**Multi-ancestry genome-wide association study incorporating gene-alcohol interactions identifies new lipid loci.**

Paul S de Vries, Michael R Brown, Amy R Bentley, Yun J Sung, Thomas W Winkler, Ioanna Ntalla, Karen Schwander, Aldi T Kraja, Xiuqing Guo, Nora Franceschini, Ching-Yu Cheng, Xueling Sim, Dina Vojinovic, Jennifer E Huffman, Solomon K Musani, Changwei Li, Mary F Feitosa, Melissa A Richard, Raymond Noordam, Hugues Aschard, Traci M Bartz, Lawrence F Bielak, Xuan Deng, Rajkumar Dorajoo, Kurt K Lohman, Alisa K Manning, Tuomo Rankinen, Albert V Smith, Salman M Tajuddin, Evangelos Evangelou, Mariaelisa Graff, Maris Alver, Mathilde Boissel, Jin Fang Chai, Xu Chen, Jasmin Divers, Ilaria Gandin, Chuan Gao, Anuj Goel, Yanick Hagemeijer, Sarah E Harris, Fernando P Hartwig, Meian He, Andrea RVR Horimoto, Fang-Chi Hsu, Anne U Jackson, Anuradhani Kasturiratne, Pirjo Komulainen, Brigitte Kühnel, Federica Laguzzi, Joseph H Lee, Jian'an Luan, Leo-Pekka Lyytikäinen, Nana Matoba, Ilja M Nolte, Maik Pietzner, Muhammad Riaz, M Abdullah Said, Robert A Scott, Tamar Sofer, Alena Stančáková, Fumihiko Takeuchi, Bamidele O Tayo, Peter J van der Most, Tibor V Varga, Yajuan Wang, Erin B Ware, Wanqing Wen, Yong-Bing Xiang, Lisa R Yanek, Weihua Zhang, Jing Hua Zhao, Saima Afaq, Najaf Amin, Marzyeh Amini, Dan E Arking, Tin Aung, Christie Ballantyne, Eric Boerwinkle, Ulrich Broeckel, Archie Campbell, Mickaël Canouil, Sabanayagam Charumathi, Yii-Der Ida Chen, John M Connell, Ulf de Faire, Lisa de las Fuentes, Renée de Mutsert, H Janaka de Silva, Jingzhong Ding, Anna F Dominiczak, Qing Duan, Charles B Eaton, Ruben N Eppinga, Jessica D Faul, Virginia Fisher, Nita G Forouhi, Terrence Forrester, Oscar H Franco, Yechiel Friedlander, Mohsen Ghanbari, Franco Giulianini, Hans J Grabe, Megan L Grove, C Charles Gu, Göran Hallmans, Tamara B Harris, Sami Heikkinen, Chew-Kiat Heng, Makoto Hirata, James E Hixson, Barbara V Howard, M Arfan Ikram, InterAct Consortium,

David R Jacobs Jr, Craig Johnson, Jost Bruno Jonas, Candace M Kammerer, Tomohiro Katsuya, Chiea Chuen Khor, Tuomas O Kilpeläinen, Woon-Puay Koh, Heikki A Koistinen, Ivana Kolcic, Charles Kooperberg, Jose E Krieger, Steve B Kritchevsky, Michiaki Kubo, Johanna Kuusisto, Timo A Lakka, Carl D Langefeld, Claudia Langenberg, Lenore J Launer, Benjamin Lehne, Rozenn N Lemaitre, Yize Li, Jingjing Liang, Jianjun Liu, Kiang Liu, Marie Loh, Tin Louie, Reedik Mägi, Ani W Manichaikul, Colin A McKenzie, Thomas Meitinger, Andres Metspalu, Yuri Milaneschi, Lili Milani, Karen L Mohlke, Yukihide Momozawa, Thomas H Mosley Jr, Kenneth J Mukamal, Mike A Nalls, Matthias Nauck, Christopher P Nelson, Nona Sotoodehnia, Jeff R O'Connell, Nicholette D Palmer, Raha Pazoki, Nancy L Pedersen, Annette Peters, Patricia A Peyser, Ozren Polasek, Neil Poulter, Leslie J Raffel, Olli T Raitakari, Alex P Reiner, Treva K Rice, Stephen S Rich, Antonietta Robino, Jennifer G Robinson, Lynda M Rose, Igor Rudan, Carsten O Schmidt, Pamela J Schreiner, William R Scott, Peter Sever, Yuan Shi, Stephen Sidney, Mario Sims, Blair H Smith, Jennifer A Smith, Harold Snieder, John M Starr, Konstantin Strauch, Nicholas Tan, Kent D Taylor, Yik Ying Teo [-], Yih Chung Tham, André G Uitterlinden, Diana van Heemst, Dragana Vuckovic, Melanie Waldenberger, Lihua Wang, Yujie Wang, Zhe Wang, Wen Bin Wei, Christine Williams, Gregory Wilson Sr, Mary K Wojczynski, Jie Yao, Bing Yu, Caizheng Yu, Jian-Min Yuan, Wei Zhao, Alan B Zonderman, Diane M Becker, Michael Boehnke, Donald W Bowden, John C Chambers , Richard S Cooper, Ian J Deary, Tõnu Esko, Martin Farrall, Paul W Franks [-], Barry I Freedman, Philippe Froguel, Paolo Gasparini, Christian Gieger, Bernardo L Horta, Yoichiro Kamatani, Norihiro Kato, Jaspal S Kooner, Markku Laakso, Cathy C Laurie, Karin Leander, Terho Lehtimäki, Lifelines Cohort Study, Patrik KE Magnusson, Thomas Perls, Brenda Penninx, Alexandre C Pereira, Rainer Rauramaa, Nilesh J Samani, James Scott, Xiao-Ou Shu, Pim van der Harst, Lynne E Wagenknecht, Ya Xing

Wang, Nicholas J Wareham, Hugh Watkins, David R Weir, Ananda R Wickremasinghe, Tangchun Wu, Wei Zheng, Paul Elliott, Kari E North, Claude Bouchard, Michele K Evans, Vilmundur Gudnason, Ching-Ti Liu, Yongmei Liu, Bruce M Psaty, Paul M Ridker, Rob M van Dam, Sharon LR Kardia, Xiaofeng Zhu, Charles N Rotimi, Dennis O Mook-Kanamori, Myriam Fornage, Tanika N Kelly, Ervin R Fox, Caroline Hayward, Cornelia M van Duijn, E Shyong Tai, Tien Yin Wong, Jingmin Liu, Jerome I Rotter, W James Gauderman, Michael A Province, Patricia B Munroe, Kenneth Rice, Daniel I Chasman, L Adrienne Cupples, Dabeeru C Rao, and Alanna C Morrison.

Abbreviations:

LDL-C: Low-density lipoprotein cholesterol

HDL-C: High-density lipoprotein cholesterol

TG: Triglycerides

DF: Degrees of freedom

GWAS: Genome-wide association study

FDR: False discovery rate

Correspondence should be addressed to:

Paul S de Vries: Human Genetics Center, Department of Epidemiology, Human Genetics, and Environmental Sciences, School of Public Health, The University of Texas Health Science Center at Houston, 1200 Pressler St., Suite E-429, Houston, TX 77030, Tel 713-500-9809, Fax: 713-500-0900, Email: paul.s.devries@uth.tmc.edu.

Dabeeru C Rao: Division of Biostatistics, Washington University School of Medicine, 660 S. Euclid Avenue, Campus Box 8067, St. Louis, MO 63110, Tel: 314-362-3608, Fax: 314-362-2693, Email: rao@wustl.edu.

Alanna C Morrison: Human Genetics Center, Department of Epidemiology, Human Genetics, and Environmental Sciences, School of Public Health, The University of Texas Health Science Center at Houston, 1200 Pressler St., Suite E-447, Houston, TX 77030, Tel 713-500-9913, Fax: 713-500-0900, Email: alanna.c.morrison@uth.tmc.edu.

Running head: Gene-alcohol interactions and lipid levels

ABSTRACT

An individual's lipid profile is influenced by genetic variants and alcohol consumption, but the contribution of interactions between these exposures has not been studied. We therefore incorporated gene-alcohol interactions into a multi-ancestry genome-wide association study of levels of high-density lipoprotein cholesterol, low-density lipoprotein cholesterol, and triglycerides. We included 45 studies in Stage 1 (genome-wide discovery), and 66 studies in Stage 2 (focused follow-up), for a total of 394,584 individuals from five ancestry group. Genetic main and interaction effects were jointly assessed by a 2 degrees of freedom (DF) test, and a 1 DF test was used to assess the interaction effects alone. Variants at 495 loci were suggestively associated ($P<1\times10^{-6}$) with lipid levels in Stage 1 and were evaluated in Stage 2, followed by combined analyses of Stage 1 and Stage 2. In the combined analysis of Stage 1 and Stage 2, 147 independent loci were associated with lipid levels at $P<5\times10^{-8}$ using 2 DF tests, of which 18 were novel. No genome-wide significant associations were found testing the interaction effect alone. The novel loci included several genes (*PCSK5*, *VEGFB*, and *A1CF*) with a putative role in lipid metabolism based on existing evidence from cellular and experimental models.

KEYWORDS

Alcohol consumption, gene-environment interactions, gene-lifestyle interactions, genome-wide association study, lipid levels, cholesterol, triglycerides

Serum concentrations of high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C), and triglycerides (TG) are modifiable risk factors for cardiovascular disease, the leading cause of death globally (1). Lipid levels are influenced by multiple exposures, including genetic and lifestyle factors. The genetic factors influencing lipid levels have been widely studied (2-8), and large-scale genome-wide association studies (GWAS) have identified 236 loci associated with HDL-C, LDL-C, and TG, which account for up to ~12 percent of the total trait variance in the studied populations (5, 7).

Lifestyle factors, such as alcohol consumption, also associate considerably with lipid levels: in epidemiologic studies, higher alcohol consumption is associated with improved lipid profile, including associations with HDL-C levels, HDL particle concentration, and HDL-C subfractions (9, 10). The relationship between alcohol use and LDL-C or TG is less clear, with some studies reporting positive while others reported negative associations (11-20). Recent Mendelian randomization studies support the causal role of regular low-to-moderate alcohol consumption in improving overall lipid profile (21, 22).

Potential modification of genetic effects on lipid levels by lifestyle exposures, including alcohol consumption, is relatively unexplored (23). Genetic association studies accounting for potential gene-alcohol interactions may lead to the identification of novel lipid loci, and may reveal new biological insights that can potentially be explored for treatment or prevention of dyslipidemia. We hypothesize that alcohol consumption modifies the effect of genetic variants on lipid levels. In order investigate the potential modulating role of alcohol consumption in the genetic architecture of lipid levels, and identify novel HDL-C, LDL-C, and TG loci , we performed genome-wide gene-alcohol interaction meta-analyses of LDL-C, HDL-C and TG.

## METHODS

### Overall design

**Figure 1** shows the overall design of this study, conducted within the setting of the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium Gene-Lifestyle Interactions Working Group (24, 25). In order to decrease the computational burden we carried out genome-wide analyses in Stage 1, and followed up suggestively associated variants in Stage 2, with the combined analysis of Stage 1 and Stage 2 serving as the primary analysis (26). We used two complementary approaches to model interactions: 1) a 2 degrees of freedom (DF) test was used to jointly assess both the genetic main effect and interaction effect on lipid levels, and 2) a 1 DF test was used to assess the effect of interactions alone. The 2 DF test is more powerful when there is both a genetic main and interaction effect, and it may thus help identify interaction effects for which the 1 DF test is underpowered (27).

### Overview of participating studies

This study includes men and women between the ages of 18-80 from five ancestry groups: European, African, Asian, Hispanic, and Brazilian. The participating studies are described in the **Web Appendix**. Each study obtained informed consent from participants and approval from the appropriate institutional review boards. Although the participating studies are based on different study designs and populations, all of them have data on lipid levels, alcohol consumption, and genotypes across the genome.

In total, this study comprises 394,584 individuals. A total of 45 studies participated in Stage 1 (**Web Table 1**), including 89,893 European, 20,989 African, 12,450 Asian, and 3,994 Hispanic ancestry participants for an overall total of 127,326 individuals. A total of 66 studies participated

in Stage 2 (**Web Table 2**), including 136,986 European, 4,475 African, 108,431 Asian, 13,714 Hispanic, and 3,652 Brazilian ancestry individuals for an overall total of 267,258 individuals.

Phenotype and lifestyle variables

Three lipids traits were analyzed separately: HDL-C (mg/dL), LDL-C (mg/dL), and TG (mg/dL). HDL-C and TG were directly assayed, while LDL-C was either directly assayed or estimated using the Friedewald equation: LDL-C = TC – HDL-C – (TG / 5) (28). Only fasting samples (≥ 8 hours) were used to assay TG, and the Friedewald equation was only used in samples with fasting TG ≤ 400 mg/dL. LDL-C values were adjusted for use of statins (**Web Appendix**). HDL-C and TG were natural log transformed prior to analyses.

Alcohol consumption was assessed using two dichotomized alcohol consumption variables: 'current drinking' status, defined as any recurrent drinking behavior, and 'regular drinking' status, as the subset of current drinkers who consume at least two drinks per week. Because the standard pure ethanol content in one alcoholic drink may vary among countries, for this study a standard drink was defined to contain approximately 13g of pure ethanol, and this measure was used to standardize the definitions across studies.

Genotyping and imputation

Information on genotyping and imputation for each of the Stage 1 and Stage 2 studies are presented in **Web Table 3** and **Web Table 4** respectively. For imputation, most studies used the 1000 Genomes Project Phase I Integrated Release Version 3 Haplotypes (2010-11 data freeze, 2012-03-14 haplotypes), which contain haplotypes for 1,092 individuals of all ethnic backgrounds (29).

Study-specific analysis

Study-specific regression analyses were performed for each variant, using models containing the genetic variant, alcohol consumption variable (current drinking or regular drinking status), and their interaction. Variants were coded according to the additive model, so that the beta coefficient represents the effect size per copy of the coded allele. These regressions were adjusted for age, sex, ancestry-informative principal components, and study-specific variables where appropriate (such as center for multi-center studies). Information on principal components and study-specific variables adjusted for in each study-specific analysis is provided in **Web Tables 3-4**.

Each study in Stage 1 performed genome-wide association analyses within each ancestry and provided the estimated genetic main effect, estimated interaction effect and a robust estimate of the corresponding covariance matrix. Each study in Stage 2 performed analyses only for the selected variants identified in Stage 1. Study-specific association analyses were performed using various software (**Web Appendix** and **Web Tables 3-4**). Extensive quality control using the R package EasyQC was performed for all study-specific GWAS results, as described in the **Web Appendix** (30).

Meta-analysis

We implemented METAL to meta-analyze the genetic main and interaction effects jointly using the 2 DF approach by Manning *et al*. (27, 31), and to meta-analyze the interaction coefficients alone using inverse-variance weighted meta-analysis (1 DF test). For each meta-analysis, results were obtained from Wald tests, calculated using genetic main effect estimates, interaction effect estimates, and robust estimates of the corresponding covariance matrix.

In Stage 1 ancestry-specific meta-analyses were performed for each of the 12 analyses (3 lipids × 2 alcohol consumption exposures × 2 tests). Genomic control correction was applied twice (32), first to the study-specific GWAS results (**Web Table 5**), and then to the ancestry-specific meta-analysis results. The results from each ancestry group were then combined in a trans-ancestry meta-analysis.

The variants that were suggestively associated (P-value < $1\times10^{-6}$) in any of Stage 1 interaction analyses were pursued for Stage 2 analysis. In Stage 2, we used the same approaches as in Stage 1 to perform ancestry-specific and trans-ancestry meta-analyses. Finally, ancestry-specific and trans-ancestry meta-analyses were performed to combine Stage 1 results with Stage 2 results. Variants with *P*-value < $5\times10^{-8}$ for either the 2 DF joint test of genetic main and interaction effects or the 1 DF test of interaction effects were considered genome-wide significant. False discovery rate (FDR) *q*-values were calculated using the Benjamini and Hochberg method implemented in the "p.adjust" function in R, correcting for the number of tests performed in Stage 1. FDR *q*-values < 0.05 thus indicate a < 5% false discovery rate even after considering the multiple testing introduced by performing genome-wide analyses on multiple outcomes using multiple models. An independent locus was defined as the ±1 Mbp region surrounding an index variant. For each locus the closest genes were determined based on proximity to the index variant. For loci with intergenic index variants we provided the closest gene in each direction.

Additional analyses

The percent of variance explained in HDL-C, LDL-C, and TG by all previously known and novel variants was evaluated in ten studies from multiple ancestries (**Web Appendix**). HaploReg, RegulomeDB, and GTEx were used to annotate variants at significant loci (33-35). We also used the Data-driven Expression Prioritized Integration for Complex Traits (DEPICT)

software to prioritize genes at the 147 loci associated in the combined analysis of Stage 1 and 2. More details on gene prioritization using DEPICT can be found in the **Web Appendix**.

Lastly, we examined the association of index variants at the 147 significant loci with coronary artery disease and myocardial infarction using publicly available summary association results from a large GWAS of these phenotypes performed by the CARDIoGRAMplusC4D consortium (36).

RESULTS

Descriptive statistics of the studies participating in Stage 1 are shown in **Web Table 1**: 56.1 percent of Stage 1 participants were current drinkers and 39.9 percent were regular drinkers. The Stage 1 genome-wide analyses identified 25,115 variants in 495 independent loci that were suggestively associated ($P$-value < $1\times10^{-6}$) with HDL-C, LDL-C, or TG using either the 1 DF test of the interaction or the 2 DF test that jointly assesses genetic main and interaction effects. The 1 DF interaction test identified 356 suggestively associated variants, while the 2 DF joint test identified an additional 24,759. Manhattan and QQ plots are shown in **Web Figures 1 and 2, respectively**.

The 25,115 variants were then evaluated in Stage 2. Descriptive statistics of the studies participating in Stage 2 are shown in **Web Table 2**: 58.5 percent of Stage 2 participants were current drinkers and 41.0 percent were regular drinkers. The combined analysis of Stage 1 and Stage 2 identified 22,590 variants at 147 independent loci at genome-wide significance ($P$-value < $5\times10^{-8}$, **Web Table 6**). All genome-wide significant associations were identified through the 2 DF joint tests of main and interaction effects. There were no genome-wide significant 1 DF interaction associations in the combined analysis of Stage 1 and Stage 2. At genome-wide

significance, 95 of the 147 loci were associated with HDL-C, 66 were associated with LDL-C, and 58 were associated with TG. Out of the 147 loci, 60 loci were associated with more than one lipid trait, as shown in a Venn diagram in **Figure 2**.

Novel loci

Of the 147 identified loci, 18 are novel lipid loci that have not been previously identified by other association studies for HDL-C, LDL-C, TG, or total cholesterol (**Table 1** and **Web Figure 3**) (2-8). A concurrent genetic association study of exonic variants also identified four of these 18 novel loci (37), as indicated in **Table 1**. Eight of the novel loci involved HDL-C, eight involved LDL-C, and seven involved TG, as shown in the heatmap in **Figure 3**. The most significant associations at each of the 18 novel loci all had FDR *q*-values < 0.05 (**Table 1**), indicating that they are unlikely to be false positives introduced by multiple testing. As shown in forest plots (**Web Figure 4**), the 2 DF associations at the novel loci were predominantly driven by genetic main effects, with a smaller contribution from interaction effects. Furthermore, of the 18 index variants, 15 had suggestively significant ($P < 1\times10^{-6}$) genetic main effects in Stage 1 (**Web Table 7**). None of the associations at the 18 novel loci displayed heterogeneity across ancestry groups (**Table 1**).

Known loci

The remaining 129 of the 147 significant loci had been identified in previous GWAS studies of lipid traits (**Web Table 6**) (2-8). This is a subset of all known lipid loci: **Web Table 8** shows the significance of 314 reported index variants in all 236 known lipid loci among all 2 DF joint tests and 1 DF interaction tests of the combined analysis of Stage 1 and Stage 2, or Stage 1 alone for variants not meeting the Stage 2 inclusion criteria (2-8). Considering only the 314 known variants, no 1 DF interactions were significant in the European, African, or trans-ancestry meta-

analyses ($P$-value < $8.8\times10^{-6}$, corresponding to 0.05 / [314 variants × 3 lipid traits × 2 alcohol consumption variables × 3 ancestry groups]).

Additional analyses

The percentage of variance in LDL-C, HDL-C, and TG explained by various loci was calculated in individual studies from multiple ancestries. Across ancestry groups, the mean variance explained by known lipid loci was 9.1 percent for HDL-C, 10.4 percent for LDL-C, and 7.5 percent for TG. The total percentage of additional variance explained by the 18 novel loci, including both genetic main and interaction effects, was 0.2 for HDL-C, 0.3 for LDL-C, and 0.4 for TG. Ancestry-specific and study-specific estimates are shown in **Web Table 9**.

Functional annotations using HaploReg (33) and RegulomeDB (34) for variants at the 147 loci that were associated in the combined analysis of Stage 1 and 2 are presented in **Web Table 10**, and associations of these variants with gene expression levels from the GTEx database (35) in a variety of tissues are shown in **Web Table 11**. A total of 443 variants were associated with gene expression levels, of which 27 variants were indicated by RegulomeDB as having strong evidence for an effect on enhancer function.

Our gene prioritization analyses with DEPICT highlighted (FDR $q$-values < 5 percent) 165 genes at HDL-C associated loci, 110 genes at LDL-C associated loci and 87 genes at TG associated loci (**Web Tables 11-14**). Thus, at some loci multiple potential causal genes were prioritized. DEPICT identified 656, 877 and 497 reconstituted gene sets that were significantly enriched (FDR $q$-values < 5 percent) for genes at HDL-C, LDL-C and TG loci, respectively (**Web Table 15**). This large number of processes and enriched gene sets underscores the complex genetic, biological and physiological mechanisms underlying lipid traits. Among the most significantly

enriched gene sets were processes related to “total amount of body fat” and “abnormal liver morphology”. Finally, DEPICT revealed that genes at associated HDL-C, LDL-C or TG loci were significantly enriched (FDR *q*-values < 5 percent) for expression effects in 23 tissues, 14 cell-types and 12 physiological systems (**Web Table 16**). We found a compelling enrichment of genes acting in hepatocytes and liver processes at associated loci for each of the three traits and of genes acting in adipose tissues for HDL-C and TG loci (**Web Table 16, Web Figure 5**).

Fourteen index variants at known lipid loci were associated with coronary artery disease with P-value < $1.7\times10^{-4}$ (0.05 / [147 variants × 2 disease outcomes]), and eleven of these were also associated with myocardial infarction (**Web Table 17**) (36). None of the index variants at novel loci were significantly associated with these clinical endpoints.

DISCUSSION

We performed a genome-wide association study of lipid levels incorporating interactions with alcohol consumption and identified 147 genome-wide significant lipid loci of which 18 are novel.

Despite the large sample of 394,584 individuals, which is comparable to other successful genetic interaction studies (38, 39), genome-wide significant interactions were not found in the present study. Gene-alcohol interactions also do not appear to have contributed substantially to the discovery of the 18 novel loci, given that the genetic main effect of index variants at 15 of the 18 novel loci passed the Stage 1 suggestive significance threshold. We highlight three of the novel loci below that harbor especially promising candidate genes with putative roles in lipid metabolism based on existing evidence from cellular and experimental models.

One of the newly identified associations for LDL-C maps to *PCSK5*, a member of the same gene family as *PCSK9.* Previous association studies identified loss-of-function variants in *PCSK9* associated with decreased LDL-C levels (40), and *PCSK9* has subsequently been targeted to create new drugs that successfully lower LDL-C levels (41, 42). Although *PCSK5* has not been previously implicated in the regulation of lipid levels by hypothesis-free genetic association studies, several independent lines of evidence support its involvement. First, a candidate gene study found that variants in *PCSK5* were associated with levels of HDL-C levels (43). Additionally, *in vitro* studies of cell lines show that PCSK5 inactivates endothelial lipase directly through cleavage, and that it may also inactivate endothelial lipase and lipoprotein lipase indirectly through activation of Angiopoietin-like 3 (44). Endothelial lipase, lipoprotein lipase, and Angiopoietin-like 3 have all been robustly implicated in the regulation of lipid levels, likely with primary roles in the metabolism of HDL-C and triglycerides (3, 45-47). In our study, the *PCSK5* locus was only associated at genome-wide significance with LDL-C levels, and at nominal significance ($P$-value < 0.05) with TG levels.

One novel locus for TG mapped to the *A1CF* gene, which encodes APOBEC1 Complementation Factor. Liu *et al*. also identified the same index variant (rs41274050) in association with TG in a concurrent study (37). They also showed that introducing the minor allele of rs41274050 in mice led to increased TG levels, confirming the functional role of this missense variant in the regulation of TG levels (37). APOBEC1 Complementation Factor forms an enzymatic complex with APOBEC1 and deaminates Apolipoprotein B mRNA (48). This site specific deamination of C to U results in the production of the apoB48 isoform as opposed to the apoB100 isoform (48). The apoB48 isoform is critical in the assembly and secretion of chylomicrons, which mainly carry dietary-derived triglycerides (49). Interestingly, a recent GWAS in individuals of East

Asian ancestry identified the *APOBEC1* locus for HDL-C levels (5), an association that we confirmed in our analysis (index variant: 12:7725904:ID). Thus, while both of the genes encoding the proteins that catalyze Apolipoprotein B mRNA deamination are associated with lipid levels, they are associated (at genome-wide significance) with levels of different lipids. Nevertheless, at nominal significance both the index variants near *A1CF* and *APOBEC1* were associated with all three lipid traits ($P$-value < 0.05). Given the role of apoB100 in atherosclerosis, promoting the synthesis of apoB48 instead of apoB100 may represent a possible therapeutic strategy for the prevention of cardiovascular disease (50). Neither the index variant at *A1CF* nor *APOBEC1* is significantly associated with coronary artery disease or myocardial infarction in the largest GWAS of these outcomes. However, further studies are needed to characterize their role in cardiovascular disease, given our multi-ancestry design and the European-driven design of the available GWAS data on cardiovascular disease outcomes **(Web Table 17)**.

Variants closest to the *MACROD1* gene were associated with HDL-C levels and TG levels. This locus was also reported in the concurrent study by Liu *et al.* (37), although the index variant in their study was located in the *PLCB3* gene, around 120 Kbp away from the index variant in the present study. In contrast, we found that variants at this locus were associated with expression levels of another nearby gene, *VEGFB*, in adipose and heart tissue (**Web Table 11**). VEGF-B is reportedly involved in endothelial fatty acid transport, with $Vegfb^{-/-}$ mice showing less accumulation of lipids in muscle, heart and brown adipose tissue, but a greater uptake of fatty acids in white adipose tissue, and higher body weight (51). Additionally, inhibition of VEGF-B in a mouse model of type 2 diabetes resulted in improved glycemic profile as well as a reduction of dyslipidemia (52). Mice lacking VEGF-B had lower levels of TG and LDL-C accompanied by

higher levels of HDL-C. Subsequent studies on other mouse models did not corroborate these findings: Dijkstra *et al.* found in an independent strain of mice that knocking out VEGF-B had no effect on TG and total cholesterol levels (53), while Rubciuc *et al.* reported that transduction of the human *VEGFB* gene into mice led to increased vascularity in adipose tissue, and improved lipid profile (54, 55). Our results provide insight into the effects of VEGF-B in humans to complement the divergent reports from rodent studies. The A allele of index variant rs190528931 is associated with decreased expression of *VEGFB* in adipose and heart tissue, decreased levels of HDL-C and increased levels of TG. Additionally, rs190528931 was also associated with nominally significant increased levels of LDL-C ($P$-value < 0.05). Hence, evidence from our study suggests that inhibition of VEGF-B does not improve lipid profile, but instead promotes dyslipidemia.

In addition to *PCSK5*, *A1CF*, and *VEGFB*, we identified other loci with potential biological mechanisms influencing lipid metabolism. Three of the novel loci harbored genes from the Diacylglycerol Kinase gene family: *DGKG* near *ETV5*, *DGKI* near *CREB3L2*, and *DGKQ* near *TMEM175*. The proteins encoded by these genes phosphorylate diacylglycerol to generate phosphatic acid (56). *In vitro* studies have specifically linked *DGKQ* to cholesterol metabolism (57). Silencing of *DGKQ* expression in human adrenocortical cells influenced the expression of many genes related to cholesterol utilization, including *HMGCR*, which is the target of lipid-lowering statins (57). The TG decreasing and HDL-C increasing C allele of the index variant near *DGKQ* was associated with decreased expression levels of the *DGKQ* gene across a variety of tissues.

The strengths of this study include the large sample size and diverse ancestral composition of the sample, and the use of a dense reference panel for genotype imputation (58). A limitation of this

study is the imbalance in ancestry groups between Stage 1 and Stage 2. African ancestry individuals were well-represented in Stage 1, but underrepresented in Stage 2. In contrast, Asian and Hispanic ancestry individuals were relatively underrepresented in Stage 1 compared to Stage 2. A more balanced division of participants across Stage 1 and Stage 2 may have led to the identification of additional loci. Additionally, alcohol consumption may be underreported in both self-administered questionnaires and interviews, leading to a loss of statistical power due to misclassification (59). Similarly, the classification of alcohol consumption into categories such as regular drinkers and current drinkers may have reduced power relative to treating it as a fully quantitative variable (60). Nevertheless, the use of such categories was necessary for harmonizing data from 111 studies with heterogeneous measurement of alcohol consumption. It is plausible that more comprehensive characterization of alcohol consumption could reveal interactions that were missed in our study.

In conclusion, we identified 18 novel loci that were significantly associated with lipid traits, and these include several loci with genes (*PCSK5*, *VEGFB*, and *A1CF*) that have a putative role in lipid metabolism based on existing evidence from cellular and experimental models. The associations identified in this study appear to be driven by genetic main effects and it remains uncertain whether alcohol consumption modifies the effect of genetic variants on lipid levels.

ACKNOWLEDGEMENTS

Affiliations: 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, Texas (Paul S de Vries, Michael R Brown, Megan L Grove, Zhe Wang, Bing Yu, Myriam Fornage, Alanna C Morrison); Center for Research on Genomics and Global Health, National Human Genome Research Institute, National Institutes of Health,

Bethesda, Maryland (Amy R Bentley, Charles M Rotimi); Division of Biostatistics, Washington University School of Medicine, St. Louis, Missouri (Yun Ju Sung, Karen Schwander, Lisa de las Fuentes, C Charles Gu, Yize Li, Treva K Rice, Dabeeru C Rao); Department of Genetic Epidemiology, University of Regensburg, Regensburg, Germany (Thomas W Winkler); Clinical Pharmacology, William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, United Kingdom (Ioanna Ntalla, Patricia B Munroe); Division of Statistical Genomics, Department of Genetics, Washington University School of Medicine, St. Louis, Missouri (Aldi T Kraja, Mary F Feitosa, Lihua Wang, Christine Williams, Mary K Wojczynski, Michael A Province); Genomic Outcomes, Pediatrics, Institute for Translational Genomics and Population Sciences, LABioMed at Harbor-UCLA Medical Center, Torrance, California (Xiuqing Guo, Yii-Der Ida Chen, Kent D Taylor, Jie Yao, Jerome I Rotter); Epidemiology, University of North Carolina Gilling School of Global Public Health, Chapel Hill, North Carolina (Nora Franceschini, Mariaelisa Graff, Yujie Wang, Kari E North); Singapore Eye Research Institute, Singapore National Eye Centre, Singapore, Singapore (Ching-Yu Cheng, Tin Aung, Sabanayagam Charumathi, Yuan Shi, Nicholas Tan, Yih Chung Tham, Tien Yin Wong); Centre for Quantitative Medicine, Academic Medicine Research Institute, Ophthalmology & Visual Sciences Academic Clinical Program (Eye ACP), Duke-NUS Medical School, Singapore, Singapore, Singapore (Ching-Yu Cheng, Sabanayagam Charumathi); Department of Ophthalmology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore (Ching-Yu Cheng, Tin Aung, Nicholas Tan, Tien Yin Wong); Saw Swee Hock School of Public Health, National University Health System and National University of Singapore, Singapore, Singapore (Xueling Sim, Jin Fang Chai, Woon-Puay Koh, Jianjun Liu, Yik Ying Teo, Rob M van Dam, E Shyong Tai); Department of

Epidemiology, Erasmus University Medical Center, Rotterdam, The Netherlands (Dina Vojinovic, Najaf Amin, Oscar H Franco, Mohsen Ghanbari, M Arfan Ikram, André G Uitterlinden, Cornelia M van Duijn); Medical Research Council Human Genetics Unit, Institute of Genetics and Molecular Medicine, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, United Kingdom (Jennifer E Huffman, Caroline Hayward); Jackson Heart Study, Department of Medicine, University of Mississippi Medical Center, Jackson, Mississippi (Solomon K Musani, Mario Sims); Epidemiology and Biostatistics, University of Giorgia at Athens College of Public Health, Athens, Georgia (Changwei Li); Brown Foundation Institute of Molecular Medicine, The University of Texas Health Science Center at Houston, Houston, Texas (Melissa A Richard, Myriam Fornage); Internal Medicine, Gerontology and Geriatrics, Leiden University Medical Center, Leiden, The Netherlands (Raymond Noordam, Diana van Heemst); Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts (Hugues Aschard); Centre de Bioinformatique, Biostatistique et Biologie Intégrative (C3BI), Institut Pasteur, Paris, France (Hugues Aschard) ; Cardiovascular Health Research Unit, Biostatistics and Medicine, University of Washington, Seattle, Washington (Traci M Bartz); School of Public Health, Epidemiology, University of Michigan, Ann Arbor, Michigan (Lawrence F Bielak, Patricia A Peyser, Jennifer A Smith, Wei Zhao, Sharon LR Kardia); Biostatistics, Boston University School of Public Health, Boston, Massachusetts (Xuan Deng, Virginia Fisher, Ching-Ti Liu, L Adrienne Cupples); Genome Institute of Singapore, Agency for Science Technology and Research, Singapore, Singapore (Rajkumar Dorajoo, Chiea Chuen Khor, Jianjun Liu, Yik Ying Teo); Public Health Sciences, Biostatistical Sciences, Wake Forest University Health Sciences, Winston-Salem, North Carolina (Kurt K Lohman); Clinical and Translational Epidemiology Unit, Massachusetts General Hospital, Boston, Massachusetts

(Alissa K Manning); Department of Medicine, Harvard Medical School, Boston, Massachusetts (Alisa K Manning, Tamar Sofer); Human Genomics Laboratory, Pennington Biomedical Research Center, Baton Rouge, Louisiana (Tuomo Rankinen, Claude Bouchard); Icelandic Heart Association, Kopavogur, Iceland (Albert V Smith, Vilmundur Gudnason); Faculty of Medicine, University of Iceland, Reykjavik, Iceland (Albert V Smith, Vilmundur Gudnason); Health Disparities Research Section, Laboratory of Epidemiology and Population Sciences, National Institute on Aging, National Institutes of Health, Baltimore, Maryland (Salman M Tajuddin, Michele K Evans); Department of Epidemiology and Biostatistics, Imperial College London, London, United Kingdom (Evangelos Evangelou, Weihua Zhang, Saima Afaq, Benjamin Lehne, Marie Loh, Raha Pazoki, William R Scott, John C Chambers, Paul Elliott); Department of Hygiene and Epidemiology, University of Ioannina Medical School, Ioannina, Greece (Evangelos Evangelou); Estonian Genome Center, University of Tartu, Tartu, Estonia (Maris Alver, Reedik Mägi, Andres Metspalu, Lili Milani, Tõnu Esko); CNRS UMR 8199, European Genomic Institute for Diabetes (EGID), Institut Pasteur de Lille, University of Lille, Lille, France (Mathilde Boissel, Mickaël Canouil, Philippe Froguel); Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Stockholm, Sweden (Xu Chen, Nancy L Pedersen, Patrik KE Magnusson); Biostatistical Sciences, Public Health Sciences, Wake Forest School of Medicine, Winston-Salem, North Carolina (Jasmin Divers, Fang-Chi Hsu, Carl D Langefeld); Department of Medical Sciences, University of Trieste, Trieste, Italy (Ilaria Gandin, Dragana Vuckovic, Paolo Gasparini); Molecular Genetics and Genomics Program, Molecular Genetics and Genomics Program, Wake Forest School of Medicine, Winston-Salem, North Carolina (Chuan Gao); Division of Cardiovascular Medicine, Radcliffe Department of Medicine, University of Oxford, Oxford, Oxfordshire, United Kingdom (Anuj

Goel, Martin Farrall, Hugh Watkins); Wellcome Centre for Human Genetics, University of Oxford, Oxford, Oxfordshire, United Kingdom (Anuj Goel, Martin Farrall, Hugh Watkins); Department of Cardiology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands (Yanick Hagemeijer, M Abdullah Said, Ruben N Eppinga, Pim van der Harst); Centre for Cognitive Ageing and Cognitive Epidemiology, The University of Edinburgh, Edinburgh, United Kingdom (Sarah E Harris, John M Starr, Ian J Deary); Medical Genetics Section, University of Edinburgh Centre for Genomic and Experimental Medicine and MRC Institute of Genetics and Molecular Medicine, The University of Edinburgh, Edinburgh, United Kingdom (Sarah E Harris); Postgraduate Programme in Epidemiology, Federal University of Pelotas, Pelotas, RS, Brazil (Fernando P Hartwig, Bernardo L Horta); Medical Research Council Integrative Epidemiology Unit, University of Bristol, Bristol, United Kingdom (Fernando P Hartwig); Department of Occupational and Environmental Health and State Key Laboratory of Environmental Health for Incubating, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China (Meian He, Caizheng Yu, Tangchun Wu); Laboratory of Genetics and Molecular Cardiology, Heart Institute (InCor), University of São Paulo Medical School, São Paulo, SP, Brazil (Andrea RVR Horimoto, Jose E Krieger, Alexandre C Pereira); Department of Biostatistics and Center for Statistical Genetics, University of Michigan, Ann Arbor, Michigan (Anne U Jackson, Michael Boehnke); Department of Public Health, Faculty of Medicine, University of Kelaniya, Ragama, Sri Lanka (Anuradhani Kasturiratne, Ananda R Wickremasinghe); Foundation for Research in Health Exercise and Nutrition, Kuopio Research Institute of Exercise Medicine, Kuopio, Finland (Pirjo Komulainen, Timo A Lakka, Rainer Rauramaa); Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany

(Brigitte Kühnel, Melanie Waldenberger, Christian Gieger); Institute of Epidemiology II, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany (Brigitte Kühnel, Annette Peters, Melanie Waldenberger); Unit of Cardiovascular Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden (Federica Laguzzi, Ulf de Faire, Karin Leander); Sergievsky Center, College of Physicians and Surgeons, Columbia University Mailman School of Public Health, New York, New York (Joseph H Lee); MRC Epidemiology Unit, University of Cambridge, Cambridge, United Kingdom (Jian'an Luan, Robert A Scott, Jing Hua Zhao, Nita G Forouhi, Claudia Langenberg, Nicholas J Wareham); Department of Clinical Chemistry, Fimlab Laboratories, Tampere, Finland (Leo-Pekka Lyytikäinen, Terho Lehtimäki); Department of Clinical Chemistry, Finnish Cardiovascular Research Center - Tampere, Faculty of Medicine and Life Sciences, University of Tampere, Tampere, Finland (Leo-Pekka Lyytikäinen, Terho Lehtimäki); Laboratory for Statistical Analysis, Center for Integrative Medical Sciences, RIKEN, Yokohama, Japan (Nana Matoba, Yoichiro Kamatani); Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands (Ilja M Nolte, Peter J van der Most, Marzyeh Amini, Harold Snieder); Institute of Clinical Chemistry and Laboratory Medicine, University Medicine Greifswald, Greifswald, Germany (Maik Pietzner, Matthias Nauck); DZHK (German Centre for Cardiovascular Health), Partner Site Greifswald, Greifswald, Germany (Maik Pietzner, Matthias Nauck); Department of Cardiovascular Sciences, University of Leicester, Leicester, United Kingdom (Muhammad Riaz, Christopher P Nelson, Nilesh J Samani); NIHR Leicester Biomedical Research Centre, Glenfield Hospital, Leicester, United Kingdom (Muhammad Riaz, Christopher P Nelson, Nilesh J Samani); Division of Sleep and Circadian Disorders, Brigham and Women's Hospital, Boston, Massachusetts (Tamar Sofer);

Institute of Clinical Medicine, Internal Medicine, University of Eastern Finland, Kuopio, Finland (Alena Stančáková, Johanna Kuusisto, Markku Laakso); Department of Gene Diagnostics and Therapeutics, Research Institute, National Center for Global Health and Medicine, Tokyo, Japan (Fumihiko Takeuchi, Norihiro Kato); Department of Public Health Sciences, Loyola University Chicago, Maywood, Illinois (Bamidele O Tayo, Richard S Cooper); Department of Clinical Sciences, Genetic and Molecular Epidemiology Unit, Lund University Diabetes Centre, Skåne University Hospital, Malmö, Sweden (Tibor V Varga, Paul W Franks); Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, Ohio (Yajuan Wang, Jingjing Liang, Xiaofeng Zhu); Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, Michigan (Erin B Ware, Jessica D Faul, Jennifer A Smith, David R Weir); Division of Epidemiology, Department of Medicine, Vanderbilt University School of Medicine, Nashville, Tennessee (Wanqing Wen, Xiao-Ou Shu, Wei Zheng); SKLORG & Department of Epidemiology, Shanghai Cancer Institute, Renji Hospital, Shanghai Jiaotong University School of Medicine, Shanghai, P.R. China (Yong-Bing Xiang); Division of General Internal Medicine, Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland (Lisa R Yanek, Diane M Becker); Department of Cardiology, Ealing Hospital, Middlesex, United Kingdom (Weihua Zhang, John C Chambers, Jaspal S Kooner); McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland (Dan E Arking); Ophthalmology & Visual Sciences Academic Clinical Program (Eye ACP), Duke-NUS Medical School, Singapore, Singapore (Tin Aung, Tien Yin Wong); Section of Cardiovascular Research, Baylor College of Medicine, Houston, Texas (Christie Ballantyne); Houston Methodist Debakey Heart and Vascular Center, Houston, Texas (Christie Ballantyne); Department of Epidemiology, Human Genetics, and Environmental

Sciences, The University of Texas School of Public Health, Houston, Texas (Eric Boerwinkle, James E Hixson); Human Genome Sequencing Center, Baylor College of Medicine, Houston, Texas (Eric Boerwinkle); Section of Genomic Pediatrics, Department of Pediatrics, Medicine and Physiology, Medical College of Wisconsin, Milwaukee, Wisconsin (Ulrich Broeckel); Centre for Genomic & Experimental Medicine, Institute of Genetics & Molecular Medicine, University of Edinburgh, Edinburgh, United Kingdom (Archie Campbell); Ninewells Hospital & Medical School, University of Dundee, Dundee, Scotland, United Kingdom (John M Connell); Cardiovascular Division, Department of Medicine, Washington University, St. Louis, Missouri (Lisa de las Fuentes); Clinical Epidemiology, Leiden University Medical Center, Leiden, Netherlands (Renée de Mutsert, Dennis O Mook-Kanamori); Department of Medicine, Faculty of Medicine, University of Kelaniya, Ragama, Sri Lanka (H Janaka de Silva); Center on Diabetes, Obesity, and Metabolism, Gerontology and Geriatric Medicine, Wake Forest University Health Sciences, Winston-Salem, North Carolina (Jingzhong Ding); Institute of Cardiovascular and Medical Sciences, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, Scotland, United Kingdom (Anna F Dominiczak); Department of Genetics, University of North Carolina, Chapel Hill, North Carolina (Qing Duan, Karen L Mohlke); Department of Family Medicine and Epidemiology, Alpert Medical School of Brown University, Providence, Rhode Island (Charles B Eaton); Tropical Metabolism Research Unit, Tropical Medicine Research Institute, University of the West Indies, Mona, Jamaica (Terrence Forrester, Colin A McKenzie); Braun School of Public Health, Hebrew University-Hadassah Medical Center, Jerusalem, Israel (Yechiel Friedlander); Department of Genetics, School of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran (Mohsen Ghanbari); Preventive Medicine, Brigham and Women's Hospital, Boston, Massachusetts (Franco

Giulainini, Lynda M Rose, Paul M Ridker, Daniel I Chasman); Department Psychiatry and Psychotherapy, University Medicine Greifswald, Greifswald, Germany (Hans J Grabe); Department of Public Health and Clinical Medicine, Nutritional Research, Umeå University, Umeå, Västerbotten, Sweden (Göran Hallmans); Laboratory of Epidemiology and Population Sciences, National Institute on Aging, National Institutes of Health, Bethesda, Maryland (Tamara B Harris, Lenore J Launer); Institute of Biomedicine, School of Medicine, University of Eastern Finland, Kuopio Campus, Finland (Sami Heikkinen, Timo A Lakka); Department of Paediatrics, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore (Chew-Kiat Heng); Khoo Teck Puat – National University Children's Medical Institute, National University Health System, Singapore, Singapore (Chew-Kiat Heng); Laboratory of Genome Technology, Human Genome Center, Institute of Medical Science, The University of Tokyo, Minato-ku, Japan (Makoto Hirata); MedStar Health Research Institute, Hyattsville, Maryland (Barbara V Howard); Center for Clinical and Translational Sciences and Department of Medicine, Georgetown-Howard Universities, Washington, District of Columbia (Barbara V Howard); Department of Radiology and Nuclear Medicine, Erasmus University Medical Center, Rotterdam, The Netherlands (M Arfan Ikram); Department of Neurology, Erasmus University Medical Center, Rotterdam, The Netherlands (M Arfan Ikram); Epidemiology & Community Health, School of Public Health, University of Minnesota, Minneapolis, Minnesota (David R Jacobs Jr, Pamela J Schreiner); Collaborative Health Studies Coordinating Center, University of Washington, Seattle, Washington (Craig Johnson); Department of Ophthalmology, Medical Faculty Mannheim, University Heidelberg, Mannheim, Germany (Jost Bruno Jonas); Beijing Institute of Ophthalmology, Beijing Ophthalmology and Visual Science Key Lab, Beijing Tongren Eye Center, Capital Medical University, Beijing,

China (Jost Bruno Jonas); Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, Pennsylvania (Candace M Kammerer); Department of Clinical Gene Therapy, Osaka University Graduate School of Medicine, Suita, Japan (Tomohiro Katsuya); Department of Geriatric Medicine and Nephrology, Osaka University Graduate School of Medicine, Suita, Japan (Tomohiro Katsuya); Department of Biochemistry, National University of Singapore, Singapore, Singapore (Chiea Chuen Khor); Novo Nordisk Foundation Center for Basic Metabolic Research, Section of Metabolic Genetics, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark (Tuomas O Kilpeläinen); Department of Environmental Medicine and Public Health, The Icahn School of Medicine at Mount Sinai, New York, New York (Tuomas O Kilpeläinen); Duke-NUS Medical School, Singapore, Singapore (Woon-Puay Koh, E Shyong Tai); Department of Health, National Institute for Health and Welfare, Helsinki, Finland (Heikki A Koistinen); Department of Medicine and Abdominal Center: Endocrinology, University of Helsinki and Helsinki University Central Hospital, Helsinki, Finland (Heikki A Koistinen); Minerva Foundation Institute for Medical Research, Helsinki, Finland (Heikki A Koistinen); Department of Public Health, Department of Medicine, University of Split, Split, Croatia (Ivan Kolcic, Ozren Polasek); Fred Hutchinson Cancer Research Center, University of Washington School of Public Health, Seattle, Washington (Charles Kooperberg, Alex P Reiner); Sticht Center for Health Aging and Alzheimer's Prevention, Department of Internal Medicine, Wake Forest School of Medicine, Winston-Salem, North Carolina (Steve B Kritchevsky); Center for Integrative Medical Sciences, RIKEN, Yokohama, Japan (Michiaki Kubo); Department of Clinical Physiology and Nuclear Medicine, Kuopio University Hospital, Kuopio, Finland (Timo A Lakka); Cardiovascular Health Research Unit, Medicine, University of Washington, Seattle, Washington (Rozenn N Lemaitre);

Epidemiology, Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, Illinois (Kiang Liu); Translational Laboratory in Genetic Medicine, Agency for Science, Technology and Research, Singapore (Marie Loh); Department of Biostatistics, University of Washington, Seattle, Washington (Tin Louie, Cathy C Laurie, Kenneth Rice); Biostatistics Section, Center for Public Health Genomics, University of Virginia, School of Medicine, West Complex, Charlottesville, Virginia (Ani W Manichaikul); Institute of Human Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany (Thomas Meitinger); Institute of Human Genetics, Technische Universität München, Munich, Germany (Thomas Meitinger); Department of Psychiatry, Amsterdam Neuroscience and Amsterdam Public Health Research Institute, VU University Medical Center, Amsterdam, The Netherlands (Yuri Milaneschi, Brenda Penninx); Laboratory for Genotyping Development, Center for Integrative Medical Sciences, RIKEN, Yokohama, Japan (Yukihide Momozawa); Geriatrics, Medicine, University of Mississippi, Jackson, Mississippi (Thomas H Mosley Jr); General Medicine & Primary Care, Medicine, Beth Israel Deaconess Medical Center, Boston, Massachusetts (Kenneth J Mukamal); Data Tecnica International, Glen Echo, Maryland (Mike A Nalls); Laboratory of Neurogenetics, National Institute on Aging, Bethesda, Maryland (Mike A Nalls); Cardiovascular Health Research Unit, Division of Cardiology, University of Washington, Seattle, Washington (Nona Sotoodehnia); Division of Endocrinology, Diabetes, and Nutrition, University of Maryland School of Medicine, Baltimore, Maryland (Jeff R O'Connell); Program for Personalized and Genomic Medicine, University of Maryland School of Medicine, Baltimore, Maryland (Jeff R O'Connell); Biochemistry, Wake Forest School of Medicine, Winston-Salem, North Carolina (Nicholette D Palmer, Donald W Bowden); DZHK (German Centre for Cardiovascular Research), partner site Munich Heart Alliance, Neuherberg,

Germany (Annette Peters); Psychiatric Hospital "Sveti Ivan", Zagreb, Croatia (Ozren Polasek); Gen-info Ltd, Zagreb, Croatia (Ozren Polasek); School of Public Health, Imperial College London, London, United Kingdom (Neil Poulter); Division of Genetic and Genomic Medicine, Department of Pediatrics, University of California, Irvine, California (Leslie J Raffel); Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku, Finland (Olli T Raitakari); Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, Finland (Olli T Raitakari); Center for Public Health Genomics, University of Virginia, School of Medicine, West Complex, Charlottesville, Virginia (Stephen S Rich); Institute for Maternal and Child Health - IRCCS "Burlo Garofolo", Trieste, Italy (Antonietta Robino, Paolo Gasparini); Department of Epidemiology and Medicine, University of Iowa, Iowa City, Iowa (Jennifer G Robinson); Centre for Global Health Research, Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, United Kingdom (Igor Rudan); Institute for Community Medicine, University Medicine Greifswald, Greifswald, Germany (Carsten O Schmidt); National Heart and Lung Institute, Imperial College London, London, United Kingdom (William R Scott, Peter Sever, Jaspal S Kooner, James Scott); Division of Research, Kaiser Permanente of Northern California, Oakland, California (Stephen Sidney); Division of Population Health Sciences, Ninewells Hospital and Medical School, University of Dundee, Dundee, United Kingdom (Blair H Smith); Alzheimer Scotland Dementia Research Centre, The University of Edinburgh, Edinburgh, United Kingdom (John M Starr); Institute of Genetic Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany (Konstantin Strauch); Chair of Genetic Epidemiology, IBE, Faculty of Medicine, LMU, Munich, Germany (Konstantin Strauch); Life Sciences Institute, National University of Singapore, Singapore, Singapore (Yik Ying Teo); NUS

Graduate School for Integrative Science and Engineering, National University of Singapore, Singapore, Singapore (Yik Ying Teo); Department of Statistics and Applied Probability, National University of Singapore, Singapore, Singapore (Yik Ying Teo); Department of Internal Medicine, Erasmus University Medical Center, Rotterdam, The Netherlands (André G Uitterlinden); Beijing Tongren Eye Center, Beijing Tongren Hospital, Capital Medical University, Beijing, China (Wen Bin Wei); Jackson Heart Study, School of Public Health, Jackson State University, Jackson, Mississippi (Gregory Wilson Sr); Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, Pennsylvania (Jian-Min Yuan); Division of Cancer Control and Population Sciences, UPMC Hillman Cancer, University of Pittsburgh, Pittsburgh, Pennsylvania (Jian-Min Yuan); Behavioral Epidemiology Section, Laboratory of Epidemiology and Population Sciences, National Institute on Aging, National Institutes of Health, Baltimore, Maryland (Alan B Zonderman); Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore (John C Chambers); Imperial College Healthcare NHS Trust, London, United Kingdom (John C Chambers, Jaspal S Kooner); MRC-PHE Centre for Environment and Health, Imperial College London, London, United Kingdom (John C Chambers, Jaspal S Kooner, Paul Elliott); Psychology, The University of Edinburgh, Edinburgh, United Kingdom (Ian J Deary); Broad Institute of the Massachusetts Institute of Technology and Harvard University, Boston, Massachusetts (Tõnu Esko); Harvard T. H. Chan School of Public Health, Department of Nutrition, Harvard University, Boston, Massachusetts (Paul W Franks); Department of Public Health & Clinical Medicine, Umeå University, Umeå, Västerbotten, Sweden (Paul W Franks); OCDEM, Radcliffe Department of Medicine, University of Oxford, Oxford, United Kingdom (Paul W Franks); Nephrology, Internal Medicine, Wake Forest School of Medicine, Winston-

Salem, North Carolina (Barry I Freedman); Department of Genomics of Common Disease, Imperial College London, London, United Kingdom (Philippe Froguel); German Center for Diabetes Research (DZD e.V.), Neuherberg, Germany (Christian Gieger); Lifelines Cohort, Groningen, The Netherlands (Lifelines Cohort Study); Geriatrics Section, Boston University Medical Center, Boston, Massachusetts (Thomas Perls); Department of Genetics, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands (Pim van der Harst); Public Health Sciences, Wake Forest School of Medicine, Winston-Salem, North Carolina (Lynne E Wagenknecht); Beijing Institute of Ophthalmology, Beijing Tongren Eye Center, Beijing Ophthalmology and Visual Science Key Lab, Beijing Tongren Hospital, Capital Medical University, Beijing, China (Ya Xing Wang); Carolina Center of Genome Sciences, University of North Carolina, Chapel Hill, North Carolina (Kari E North); Public Health Sciences, Epidemiology and Prevention, Wake Forest University Health Sciences, Winston-Salem, North Carolina (Yongmei Liu); Cardiovascular Health Research Unit, Epidemiology, Medicine and Health Services, University of Washington, Seattle, Washington (Bruce Psaty); Kaiser Permanente Washington, Health Research Institute, Seattle, Washington Bruce Psaty); Harvard Medical School, Boston, Massachusetts (Paul M Ridker, Daniel I Chasman); Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore (Rob M van Dam, E Shyong Tai); Public Health and Primary Care, Leiden University Medical Center, Leiden, Leiden (Dennis O Mook-Kanamori); Epidemiology, Tulane University School of Public Health and Tropical Medicine, New Orleans, Louisiana (Tanika N Kelly); Cardiology, Medicine, University of Mississippi Medical Center, Jackson, Mississippi (Ervin R Fox); WHI CCC, Fred Hutchinson Cancer Research Center, Seattle, Washington Jingmin Liu); Biostatistics, Preventive Medicine, University of Southern California, Los

Angeles, California (W James Gauderman); NIHR Barts Cardiovascular Biomedical Research Unit, Queen Mary University of London, London, United Kingdom (Patricia B Munroe); and NHLBI Framingham Heart Study, Framingham, Massachusetts (L Adrienne Cupples).

Equal contributions: Paul S de Vries, Michael R Brown, Amy R Bentley, Yun J Sung, Thomas W Winkler, Ioanna Ntalla contributed equally to the manuscript as junior authors. Patricia B Munroe, Kenneth Rice, Daniel I Chasman, L Adrienne Cupples, Dabeeru C Rao, Alanna C Morrison contributed equally to the manuscript as senior authors.

Funding: Infrastructure for the CHARGE Consortium is supported in part by the National Heart, Lung, and Blood Institute grant R01HL105756. Infrastructure for the Gene-Lifestyle Working Group in particular is supported by the National Heart, Lung, and Blood Institute grant R01HL118305. American Heart Association Grant #17POST33350042 made it possible for Dr. de Vries to lead this project. We thank the participants and staff of all of the included studies. We thank Matthew Brown for his help in tracking author affiliations and contributions. Study-specific funding and acknowledgements are provided below, starting with Stage 1 studies followed by Stage 2.

*AGES (Age Gene/Environment Susceptibility Reykjavik Study):* This study has been funded by NIH contract N01-AG012100, the NIA Intramural Research Program, an Intramural Research Program Award (ZIAEY000401) from the National Eye Institute, an award from the National Institute on Deafness and Other Communication Disorders (NIDCD) Division of Scientific Programs (IAA Y2-DC_1004-02), Hjartavernd (the Icelandic Heart Association), and the Althingi (the Icelandic Parliament). The study is approved by the Icelandic National Bioethics Committee, VSN: 00-063. The researchers are indebted to the participants for their willingness to participate in the study.

*ARIC (Atherosclerosis Risk in Communities) Study:* The ARIC study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts (HHSN268201100005C, HHSN268201100006C, HHSN268201100007C, HHSN268201100008C, HHSN268201100009C, HHSN268201100010C, HHSN268201100011C, and HHSN268201100012C), R01HL087641, R01HL59367 and R01HL086694; National Human Genome Research Institute contract U01HG004402; and National Institutes of Health contract HHSN268200625226C. The authors thank the staff and participants of the ARIC study for their important contributions. Infrastructure was partly supported by Grant Number UL1RR025005, a component of the National Institutes of Health and NIH Roadmap for Medical Research.

*Baependi Heart Study (Brazil):* The Baependi Heart Study was supported by Fundação de Amparo a Pesquisa do Estado de São Paulo (FAPESP) (Grant 2013/17368-0), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and Hospital Samaritano Society (Grant 25000.180.664/2011-35), through Ministry of Health to Support Program Institutional Development of the Unified Health System (SUS-PROADI).

*CARDIA (Coronary Artery Risk Development in Young Adults):* The CARDIA Study is conducted and supported by the National Heart, Lung, and Blood Institute in collaboration with the University of Alabama at Birmingham (HHSN268201300025C & HHSN268201300026C), Northwestern University (HHSN268201300027C), University of Minnesota (HHSN268201300028C), Kaiser Foundation Research Institute (HHSN268201300029C), and Johns Hopkins University School of Medicine (HHSN268200900041C). CARDIA is also partially supported by the Intramural Research Program of the National Institute on Aging. Genotyping was funded as part of the NHLBI Candidate-gene Association Resource (N01-HC-65226) and the NHGRI Gene Environment Association Studies (GENEVA) (U01-HG004729, U01-HG04424, and U01-HG004446). This manuscript has been reviewed and approved by CARDIA for scientific content.

*CHS (Cardiovascular Health Study):* This CHS research was supported by NHLBI contracts HHSN268201200036C, HHSN268200800007C, HHSN268200960009C, N01HC55222, N01HC85079, N01HC85080, N01HC85081, N01HC85082, N01HC85083, N01HC85086; and NHLBI grants U01HL080295, R01HL085251, R01HL087652, R01HL105756, R01HL103612, R01HL120393 and R01HL130114 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. The provision of genotyping data was supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1TR000124, 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. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

*ERF (Erasmus Rucphen Family study):* The ERF study as a part of EUROSPAN (European Special Populations Research Network) was supported by European Commission FP6 STRP grant number 018947 (LSHG-CT-2006-01947) and also received funding from the European Community's Seventh Framework Programme (FP7/2007-2013)/grant agreement HEALTH-F4-2007-201413 by the European Commission under the programme "Quality of Life and Management of the Living Resources" of 5th Framework Programme (no. QLG2-CT-2002-01254). The ERF study was further supported by ENGAGE consortium and CMSB. High-throughput analysis of the ERF data was supported by joint grant from Netherlands Organisation for Scientific Research and the Russian Foundation for Basic Research (NWO-RFBR 047.017.043). ERF was further supported by the ZonMw grant (project 91111025). We are grateful to all study participants and their relatives, general practitioners and neurologists for their contributions and to P. Veraart for her help in genealogy, J. Vergeer for the supervision of the laboratory work, P. Snijders for his help in data collection and E.M. van Leeuwen for genetic imputation.

*FamHS (Family Heart Study):* The FamHS is funded by R01HL118305 and R01HL117078 NHLBI grants, and 5R01DK07568102 and 5R01DK089256 NIDDK grant.

*FHS (Framingham Heart Study):* This research was conducted in part using data and resources from the Framingham Heart Study of the National Heart Lung and Blood Institute of the National Institutes of Health and Boston University School of Medicine. The analyses reflect intellectual input and resource development from the Framingham Heart Study investigators participating in the SNP Health Association Resource (SHARe) project. This work was partially supported by the National Heart, Lung and Blood Institute's Framingham Heart Study (Contract Nos. N01-HC-25195 and HHSN268201500001I) and its contract with Affymetrix, Inc for genotyping services (Contract No. N02-HL-6-4278). A portion of this research utilized the Linux Cluster for Genetic Analysis (LinGA-II) funded by the Robert Dawson Evans Endowment of the Department of Medicine at Boston University School of Medicine and Boston Medical Center. This research was partially supported by grant R01-DK089256 from the National Institute of Diabetes and Digestive and Kidney Diseases (MPIs: Ingrid B. Borecki, L. Adrienne Cupples, Kari North).

*GENOA (Genetic Epidemiology Network of Arteriopathy):* Support for GENOA was provided by the National Heart, Lung and Blood Institute (HL119443, HL118305, HL054464, HL054457, HL054481, HL071917 and HL087660) of the National Institutes of Health. Genotyping was performed at the Mayo Clinic (Stephen T. Turner, MD, Mariza de Andrade PhD, Julie Cunningham, PhD). We thank Eric Boerwinkle, PhD and Megan L. Grove from the Human Genetics Center and Institute of Molecular Medicine and Division of Epidemiology, University of Texas Health Science Center, Houston, Texas, USA for their help with genotyping. We would also like to thank the families that participated in the GENOA study.

*GenSalt (Genetic Epidemiology Network of Salt Sensitivity):* The Genetic Epidemiology Network of Salt Sensitivity is supported by research grants (U01HL072507, R01HL087263, and R01HL090682) from the National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD.

*HANDLS (Healthy Aging in Neighborhoods of Diversity across the Life Span):* The Healthy Aging in Neighborhoods of Diversity across the Life Span (HANDLS) study was supported by the Intramural Research Program of the NIH, National Institute on Aging and the National Center on Minority Health and Health Disparities (project # Z01-AG000513 and human subjects protocol number 09-AG-N248). Data analyses for the HANDLS study utilized the high-performance computational resources of the Biowulf Linux cluster at the National Institutes of Health, Bethesda, MD (http://biowulf.nih.gov; http://hpc.nih.gov)).

*Health ABC (Health, Aging, and Body Composition):* Health ABC was funded by the National Institutes of Aging. This research was supported by NIA contracts N01AG62101, N01AG62103, and N01AG62106. The GWAS was funded by NIA grant 1R01AG032098-01A1 to Wake Forest University Health Sciences and genotyping services were provided by the Center for Inherited Disease Research (CIDR). CIDR is fully funded through a federal contract from the National Institutes of Health to The Johns Hopkins University, contract number HHSN268200782096C. This research was supported in part by the Intramural Research Program of the NIH, National Institute on Aging.

*HERITAGE (Health, Risk Factors, Exercise Training and Genetics):* The HERITAGE Family Study was supported by National Heart, Lung, and Blood Institute grant HL-45670.

*HUFS (Howard University Family Study):* The Howard University Family Study was supported by National Institutes of Health grants S06GM008016-320107 to Charles Rotimi and S06GM008016-380111 to Adebowale Adeyemo. We thank the participants of the study, for which enrollment was carried out at the Howard University General Clinical Research Center, supported by National Institutes of Health grant 2M01RR010284. The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official view of the National Institutes of Health. This research was supported in part by the Intramural Research Program of the Center for Research on Genomics and Global Health (CRGGH). The CRGGH is supported by the National Human Genome Research Institute,

the National Institute of Diabetes and Digestive and Kidney Diseases, the Center for Information Technology, and the Office of the Director at the National Institutes of Health (Z01HG200362). Genotyping support was provided by the Coriell Institute for Medical Research.

*HyperGEN (Hypertension Genetic Epidemiology Network):* The hypertension network is funded by cooperative agreements (U10) with NHLBI: HL54471, HL54472, HL54473, HL54495, HL54496, HL54497, HL54509, HL54515, and 2 R01 HL55673-12. The study involves: University of Utah: (Network Coordinating Center, Field Center, and Molecular Genetics Lab); Univ. of Alabama at Birmingham: (Field Center and Echo Coordinating and Analysis Center); Medical College of Wisconsin: (Echo Genotyping Lab); Boston University: (Field Center); University of Minnesota: (Field Center and Biochemistry Lab); University of North Carolina: (Field Center); Washington University: (Data Coordinating Center); Weil Cornell Medical College: (Echo Reading Center); National Heart, Lung, & Blood Institute. For a complete list of HyperGEN Investigators: http://www.biostat.wustl.edu/hypergen/Acknowledge.html

*CROATIA-Korcula:* We would like to acknowledge the staff of several institutions in Croatia that supported the field work, including but not limited to The University of Split and Zagreb Medical Schools and the Croatian Institute for Public Health. We would like to acknowledge the invaluable contributions of the recruitment team in Korcula, the administrative teams in Croatia and Edinburgh and the participants. The SNP genotyping for the CROATIA-Korcula cohort was performed in Helmholtz Zentrum München, Neuherberg, Germany. CROATIA-Korcula (CR-Korcula) was funded by the Medical Research Council UK, The Croatian Ministry of Science, Education and Sports (grant 216-1080315-0302), the European Union framework program 6 EUROSPAN project (contract no. LSHG-CT-2006-018947) and the Croatian Science Foundation (grant 8875).

*CROATIA-Vis:* We would like to acknowledge the staff of several institutions in Croatia that supported the field work, including but not limited to The University of Split and Zagreb Medical Schools, the

Institute for Anthropological Research in Zagreb and Croatian Institute for Public Health. The SNP genotyping for the CROATIA-Vis cohort was performed in the core genotyping laboratory of the Wellcome Trust Clinical Research Facility at the Western General Hospital, Edinburgh, Scotland. CROATIA-Vis (CR-Vis) was funded by the Medical Research Council UK, The Croatian Ministry of Science, Education and Sports (grant 216-1080315-0302), the European Union framework program 6 EUROSPAN project (contract no. LSHG-CT-2006-018947) and the Croatian Science Foundation (grant 8875).

*GS:SFHS:* Generation Scotland received core support from the Chief Scientist Office of the Scottish Government Health Directorates [CZD/16/6] and the Scottish Funding Council [HR03006]. Genotyping of the GS:SFHS samples was carried out by the Genetics Core Laboratory at the Wellcome Trust Clinical Research Facility, Edinburgh, Scotland and was funded by the Medical Research Council UK and the Wellcome Trust (Wellcome Trust Strategic Award "STratifying Resilience and Depression Longitudinally" (STRADL) Reference 104036/Z/14/Z). Ethics approval for the study was given by the NHS Tayside committee on research ethics (reference 05/S1401/89). We are grateful to all the families who took part, the general practitioners and the Scottish School of Primary Care for their help in recruiting them, and the whole Generation Scotland team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists, healthcare assistants and nurses.

*JHS (Jackson Heart Study):* The Jackson Heart Study is supported by contracts HSN268201300046C, HHSN268201300047C, HHSN268201300048C, HHSN268201300049C, HHSN268201300050C from the National Heart, Lung, and Blood Institute on Minority Health and Health Disparities. The authors acknowledge the Jackson Heart Study team institutions (University of Mississippi Medical Center, Jackson State University and Tougaloo College) and participants for their long-term commitment that continues to improve our understanding of the genetic epidemiology of cardiovascular and other chronic diseases among African Americans.

*MESA (Multi-Ethnic Study of Atherosclerosis):* This research was supported by the Multi-Ethnic Study of Atherosclerosis (MESA) contracts N01-HC-95159, N01-HC-95160, N01-HC-95161, N01-HC-95162, N01-HC-95163, N01-HC-95164, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168, N01-HC-95169, by grant HL071205 and by UL1-DR-001079 from NCRR . Funding for MESA SHARe genotyping was provided by NHLBI Contract N02-HL-6-4278. This publication was partially developed under a STAR research assistance agreement, No. RD831697 (MESA Air), awarded by the U.S Environmental Protection Agency. It has not been formally reviewed by the EPA. The views expressed in this document are solely those of the authors and the EPA does not endorse any products or commercial services mentioned in this publication. The provision of genotyping data was supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1TR000124, 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. The authors thank the participants of the MESA study, the Coordinating Center, MESA investigators, and study staff for their valuable contributions. A full list of participating MESA investigators and institutions can be found at http://www.mesa-nhlbi.org.

*NEO (The Netherlands Epidemiology of Obesity study):* The authors of the NEO study thank all individuals who participated in the Netherlands Epidemiology in Obesity study, all participating general practitioners for inviting eligible participants and all research nurses for collection of the data. We thank the NEO study group, Petra Noordijk, Pat van Beelen and Ingeborg de Jonge for the coordination, lab and data management of the NEO study. The genotyping in the NEO study was supported by the Centre National de Génotypage (Paris, France), headed by Jean-Francois Deleuze. The NEO study is supported by the participating Departments, the Division and the Board of Directors of the Leiden University Medical Center, and by the Leiden University, Research Profile Area Vascular and Regenerative Medicine. Dennis Mook-Kanamori is supported by Dutch Science Organization (ZonMW-VENI Grant 916.14.023).

*Pelotas Birth Cohort Study (The 1982 Pelotas Birth Cohort Study, Brazil):* The 1982 Pelotas Birth Cohort Study is conducted by the Postgraduate Program in Epidemiology at Universidade Federal de Pelotas with the collaboration of the Brazilian Public Health Association (ABRASCO). From 2004 to 2013, the Wellcome Trust supported the study. The International Development Research Center, World Health Organization, Overseas Development Administration, European Union, National Support Program for Centers of Excellence (PRONEX), the Brazilian National Research Council (CNPq), and the Brazilian Ministry of Health supported previous phases of the study.

Genotyping of 1982 Pelotas Birth Cohort Study participants was supported by the Department of Science and Technology (DECIT, Ministry of Health) and National Fund for Scientific and Technological Development (FNDCT, Ministry of Science and Technology), Funding of Studies and Projects (FINEP, Ministry of Science and Technology, Brazil), Coordination of Improvement of Higher Education Personnel (CAPES, Ministry of Education, Brazil).

*RS (Rotterdam Study):* The Rotterdam Study is funded by Erasmus Medical Center and Erasmus University, Rotterdam, Netherlands Organization for the Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam. The authors are grateful to the study participants, the staff from the Rotterdam Study and the participating general practitioners and pharmacists.

The generation and management of GWAS genotype data for the Rotterdam Study was executed by the Human Genotyping Facility of the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, Rotterdam, The Netherlands. The GWAS datasets are supported by the Netherlands Organisation of Scientific Research NWO Investments (nr. 175.010.2005.011, 911-03-012), the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, the Research Institute for Diseases in the Elderly (014-93-015; RIDE2), the Netherlands Genomics Initiative (NGI)/Netherlands Organisation for Scientific Research (NWO) Netherlands Consortium for Healthy Aging (NCHA), project nr. 050-060-810. We

thank Pascal Arp, Mila Jhamai, Marijn Verkerk, Lizbeth Herrera, Marjolein Peters and Carolina Medina-Gomez for their help in creating the GWAS database, and Karol Estrada, Yurii Aulchenko and Carolina Medina-Gomez for the creation and analysis of imputed data.

*SCHS-CHD (Singapore Chinese Health Study - Coronary Heart Disease):* The Singapore Chinese Health Study is supported by the National Institutes of Health, USA (RO1 CA144034 and UM1 CA182876), the nested case-control study of myocardial infarction by the Singapore National Medical Research Council (NMRC 1270/2010) and genotyping by the HUJ-CREATE Programme of the National Research Foundation, Singapore (Project Number 370062002).

*SCES (Singapore Chinese Eye Study), SiMES (Singapore Malay Eye Study), (SINDI) Singapore Indian Eye Study:* The Singapore Malay Eye Study (SiMES), the Singapore Indian Eye Study (SINDI), and the Singapore Chinese Eye Study (SCES) are supported by the National Medical Research Council (NMRC), Singapore (grants 0796/2003, 1176/2008, 1149/2008, STaR/0003/2008, 1249/2010, CG/SERI/2010, CIRG/1371/2013, and CIRG/1417/2015), and Biomedical Research Council (BMRC), Singapore (08/1/35/19/550 and 09/1/35/19/616). Ching-Yu Cheng is supported by an award from NMRC (CSA/033/2012). The Singapore Tissue Network and the Genome Institute of Singapore, Agency for Science, Technology and Research, Singapore provided services for tissue archival and genotyping, respectively.

*SP2 (Singapore Prospective Study Program):* SP2 is supported by the individual research grant and clinician scientist award schemes from the National Medical Research Council and the Biomedical Research Councils of Singapore.

*WGHS (Women's Genome Health Study):* The WGHS is supported by the National Heart, Lung, and Blood Institute (HL043851 and HL080467) and the National Cancer Institute (CA047988 and UM1CA182913), with collaborative scientific support and funding for genotyping provided by Amgen.

*WHI (Women's Health Initiative):* The WHI program is funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, U.S. Department of Health and Human Services through contracts HHSN268201100046C, HHSN268201100001C, HHSN268201100002C, HHSN268201100003C, HHSN268201100004C, and HHSN271201100004C. The authors thank the WHI investigators and staff for their dedication, and the study participants for making the program possible. A full listing of WHI investigators can be found at: http://www.whi.org/researchers/Documents%20%20Write%20a%20Paper/WHI%20Investigator%20Short%20List.pdf

*AA-DHS (African American Diabetes Heart Study):* The investigators acknowledge the cooperation of our Diabetes Heart Study (DHS) and AA-DHS participants. This work was supported by NIH R01 DK071891, R01 HL092301 and the General Clinical Research Center of Wake Forest School of Medicine M01-RR-07122.

*Airwave (The Airwave Health Monitoring Study):* We thank all participants in the Airwave Health Monitoring Study. The study is funded by the Home Office (Grant number 780-TETRA) with additional support from the National Institute for Health Research (NIHR), Imperial College Healthcare NHS Trust (ICHNT) and Imperial College Biomedical Research Centre (BRC). The study has ethical approval from the National Health Service Multi-site Research Ethics Committee (MREC/13/NW/0588). This work used computing resources provided by the MRC- funded UK MEDical Bioinformatics partnership programme (UK MED-BIO) (MR/L01632X/1). P.E. would like to acknowledge support from the Medical Research Council and Public Health England for the MRC-PHE Centre for Environment and Health (MR/L01341X/1) and from the NIHR NIHR Health Protection Research Unit in Health Impact of Environmental Hazards (HPRU-2012-10141).

*ASCOT (Anglo-Scandinavian Cardiac Outcomes Trial):* The ASCOT study was supported by Pfizer, New York, NY, USA for the ASCOT study and the collection of the ASCOT DNA repository; by Servier

Research Group, Paris, France; and by Leo Laboratories, Copenhagen, Denmark. We thank all ASCOT trial participants, physicians, nurses, and practices in the participating countries for their important contribution to the study. In particular we thank Clare Muckian and David Toomey for their help in DNA extraction, storage, and handling. Genotyping was funded by the CNG, MRC and the National Institutes of Health Research (NIHR). We would also like to acknowledge the Barts and The London Genome Centre staff for genotyping. This work forms part of the research programme of the NIHR Cardiovascular Biomedical Research Unit at Barts and The London, QMUL. H.R.W, M.J.C and P.B.M. wishes to acknowledge the NIHR Cardiovascular Biomedical Research Unit at Barts and The London, Queen Mary University of London, UK for support.

*BBJ (Biobank Japan Project):* BioBank Japan project is supported by the Japan Agency for Medical Research and Development and by the Ministry of Education, Culture, Sports, Sciences and Technology of the Japanese government.

*BES (Beijing Eye Study):* BES was supported by the National Key Laboratory Fund, Beijing, China.

*BRIGHT (British Genetics of Hypertension):* This work was supported by the Medical Research Council of Great Britain (grant number G9521010D) and the British Heart Foundation (grant number PG/02/128). The BRIGHT study is extremely grateful to all the patients who participated in the study and the BRIGHT nursing team. This work forms part of the research program of the National Institutes of Health Research (NIHR Cardiovascular Biomedical Research) Cardiovascular Biomedical Unit at Barts and The London, QMUL.

*CAGE-Amagasaki (Cardio-metabolic Genome Epidemiology Network, Amagasaki Study):* The CAGE Network studies were supported by grants for the Core Research for Evolutional Science and Technology (CREST) from the Japan Science Technology Agency; the Program for Promotion of Fundamental Studies in Health Sciences, National Institute of Biomedical Innovation Organization (NIBIO); and the Grant of National Center for Global Health and Medicine (NCGM).

*DESIR (Data from an Epidemiological Study on the Insulin Resistance):* The DESIR Study Group is composed of Inserm-U1018 (Paris: B. Balkau, P. Ducimetière, E. Eschwège), Inserm-U367 (Paris: F. Alhenc-Gelas), CHU d'Angers (A. Girault), Bichat Hospital (Paris: F. Fumeron, M. Marre, R. Roussel), CHU de Rennes (F. Bonnet), CNRS UMR-8199 (Lille: A. Bonnefond, P. Froguel), Medical Examination Services (Alençon, Angers, Blois, Caen, Chartres, Chateauroux, Cholet, LeMans, Orléans and Tours), Research Institute for General Medicine (J. Cogneau), the general practitioners of the region and the Cross- Regional Institute for Health (C. Born, E. Caces, M. Cailleau, N. Copin, J.G. Moreau, F. Rakotozafy, J. Tichet, S. Vol). The DESIR study was supported by Inserm contracts with CNAMTS, Lilly, Novartis Pharma and Sanofi-aventis, and by Inserm (Réseaux en Santé Publique, Interactions entre les déterminants de la santé, Cohortes Santé TGIR 2008), the Association Diabète Risque Vasculaire, the Fédération Française de Cardiologie, La Fondation de France, ALFEDIAM, ONIVINS, Société Francophone du Diabète, Ardix Medical, Bayer Diagnostics, Becton Dickinson, Cardionics, Merck Santé, Novo Nordisk, Pierre Fabre, Roche and Topcon.

*DFTJ (Dongfeng-Tongji Cohort Study):* This work was supported by grants from the National Basic Research Program grant (2011CB503800), the Programme of Introducing Talents of Discipline, the grants from the National Natural Science Foundation (grant NSFC-81473051, 81522040 and 81230069), and the Program for the New Century Excellent Talents in University (NCET-11-0169).

*DHS (Diabetes Heart Study):* The authors thank the investigators, staff, and participants of the DHS for their valuable contributions. This study was supported by the National Institutes of Health through HL67348 and HL092301.

*DR's EXTRA (Dose Responses to Exercise Training):* The study was supported by grants from Ministry of Education and Culture of Finland (722 and 627; 2004-2010); Academy of Finland (102318, 104943, 123885, 211119); European Commission FP6 Integrated Project (EXGENESIS), LSHM-CT-2004-005272; City of Kuopio; Juho Vainio Foundation; Finnish Diabetes Association; Finnish Foundation for

Cardiovascular Research; Kuopio University Hospital; Päivikki and Sakari Sohlberg Foundation; Social Insurance Institution of Finland 4/26/2010.

*EGCUT (Estonian Genome Center - University of Tartu (Estonian Biobank)):* This study was supported by EU H2020 grants 692145, 676550, 654248, Estonian Research Council Grant IUT20-60, NIASC, EIT – Health and NIH-BMI Grant No: 2R01DK075787-06A1 and EU through the European Regional Development Fund (Project No. 2014-2020.4.01.15-0012 GENTRANSMED.

*EPIC (European Prospective Investigation into Cancer and Nutrition):* The EPIC Norfolk Study is funded by Cancer Research, United Kingdom, British Heart Foundation, the Medical Research Council, the Ministry of Agriculture, Fisheries and Food, and the Europe against Cancer Programme of the Commission of the European Communities. We thank all EPIC participants and staff for their contribution to the study. We thank staff from the Technical, Field Epidemiology and Data Functional Group Teams of the Medical Research Council Epidemiology Unit in Cambridge, UK, for carrying out sample preparation, DNA provision and quality control, genotyping and data handling work. We specifically thank Sarah Dawson for coordinating the sample provision for biomarker measurements, Abigail Britten for coordinating DNA sample provision and genotyping of candidate markers, Nicola Kerrison, Chris Gillson and Abigail Britten for data provision and genotyping quality control, Matt Sims for writing the technical laboratory specification for the intermediate pathway biomarker measurements and for overseeing the laboratory work.

*FENLAND (The Fenland Study):* The Fenland Study is funded by the Wellcome Trust and the Medical Research Council (MC_U106179471). We are grateful to all the volunteers for their time and help, and to the General Practitioners and practice staff for assistance with recruitment. We thank the Fenland Study Investigators, Fenland Study Co-ordination team and the Epidemiology Field, Data and Laboratory teams. We further acknowledge support from the Medical research council (MC_UU_12015/1).

*FUSION (Finland-United States Investigation of NIDDM Genetics):* The FUSION study was supported by DK093757, DK072193, DK062370, and ZIA-HG000024. Genotyping was conducted at the Genetic Resources Core Facility (GRCF) at the Johns Hopkins Institute of Genetic Medicine.

*GeneSTAR (Genetic Studies of Atherosclerosis Risk):* GeneSTAR was supported by National Institutes of Health grants from the National Heart, Lung, and Blood Institute (HL49762, HL59684, HL58625, HL071025, U01 HL72518, HL087698, HL092165, HL099747, and K23HL105897), National Institute of Nursing Research (NR0224103), National Institute of Neurological Disorders and Stroke (NS062059), and by grants from the National Center for Research Resources to the Johns Hopkins General Clinical Research Center (M01-RR000052) and the Johns Hopkins Institute for Clinical & Translational Research (UL1 RR 025005).

*GLACIER (Gene x Lifestyle Interactions and Complex Traits Involved in Elevated Disease Risk):* We thank the participants, health professionals and data managers involved in the Västerbottens Intervention Project. We are also grateful to the staff of the Northern Sweden Biobank for preparing materials and to K Enqvist and T Johansson (Västerbottens County Council, Umeå, Sweden) for DNA preparation. The current study was supported by Novo Nordisk (PWF), the Swedish Research Council (PWF), the Swedish Heart Lung Foundation (PWF), the European Research Council (PWF), and the Skåne Health Authority (PWF).

*GRAPHIC (Genetic Regulation of Arterial Pressure of Humans in the Community):* The GRAPHIC Study was funded by the British Heart Foundation (BHF/RG/2000004). This work falls under the portfolio of research supported by the NIHR Leicester Cardiovascular Biomedical Research Unit. CPN and NJS are funded by the BHF and NJS is a NIHR Senior Investigator.

*HCHS/SOL (Hispanic Community Health Study/ Study of Latinos):* The baseline examination of HCHS/SOL was 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 National Institute on Minority Health and Health Disparities, National Institute on Deafness and Other Communication Disorders, National Institute of Dental and Craniofacial Research (NIDCR), National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institute of Neurological Disorders and Stroke, and NIH Office of Dietary Supplements additionally contributed funding to HCHS/SOL. The Genetic Analysis Center at the University of Washington was supported by NHLBI and NIDCR contracts (HHSN268201300005C AM03 and MOD03). Additional analysis support was provided by 1R01DK101855-01 and 13GRNT16490017. Genotyping was also supported by National Center for Advancing Translational Sciences UL1TR000124 and NIDDK DK063491 to the Southern California Diabetes Endocrinology Research Center. This research was also supported in part by the Intramural Research Program of the NIDDK, contract no. HHSB268201200054C, and Illumina.

*HRS (Health & Retirement Study):* HRS is supported by the National Institute on Aging (NIA U01AG009740 and R03 AG046389). Genotyping was funded separately by NIA (RC2 AG036495, RC4 AG039029). Our genotyping was conducted by the NIH Center for Inherited Disease Research (CIDR) at Johns Hopkins University. Genotyping quality control and final preparation of the data were performed by the Genetics Coordinating Center at the University of Washington.

*HyperGEN-AXIOM (Hypertension Genetic Epidemiology Network):* The study was support by the National Institutes of Health, the National Heart, Lung, Blood Institute grant HL086718.

*INGI-CARL (Italian Network Genetic Isolates):* This study was partially supported by Regione FVG (L.26.2008) and Italian Ministry of Health (GR-2011-02349604).

*INGI-FVG (Italian Network Genetic Isolates):* This study was partially supported by Regione FVG (L.26.2008) and Italian Ministry of Health (GR-2011-02349604).

*InterAct (The EPIC-InterAct Case-Cohort Study):* We thank all EPIC participants and staff for their contribution to the study. The InterAct study received funding from the European Union (Integrated Project LSHM-CT-2006-037197 in the Framework Programme 6 of the European Community).

*IRAS (Insulin Resistance Atherosclerosis Study):* The IRAS is supported by the National Heart Lung Institute (HL047887, HL047889, HL047890, and HL47902). Genotyping for this study was supported by the GUARDIAN Consortium with grant support from the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK; DK085175) and in part by UL1TR000124 (CTSI) and DK063491 (DRC). The authors thank study investigators, staff, and participants for their valuable contributions.

*IRAS Family Study (Insulin Resistance Atherosclerosis Study):* The IRASFS is supported by the National Heart Lung and Blood Institute (HL060944, HL061019, and HL060919). Genotyping for this study was supported by the GUARDIAN Consortium with grant support from the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK; DK085175) and in part by UL1TR000124 (CTSI) and DK063491 (DRC). The authors thank study investigators, staff, and participants for their valuable contributions.

*KORA (Cooperative Health Research in the Augsburg Region):* 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.

*LBC1936 (Lothian Birth Cohort 1936):* We thank the LBC1936 cohort participants and team members who contributed to these studies. Phenotype collection was supported by Age UK (The Disconnected Mind project). Genotyping was funded by the BBSRC (BB/F019394/1). The work was undertaken by The University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology, part of the cross

council Lifelong Health and Wellbeing Initiative (MR/K026992/1). Funding from the BBSRC and Medical Research Council (MRC) is gratefully acknowledged.

*LifeLines (Netherlands Biobank):* The Lifelines Cohort Study, and generation and management of GWAS genotype data for the Lifelines Cohort Study is supported by the Netherlands Organization of Scientific Research NWO (grant 175.010.2007.006), the Economic Structure Enhancing Fund (FES) of the Dutch government, the Ministry of Economic Affairs, the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the Northern Netherlands Collaboration of Provinces (SNN), the Province of Groningen, University Medical Center Groningen, the University of Groningen, Dutch Kidney Foundation and Dutch Diabetes Research Foundation.

The authors wish to acknowledge the services of the Lifelines Cohort Study, the contributing research centers delivering data to Lifelines, and all the study participants.

*LLFS (The Long Life Family Study):* The study is supported by the National Institute on Aging (NIA) grant U01AG023746.

*LOLIPOP (London Life Sciences Prospective Population Study):* The LOLIPOP study is supported by the National Institute for Health Research (NIHR) Comprehensive Biomedical Research Centre Imperial College Healthcare NHS Trust, the British Heart Foundation (SP/04/002), the Medical Research Council (G0601966, G0700931), the Wellcome Trust (084723/Z/08/Z), the NIHR (RP-PG-0407-10371), European Union FP7 (EpiMigrant, 279143), and Action on Hearing Loss (G51). We thank the participants and research staff who made the study possible.

*Loyola GxE (Kingston Gene-by-environment; subset of International Collaborative Study of Hypertension in Blacks (ICSHIB)):* The Loyola GxE project was supported by NIH Grant R01HL53353.

*Loyola SPT (Spanish Town; subset of International Collaborative Study of Hypertension in Blacks (ICSHIB)):* The Loyola SPT project was supported by NIH Grant R01HL53353.

*METSIM (Metabolic Syndrome In Men):* The METSIM study was supported by the Academy of Finland (contract 124243), the Finnish Heart Foundation, the Finnish Diabetes Foundation, Tekes (contract 1510/31/06), and the Commission of the European Community (HEALTH-F2-2007 201681), and the US National Institutes of Health grants DK093757, DK072193, DK062370, and ZIA- HG000024. Genotyping was conducted at the Genetic Resources Core Facility (GRCF) at the Johns Hopkins Institute of Genetic Medicine.

*NESDA (Netherlands Study of Depression and Anxiety):* The infrastructure for the NESDA study is funded through the Geestkracht programme of the Dutch Scientific Organization (ZON-MW, grant number 10-000-1002) and matching funds from participating universities and mental health care organizations. Genotyping in NESDA was funded by the Genetic Association Information Network (GAIN) of the Foundation for the US National Institutes of Health. Statistical analyses were carried out on the Genetic Cluster Computer (http://www.geneticcluster.org), which is financially supported by the Netherlands Scientific Organization (NWO 480-05-003) along with a supplement from the Dutch Brain Foundation.

*OBA (French obese cases):* The obese French adults were recruited by the laboratory "Integrated Genomics and Metabolic Diseases Modeling" (UMR 8199 CNRS / Université de Lille 2 / Institut Pasteur de Lille) of Pr. Philippe Froguel.

*PREVEND (The Prevention of REnal and Vascular ENd stage Disease study):* PREVEND genetics is supported by the Dutch Kidney Foundation (Grant E033), the EU project grant GENECURE (FP-6 LSHM CT 2006 037697), the National Institutes of Health (grant 2R01LM010098), The Netherlands organization for health research and development (NWO-Groot grant 175.010.2007.006, NWO VENI grant 916.761.70, ZonMw grant 90.700.441).

*PROCARDIS (Precocious Coronary Artery Disease):* PROCARDIS was supported by the European Community Sixth Framework Program (LSHM-CT- 2007-037273), AstraZeneca, the British Heart

Foundation, the Swedish Research Council, the Knut and Alice Wallenberg Foundation, the Swedish Heart-Lung Foundation, the Torsten and Ragnar Söderberg Foundation, the Strategic Cardiovascular Program of Karolinska Institutet and Stockholm County Council, the Foundation for Strategic Research and the Stockholm County Council (560283). M.F and H.W acknowledge the support of the Wellcome Trust core award (090532/Z/09/Z) and the BHF Centre of Research Excellence. A.G and H.W acknowledge European Union Seventh Framework Programme FP7/2007-2013 under grant agreement no. HEALTH-F2-2013-601456 (CVGenes@Target) & and A.G, the Wellcome Trust Institutional strategic support fund.

*RHS (Ragama Health Study):* The RHS was supported by the Grant of National Center for Global Health and Medicine (NCGM).

*SHEEP (Stockholm Heart Epidemiology Program):* This study was supported by grants from the Swedish Research Council for Health, Working Life and Welfare (http://www.forte.se/en/), the Stockholm County Council (http://www.sll.se/om-landstinget/Information-in-English1/), the Swedish Research Council (http://www.vr.se/inenglish.4.12fff4451215cbd83e4800015152.html), the Swedish Heart and Lung Foundation (https://www.hjart-lungfonden.se/HLF/Om-Hjart-lungfonden/About-HLF/), and the Cardiovascular Programme at Karolinska Institutet (http://ki.se/en/mmk/cardiovascular-research-networks).

*SHIP (Study of Health in Pomerania):* SHIP is part of the Community Medicine Research net of the University of Greifswald, Germany, which is funded by the Federal Ministry of Education and Research (grants no. 01ZZ9603, 01ZZ0103, and 01ZZ0403), the Ministry of Cultural Affairs as well as the Social Ministry of the Federal State of Mecklenburg-West Pomerania, and the network 'Greifswald Approach to Individualized Medicine (GANI_MED)' funded by the Federal Ministry of Education and Research (grant 03IS2061A). Genome-wide data were supported by the Federal Ministry of Education and Research (grant no. 03ZIK012) and a joint grant from Siemens Healthcare, Erlangen, Germany and the

Federal State of Mecklenburg- West Pomerania. The University of Greifswald is a member of the 'Center of Knowledge Interchange' program of the Siemens AG.

*SWHS/SMHS (Shanghai Women's Health Study/ Shanghai Men's Health Study):* We thank all the individuals who took part in these studies and all the researchers who have enabled this work to be carried out. The Shanghai Women's Health Study and the Shanghai Men's Health Study are supported by research grants UM1CA182910 and UM1CA173640 from the U.S. National Cancer Institute, respectively.

*TWINGENE (TwinGene of the Swedish Twin Registry):* The Swedish Twin Registry is financially supported by Karolinska Institutet. TwinGene project received funding from the Swedish Research Council (M-2005-1112), GenomEUtwin (EU/QLRT-2001-01254; QLG2-CT-2002-01254), NIH DK U01-066134, The Swedish Foundation for Strategic Research (SSF) and the Heart and Lung foundation no. 20070481

*YFS (The Cardiovascular Risk in Young Finns Study):* The Young Finns Study has been financially supported by the Academy of Finland: grants 286284, 134309 (Eye), 126925, 121584, 124282, 129378 (Salve), 117787 (Gendi), and 41071 (Skidi); the Social Insurance Institution of Finland; Kuopio, Tampere and Turku University Hospital Medical Funds (grant X51001); Juho Vainio Foundation; Paavo Nurmi Foundation; Finnish Foundation for Cardiovascular Research ; Finnish Cultural Foundation; Tampere Tuberculosis Foundation; Emil Aaltonen Foundation; Yrjö Jahnsson Foundation; Signe and Ane Gyllenberg Foundation; and Diabetes Research Foundation of Finnish Diabetes Association. The expert technical assistance in the statistical analyses by Leo-Pekka Lyytikäinen and Irina Lisinen is gratefully acknowledged.

Study group membership: The following individuals are members of the LifeLines Cohort Study: Behrooz Z Alizadeh, H Marike Boezen, Lude Frank, Pim van der Harst, Gerjan Navis, Marianne Rots, Harold Snieder, Morris Swertz, Bruce HR Wolffenbuttel, and Cisca Wijmenga. Department of

Epidemiology, University of Groningen, University Medical Center Groningen, the Netherlands (Behrooz Z Alizadeh, H Marike Boezen, and Harold Snieder); Department of Genetics, University of Groningen, University Medical Center Groningen, the Netherlands (Lude Franke, Morris Swertz, and Cisca Wijmenga); Department of Cardiology, University of Groningen, University Medical Center Groningen, the Netherlands (Pim van der Harst); Department of Internal Medicine, Division of Nephrology, University of Groningen, University Medical Center Groningen, the Netherlands (Gerjan Navis); Department of Medical Biology, University of Groningen, University Medical Center Groningen, the Netherlands (Marianne Rots); and Department of Endocrinology, University of Groningen, University Medical Center Groningen, the Netherlands (Bruce HR Wolffenbuttel).

Conflict of interest statement: The authors declare no competing financial interests except for the following. Bruce M Psaty serves on the DSMB of a clinical trial funded by the manufacturer (Zoll LifeCor) and on the Steering Committee of the Yale Open Data Access Project funded by Johnson & Johnson. Oscar H Franco received grants from Metagenics (on women's health and epigenetics) and from Nestle (on child health). Mike A. Nalls' participation is suppported by a consulting contract between Data Tecnica International and the National Institute on Aging, National Institutes of Health, Bethesda, MD, USA, and Dr. Nalls also consults for Illumina Inc, the Michael J. Fox Foundation and University of California Healthcare among others. Neil Poulter has received financial support and consultancy fees from several pharmaceutical companies that manufacture either blood pressure-lowering or lipid lowering agents or both. Peter Sever has received research awards from Pfizer Inc. Jost Bruno Jonas serves as a consultant for Mundipharma Co. (Cambridge, UK), patent holder with Biocompatibles UK Ltd. (Franham, Surrey, UK) (Title: Treatment of eye diseases using encapsulated cells encoding and secreting neuroprotective factor and / or anti-angiogenic factor; Patent number: 20120263794), and Patent applicant with University of Heidelberg (Heidelberg, Germany) (Title: Agents for use in the

therapeutic or prophylactic treatment of myopia or hyperopia; Europäische Patentanmeldung 15 000 771.4).

REFERENCES


1. Roth GA, Johnson C, Abajobir A, et al. Global, Regional, and National Burden of Cardiovascular Diseases for 10 Causes, 1990 to 2015. *J Am Coll Cardiol* 2017.
2. Asselbergs FW, Guo Y, van Iperen EP, et al. Large-scale gene-centric meta-analysis across 32 studies identifies multiple lipid loci. *Am J Hum Genet* 2012;91(5):823-38.
3. Global Lipids Genetics Consortium, Willer CJ, Schmidt EM, et al. Discovery and refinement of loci associated with lipid levels. *Nat Genet* 2013;45(11):1274-83.
4. Peloso GM, Auer PL, Bis JC, et al. Association of low-frequency and rare coding-sequence variants with blood lipids and coronary heart disease in 56,000 whites and blacks. *Am J Hum Genet* 2014;94(2):223-32.
5. Spracklen CN, Chen P, Kim YJ, et al. Association analyses of East Asian individuals and trans-ancestry analyses with European individuals reveal new loci associated with cholesterol and triglyceride levels. *Hum Mol Genet* 2017;26(9):1770-84.
6. Surakka I, Horikoshi M, Magi R, et al. The impact of low-frequency and rare variants on lipid levels. *Nat Genet* 2015;47(6):589-97.
7. Teslovich TM, Musunuru K, Smith AV, et al. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* 2010;466(7307):707-13.
8. van Leeuwen EM, Sabo A, Bis JC, et al. Meta-analysis of 49 549 individuals imputed with the 1000 Genomes Project reveals an exonic damaging variant in ANGPTL4 determining fasting TG levels. *J Med Genet* 2016;53(7):441-9.
9. Muth ND, Laughlin GA, von Muhlen D, et al. High-density lipoprotein subclasses are a potential intermediary between alcohol intake and reduced risk of cardiovascular disease: the Rancho Bernardo Study. *Br J Nutr* 2010;104(7):1034-42.
10. Gardner CD, Tribble DL, Young DR, et al. Associations of HDL, HDL(2), and HDL(3) cholesterol and apolipoproteins A-I and B with lifestyle factors in healthy women and men: the Stanford Five City Project. *Prev Med* 2000;31(4):346-56.
11. Wakabayashi I. Relationship between alcohol intake and lipid accumulation product in middle-aged men. *Alcohol Alcohol* 2013;48(5):535-42.
12. Klop B, do Rego AT, Cabezas MC. Alcohol and plasma triglycerides. *Curr Opin Lipidol* 2013;24(4):321-6.
13. Whitfield JB, Heath AC, Madden PA, et al. Metabolic and biochemical effects of low-to-moderate alcohol consumption. *Alcohol Clin Exp Res* 2013;37(4):575-86.
14. Rifler JP, Lorcerie F, Durand P, et al. A moderate red wine intake improves blood lipid parameters and erythrocytes membrane fluidity in post myocardial infarct patients. *Mol Nutr Food Res* 2012;56(2):345-51.
15. Brinton EA. Effects of ethanol intake on lipoproteins. *Curr Atheroscler Rep* 2012;14(2):108-14.
16. Kechagias S, Zanjani S, Gjellan S, et al. Effects of moderate red wine consumption on liver fat and blood lipids: a prospective randomized study. *Ann Med* 2011;43(7):545-54.
17. Wakabayashi I. Associations of alcohol drinking and cigarette smoking with serum lipid levels in healthy middle-aged men. *Alcohol Alcohol* 2008;43(3):274-80.
18. Rakic V, Puddey IB, Dimmitt SB, et al. A controlled trial of the effects of pattern of alcohol intake on serum lipid levels in regular drinkers. *Atherosclerosis* 1998;137(2):243-52.

19. Brien SE, Ronksley PE, Turner BJ, et al. Effect of alcohol consumption on biological markers associated with risk of coronary heart disease: systematic review and meta-analysis of interventional studies. *BMJ* 2011;342:d636.
20. Rimm EB, Williams P, Fosher K, et al. Moderate alcohol intake and lower risk of coronary heart disease: meta-analysis of effects on lipids and haemostatic factors. *BMJ* 1999;319(7224):1523-8.
21. Tabara Y, Ueshima H, Takashima N, et al. Mendelian randomization analysis in three Japanese populations supports a causal role of alcohol consumption in lowering low-density lipid cholesterol levels and particle numbers. *Atherosclerosis* 2016;254:242-8.
22. Vu KN, Ballantyne CM, Hoogeveen RC, et al. Causal Role of Alcohol Consumption in an Improved Lipid Profile: The Atherosclerosis Risk in Communities (ARIC) Study. *PLoS One* 2016;11(2):e0148765.
23. Kirk EP. Genes, Environment, and the Heart: Putting the Pieces Together. *Circ Cardiovasc Genet* 2017;10(3).
24. Psaty BM, O'Donnell CJ, Gudnason V, et al. Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium: Design of prospective meta-analyses of genome-wide association studies from 5 cohorts. *Circ Cardiovasc Genet* 2009;2(1):73-80.
25. Rao DC, Sung YJ, Winkler TW, et al. Multiancestry Study of Gene-Lifestyle Interactions for Cardiovascular Traits in 610 475 Individuals From 124 Cohorts: Design and Rationale. *Circ Cardiovasc Genet* 2017;10(3).
26. Skol AD, Scott LJ, Abecasis GR, et al. Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. *Nat Genet* 2006;38(2):209-13.
27. Manning AK, LaValley M, Liu CT, et al. Meta-analysis of gene-environment interaction: joint estimation of SNP and SNP x environment regression coefficients. *Genet Epidemiol* 2011;35(1):11-8.
28. Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. *Clin Chem* 1972;18(6):499-502.
29. 1000 Genomes Project Consortium, Auton A, Brooks LD, et al. A global reference for human genetic variation. *Nature* 2015;526(7571):68-74.
30. Winkler TW, Day FR, Croteau-Chonka DC, et al. Quality control and conduct of genome-wide association meta-analyses. *Nat Protoc* 2014;9(5):1192-212.
31. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* 2010;26(17):2190-1.
32. Devlin B, Roeder K. Genomic control for association studies. *Biometrics* 1999;55(4):997-1004.
33. Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res* 2012;40(Database issue):D930-4.
34. Boyle AP, Hong EL, Hariharan M, et al. Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res* 2012;22(9):1790-7.
35. GTEx Consortium. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* 2015;348(6235):648-60.
36. Nikpay M, Goel A, Won HH, et al. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet* 2015;47(10):1121-30.
37. Liu DJ, Peloso GM, Yu H, et al. Exome-wide association study of plasma lipids in >300,000 individuals. *Nat Genet* 2017;49(12):1758-66.

38. Justice AE, Winkler TW, Feitosa MF, et al. Genome-wide meta-analysis of 241,258 adults accounting for smoking behaviour identifies novel loci for obesity traits. *Nat Commun* 2017;8:14977.
39. Winkler TW, Justice AE, Graff M, et al. The Influence of Age and Sex on Genetic Associations with Adult Body Size and Shape: A Large-Scale Genome-Wide Interaction Study. *PLoS Genet* 2015;11(10):e1005378.
40. Cohen J, Pertsemlidis A, Kotowski IK, et al. Low LDL cholesterol in individuals of African descent resulting from frequent nonsense mutations in PCSK9. *Nat Genet* 2005;37(2):161-5.
41. Lepor NE, Kereiakes DJ. The PCSK9 Inhibitors: A Novel Therapeutic Target Enters Clinical Practice. *Am Health Drug Benefits* 2015;8(9):483-9.
42. Stein EA, Mellis S, Yancopoulos GD, et al. Effect of a monoclonal antibody to PCSK9 on LDL cholesterol. *N Engl J Med* 2012;366(12):1108-18.
43. Iatan I, Dastani Z, Do R, et al. Genetic variation at the proprotein convertase subtilisin/kexin type 5 gene modulates high-density lipoprotein cholesterol levels. *Circ Cardiovasc Genet* 2009;2(5):467-75.
44. Jin W, Wang X, Millar JS, et al. Hepatic proprotein convertases modulate HDL metabolism. *Cell Metab* 2007;6(2):129-36.
45. Shimamura M, Matsuda M, Yasumo H, et al. Angiopoietin-like protein3 regulates plasma HDL cholesterol through suppression of endothelial lipase. *Arterioscler Thromb Vasc Biol* 2007;27(2):366-72.
46. Jin W, Millar JS, Broedl U, et al. Inhibition of endothelial lipase causes increased HDL cholesterol levels in vivo. *J Clin Invest* 2003;111(3):357-62.
47. Tsutsumi K. Lipoprotein lipase and atherosclerosis. *Curr Vasc Pharmacol* 2003;1(1):11-7.
48. Chester A, Scott J, Anant S, et al. RNA editing: cytidine to uridine conversion in apolipoprotein B mRNA. *Biochim Biophys Acta* 2000;1494(1-2):1-13.
49. Kendrick JS, Chan L, Higgins JA. Superior role of apolipoprotein B48 over apolipoprotein B100 in chylomicron assembly and fat absorption: an investigation of apobec-1 knock-out and wild-type mice. *Biochem J* 2001;356(Pt 3):821-7.
50. Khoo B, Krainer AR. Splicing therapeutics in SMN2 and APOB. *Curr Opin Mol Ther* 2009;11(2):108-15.
51. Hagberg CE, Falkevall A, Wang X, et al. Vascular endothelial growth factor B controls endothelial fatty acid uptake. *Nature* 2010;464(7290):917-21.
52. Hagberg CE, Mehlem A, Falkevall A, et al. Targeting VEGF-B as a novel treatment for insulin resistance and type 2 diabetes. *Nature* 2012;490(7420):426-30.
53. Dijkstra MH, Pirinen E, Huusko J, et al. Lack of cardiac and high-fat diet induced metabolic phenotypes in two independent strains of Vegf-b knockout mice. *Sci Rep* 2014;4:6238.
54. Robciuc MR, Kivela R, Williams IM, et al. VEGFB/VEGFR1-Induced Expansion of Adipose Vasculature Counteracts Obesity and Related Metabolic Complications. *Cell Metab* 2016;23(4):712-24.
55. Rafii S, Carmeliet P. VEGF-B Improves Metabolic Health through Vascular Pruning of Fat. *Cell Metab* 2016;23(4):571-3.
56. Choi H, Allahdadi KJ, Tostes RC, et al. Diacylglycerol Kinase Inhibition and Vascular Function. *Curr Enzym Inhib* 2009;5(3):148-52.
57. Cai K, Lucki NC, Sewer MB. Silencing diacylglycerol kinase-theta expression reduces steroid hormone biosynthesis and cholesterol metabolism in human adrenocortical cells. *Biochim Biophys Acta* 2014;1841(4):552-62.

58. de Vries PS, Sabater-Lleal M, Chasman DI, et al. Comparison of HapMap and 1000 Genomes Reference Panels in a Large-Scale Genome-Wide Association Study. *PLoS One* 2017;12(1):e0167742.
59. Livingston M, Callinan S. Underreporting in alcohol surveys: whose drinking is underestimated? *J Stud Alcohol Drugs* 2015;76(1):158-64.
60. Altman DG, Royston P. The cost of dichotomising continuous variables. *BMJ* 2006;332(7549):1080.

TABLES

**Table 1.** Novel Loci Discovered in the Combined Analysis of Stages 1 and 2 Using the 2 DF Model that Jointly Tests Main and Interaction Effects.

| rsID | Chr:Position | Alleles | Freq | Closest Gene(s) | Main Effect[a] | Interaction Effect[a] | Joint *P*-value[a] | Joint FDR *q*-value | Interaction *P*-value[b] | Heterogeneity *P*-value | Most Significant 2 DF Model |
|---|---|---|---|---|---|---|---|---|---|---|---|
| rs190528931[c] | 11:63911273 | A/C | 0.04 | *MACROD1* | 0.0109 | -0.0023 | $1.9\times10^{-16}$ | $3.6\times10^{-11}$ | 0.32 | 0.96 | META - HDL-C – CURDRINK |
| rs7904973[c] | 10:124693587 | T/G | 0.55 | *C10orf88* | 0.92 | -0.15 | $1.9\times10^{-15}$ | $3.5\times10^{-10}$ | 0.38 | 0.89 | META - LDL-C - CURDRINK |
| rs73729083 | 7:137559799 | C/T | 0.91 | *CREB3L2* | 4.01 | 0.65 | $8.2\times10^{-15}$ | $1.4\times10^{-9}$ | 0.57 | 0.22 | META - LDL-C - CURDRINK |
| rs80080062 | 3:185812169 | G/C | 0.87 | *ETV5* | 0.0061 | 0.0031 | $1.1\times10^{-12}$ | $1.7\times10^{-7}$ | 0.38 | 0.85 | META - HDL-C - REGDRINK |
| rs7140110 | 13:114544024 | C/T | 0.73 | *GAS6-AS1* | -0.01 | -0.004 | $3.4\times10^{-12}$ | $5.1\times10^{-7}$ | 0.19 | 0.42 | META - TG - CURDRINK |
| rs34311866 | 4:951947 | C/T | 0.83 | *TMEM175* | -0.02 | 0.004 | $1.5\times10^{-11}$ | $2.1\times10^{-6}$ | 0.42 | 0.90 | EUR - TG – CURDRINK |
| rs2911971 | 8:6607634 | G/C | 0.34 | *AGPAT5* | -0.75 | 0.01 | $7.5\times10^{-11}$ | $1.1\times10^{-5}$ | 0.53 | 0.49 | META - LDL-C - CURDRINK |
| rs56076449 | 5:132442190 | G/T | 0.79 | *HSPA4 / FSTL4* | 0.013 | -0.002 | $9.3\times10^{-11}$ | $1.3\times10^{-5}$ | 0.80 | 0.80 | META - TG - REGDRINK |
| rs41274050[c] | 10:52573772 | T/C | 0.01 | *A1CF* | 0.108 | -0.031 | $9.6\times10^{-10}$ | $1.3\times10^{-4}$ | 0.62 | 1 | EUR - TG – REGDRINK |
| rs7035578 | 9:78745177 | A/G | 0.15 | *PCSK5* | -1.27 | 0.08 | $1.2\times10^{-9}$ | $1.6\times10^{-4}$ | 0.70 | 0.82 | EUR - LDL-C - CURDRINK |
| rs201445483 | 2:17890087 | I/D | 0.83 | *SMC6* | 1.43 | 0.68 | $4.7\times10^{-9}$ | $6.0\times10^{-4}$ | 0.17 | 0.46 | META - LDL-C - CURDRINK |
| rs72729610 | 4:154190965 | G/A | 0.86 | *TRIM2* | 0.0075 | -0.0036 | $5.6\times10^{-9}$ | $7.2\times10^{-4}$ | 0.077 | 0.26 | META - HDL-C - REGDRINK |
| rs143528679 | 4:124558378 | G/A | 0.1 | *SPRY1 / LINC01091* | -1.2 | -5.63 | $6.3\times10^{-9}$ | $8.0\times10^{-4}$ | $6.4\times10^{-4}$ | 0.096 | AFR - LDL-C - CURDRINK |
| rs2111622[c] | 2:53984823 | G/A | 0.77 | *ASB3 / GPR75-ASB3* | 0.0008 | -0.0072 | $7.9\times10^{-9}$ | $9.9\times10^{-4}$ | 0.013 | 0.12 | EUR - HDL-C - CURDRINK |
| rs13284665 | 9:131513370 | G/A | 0.88 | *ZER1* | 1.99 | -0.89 | $1.1\times10^{-8}$ | $1.3\times10^{-3}$ | 0.35 | 0.89 | EUR - LDL-C - CURDRINK |
| rs4898521 | 12:49755162 | A/T | 0.95 | *DNAJC22 / SPATS2* | 0.0179 | -0.0107 | $1.3\times10^{-8}$ | $1.7\times10^{-3}$ | 0.060 | 1 | EUR - HDL-C - REGDRINK |
| rs6063050 | 20:45604240 | C/T | 0.75 | *EYA2* | 0.011 | 0 | $2.9\times10^{-8}$ | $3.6\times10^{-3}$ | 0.30 | 0.39 | META - TG - CURDRINK |
| rs2963472 | 5:157999022 | A/G | 0.21 | *LOC101927697 / EBF1* | 0.014 | -0.002 | $3.5\times10^{-8}$ | $4.2\times10^{-3}$ | 0.96 | 0.23 | EUR - TG – REGDRINK |

[a]These estimates pertain to the 2 DF joint test of main and interaction effects. [b]These *P*-values pertain to 1 DF tests of interaction effects. [c]These loci were also discovered by a concurrent association study focused on exonic variants (37). The Alleles column

reports the coded/non-coded alleles. The Freq column reports the frequency of the coded allele. The Heterogeneity *P*-value column indicates the significance of the Stage 1 heterogeneity across ancestry groups in the most significant 2 DF Model. In the Most Significant 2 DF Model column, META refers to the trans-ancestry meta-analysis, EUR refers to the European-ancestry meta-analysis, and AFR refers to the African-ancestry meta-analysis. CURDRINK means that alcohol consumption was categorized into drinkers and non-drinkers, while REGDRINK means that it was categorized into regular drinkers and those that were not regular drinkers.

FIGURE LEGENDS

**Figure 1.** Flow Chart of the Overall Study Design.

For each lipid trait, association analyses were performed, accounting for the two alcohol consumption variables: “current drinker” status and “regular drinker”. For each ancestry group, study-specific results were combined to perform the 1 degree of freedom (DF) test for an interaction effect and the 2 DF joint test of genetic main effect and interaction with drinking exposure. Individuals from five ancestry groups were included: European (EUR), African (AFR), Asian (ASN), Hispanic (HIS), and Brazilian (BRA).

**Figure 2.** Venn Diagram Showing the Distribution of Genome-wide Significant Associations at the 147 Identified Loci Among Lipid Traits.

**Figure 3.** Heat Map of the Significance and Effect Direction of Index Variants at the 18 Novel Loci for the Three Lipid Traits.

For each combination of index variant and lipid trait, the effect direction and *P*-value of the most significant association is shown. For example: the 11:63911273 variant was most significantly associated with HDL-C in the trans-ancestry meta-analysis, using the current drinker alcohol consumption variable. Shades of purple and yellow represent negative and positive directions of effect, respectively, while associations of either direction with *P*-value > 0.05 are white.