by early onset hypertension (i.e., before 60 years); to participate, they had to have at least one hypertensive sibling who was also willing to participate. Data from detailed clinical exams as well as genotyping data for linkage studies, candidate gene studies and GWAS have been collected and is shared between HyperGEN and the ancillary HyperGEN - Genetics of LV Hypertrophy study.

The HyperGEN Study is part of the National Heart, Lung, and Blood Institute (NHLBI) Family Blood Pressure Program; collection of the data represented here was supported by grants U01 HL054472 (MN Lab), U01 HL054473 (DCC), U01 HL054495 (AL FC), and U01 HL054509 (NC FC). The HyperGEN: Genetics of Left Ventricular Hypertrophy Study was supported by NHLBI grant R01 HL055673 with whole-genome sequencing made possible by supplement -18S1. WGS for "NHLBI TOPMed: Hypertension Genetic Epidemiology Network" (phs001293) was performed at the University of Washington Northwest Genomics Center (3R01HL055673-18S1).

Jackson Heart Study (JHS, 1722)

TOPMed dbGaP accession#: phs000964, Parent dbGaP accession#: phs000286

The purpose of the Jackson Heart Study (JHS) is to explore the reasons for heightened cardiovascular disease prevalence among African Americans and to uncover new approaches to reduce it. The JHS is a large, community-based, observational study whose 5,306 participants were recruited from among the non-institutionalized African-American adults from urban and rural areas of the three counties (Hinds, Madison, and Rankin) that make up the Jackson, MS, metropolitan statistical area (MSA).^{4,39,40} The JHS design included participants from the Jackson ARIC study who had originally been recruited through random selection from a drivers' license registry. New JHS participants were chosen randomly from the Accudata America commercial listing, which provides householder name, address, zip code, phone number (if available), age group in decades, and family components. In addition, a family component was included in the JHS. The sampling frame for the family study was a participant in any one of the ARIC, random, or volunteer samples whose family size met eligibility requirements. Recruitment was limited to persons 35-84 years old except in the family cohort, where those 21 years old and above were eligible.

The Jackson Heart Study (JHS) is supported and conducted in collaboration with Jackson State University (HHSN268201800013I), Tougaloo College (HHSN268201800014I), the Mississippi State Department of Health (HHSN268201800015I) and the University of Mississippi Medical Center (HHSN268201800010I, HHSN268201800011I and HHSN268201800012I) contracts from the National Heart, Lung, and Blood Institute (NHLBI) and the National Institute on Minority Health and Health Disparities (NIMHD). WGS for "NHLBI TOPMed: The Jackson Heart Study" (phs000964) was performed at the University of Washington

Northwest Genomics Center (HHSN268201100037C). The authors also wish to thank the staffs and participants of the JHS.

Multi-Ethnic Study of Atherosclerosis (MESA, 5,185)

TOPMed dbGaP accession#: phs001416, Parent dbGaP accession#: phs000209

The Multi-Ethnic Study of Atherosclerosis (MESA) is a study of the characteristics of subclinical cardiovascular disease (disease detected non-invasively before it has produced clinical signs and symptoms) and the risk factors that predict progression to clinically overt cardiovascular disease or progression of the subclinical disease.⁴¹ MESA researchers study a diverse, population-based sample of 6,814 asymptomatic men and women aged 45-84. Thirty-eight percent of the recruited participants are white, 28 percent African-American, 22 percent Hispanic, and 12 percent Asian, predominantly of Chinese descent. Participants were recruited from six field centers across the United States: Wake Forest University, Columbia University, Johns Hopkins University, University of Minnesota, Northwestern University and University of California - Los Angeles. Each participant received an extensive exam and determination of coronary calcification, ventricular mass and function, flow-mediated endothelial vasodilation, carotid intimal-medial wall thickness and presence of echogenic lucencies in the carotid artery, lower extremity vascular insufficiency, arterial wave forms, electrocardiographic (ECG) measures, standard coronary risk factors, sociodemographic factors, lifestyle factors, and psychosocial factors. Selected repetition of subclinical disease measures and risk factors at follow-up visits allows study of the progression of disease. Blood samples have been assayed for putative biochemical risk factors and stored for case-control studies. DNA has been extracted and lymphocytes cryopreserved (for possible immortalization) for study of candidate genes and possibly, genome-wide scanning, expression, and other genetic techniques. Participants are being followed for identification and characterization of cardiovascular disease events, including acute myocardial infarction and other forms of coronary heart disease (CHD), stroke, and congestive heart failure; for cardiovascular disease interventions; and for mortality.

Whole genome sequencing (WGS) for the Trans-Omics in Precision Medicine (TOPMed) program was supported by the National Heart, Lung and Blood Institute (NHLBI). WGS for "NHLBI TOPMed: Multi-Ethnic Study of Atherosclerosis (MESA)" (phs001416.v1.p1) was performed at the Broad Institute of MIT and Harvard (3U54HG003067-13S1). Centralized read mapping and genotype calling, along with variant quality metrics and filtering were provided by the TOPMed Informatics Research Center (3R01HL-117626-02S1). Phenotype harmonization, data management, sample-identity QC, and general study coordination, were provided by the TOPMed Data Coordinating Center (3R01HL-120393-02S1). MESA and the MESA SHARe project are conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with MESA investigators. The MESA project is conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with MESA investigators. Support for MESA is provided by contracts 75N92020D00001, HHSN268201500003I, N01-HC-95159, 75N92020D00005, N01-HC-95160, 75N92020D00002, N01-HC-95161, 75N92020D00003, N01-HC-95162, 75N92020D00006, N01-HC-95163, 75N92020D00004, N01-HC-95164, 75N92020D00007, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168, N01-HC-95169, UL1-TR-000040, UL1-TR-001079, UL1-TR-001420. Support is

provided by grants and contracts R01HL071051, R01HL071205, R01HL071250, R01HL071251, R01HL071258, R01HL071259, by the National Center for Research Resources, Grant UL1RR033176. The provision of genotyping data was supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1TR001881, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center.

Massachusetts General Hospital Atrial Fibrillation Study (MGH_AF, 682)

TOPMed dbGaP accession#: phs001062, Parent dbGaP accession#: phs001001

The Massachusetts General Hospital (MGH) Atrial Fibrillation Study was initiated in 2001.^{42,43} The study has enrolled serial probands, unaffected and affected family members with atrial fibrillation. At enrollment participants undergo a structured interview to systematically capture their past medical history, AF treatments, and family history. An electrocardiogram is performed; the results of an echocardiogram are obtained; and blood samples are obtained. For the TOPMed whole genome sequencing project only early-onset atrial fibrillation cases were sequenced. Early-onset atrial fibrillation was defined as an age of onset prior to 66 years of age.

The MGH AF Study was supported by grants to Dr. Ellinor from the Fondation Leducq (14CVD01), the National Institutes of Health to Dr. Ellinor (1R01HL092577, R01HL128914, K24HL105780) and Dr. Lubitz (1R01HL139731) and by grants from the American Heart Association to Dr. Ellinor (18SFRN34110082) and to Dr. Lubitz (18SFRN34250007). WGS for “NHLBI TOPMed: Massachusetts General Hospital Atrial Fibrillation Study” (phs001062) was performed at the Broad Institute of MIT and Harvard (3R01HL092577-06S1, 3U54HG003067-12S2, 3U54HG003067-13S1, and 3UM1HG008895-01S2)

San Antonio Family Study (SAFS, 575)

TOPMed dbGaP accession#: phs001215, Parent dbGaP accession#: phs000462

The San Antonio Family Heart Study is a complex pedigree-based mixed longitudinal study designed to identify low frequency or rare variants influencing susceptibility to cardiovascular disease, using whole genome sequence (WGS) information from 3,000 individuals in large Mexican American pedigrees from San Antonio, Texas.⁴⁴ The major objectives of this study are to identify low frequency or rare variants in and around known common variant signals for CVD, as well as to find novel low frequency or rare variants influencing susceptibility to CVD. The study began in 1991, and included 1,431 individuals in 42 extended families at baseline. Probands were 40 to 60 year old low-income Mexican Americans selected at random without regard to presence or absence of disease, almost exclusively from Mexican American census tracts in San Antonio, Texas. All first, second, and third -degree relatives of the proband and of the proband's spouse, aged 16 years or above, were eligible to participate in the study. 1,200 WGS at 30X WGS were obtained through Illumina funded by a supplement as part of the NHLBI's TOPMed program.

Collection of the San Antonio Family Study data was supported in part by National Institutes of Health (NIH) grants R01 HL045522, MH078143, MH078111 and MH083824; and whole genome sequencing of SAFS subjects was supported by

U01 DK085524 and R01 HL113323. We are very grateful to the participants of the San Antonio Family Study for their continued involvement in our research programs. WGS for “NHLBI TOPMed: Whole Genome Sequencing to Identify Causal Genetic Variants Influencing CVD Risk - San Antonio Family Studies” (phs001215) was performed at Illumina (3R01HL113323-03S1 and R01HL113322).

Samoan Adiposity Study (Samoan, 1,182)

TOPMed dbGaP accession#: phs000972, Parent dbGaP accession#: phs000914

The research goal of the Samoan Adiposity Study is to identify genetic variation that increases susceptibility to obesity and cardiometabolic phenotypes among adult Samoans using genome-wide association (GWAS) methods.^{45,46} DNA from peripheral blood and phenotypic information were collected from 3,119 adult Samoans, 23 to 70 years of age. The participants reside throughout the independent nation of Samoa, which is experiencing economic development and the nutrition transition. Genotyping was performed with the Affymetrix Genome-Wide Human SNP 6.0 Array using a panel of approximately 900,000 SNPs. Anthropometric, fasting blood biomarkers and detailed dietary, physical activity, health and socio-demographic variables were collected. Whole genome sequencing of a subset was motivated by the opportunity to create a Samoan-specific reference panel for imputation into the larger parent study.

Data collection was funded by NIH grant R01-HL093093 and R01-HL133040. WGS for “NHLBI TOPMed: Samoan Adiposity Study” (phs000972) was performed at the University of Washington Northwest Genomics Center (HHSN268201100037C and HHSN268201500016C). We thank the Samoan participants of the study and local village authorities. We acknowledge the support of the Samoan Ministry of Health and the Samoa Bureau of Statistics for their support of this research.

The Samoan Obesity, Lifestyle and Genetic Adaptations Study (OLaGA) Group:

Ranjan Deka, Dept. of Environmental Health, University of Cincinnati;
Nicola L. Hawley, Dept. of Chronic Disease Epidemiology, Yale University;
Stephen T McGarvey, Dept. of Epidemiology and International Health Institute, and Dept. of Anthropology, Brown University;
Ryan L Minster, Dept. of Human Genetics, University of Pittsburgh;
Take Naseri, Ministry of Health, Government of Samoa;
Muagututi'a Sefuiva Reupena, Lutia I Puava Ae Mapu I Fagalele;
Daniel E. Weeks, Depts. of Human Genetics and Biostatistics, University of Pittsburgh

Taiwan Study of Hypertension using Rare Variants (THRV, 1,979)

TOPMed dbGaP accession#: phs001387, Parent dbGaP accession#: phs001387

The THRV-TOPMed study consists of three cohorts: The SAPPHiRe Family cohort (N=1,271), TSGH (Tri-Service General Hospital, a hospital-based cohort, N=160), and TCVGH (Taichung Veterans General Hospital, another hospital-based cohort, N=922), all based in Taiwan.^{47,48} 1,271 subjects were previously recruited as

part of the NHLBI-sponsored SAPHIRE Network (which is part of the Family Blood Pressure Program, FBPP). The SAPHIRE families were recruited to have two or more hypertensive sibs, some families also with one normotensive/hypotensive sib. The two Hospital-based cohorts (TSGH and TCVGH) both recruited unrelated subjects with different recruitment criteria (matched with SAPHIRE subjects for age, sex, and BMI category).

The Rare Variants for Hypertension in Taiwan Chinese (THRV) is supported by the National Heart, Lung, and Blood Institute (NHLBI) grant (R01HL111249) and its participation in TOPMed is supported by an NHLBI supplement (R01HL111249-04S1). THRV is a collaborative study between Washington University in St. Louis, LA BioMed at Harbor UCLA, University of Texas in Houston, Taichung Veterans General Hospital, Taipei Veterans General Hospital, Tri-Service General Hospital, National Health Research Institutes, National Taiwan University, and Baylor University. THRV is based (substantially) on the parent SAPHIRE study, along with additional population-based and hospital-based cohorts. SAPHIRE was supported by NHLBI grants (U01HL54527, U01HL54498) and Taiwan funds, and the other cohorts were supported by Taiwan funds. WGS for “NHLBI TOPMed: Taiwan Study of Hypertension using Rare Variants” (phs001387) was performed at the Baylor College of Medicine Human Genome Sequencing Center (3R01HL111249-04S1, HHSN26820150015C)

Women's Health Initiative (WHI, 8,188)

TOPMed dbGaP accession#: phs001237, Parent dbGaP accession#: phs000200

The Women's Health Initiative (WHI) is a long-term national health study that has focused on strategies for preventing heart disease, breast and colorectal cancer, and osteoporotic fractures in postmenopausal women (clinicaltrials.gov NCT00000611).⁴⁹⁻⁵¹ The original WHI study included 161,808 postmenopausal women enrolled between 1993 and 1998. The Fred Hutchinson Cancer Research Center in Seattle, WA serves as the WHI Clinical Coordinating Center for data collection, management, and analysis of the WHI. The WHI has two major parts: a partial factorial randomized Clinical Trial (CT) and an Observational Study (OS); both were conducted at 40 Clinical Centers nationwide. The CT enrolled 68,132 postmenopausal women between the ages of 50-79 into trials testing three prevention strategies. If eligible, women could choose to enroll in one, two, or all three of the trial components. The components are: hormone therapy trials, dietary modification trial, and calcium / vitamin D trial. The Observational Study (OS) examines the relationship between lifestyle, environmental, medical and molecular risk factors and specific measures of health or disease outcomes. This component involves tracking the medical history and health habits of 93,676 women not participating in the CT. Recruitment for the observational study was completed in 1998 and participants were followed annually for 8 to 12 years.

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 HHSN268201600018C, HHSN268201600001C, HHSN268201600002C, HHSN268201600003C, and HHSN268201600004C. WGS for “NHLBI TOPMed: Women's Health Initiative” (phs001237) was performed at the Broad Institute of MIT and Harvard (HHSN268201500014C)

UK Biobank (external to TOPMed)

The UK Biobank analyses were conducted using the UK Biobank resource under application 7089.

References for Study Participant Descriptions

- 1 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* **107**, 1117-1122 (2003).
- 2 Chowdhury, R. *et al.* The Bangladesh Risk of Acute Vascular Events (BRAVE) Study: objectives and design. *European journal of epidemiology* **30**, 577-587, doi:10.1007/s10654-015-0037-2 (2015).
- 3 Do, R. *et al.* Exome sequencing identifies rare LDLR and APOA5 alleles conferring risk for myocardial infarction. *Nature* **518**, 102-106, doi:10.1038/nature13917 (2015).
- 4 Taylor, H. A., Jr. *et al.* Toward resolution of cardiovascular health disparities in African Americans: design and methods of the Jackson Heart Study. *Ethnicity & disease* **15**, S6-4-17 (2005).
- 5 Tg *et al.* Loss-of-function mutations in APOC3, triglycerides, and coronary disease. *N Engl J Med* **371**, 22-31, doi:10.1056/NEJMoa1307095 (2014).
- 6 McPherson, R. *et al.* A common allele on chromosome 9 associated with coronary heart disease. *Science* **316**, 1488-1491, doi:10.1126/science.1142447 (2007).
- 7 Clarke, R. *et al.* Genetic variants associated with Lp(a) lipoprotein level and coronary disease. *N Engl J Med* **361**, 2518-2528, doi:10.1056/NEJMoa0902604 (2009).
- 8 Saleheen, D. *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. *Eur J Epidemiol* **24**, 329-338, doi:10.1007/s10654-009-9334-y (2009).
- 9 Senti, M., Tomas, M., Marrugat, J., Elosua, R. & Investigators, R. Paraoxonase1-192 polymorphism modulates the nonfatal myocardial infarction risk associated with decreased HDLs. *Arterioscler Thromb Vasc Biol* **21**, 415-420 (2001).
- 10 Samani, N. J. *et al.* Genomewide association analysis of coronary artery disease. *The New England journal of medicine* **357**, 443-453, doi:10.1056/NEJMoa072366 (2007).
- 11 Myocardial Infarction, G. *et al.* Coding Variation in ANGPTL4, LPL, and SVEP1 and the Risk of Coronary Disease. *The New England journal of medicine* **374**, 1134-1144, doi:10.1056/NEJMoa1507652 (2016).
- 12 The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. *American journal of epidemiology* **129**, 687-702 (1989).
- 13 Mitchell, B. D. *et al.* The genetic response to short-term interventions affecting cardiovascular function: rationale and design of the Heredity and Phenotype Intervention (HAPI) Heart Study. *American heart journal* **155**, 823-828, doi:10.1016/j.ahj.2008.01.019 (2008).

- 14 Nadkarni, G. N. *et al.* Apolipoprotein L1 Variants and Blood Pressure Traits in African Americans. *Journal of the American College of Cardiology* **69**, 1564-1574, doi:10.1016/j.jacc.2017.01.040 (2017).
- 15 Hughes, G. H. *et al.* Recruitment in the Coronary Artery Disease Risk Development in Young Adults (Cardia) Study. *Control Clin Trials* **8**, 68S-73S (1987).
- 16 Redline, S. *et al.* Risk factors for sleep-disordered breathing in children. Associations with obesity, race, and respiratory problems. *Am J Respir Crit Care Med* **159**, 1527-1532, doi:10.1164/ajrccm.159.5.9809079 (1999).
- 17 Dean, D. A., 2nd *et al.* Scaling Up Scientific Discovery in Sleep Medicine: The National Sleep Research Resource. *Sleep* **39**, 1151-1164, doi:10.5665/sleep.5774 (2016).
- 18 Redline, S. *et al.* The familial aggregation of obstructive sleep apnea. *Am J Respir Crit Care Med* **151**, 682-687, doi:10.1164/ajrccm.151.3.7881656 (1995).
- 19 Zhang, G. Q. *et al.* The National Sleep Research Resource: towards a sleep data commons. *J Am Med Inform Assoc* **25**, 1351-1358, doi:10.1093/jamia/ocy064 (2018).
- 20 Fried, L. P. *et al.* The Cardiovascular Health Study: design and rationale. *Annals of epidemiology* **1**, 263-276 (1991).
- 21 O'Leary, D. H. *et al.* Carotid-artery intima and media thickness as a risk factor for myocardial infarction and stroke in older adults. Cardiovascular Health Study Collaborative Research Group. *The New England journal of medicine* **340**, 14-22, doi:10.1056/NEJM199901073400103 (1999).
- 22 O'Leary, D. H. *et al.* Use of sonography to evaluate carotid atherosclerosis in the elderly. The Cardiovascular Health Study. CHS Collaborative Research Group. *Stroke; a journal of cerebral circulation* **22**, 1155-1163 (1991).
- 23 Bowden, D. W. *et al.* Review of the Diabetes Heart Study (DHS) family of studies: a comprehensively examined sample for genetic and epidemiological studies of type 2 diabetes and its complications. *Rev Diabet Stud* **7**, 188-201, doi:10.1900/RDS.2010.7.188 (2010).
- 24 Castelli, W. P. Epidemiology of coronary heart disease: the Framingham study. *Am J Med* **76**, 4-12 (1984).
- 25 Kannel, W. B., Castelli, W. P., Gordon, T. & McNamara, P. M. Serum cholesterol, lipoproteins, and the risk of coronary heart disease. The Framingham study. *Annals of internal medicine* **74**, 1-12 (1971).
- 26 Kannel, W. B., Dawber, T. R., Kagan, A., Revotskie, N. & Stokes, J., 3rd. Factors of risk in the development of coronary heart disease--six year follow-up experience. The Framingham Study. *Annals of internal medicine* **55**, 33-50 (1961).
- 27 Kannel, W. B., Feinleib, M., McNamara, P. M., Garrison, R. J. & Castelli, W. P. An investigation of coronary heart disease in families. The Framingham offspring study. *American journal of epidemiology* **110**, 281-290 (1979).
- 28 FBPP Investigators. Multi-center genetic study of hypertension: The Family Blood Pressure Program (FBPP). *Hypertension* **39**, 3-9 (2002).
- 29 Daniels, P. R. *et al.* Familial aggregation of hypertension treatment and control in the Genetic Epidemiology Network of Arteriopathy (GENOA) study. *Am J Med* **116**, 676-681, doi:10.1016/j.amjmed.2003.12.032 (2004).
- 30 Liu, Y. *et al.* Pharmacogenetic association of the APOA1/C3/A4/A5 gene cluster and lipid responses to fenofibrate: the genetics of lipid-lowering drugs

- and diet network study. *Pharmacogenet Genomics* **19**, 161-169, doi:10.1097/FPC.0b013e32831e030e (2009).
- 31 Higgins, M. *et al.* NHLBI Family Heart Study: objectives and design. *American journal of epidemiology* **143**, 1219-1228 (1996).
 - 32 Patsch, J. R. *et al.* Relation of triglyceride metabolism and coronary artery disease. Studies in the postprandial state. *Arteriosclerosis and thrombosis : a journal of vascular biology / American Heart Association* **12**, 1336-1345 (1992).
 - 33 GenSalt Collaborative Research, G. GenSalt: rationale, design, methods and baseline characteristics of study participants. *J Hum Hypertens* **21**, 639-646, doi:10.1038/sj.jhh.1002207 (2007).
 - 34 Kral, B. G. *et al.* A common variant in the CDKN2B gene on chromosome 9p21 protects against coronary artery disease in Americans of African ancestry. *J Hum Genet* **56**, 224-229, doi:10.1038/jhg.2010.171 (2011).
 - 35 Bray, P. F. *et al.* Heritability of platelet function in families with premature coronary artery disease. *J Thromb Haemost* **5**, 1617-1623, doi:10.1111/j.1538-7836.2007.02618.x (2007).
 - 36 Sorlie, P. D. *et al.* Design and implementation of the Hispanic Community Health Study/Study of Latinos. *Annals of epidemiology* **20**, 629-641, doi:10.1016/j.annepidem.2010.03.015 (2010).
 - 37 Arnett, D. K. *et al.* Sibling correlation of left ventricular mass and geometry in hypertensive African Americans and whites: the HyperGEN study. Hypertension Genetic Epidemiology Network. *Am J Hypertens* **14**, 1226-1230, doi:10.1016/s0895-7061(01)02200-2 (2001).
 - 38 Williams, R. R. *et al.* NHLBI family blood pressure program: methodology and recruitment in the HyperGEN network. Hypertension genetic epidemiology network. *Annals of epidemiology* **10**, 389-400 (2000).
 - 39 Fuqua, S. R. *et al.* Recruiting African-American research participation in the Jackson Heart Study: methods, response rates, and sample description. *Ethnicity & disease* **15**, S6-18-29 (2005).
 - 40 Wilson, J. G. *et al.* Study design for genetic analysis in the Jackson Heart Study. *Ethnicity & disease* **15**, S6-30-37 (2005).
 - 41 Bild, D. E. *et al.* Multi-Ethnic Study of Atherosclerosis: objectives and design. *American journal of epidemiology* **156**, 871-881 (2002).
 - 42 Ellinor, P. T., Low, A. F., Patton, K. K., Shea, M. A. & Macrae, C. A. Discordant atrial natriuretic peptide and brain natriuretic peptide levels in lone atrial fibrillation. *Journal of the American College of Cardiology* **45**, 82-86, doi:10.1016/j.jacc.2004.09.045 (2005).
 - 43 Ellinor, P. T., Yoerger, D. M., Ruskin, J. N. & MacRae, C. A. Familial aggregation in lone atrial fibrillation. *Human genetics* **118**, 179-184, doi:10.1007/s00439-005-0034-8 (2005).
 - 44 Mitchell, B. D. *et al.* Genetic and environmental contributions to cardiovascular risk factors in Mexican Americans. The San Antonio Family Heart Study. *Circulation* **94**, 2159-2170 (1996).
 - 45 Hawley, N. L. *et al.* Prevalence of adiposity and associated cardiometabolic risk factors in the Samoan genome-wide association study. *Am J Hum Biol* **26**, 491-501, doi:10.1002/ajhb.22553 (2014).
 - 46 Minster, R. L. *et al.* A thrifty variant in CREBRF strongly influences body mass index in Samoans. *Nat Genet* **48**, 1049-1054, doi:10.1038/ng.3620 (2016).

- 47 Ranade, K. *et al.* A genome scan for hypertension susceptibility loci in
populations of Chinese and Japanese origins. *Am J Hypertens* **16**, 158-162,
doi:10.1016/s0895-7061(02)03245-4 (2003).
- 48 Wu, K. D. *et al.* Clustering and heritability of insulin resistance in Chinese and
Japanese hypertensive families: a Stanford-Asian Pacific Program in
Hypertension and Insulin Resistance sibling study. *Hypertens Res* **25**, 529-
536 (2002).
- 49 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).
- 50 Anderson, G. L. *et al.* Effects of conjugated equine estrogen in
postmenopausal women with hysterectomy: the Women's Health Initiative
randomized controlled trial. *JAMA : the journal of the American Medical
Association* **291**, 1701-1712, doi:10.1001/jama.291.14.1701 (2004).
- 51 Manson, J. E. *et al.* Menopausal Hormone Therapy and Long-term All-Cause
and Cause-Specific Mortality: The Women's Health Initiative Randomized
Trials. *JAMA : the journal of the American Medical Association* **318**, 927-938,
doi:10.1001/jama.2017.11217 (2017).

Acknowledgements

This work was supported by a grant from the Swedish Research Council (2016-06830) to GH. This work was supported by grants from the National Heart, Lung, and Blood Institute (NHLBI): R01HL142711 and R01HL127564 (P.N. and G.M.P.), R03HL141439 (G.M.P.), R01HL148050 and R01HL148565 (PN). P.N. is also supported by a Hassenfeld Scholar Award from the Massachusetts General Hospital, Foundation Leducq (TNE-18CVD04), and additional grants from the National Heart, Lung, and Blood Institute (R01HL148565 and R01HL148050). PSdV is supported by American Heart Association grant number 18CDA34110116. BEC and JL are supported by R35HL135818, HL113338, and HL46380. BEC is also supported by K01HL135405. S.A.L. is supported by NIH grant 1R01HL139731 and American Heart Association 18SFRN34250007. AVK is supported by K08HG010155. JF is supported by R01DK125490. DJL is supported by R01HG008983 and R01GM126479. CJW is supported by R35HL135824, R01HL142023, and R01HL127564. MJB is supported British Heart Foundation CS/14/2/30841. OM is supported by European Research Council ERC-AdG-2019-885003, NNF17OC0026936, and VR 2018-02760. MOM is supported by the European Research Council ERC-CoG-2014-649021 and NNF18OC0034386. RM is supported by Canadian Institutes Health Research FDN 154308. SR is supported by R35HL135818, HL113338, HL46380, and K01HL135405. NDP is supported by R01 HL92301, R01 HL67348, R01 NS058700, R01 AR48797, R01 DK071891, R01 AG058921, the General Clinical Research Center of the Wake Forest University School of Medicine (M01 RR07122, F32 HL085989), the American Diabetes Association, and a pilot grant from the Claude Pepper Older Americans Independence Center of Wake Forest University Health Sciences (P60 AG10484). BIF is supported by R01 NS058700 and R01 DK071891.

QQ is supported by R01-DK119268, R01-HL060712, and R01-HL140976. PTE is supported by the National Institutes of Health (1R01HL092577, R01HL128914, K24HL105780), the American Heart Association (18SFRN34110082), and by the Foundation Leducq (14CVD01). NLHC is supported by HHSN268201500001. JCF is support by NIDDK U01 DK105554 and NIDDK K24 DK110550. MIM is supported by NIDDK U01-DK105535 and Wellcome 090532, 098381, 106130, 203141, 212259. NPB is supported by NIDDK U01-DK105554. NG is supported by Novo Nordisk Foundation Center for Basic Metabolic Research is an independent Research Center, based at the University of Copenhagen, Denmark and partially funded by an unconditional donation from the Novo Nordisk Foundation (www.cbmr.ku.dk) (Grant number NNF18CC0034900). TLT and HMM are supported by CONACYT, 312688. KLM is supported by US National Institutes of Health DK072193 and DK093757. CLH is supported by DK085591. EST is supported by the National Medical Research Council of Singapore Clinician Scientist Award. NB is supported by The American Federation for Aging Research, the Einstein Glenn Center and the NIA (P01AG027734, R01AG 046949, 1R01AG042188 and P30AG038072). RCWM is supported by the Hong Kong Research Grants Council Theme-based Research Scheme (T12-402/13N) and Research Impact Fund (R4012-18).

The KARE cohort was supported by grants from Korea Centers for Disease Control and Prevention (4845–301, 4851–302, 4851–307) and intramural grants from the Korea National Institute of Health (2019-NG-053-01). Collection of the San Antonio Family Study data were supported in part by National Institutes of Health (NIH) grants P01 HL045522, R01 MH078143, R01 MH078111 and R01 MH083824; and whole genome sequencing of SAFS subjects was supported by U01 DK085524 and R01 HL113323. We are very grateful to the participants of the San Antonio Family

Study for their continued involvement in our research programs. DGI (Principal investigators Leif Groop and Tiinamaija Tuomi) and the Botnia Study is supported by the Sigrid Juselius Foundation, The Folkhalsan Research Foundation, Nordic Center of Excellence in Disease Genetics, EU (EXGENESIS), Finnish Diabetes Research Foundation, Foundation for Life and Health in Finland, Finnish Medical Society, Helsinki University Central Hospital Research Foundation, Perklén Foundation, Ollqvist Foundation, Narpes Health Care Foundation as well as the Municipal Health Care Center and Hospital in Jakobstad and Health Care Centers in Vasa, Narpes and Korsholm. The work in Malmö, Sweden, was also funded by a Linné grant from the Swedish Research Council (349-2006-237). The contribution of the Botnia and Skara research teams is gratefully acknowledged. The TwinsUK study was funded by the Wellcome Trust and European Community's Seventh Framework Programme (FP7/2007-2013). The TwinsUK study also receives support from the National Institute for Health Research (NIHR)- funded BioResource, Clinical Research Facility and Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust in partnership with King's College London. The KORA study 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 (BMBF) 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. Funded by the Bavarian State Ministry of Health and Care through the research project DigiMed Bayern (www.digimed-bayern.de).

Molecular data for the Trans-Omics in Precision Medicine (TOPMed) program was supported by the National Heart, Lung and Blood Institute (NHLBI). See the TOPMed Omics Support information in the **Supplementary Text** for study specific omics support information. Core support including centralized genomic read mapping and genotype calling, along with variant quality metrics and filtering were provided by the TOPMed Informatics Research Center (3R01HL-117626-02S1; contract HHSN268201800002I). Core support including phenotype harmonization, data management, sample-identity QC, and general program coordination were provided by the TOPMed Data Coordinating Center (R01HL-120393; U01HL-120393; contract HHSN268201800001I). We gratefully acknowledge the studies and participants who provided biological samples and data for TOPMed. The views expressed in this manuscript are those of the authors and do not necessarily represent the views of the National Heart, Lung, and Blood Institute; the National Institutes of Health; or the U.S. Department of Health and Human Services.

AMP-T2D-GENES Consortium

Carlos A. Aguilar-Salinas¹, Gil Atzmon^{2,3}, Francisco Barajas-Olmos⁴, Nir Barzilai², John Blangero⁵, Michael Boehnke⁶, Eric Boerwinkle^{7,8}, Lori L. Bonnycastle⁹, Erwin Bottinger^{10,11}, Donald W. Bowden¹², Noël P. Burt¹³, Federico Centeno-Cruz⁴, John C. Chambers¹⁴⁻¹⁶, Edmund Chan¹⁷, Juliana Chan¹⁸⁻²⁰, Ching-Yu Cheng²¹⁻²³, Yoon Shin Cho²⁴, Cecilia Contreras-Cubas⁴, Emilio J. Córdova⁴, Adolfo Correa²⁵, Ralph A. DeFronzo²⁶, Ravindranath Duggirala⁵, Josée Dupuis²⁷, Jason A. Flannick^{13,28,29}, Jose C. Florez³⁰⁻³³, Ma. Eugenia Garay-Sevilla³⁴, Humberto García-Ortiz⁴, Christian Gieger³⁵⁻³⁷, Benjamin Glaser³⁸, Clicerio Gonzalez³⁹, Maria Elena Gonzalez-Villalpando⁴⁰, Niels Grarup⁴¹, Leif C. Groop^{42,43}, Myron Gross⁴⁴, Christopher Haiman⁴⁵, Sohee Han⁴⁶, Craig L. Hanis⁴⁷, Torben Hansen^{41,48}, Nancy L. Heard-Costa^{49,50}, Brian E. Henderson⁴⁵, Mi Yeong Hwang⁴⁶, Sergio Islas-Andrade⁵¹, Marit E. Jørgensen⁵²⁻⁵⁴, Hyun Min Kang⁶, Bong-Jo Kim⁴⁶, Young Jin Kim⁴⁶, Heikki A. Koistinen⁵⁵⁻⁵⁷, Jaspal Singh Kooner⁵⁸, Johanna Kuusisto⁵⁹, Soo Heon Kwak⁶⁰, Markku Laakso⁵⁹, Leslie A. Lange⁶¹, Jong-Young Lee⁶², Juyoung Lee⁴⁶, Donna M. Lehman²⁶, Allan Linneberg^{63,64}, Jianjun Liu^{17,65}, Ruth J.F. Loos^{66,67}, Valeriya Lyssenko^{42,68}, Ronald C.W. Ma¹⁸⁻²⁰, Juan Manuel Malacara Hernandez⁶⁹, Angélica Martínez-Hernández⁴, Mark I. McCarthy^{70,71}, James B. Meigs^{33,72,73}, Thomas Meitinger^{74,75}, Elvia Mendoza-Caamal⁴, Karen L. Mohlke⁷⁶, Andrew D. Morris⁷⁷, Alanna C. Morrison⁷, Maggie C.Y. Ng^{12,78,79}, Peter M. Nilsson⁸⁰, Christopher J. O'Donnell⁸¹⁻⁸⁴, Lorena Orozco⁴, Colin N.A. Palmer⁸⁵, Kyong Soo Park^{60,86,87}, Oluf Pedersen⁴¹, Wendy S. Post⁸⁸, Michael Preuss⁶⁶, Bruce M. Psaty⁸⁹⁻⁹¹, Alexander P. Reiner⁹², Cristina Revilla-Monsalve⁹³, Stephen S. Rich⁹⁴, Jerome I. Rotter⁹⁵, Danish Saleheen^{96,97}, Claudia Schurmann^{10,11,66}, Xueling Sim⁹⁸, Rob Sladek⁹⁹⁻¹⁰¹, Kerrin S. Small¹⁰², Wing Yee So¹⁸⁻²⁰, Xavier Soberón⁴, Timothy D. Spector¹⁰², Konstantin Strauch^{36,103}, Tim M. Strom^{75,104}, E Shyong Tai^{17,98,105}, Claudia H.T. Tam¹⁸⁻²⁰, Yik Ying Teo^{98,106,107}, Farook Thameem¹⁰⁸, Brian Tomlinson¹⁰⁹, Russell P. Tracy^{110,111}, Tiinamaija Tuomi¹¹²⁻¹¹⁴, Jaakko Tuomilehto¹¹⁵⁻¹¹⁷, Teresa Tusié-Luna^{1,118}, Rob M. van Dam^{98,119}, Ramachandran S. Vasan^{49,120}, James G. Wilson¹²¹, Daniel R. Witte^{122,123}, Tien-Yin Wong²¹⁻²³

1 - Instituto Nacional de Ciencias Medicas y Nutricion , Mexico City, Mexico.; 2 - Departments of Medicine and Genetics, Albert Einstein College of Medicine, New York, USA.; 3 - University of Haifa, Faculty of natural science, Haifa, Isarel.; 4 - Instituto Nacional de Medicina Genómica, Mexico City, Mexico.; 5 - Department of Human Genetics and South Texas Diabetes and Obesity Institute, University of Texas Rio Grande Valley School of Medicine, Brownsville, TX 78520, USA.; 6 - Department of Biostatistics and Center for Statistical Genetics, University of Michigan, Ann Arbor, Michigan, USA.; 7 - 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, 77030.; 8 - Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX, 77030, USA.; 9 - Medical Genomics and Metabolic Genetics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland, USA.; 10 - Hasso Plattner Institute for Digital Health at Mount Sinai, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA.; 11 - Digital Health Center, Hasso Plattner Institute, University of Potsdam, Potsdam, Germany.; 12 - Department of Biochemistry, Wake Forest School of Medicine, Winston-Salem, NC, USA, 27157.; 13 - Programs in Metabolism and Medical & Population Genetics, Broad Institute, Cambridge, Massachusetts, USA.; 14 - Department of Epidemiology and Biostatistics, Imperial College London, London, UK.; 15 - Imperial College Healthcare NHS Trust, Imperial College London, London, UK.; 16 - Department of Cardiology, Ealing Hospital NHS Trust, Southall, Middlesex, UK.; 17 - Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore, National University Health System, , Singapore.; 18 - Department of Medicine and Therapeutics, The Chinese University of Hong Kong, Hong Kong, China.; 19 - Hong Kong Institute of Diabetes and Obesity, The Chinese University of Hong Kong, Hong Kong, China.; 20 - Li Ka Shing Institute of Health Sciences, The Chinese University of Hong Kong, Hong Kong, China.; 21 - Department of Ophthalmology, Yong Loo Lin School of Medicine, National University of Singapore, National University Health System, , Singapore.; 22 - Ophthalmology & Visual Sciences Academic Clinical Program (Eye ACP), Duke-NUS Medical School, Singapore.; 23 - Singapore Eye Research Institute, Singapore National Eye Centre, Singapore.; 24 - Department of Biomedical Science, Hallym University, Chuncheon, Republic of Korea.; 25 - Department of Medicine, University of Mississippi Medical Center, Jackson, MS, USA, 39216.; 26 - Department of Medicine, University of Texas Health Science Center, San Antonio, Texas, USA.; 27 - Department of Biostatistics, Boston University School of Public Health, Boston, Massachusetts, USA.; 28 - Department of Pediatrics, Harvard Medical School, Boston, MA, USA, 02115.; 29 - Division of Genetics and Genomics, Boston Children's Hospital, Boston, MA, USA, 02115.; 30 - Center for Genomic Medicine, Department of Medicine, Massachusetts General Hospital, Boston, Massachusetts, USA, 02114.; 31 - Diabetes Research Center (Diabetes Unit), Department of Medicine, Massachusetts General Hospital, Boston, Massachusetts, USA.; 32 - Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA, USA, 02142.; 33 - Department of Medicine, Harvard Medical School, Boston, MA, USA, 02115.; 34 - Department of Medical Science, Division of Health Science. University of Guanajuato.; 35 - Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany.; 36 - Institute of Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany.; 37 - German Center for Diabetes Research (DZD), Neuherberg, Germany.; 38 - Endocrinology and Metabolism Service, Hadassah-Hebrew University Medical Center, Jerusalem, Israel.; 39 - Unidad de Diabetes y Riesgo Cardiovascular, Instituto Nacional de Salud Pública, Cuernavaca, Morelos, Mexico.; 40 - Centro de Estudios en Diabetes, Mexico City, Mexico.; 41 - Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark.; 42 - Department of Clinical Sciences, Diabetes and Endocrinology, Lund University Diabetes Centre, Malmö, Sweden.; 43 - Finnish Institute for Molecular Genetics, University of Helsinki, Helsinki, Finland.; 44 - Department of Laboratory Medicine and Pathology, University of Minnesota, Minneapolis, Minnesota, USA.; 45 - Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, California, USA.; 46 - Division of Genome Science, Department of Precision Medicine, Chungcheongbuk-do, Republic of Korea.; 47 - Human Genetics Center, School of Public Health, The University of Texas Health Science Center at Houston, Houston, Texas, USA.; 48 - Faculty of Health Sciences, University of Southern Denmark, Odense,

Denmark.; 49 - NHLBI Framingham Heart Study, Framingham, Massachusetts, USA.; 50 - Department of Neurology, Boston University School of Medicine, Boston, Massachusetts, USA.; 51 - Dirección de Investigación, Hospital General de México "Dr. Eduardo Liceaga", Secretaría de Salud, Mexico City, Mexico.; 52 - National Institute of Public Health, University of Southern Denmark, Copenhagen, Denmark.; 53 - Steno Diabetes Center Copenhagen, Gentofte, Denmark.; 54 - Greenland Centre for Health Research, University of Greenland, Nuuk, Greenland.; 55 - Department of Public Health Solutions, Finnish Institute for Health and Welfare, Helsinki, Finland.; 56 - Minerva Foundation Institute for Medical Research, Helsinki, Finland.; 57 - University of Helsinki and Department of Medicine, Helsinki University Central Hospital, Helsinki, Finland.; 58 - National Heart and Lung Institute, Cardiovascular Sciences, Hammersmith Campus, Imperial College London, London, UK.; 59 - Institute of Clinical Medicine, University of Eastern Finland, and Kuopio University Hospital, Kuopio, Finland.; 60 - Department of Internal Medicine, Seoul National University Hospital, Seoul, Republic of Korea.; 61 - Department of Medicine, University of Colorado Denver, Anschutz Medical Campus, Aurora, CO, USA 80045.; 62 - Ministry of Health and Welfare, Seoul, Republic of Korea.; 63 - Center for Clinical Research and Prevention, Bispebjerg and Frederiksberg Hospital, The Capital Region, Copenhagen, Denmark.; 64 - Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark.; 65 - Genome Institute of Singapore, Agency for Science, Technology and Research, Singapore.; 66 - Charles R. Bronfman Institute of Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, New York, USA.; 67 - The Mindich Child Health and Development Institute, Icahn School of Medicine at Mount Sinai, New York, NY 10029.; 68 - University of Bergen, Norway.; 69 - Departments of Medicine and Human Genetics, The University of Chicago, Chicago, Illinois, USA.; 70 - Oxford Centre for Diabetes, Endocrinology and Metabolism, Radcliffe Department of Medicine, University of Oxford, Oxford, UK.; 71 - Wellcome Centre for Human Genetics, University of Oxford, Oxford, UK.; 72 - General Medicine Division, Massachusetts General Hospital, Boston, Massachusetts, USA.; 73 - Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA.; 74 - Deutsches Forschungszentrum für Herz-Kreislaufkrankungen (DZHK), Partner Site Munich Heart Alliance, Munich, Germany.; 75 - Institute of Human Genetics, Technische Universität München, Munich, Germany.; 76 - Department of Genetics, University of North Carolina, Chapel Hill, North Carolina, USA.; 77 - Clinical Research Centre, Centre for Molecular Medicine, Ninewells Hospital and Medical School, Dundee, UK.; 78 - Center for Diabetes Research, Wake Forest School of Medicine, Winston-Salem, North Carolina, USA.; 79 - Center for Genomics and Personalized Medicine Research, Wake Forest School of Medicine, Winston-Salem, North Carolina, USA.; 80 - Department of Clinical Sciences, Lund University, Malmö, Sweden.; 81 - Section of Cardiology, Department of Medicine, VA Boston Healthcare, Boston, Massachusetts, USA.; 82 - Harvard Medical School, Boston, Massachusetts, USA.; 83 - Department of Medicine, Brigham and Women's Hospital, Boston, MA 02115.; 84 - Intramural Administration Management Branch, National Heart, Lung, and Blood Institute, NIH, Framingham, Massachusetts, USA.; 85 - Pat Macpherson Centre for Pharmacogenetics and Pharmacogenomics, Division of Population Health and Genomics, University of Dundee, Ninewells Hospital and Medical School, Dundee, UK.; 86 - Department of Internal Medicine, Seoul National University College of Medicine, Seoul, Republic of Korea.; 87 - Department of Molecular Medicine and Biopharmaceutical Sciences, Graduate School of Convergence Science and Technology, Seoul National University, Seoul, Republic of Korea.; 88 - Division of Cardiology, Department of Medicine, Johns Hopkins University, Baltimore, Maryland, USA.; 89 - Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, WA, USA, 98101.; 90 - Department of Epidemiology, University of Washington, Seattle, WA, USA 98101.; 91 - Department of Health Services, University of Washington, Seattle, WA, USA 98101.; 92 - University of Washington, Seattle, Washington, USA.; 93 - Instituto Mexicano del Seguro Social SXXI, Mexico City, Mexico.; 94 - Center for Public Health Genomics, University of Virginia, Charlottesville, VA 22908, USA.; 95 - The Institute for Translational Genomics and Population Sciences, Department of Pediatrics, The Lundquist Institute for Biomedical Innovation at Harbor-UCLA Medical Center, Torrance, CA 90502.; 96 - Department of Biostatistics and Epidemiology, University of Pennsylvania, Philadelphia, Pennsylvania, USA.; 97 - Center for Non-Communicable Diseases, Karachi, Pakistan.; 98 - Saw Swee Hock School of Public Health, National University of Singapore and National University Health System, Singapore.; 99 - Department of Human Genetics, McGill University, Montreal, Quebec, Canada.; 100 - Division of Endocrinology and Metabolism, Department of Medicine, McGill University, Montreal, Quebec, Canada.; 101 - McGill University and Génome Québec Innovation Centre, Montreal, Quebec, Canada.; 102 - Department of Twin Research and Genetic Epidemiology, King's College London, London, UK.; 103 - Institute of Genetic Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany.; 104 - Institute of Human Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany.; 105 - Duke-NUS Medical School Singapore, Singapore.; 106 - Department of Statistics and Applied Probability, National University of Singapore, Singapore.; 107 - Life Sciences Institute, National University of Singapore, Singapore.; 108 - Department of Biochemistry, Faculty of Medicine, Health Science Center, Kuwait University, Safat, Kuwait.; 109 - Faculty of Medicine, Macau University of Science & Technology, Macau, China.; 110 - Department of Biochemistry, Robert Larner, M.D. College of Medicine, University of Vermont, Burlington, Vermont, USA.; 111 - Department of Pathology and Laboratory Medicine, Robert Larner, M.D. College of Medicine, University of Vermont, Burlington, Vermont, USA.; 112 - Department of Endocrinology, Abdominal Centre, Helsinki University Hospital, Helsinki, Finland.; 113 - Folkhälsan Research Centre, Helsinki, Finland.; 114 - Research Programs Unit, Diabetes and Obesity, University of Helsinki, Helsinki, Finland.; 115 - Public Health Promotion Unit, Finnish Institute for Health and Welfare, Helsinki, Finland.; 116 - Department of Public Health, University of Helsinki, Helsinki, Finland.; 117 - Diabetes Research Group, King Abdulaziz University, Jeddah, Saudi Arabia.; 118 - Departamento de Medicina Genómica y Toxicología, Instituto de Investigaciones Biomédicas, Universidad Nacional Autónoma de México/ Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Mexico City, Mexico.; 119 - Department of Nutrition, Harvard School of Public Health, Boston, Massachusetts, USA.; 120 - Departments of Medicine & Epidemiology, Boston University Schools of Medicine & Public Health, Boston, MA.; 121 - Division of Cardiovascular Medicine, Beth Israel Deaconess Medical Center, Boston, MA 02215.; 122 - Department of Public Health, Aarhus University, Aarhus, Denmark.; 123 - Steno Diabetes Center Aarhus, Aarhus, Denmark.;

Myocardial Infarction Genetics Consortium

Diego Ardisino¹⁻³, Usman Baber⁴, Matthew J. Bown^{5,6}, Mark D. Chaffin⁷⁻⁹, Rajiv Chowdhury^{10,11}, John Danesh^{10,12}, Roberto Elosua Llanos¹³⁻¹⁵, Valentin Fuster^{4,16}, Namrata Gupta⁷, Sekar Kathiresan^{7,8,17,18}, Amit V. Khera⁷⁻⁹, Ruth McPherson¹⁹, Olle Melander^{20,21}, Marju Orho-Melander²², Danish Saleheen^{23,24}, Nilesh Samani^{5,6}, Heribert Schunkert²⁵, Minxian X. Wang^{7,8,26}, Hugh Watkins²⁷

1 - ASTC: Associazione per lo Studio Della Trombosi in Cardiologia, Pavia, Italy.; 2 - Azienda Ospedaliero-Universitaria di Parma, Parma, Italy.; 3 - Università degli Studi di Parma, Parma, Italy.; 4 - Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, NY 10029.; 5 - Department of Cardiovascular Sciences, University of Leicester, Leicester, UK.; 6 - NIHR Leicester Biomedical Research Centre, Glenfield Hospital, Leicester UK.; 7 - Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA, USA, 02142.; 8 - Center for Genomic Medicine, Department of Medicine, Massachusetts General Hospital, Boston, Massachusetts, USA, 02114.; 9 - Cardiovascular Research Center, Massachusetts General Hospital, Boston, MA, USA, 02114.; 10 - MRC/BHF Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK.; 11 - Centre for Non-Communicable disease Research (CNCR), Bangladesh.; 12 - The National Institute for Health Research Blood and Transplant Research Unit (NIHR BTRU) in Donor Health and Genomics at the University of Cambridge, Cambridge, UK.; 13 - Cardiovascular Epidemiology and Genetics, Hospital del Mar Research Institute, Barcelona, Spain.; 14 - CIBER Enfermedades Cardiovasculares (CIBERCV), Barcelona, Spain.; 15 - Facultat de Medicina, Universitat de Vic-Central de Catalunya, Vic, Spain.; 16 - Centro Nacional de Investigaciones Cardiovasculares Carlos III (CNIC), Madrid, Spain.; 17 - Department of Medicine, Harvard Medical School, Boston, MA, USA, 02115.; 18 - Verve Therapeutics, Cambridge, MA, USA 02139.; 19 - Ruddy Canadian Cardiovascular Genetics Centre, University of Ottawa Heart Institute, Ottawa, Canada.; 20 - Department of Clinical Sciences, Diabetes and Endocrinology, Lund University Diabetes Centre, Malmö, Sweden.; 21 - Department of Emergency and Internal Medicine, Skåne University Hospital, Malmö, Sweden.; 22 - Department of Clinical Sciences, Lund University, Malmö, Sweden.; 23 - Department of Biostatistics and Epidemiology, University of Pennsylvania, Philadelphia, Pennsylvania, USA.; 24 - Center for Non-Communicable Diseases, Karachi, Pakistan.; 25 - Deutsches Herzzentrum München, Technische Universität München, Deutsches Zentrum für Herz-Kreislauf-Forschung, München, Germany.; 26 - Cardiovascular Disease Initiative, Broad Institute of Harvard and MIT, Cambridge, MA, USA, 02142.; 27 - Cardiovascular Medicine, Radcliffe Department of Medicine and the Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK.;

NHLBI Trans-Omics for Precision Medicine (TOPMed) Consortium

Namiko Abe¹, Gonalo Abecasis², Francois Aguet³, Christine Albert⁴, Laura Almasy⁵, Alvaro Alonso⁶, Seth Ament⁷, Peter Anderson⁸, Pramod Anugu⁹, Deborah Applebaum-Bowden¹⁰, Kristin Ardlie³, Dan Arking¹¹, Donna K Arnett¹², Allison Ashley-Koch¹³, Stella Aslibekyan¹⁴, Tim Assimes¹⁵, Paul Auer¹⁶, Dimitrios Avramopoulos¹¹, John Barnard¹⁷, Kathleen Barnes¹⁸, R. Graham Barr¹⁹, Emily Barron-Casella¹¹, Lucas Barwick²⁰, Terri Beaty¹¹, Gerald Beck²¹, Diane Becker²², Lewis Becker¹¹, Rebecca Beer²³, Amber Beitelshes⁷, Emelia Benjamin²⁴, Takis Benos²⁵, Marcos Bezerra²⁶, Larry Bielak², Joshua Bis²⁷, Thomas Blackwell², John Blangero²⁸, Eric Boerwinkle²⁹, Donald W. Bowden³⁰, Russell Bowler³¹, Jennifer Brody⁸, Ulrich Broeckel³², Jai Broome⁸, Karen Bunting¹, Esteban Burchard³³, Carlos Bustamante³⁴, Erin Buth³⁵, Brian Cade³⁶, Jonathan Cardwell³⁷, Vincent Carey³⁸, Cara Carty³⁹, Richard Casaburi⁴⁰, James Casella¹¹, Peter Castaldi⁴¹, Mark Chaffin³, Christy Chang⁷, Yi-Cheng Chang⁴², Daniel Chasman⁴³, Sameer Chavan³⁷, Bo-Juen Chen¹, Wei-Min Chen⁴⁴, Yii-Der Ida Chen⁴⁵, Michael Cho³⁸, Seung Hoan Choi³, Lee-Ming Chuang⁴⁶, Mina Chung⁴⁷, Ren-Hua Chung⁴⁸, Clary Clish⁴⁹, Suzy Comhair⁵⁰, Matthew Conomos³⁵, Elaine Cornell⁵¹, Adolfo Correa⁵², Carolyn Crandall⁴⁰, James Crapo⁵³, L. Adrienne Cupples⁵⁴, Joanne Curran⁵⁵, Jeffrey Curtis², Brian Custer⁵⁶, Coleen Damcott⁷, Dawood Darbar⁵⁷, Sayantan Das², Sean David⁵⁸, Colleen Davis⁸, Michelle Daya³⁷, Michael DeBaun⁵⁹, Dawn DeMeo³⁸, Ranjan Deka⁶⁰, Scott Devine⁷, Qing Duan⁶¹, Ravi Duggirala⁶², Jon Peter Durda⁵¹, Susan Dutcher⁶³, Charles Eaton⁶⁴, Lynette Ekunwe⁹, Adel El Boueiz⁶⁵, Patrick Ellinor⁶⁶, Leslie Emery⁸, Serpil Erzurum¹⁷, Charles Farber⁴⁴, Tasha Fingerlin⁶⁷, Matthew Flickinger², Myriam Fornage²⁹, Nora Franceschini⁶⁸, Chris Frazar⁸, Mao Fu⁷, Stephanie M. Fullerton⁸, Lucinda Fulton⁶³, Stacey Gabriel³, Weiniu Gan²³, Shanshan Gao³⁷, Yan Gao⁹, Margery Gass⁶⁹, Bruce Gelb⁷⁰, Xiaoli (Priscilla) Geng², Mark Geraci⁷¹, Soren Germer¹, Robert Gerszten⁷², Auyon Ghosh³⁸, Richard Gibbs⁷³, Chris Gignoux¹⁵, Mark Gladwin²⁵, David Glahn⁷⁴, Stephanie Gogarten⁸, Da-Wei Gong⁷, Harald Goring⁷⁵, Sharon Graw¹⁸, Daniel Grine³⁷, C. Charles Gu⁶³, Yue Guan⁷, Xiuqing Guo⁴⁵, Namrata Gupta³, Jeff Haessler⁶⁹, Michael Hall⁷⁶, Daniel Harris⁷, Nicola L. Hawley⁷⁷, Jiang He⁷⁸, Ben Heavner³⁵, Susan Heckbert⁸, Ryan Hernandez³³, David Herrington⁷⁹, Craig Hersh⁸⁰, Bertha Hidalgo¹⁴, James Hixson²⁹, Brian Hobbs³⁸, John Hokanson³⁷, Elliott Hong⁷, Karin Hoth⁸¹, Chao (Agnes) Hsiung⁸², Yi-Jen Hung⁸³, Haley Huston⁸⁴, Chii Min Hwu⁸⁵, Marguerite Ryan Irvin¹⁴, Rebecca Jackson⁸⁶, Deepti Jain⁸, Cashell Jaquish²³, Min A Jhun², Jill Johnsen⁸⁴, Andrew Johnson²³, Craig Johnson⁸, Rich Johnston⁶, Kimberly Jones¹¹, Hyun Min Kang⁸⁷, Robert Kaplan⁸⁸, Sharon Kardia², Sekar Kathiresan³, Shannon Kelly⁵⁶, Eimear Kenny⁷⁰, Michael Kessler⁷, Alyna Khan⁸, Wonji Kim⁸⁹, Greg Kinney⁹⁰, Barbara Konkle⁸⁴, Charles Kooperberg⁶⁹, Holly Kramer⁹¹, Christoph Lange⁹², Ethan Lange³⁷, Leslie Lange³⁷, Cathy Laurie⁸, Cecelia Laurie⁸, Meryl LeBoff³⁸, Jonathon LeFaive², Jiwon Lee³⁸, Seunggeun Shawn Lee², Wen-Jane Lee⁸⁵, David Levine⁸, Dan Levy²³, Joshua Lewis⁷, Xiaohui Li⁴⁵, Yun Li⁶¹, Henry Lin⁴⁵, Honghuang Lin⁹³, Keng Han Lin², Xihong Lin⁹⁴, Simin Liu⁹⁵, Yongmei Liu⁹⁶, Yu Liu⁹⁷, Ruth J.F. Loos⁹⁸, Steven Lubitz⁶⁶, Kathryn Lunetta⁹³, James Luo²³, Michael Mahaney⁵⁵, Barry Make¹¹, Ani Manichaikul⁴⁴, JoAnn Manson³⁸, Lauren Margolin³, Lisa Martin⁹⁹, Susan Mathai³⁷, Rasika Mathias¹¹, Susanne May³⁵, Patrick McArdle⁷, Merry-Lynn McDonald¹⁴, Sean McFarland⁸⁹, Stephen McGarvey⁶⁴, Daniel McGoldrick⁸, Caitlin McHugh³⁵, Hao Mei⁹, Luisa Mestroni¹⁸, Deborah A Meyers¹⁰⁰, Julie Mikulla²³, Nancy Min⁹, Mollie Minear²³, Ryan L Minster²⁵, Braxton D. Mitchell⁷, Matt Moll⁴¹, May E. Montasser⁷, Courtney Montgomery¹⁰¹, Arden Moscati⁷⁰, Solomon Musani⁵², Stanford Mwasongwe⁹, Josyf C Mychaleckyj⁴⁴, Girish Nadkarni⁷⁰, Rakhi Naik¹¹, Take Naseri¹⁰², Pradeep Natarajan³, Sergei Nekhai¹⁰³, Sarah C. Nelson³⁵, Bonnie Neltner³⁷, Deborah Nickerson⁸, Kari North⁶¹, Jeff O'Connell¹⁰⁴, Tim O'Connor⁷, Heather Ochs-Balcom¹⁰⁵, David T. Paik¹⁰⁶, Nicholette Palmer¹⁰⁷, James Pankow¹⁰⁸, George Papanicolaou²³, Afshin Parsa⁷, Juan Manuel Peralta⁶², Marco Perez¹⁵, James Perry⁷, Ulrike Peters⁶⁹, Patricia Peyser², Lawrence S Phillips⁶, Toni Pollin⁷, Wendy Post¹⁰⁹, Julia Powers Becker¹¹⁰, Meher Preethi Boorgula³⁷, Michael Preuss⁷⁰, Bruce Psaty⁸, Pankaj Qasba²³, Dandi Qiao³⁸, Zhaohui Qin⁶, Nicholas Rafaels¹¹¹, Laura Raffield¹¹², Vasani S. Ramachandran⁹³, D.C. Rao⁶³, Laura Rasmussen-Torvik¹¹³, Aakrosh Ratan⁴⁴, Susan Redline³⁸, Robert Reed⁷, Elizabeth Regan⁵³, Alex Reiner¹¹⁴, Muagututia Sefuiva Reupena¹¹⁵, Ken Rice⁸, Stephen Rich⁴⁴, Dan Roden¹¹⁶, Carolina Roselli³, Jerome Rotter⁴⁵, Ingo Ruczinski¹¹, Pamela Russell³⁷, Sarah Ruuska⁸⁴, Kathleen Ryan⁷, Ester Cerdeira Sabino¹¹⁷, Danish Saleheen¹⁹, Shabnam Salimi⁷, Steven Salzberg¹¹, Kevin Sandow¹¹⁸, Vijay G. Sankaran¹¹⁹, Christopher Scheller², Ellen Schmidt², Karen Schwander⁶³, David Schwartz³⁷, Frank Sciurba²⁵, Christine Seidman¹²⁰, Jonathan Seidman¹²¹, Vivien Sheehan¹²², Stephanie L. Sherman¹²³, Amol Shetty⁷, Aniket Shetty³⁷, Wayne Hui-Heng Sheu⁸⁵, M. Benjamin Shoemaker¹²⁴, Brian Silver¹²⁵, Edwin Silverman³⁸, Jennifer Smith², Josh Smith⁸, Nicholas Smith¹²⁶, Tanja Smith¹, Sylvia Smoller⁸⁸, Beverly Snively¹²⁷, Michael Snyder¹⁵, Tamar Sofer³⁸, Nona Sotoodehnia⁸, Adrienne M. Stilp⁸, Garrett Storm³⁷, Elizabeth Streeten⁷, Jessica Lasky Su³⁸, Yun Ju Sung⁶³, Jody Sylvia³⁸, Adam Szpiro⁸, Carole Sztalryd⁷, Daniel Taliun², Hua Tang¹²⁸, Margaret Taub¹¹, Kent D. Taylor¹²⁹, Matthew Taylor¹⁸, Simeon Taylor⁷, Marilyn Telen¹³, Timothy A. Thornton⁸, Machiko Threlkeld¹³⁰, Lesley Tinker⁶⁹, David

Tirschwell⁸, Sarah Tishkoff¹³¹, Hemant Tiwari¹³², Catherine Tong¹³³, Russell Tracy¹³⁴, Michael Tsai¹⁰⁸, Dhananjay Vaidya¹¹, David Van Den Berg¹³⁵, Peter VandeHaar², Scott Vrieze¹⁰⁸, Tarik Walker³⁷, Robert Wallace⁸¹, Avram Walts³⁷, Fei Fei Wang⁸, Heming Wang¹³⁶, Karol Watson⁴⁰, Daniel E. Weeks²⁵, Bruce Weir⁸, Scott Weiss³⁸, Lu-Chen Weng⁶⁶, Jennifer Wessel¹³⁷, Cristen Willer¹³⁸, Kayleen Williams³⁵, L. Keoki Williams¹³⁹, Carla Wilson³⁸, James Wilson¹⁴⁰, Quenna Wong⁸, Joseph Wu¹⁰⁶, Huichun Xu⁷, Lisa Yanek¹¹, Ivana Yang³⁷, Rongze Yang⁷, Norann Zaghloul⁷, Maryam Zekavat³, Yingze Zhang¹⁴¹, Snow Xueyan Zhao⁵³, Wei Zhao¹⁴², Degui Zhi²⁹, Xiang Zhou², Xiaofeng Zhu¹⁴³, Michael Zody¹, Sebastian Zoellner², Mariza de Andrade¹⁴⁴, Lisa de las Fuentes¹⁴⁵

1 - New York Genome Center, New York, New York, 10013; 2 - University of Michigan, Ann Arbor, Michigan, 48109; 3 - Broad Institute, Cambridge, Massachusetts, 02142; 4 - Cedars Sinai, Boston, Massachusetts, 02114; 5 - Children's Hospital of Philadelphia, University of Pennsylvania, Philadelphia, Pennsylvania, 19104; 6 - Emory University, Atlanta, Georgia, 30322; 7 - University of Maryland, Baltimore, Maryland, 21201; 8 - University of Washington, Seattle, Washington, 98195; 9 - University of Mississippi, Jackson, Mississippi, 38677; 10 - National Institutes of Health, Bethesda, Maryland, 20892; 11 - Johns Hopkins University, Baltimore, Maryland, 21218; 12 - University of Kentucky, Lexington, Kentucky, 40506; 13 - Duke University, Durham, North Carolina, 27708; 14 - University of Alabama, Birmingham, Alabama, 35487; 15 - Stanford University, Stanford, California, 94305; 16 - University of Wisconsin Milwaukee, Milwaukee, Wisconsin, 53211; 17 - Cleveland Clinic, Cleveland, Ohio, 44195; 18 - University of Colorado Anschutz Medical Campus, Aurora, Colorado, 80045; 19 - Columbia University, New York, New York, 10027; 20 - The Emmes Corporation, LTRC, Rockville, Maryland, 20850; 21 - Cleveland Clinic, Quantitative Health Sciences, Cleveland, Ohio, 44195; 22 - Johns Hopkins University, Medicine, Baltimore, Maryland, 21218; 23 - National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, Maryland, 20892; 24 - Boston University, Massachusetts General Hospital, Boston University School of Medicine, Boston, Massachusetts, 02114; 25 - University of Pittsburgh, Pittsburgh, Pennsylvania, 15260; 26 - Fundao de Hematologia e Hemoterapia de Pernambuco - Hemope, Recife, 52011-000; 27 - University of Washington, Cardiovascular Health Research Unit, Department of Medicine, Seattle, Washington, 98195; 28 - University of Texas Rio Grande Valley School of Medicine, Human Genetics, Brownsville, Texas, 78520; 29 - University of Texas Health at Houston, Houston, Texas, 77225; 30 - Wake Forest Baptist Health, Department of Biochemistry, Winston-Salem, North Carolina, 27157; 31 - National Jewish Health, National Jewish Health, Denver, Colorado, 80206; 32 - Medical College of Wisconsin, Milwaukee, Wisconsin, 53226; 33 - University of California, San Francisco, San Francisco, California, 94143; 34 - Stanford University, Biomedical Data Science, Stanford, California, 94305; 35 - University of Washington, Biostatistics, Seattle, Washington, 98195; 36 - Brigham & Women's Hospital, Brigham and Women's Hospital, Boston, Massachusetts, 02115; 37 - University of Colorado at Denver, Denver, Colorado, 80204; 38 - Brigham & Women's Hospital, Boston, Massachusetts, 02115; 39 - Washington State University, Pullman, Washington, 99164; 40 - University of California, Los Angeles, Los Angeles, California, 90095; 41 - Brigham & Women's Hospital, Medicine, Boston, Massachusetts, 02115; 42 - National Taiwan University, Taipei, 10617; 43 - Brigham & Women's Hospital, Division of Preventive Medicine, Boston, Massachusetts, 02215; 44 - University of Virginia, Charlottesville, Virginia, 22903; 45 - Lundquist Institute, Torrance, California, 90502; 46 - National Taiwan University, National Taiwan University Hospital, Taipei, 10617; 47 - Cleveland Clinic, Cleveland Clinic, Cleveland, Ohio, 44195; 48 - National Health Research Institute Taiwan, Miaoli County, 350; 49 - Broad Institute, Metabolomics Platform, Cambridge, Massachusetts, 02142; 50 - Cleveland Clinic, Immunity and Immunology, Cleveland, Ohio, 44195; 51 - University of Vermont, Burlington, Vermont, 05405; 52 - University of Mississippi, Medicine, Jackson, Mississippi, 38677; 53 - National Jewish Health, Denver, Colorado, 80206; 54 - Boston University, Biostatistics, Boston, Massachusetts, 02115; 55 - University of Texas Rio Grande Valley School of Medicine, Brownsville, Texas, 78520; 56 - Vitalant Research Institute, San Francisco, California, 94118; 57 - University of Illinois at Chicago, Chicago, Illinois, 60607; 58 - University of Chicago, Chicago, Illinois, 60637; 59 - Vanderbilt University, Nashville, Tennessee, 37235; 60 - University of Cincinnati, Cincinnati, Ohio, 45220; 61 - University of North Carolina, Chapel Hill, North Carolina, 27599; 62 - University of Texas Rio Grande Valley School of Medicine, Edinburg, Texas, 78539; 63 - Washington University in St Louis, St Louis, Missouri, 63130; 64 - Brown University, Providence, Rhode Island, 02912; 65 - Harvard University, Channing Division of Network Medicine, Cambridge, Massachusetts, 02138; 66 - Massachusetts General Hospital, Boston, Massachusetts, 02114; 67 - National Jewish Health, Center for Genes, Environment and Health, Denver, Colorado, 80206; 68 - University of North Carolina, Epidemiology, Chapel Hill, North Carolina, 27599; 69 - Fred Hutchinson Cancer Research Center, Seattle, Washington, 98109; 70 - Icahn School of Medicine at Mount Sinai, New York, New York, 10029; 71 - Indiana University, Medicine, Indianapolis, Indiana, 46202; 72 - Beth Israel Deaconess Medical Center, Boston, Massachusetts, 02215; 73 - Baylor College of Medicine Human Genome Sequencing Center, Houston, Texas, 77030; 74 - Boston Children's Hospital, Harvard Medical School, Department of Psychiatry, Boston, Massachusetts, 02115; 75 - University of Texas Rio Grande Valley School of Medicine, San Antonio, Texas, 78229; 76 - University of Mississippi, Cardiology, Jackson, Mississippi, 39216; 77 - Yale University, Department of Chronic Disease Epidemiology, New Haven, Connecticut, 06520; 78 - Tulane University, New Orleans, Louisiana, 70118; 79 - Wake Forest Baptist Health, Winston-Salem, North Carolina, 27157; 80 - Brigham & Women's Hospital, Channing Division of Network Medicine, Boston, Massachusetts, 02115; 81 - University of Iowa, Iowa City, Iowa, 52242; 82 - National Health Research Institute Taiwan, Institute of Population Health Sciences, NHRI, Miaoli County, 350; 83 - Tri-Service General Hospital National Defense Medical Center; 84 - Blood Works Northwest, Seattle, Washington, 98104; 85 - Taichung Veterans General Hospital Taiwan, Taichung City, 407; 86 - Ohio State University Wexner Medical Center, Internal Medicine, Division of Endocrinology, Diabetes and Metabolism, Columbus, Ohio, 43210; 87 - University of Michigan, Biostatistics, Ann Arbor, Michigan, 48109; 88 - Albert Einstein College of Medicine, New York, New York, 10461; 89 - Harvard University, Cambridge, Massachusetts, 02138; 90 - University of Colorado at Denver, Epidemiology, Aurora, Colorado, 80045; 91 - Loyola University, Public Health Sciences, Maywood, Illinois, 60153; 92 - Harvard School of Public Health, Biostats, Boston, Massachusetts, 02115; 93 - Boston University, Boston, Massachusetts, 02215; 94 - Harvard School of Public Health, Boston, Massachusetts, 02115; 95 - Brown University, Epidemiology and Medicine, Providence, Rhode Island, 02912; 96 - Duke University, Cardiology, Durham, North Carolina, 27708; 97 - Stanford University, Cardiovascular Institute, Stanford, California, 94305; 98 - Icahn School of Medicine at Mount Sinai, The Charles Bronfman Institute for Personalized Medicine, New York, New York, 10029; 99 - George Washington University, Washington, District of Columbia, 20052; 100 - University of Arizona, Tucson, Arizona, 85721; 101 - Oklahoma Medical Research Foundation, Genes and Human Disease, Oklahoma City, Oklahoma, 73104; 102 - Ministry of Health, Government of Samoa, Apia; 103 - Howard University, Washington, District of Columbia, 20059; 104 - University of Maryland, Baltimore, Maryland, 21201; 105 - University at Buffalo, Buffalo, New York, 14260; 106 - Stanford University, Stanford Cardiovascular Institute, Stanford, California, 94305; 107 - Wake Forest Baptist Health, Biochemistry, Winston-Salem, North

Carolina, 27157; 108 - University of Minnesota, Minneapolis, Minnesota, 55455; 109 - Johns Hopkins University, Cardiology/Medicine, Baltimore, Maryland, 21218; 110 - University of Colorado at Denver, Medicine, Denver, Colorado, 80204; 111 - University of Colorado at Denver, Denver, Colorado, 80045; 112 - University of North Carolina, Genetics, Chapel Hill, North Carolina, 27599; 113 - Northwestern University, Chicago, Illinois, 60208; 114 - Fred Hutchinson Cancer Research Center, University of Washington, Seattle, Washington, 98109; 115 - Lusia I Puava Ae Mapu I Fagalele, Apia; 116 - Vanderbilt University, Medicine, Pharmacology, Biomedical Informatics, Nashville, Tennessee, 37235; 117 - Universidade de Sao Paulo, Faculdade de Medicina, Sao Paulo, 01310000; 118 - Lundquist Institute, TGPS, Torrance, California, 90502; 119 - Harvard University, Division of Hematology/Oncology, Cambridge, Massachusetts, 02138; 120 - Harvard Medical School, Genetics, Boston, Massachusetts, 02115; 121 - Harvard Medical School, Boston, Massachusetts, 02115; 122 - Baylor College of Medicine, Pediatrics, Atlanta, Georgia, 30307; 123 - Emory University, Human Genetics, Atlanta, Georgia, 30322; 124 - Vanderbilt University, Medicine/Cardiology, Nashville, Tennessee, 37235; 125 - UMass Memorial Medical Center, Worcester, Massachusetts, 01655; 126 - University of Washington, Epidemiology, Seattle, Washington, 98195; 127 - Wake Forest Baptist Health, Biostatistical Sciences, Winston-Salem, North Carolina, 27157; 128 - Stanford University, Genetics, Stanford, California, 94305; 129 - Lundquist Institute, Institute for Translational Genomics and Populations Sciences, Torrance, California, 90502; 130 - University of Washington, University of Washington, Department of Genome Sciences, Seattle, Washington, 98195; 131 - University of Pennsylvania, Genetics, Philadelphia, Pennsylvania, 19104; 132 - University of Alabama, Biostatistics, Birmingham, Alabama, 35487; 133 - University of Washington, Department of Biostatistics, Seattle, Washington, 98195; 134 - University of Vermont, Pathology & Laboratory Medicine, Burlington, Vermont, 05405; 135 - University of Southern California, USC Methylation Characterization Center, University of Southern California, California, 90033; 136 - Brigham & Women's Hospital, Mass General Brigham, Boston, Massachusetts, 02115; 137 - Indiana University, Epidemiology, Indianapolis, Indiana, 46202; 138 - University of Michigan, Internal Medicine, Ann Arbor, Michigan, 48109; 139 - Henry Ford Health System, Detroit, Michigan, 48202; 140 - Beth Israel Deaconess Medical Center, Cardiology, Cambridge, Massachusetts, 02139; 141 - University of Pittsburgh, Medicine, Pittsburgh, Pennsylvania, 15260; 142 - University of Michigan, Department of Epidemiology, Ann Arbor, Michigan, 48109; 143 - Case Western Reserve University, Department of Population and Quantitative Health Sciences, Cleveland, Ohio, 44106; 144 - Mayo Clinic, Health Sciences Research, Rochester, Minnesota, 55905; 145 - Washington University in St Louis, Department of Medicine, Cardiovascular Division, St. Louis, Missouri, 63110