Multiethnic Exome-Wide Association Study of Subclinical Atherosclerosis

Pradeep Natarajan MD MMSc1,2,*, Joshua C. Bis PhD3, Lawrence F. Bielak DDS MPH4,
Amanda J. Cox PhD5, Marcus Dörr MD6,7, Mary F. Feitosa PhD5, Nora Franceschini MD MPH9,
Xiqun Guo PhD10, Shih-Jen Hwang PhD11, Aaron Isaacs PhD12,13, Min A Jhun PhD4, Maryam
Kavoussi MD PhD12, Ruifang Li-Gao MSc14, Leo-Pekka Lyytikäinen MD15,16, Riccardo E.
Marioni PhD17,18,19, Ulf Schminke MD20, Nathan O. Stitzel MD PhD8,21,22, Hayato Tada MD23,
Jessica van Setten PhD34, Albert V. Smith PhD25,26, Dina Vojnovic MD MSc12, Lisa R. Yanek
MPH27, Jie Yao MD MS10, Laura M. Yerges-Armstrong PhD28, Najaf Amin PhD12, Usman
Baber MD29, Ingrid B. Borecki PhD30, J. Jeffrey Carr MD MSc31, Yi-Der Ida Chen PhD10, L.
Adrienne Cupples PhD32, Pim A. de Jong MD PhD33, Harry de Koning PhD34, Bob D. de Vos
MSc35, Ayse Demirkan PhD12, Valentin Fuster MD PhD29, Oscar H. Franco MD PhD12, Mark O.
Goodarzi MD PhD36, Tamara B. Harris MD37, Susan R. Heckbert MD PhD38, Gerardo Heiss MD
PhD9, Udo Hoffmann MD PhD39, Albert Hofman MD PhD1,2, Ivana Išgum PhD35, J. Wouter
Jukema PhD MD40,41,42, Mika Kähönen MD PhD43,44, Sharon L.R. Kardia PhD4, Brian G. Kral
MD MPH27, Lenore J. Launer PhD37, Joe Massaro PhD32, Roxana Mehran MD29, Braxton D.
Mitchell PhD MPH28,45, Thomas H. Mosley Jr. PhD46, Renée de Mutsert PhD14, Anne B.
Newman MD47, Khanh-dung Nguyen PhD48, Kari E. North PhD3, Jeffrey R. O’Connell PhD28,
Matthijs Oudkerk PhD49, James S. Pankow PhD MPH50, Gina M. Peloso PhD1,2,52, Wendy Post
MD MS27, Michael A. Province PhD5, Laura M. Raffield PhD5, Olli T. Raitakari MD PhD51,52,
Dermot F. Reilly PhD48, Fernando Rivadeneira MD PhD12,53,54, Frits Rosendaal MD PhD14,
Samantha Sartori PhD29, Kent D. Taylor PhD10, Alexander Teumer PhD7,53, Stella Trompet
PhD40, Stephen T. Turner MD56, Andre G. Uitterlinden PhD12,53,54, Dhananjay Vaidya PhD MPH
MBBS27, Aad van der Lugt MD PhD57, Uwe Völker PhD7,58, Joanna M. Wardlaw MD17,59,
Christina L. Wassel PhD MS60, Stefan Weiss PhD7,58, Mary K. Wojcynski PhD8, Diane M.
Becker ScD PhD27, Lewis C. Becker MD27, Eric Boerwinkle PhD61,62, Donald W. Bowden PhD5,
Ian J. Deary PhD17,63, Abbas Dehghan MD PhD12, Stephan B. Felix MD6,7, Vilmundur Gudnason
MD PhD25,26, Terho Lehtimäki MD PhD15,16, Rasika Mathias ScD7, Dennis O. Mook-Kanamori
MD PhD14,64, Bruce M. Psaty MD3,38,65,66, Daniel J. Rader MD67, Jerome I. Rotter MD12, James
G. Wilson MD68, Cornelia M. van Duijn PhD12,54, Henry Völzke MD7,55,69, Sekar Kathiresan
MD1,2, Patricia A. Peyser PhD4, Christopher J. O’Donnell MD MPH70,†, CHARGE Consortium
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
† Contributed equally
* Corresponding author

1 Center for Human Genetic Research and Cardiovascular Research Center, Massachusetts General Hospital,
Boston, MA, 02114
2 Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA 02142
3 Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, WA 98101
4 Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, MI 48109
5 Diabetes Heart Study, Wake Forest Health Sciences, Winston-Salem, NC 27104
6 Department of Internal Medicine, University Medicine Greifswald, Greifswald, Germany 17475
7 DZHK (German Centre for Cardiovascular Research), partner Site Greifswald, Germany 17475
8 Division of Statistical Genetics, Department of Genetics, Washington University School of Medicine, St. Louis,
MO 63108
9 Epidemiology, Gilling School of Global Public Health, University of North Carolina, Chapel Hill, NC 27514
10 Institute for Translational Genomics and Population Sciences, Los Angeles Biomedical Research Institute, and
Department of Pediatrics at Harbor-UCLA Medical Center, Torrance, CA 90502

1
Framingham Heart Study, National Heart Lung and Blood Institute / National Institutes of Health, Framingham, MA 01701

Genetic Epidemiology Unit, Department of Epidemiology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, NL

CARIM School for Cardiovascular Diseases, Maastricht Centre for Systems Biology, Department of Biochemistry, Maastricht University, Maastricht, NL

Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, NL

Department of Clinical Chemistry, Fimlab Laboratories, Tampere, Finland 33520

Department of Clinical Chemistry, University of Tampere School of Medicine, Finland, 33014

Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, UK, EH8 9JZ

Centre for Genomic and Experimental Medicine, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK, EH4 2XU

Queensland Brain Institute, University of Queensland, Brisbane, QLD, Australia 4072

Department of Neurology, University Medicine Greifswald, Greifswald, Germany 17475

Cardiovascular Division, Department of Medicine, Washington University School of Medicine, St. Louis, MO 63110

McDonnell Genome Institute, Washington University School of Medicine, St. Louis, MO 63110

Division of Cardiovascular Medicine, Kanazawa University Graduate School of Medicine, Kanazawa, Japan 9208641

Cardiology, University Medical Center Utrecht, Utrecht, NL

Icelandic Heart Association, Kopavogur, Iceland

Faculty of Medicine, University of Iceland, Reykjavik, Iceland

Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, MD 21287

Division of Endocrinology, Diabetes & Nutrition, Department of Medicine, University of Maryland School of Medicine, Baltimore, MD 21201

Cardiovascular Institute, Icahn School of Medicine at Mount Sinai, New York, NY 10029

Analytical Genetics, Regeneron Pharmaceuticals Inc, Tarrytown, NY 10591

Radiology, Cardiovascular Medicine and Biomedical Informatics, Vanderbilt, Nashville, TN 37203

Department of Biostatistics, Boston University School of Public Health, Boston, MA 02118

Department of Radiology, University Medical Center Utrecht, Utrecht, NL

Department of Public Health, Erasmus Medical Center Rotterdam, Rotterdam, NL

Image Sciences Institute, University Medical Center Utrecht, Utrecht, NL

Division of Endocrinology, Diabetes and Metabolism, Cedars-Sinai Medical Center, Los Angeles, CA 90048

Laboratory of Epidemiology and Population Sciences, National Institute on Aging Intramural Research Institutes of Health, Baltimore, MD 20892

Department of Epidemiology, University of Washington, Seattle, WA 98101

Department of Radiology, Massachusetts General Hospital, Boston, MA 02114

Department of Cardiology, Leiden University Medical Center, Leiden, NL

Durrer Center for Cardiogenetic Research, Netherlands Heart Institute, Amsterdam, NL

Interuniversity Cardiologice Institute of the Netherlands, Netherlands Heart Institute, Utrecht, NL

Department of Clinical Physiology, Tampere University Hospital, Tampere, Finland 33521

Department of Clinical Physiology, University of Tampere School of Medicine, Tampere, Finland, 33014

Geriatrics Research and Education Clinical Center, Baltimore Veterans Administration Medical Center, Baltimore, MD 21201

Division of Geriatric Medicine, Department of Medicine, University of Mississippi Medical Center, Jackson, MS 39216

Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA 15261

Genetics & Pharmacogenomics Department, Merck Sharpe & Dohme Corp., Boston, MA 02115

Center for Medical Imaging - North East Netherlands, University Medical Center Groningen, University of Groningen, Groningen, NL

Division of Epidemiology and Community Health, University of Minnesota, Minneapolis, MN 5545

Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku, Finland 20521

Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, Finalnd 20520

Department of Internal Medicine, Erasmus MC, University Medical Center Rotterdam, Rotterdam, NL

Netherlands Consortium for Healthy Aging, Rotterdam, NL

Institute for Community Medicine, University Medicine Greifswald, Greifswald, Germany 17475

Division of Nephrology and Hypertension, Department of Internal Medicine, Mayo Clinic, Rochester, MN 55905
References: 44

Running title: Exome association with subclinical atherosclerosis

Please address correspondence to:
Christopher J. O’Donnell, MD MPH
Associate Professor Medicine, Harvard Medical School
Chief of Cardiology, Boston Veterans Affairs Healthcare System
1400 Veterans of Foreign Wars Parkway, Building 1 5B-113
Boston, MA 02132 USA
Phone: 857-203-6840, Fax: 857-203-5550
Email: odonnellc@nhlbi.nih.gov
Abstract

Background – The burden of subclinical atherosclerosis in asymptomatic individuals is heritable and associated with elevated risk of developing clinical coronary heart disease (CHD). We sought to identify genetic variants in protein-coding regions associated with subclinical atherosclerosis and the risk of subsequent CHD.

Methods and Results – We studied a total of 25,109 European ancestry and African-American participants with coronary artery calcification (CAC) measured by cardiac computed tomography and 52,869 with common carotid intima media thickness (CIMT) measured by ultrasonography within the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium. Participants were genotyped for 247,870 DNA sequence variants (231,539 in exons) across the genome. A meta-analysis of exome-wide association studies was performed across cohorts for CAC and CIMT. APOB p.Arg3527Gln was associated with four-fold excess CAC ($P = 3 \times 10^{-10}$). The APOE ε2 allele (p.Arg176Cys) was associated with both 22.3% reduced CAC ($P = 1 \times 10^{-12}$) and 1.4% reduced CIMT ($P = 4 \times 10^{-14}$) in carriers compared with non-carriers. In secondary analyses conditioning on LDL cholesterol concentration, the ε2 protective association with CAC, although attenuated, remained strongly significant. Additionally, the presence of ε2 was associated with reduced risk for CHD (OR 0.77; $P = 1 \times 10^{-11}$).

Conclusions – Exome-wide association meta-analysis demonstrates that protein-coding variants in APOB and APOE associate with subclinical atherosclerosis. APOE ε2 represents the first significant association for multiple subclinical atherosclerosis traits across multiple ethnicities as well as clinical CHD.

Key Words: Genome Wide Association Study; exome; coronary artery calcification; carotid intima-media thickness; genomics
Background

Coronary heart disease (CHD) remains the leading cause of death and infirmity in developed countries. Atherosclerosis is the underlying pathology of CHD. The presence of atherosclerosis in individuals without clinical CHD, termed “subclinical atherosclerosis,” is associated with increased risk of developing clinical CHD independent of traditional risk factors prior to the onset of symptoms. Subclinical atherosclerosis is a heritable clinical phenotype that can be ascertained non-invasively as coronary artery calcification (CAC) by cardiac computed tomography (CT) and common carotid intima media thickness (CIMT) by carotid ultrasound.

Genome-wide association studies (GWAS) within the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium have discovered sites of common non-coding genetic variation associated with both CAC and CIMT among those of European ancestry. Non-coding single nucleotide polymorphisms (SNPs) at the 9p21 and 6p24 regions, near the CDKN2A and PHACTR1 genes, respectively, are strongly associated with both CAC burden and myocardial infarction (MI). The 8q24 (ZHX2), 19q13 (APOC1), and 8q23 (PINX1) loci are strongly associated with CIMT. Observed associations for subclinical atherosclerosis among individuals of European ancestry, however, have not been replicated in those of African ancestry. Furthermore, since the biologic implications of non-coding variation are not as readily interpreted as with coding variation, the roles of such variants in human atherosclerosis remain unclear. Protein-coding variation tends to be infrequently observed and is often inadequately catalogued on earlier GWAS arrays. Rare genomic variation is not well-imputed and exome sequencing to detect such uncommon variation across large populations remains a costly endeavor. Here, we leverage the Illumina HumanExome BeadChip
array, enriched for protein-coding variation. We investigated whether there is evidence for associations of protein-coding variation with two measures of subclinical atherosclerosis across individuals of European and of African ancestry. And we further determine whether such DNA sequence variations may influence CHD risk.

**Methods**

**Study Populations**

The Illumina HumanExome Beadchip v1.0 or v1.1 (also known as the “exome chip”) was used to genotype participants across 19 cohorts of the CHARGE Consortium (Supplement). Participants with a diagnosis of CHD at the time of CAC phenotyping were excluded from CAC analysis. Participants who underwent carotid endarterectomy prior to CIMT phenotyping were excluded from CIMT analysis. 25,109 participants had CAC measured and 52,869 participants had CIMT measured. Each study received institutional review board approval, participants provided written informed consent, and respective governing ethics committees approved each study.

**Measures**

**CAC Measurement**
CIRCCVG/2016/001572-T

Cohorts used different CT scanners to ascertain CAC scoring (Table S1). CAC scoring by multidetector CT and by electron beam CT have been previously described to be highly concordant and are both recognized as valid tools to estimate CAC score.\textsuperscript{19-21} Total CAC score was quantified by the sum of CAC area weighted by density within individual coronary arteries by the Agatston method and the continuous score was used for analysis.\textsuperscript{22}

**CIMT Measurement**

Common carotid intima media thickness was derived by bilateral longitudinal common carotid artery analysis (imaging and measurement methods are described in the Table S2). The mean of the maximum thickness for each common carotid artery was the analytical variable.

**Statistical Analyses**

According to prespecified analysis plans, association analyses and meta-analyses were performed using the seqMeta package (http://cran.r-project.org/web/packages/seqMeta/index.html) in the R statistical software as has previously been performed for exome chip-based analyses.\textsuperscript{23} To reduce skewness, CAC was natural log transformed after adding 1 and CIMT was natural log transformed. Each cohort performed an analysis for each genomic variant with the trait of interest independently and separately for individuals of European and African ancestry to minimize population biases. Covariates in the models included age, sex, and principal components of ancestry derived using EIGENSTRAT.\textsuperscript{24} For studies with related samples, the pairwise kinship matrix was computed and accounted for in
the regression model. Score statistics and genotypic covariance matrices were computed for each cohort and used for additive single variant and gene-based analyses, respectively.

For our primary analyses, we tested the association of each genomic variant with CAC and with CIMT across all samples by meta-analysis that included all cohorts, irrespective of ancestry. We performed single variant analyses on variants that had a minor allele count of at least 20 and gene-based analyses for genes with combined minor allele frequency (MAF) of nonsynonymous variants at least 0.2% to reduce the likelihood of false positive results. We also performed two gene-based tests: 1) T1, where nonsynonymous variants with minor allele frequency (MAF) <1% were collapsed into a gene-based statistic, and 2) sequence kernel association test (SKAT) with MAF <5% for nonsynonymous variants to better account for collapsed variants with bidirectional phenotypic consequences. Regional association plots were generated using LocusZoom. For our secondary analyses, we tested the association of each genomic variant with CAC and CIMT by meta-analysis separately among cohorts of European and African ancestry.

Given the 238,065 variants on the array that passed quality control, the Bonferroni-adjusted level of significance for single variant tests was $0.05/238,065 = 2.10 \times 10^{-7}$. Given the 17,574 genes with nonsynonymous variants on the array, the Bonferroni-adjusted level of significance for gene-based tests was $0.05/17,574 = 2.85 \times 10^{-6}$. For CAC, we had >90% power to detect a variant (MAF <1%) with effect size 0.31 standard deviations, or a gene (combined MAF <1%) with effect size 0.28 standard deviations at a sample size of 25,000. For CIMT, we had >90% power to detect a variant (MAF <1%) with effect size 0.21 standard deviations or a gene (combined
MAF <1%) with effect size 0.20 standard deviations with a sample size of 52,000. Power calculations were performed using the Genetic Power Calculator.\textsuperscript{26}

Methods for the secondary analyses are presented in the Supplement.

Results

Study Participants

19 cohorts participated in the meta-analyses of these two subclinical atherosclerotic traits and the clinical characteristics are summarized in Table S1 and Table S2. A total of 25,109 participants were genotyped with the array and had CAC assessed; of these participants, 19,980 were of European ancestry and 5,129 were of African ancestry. 52,869 participants were genotyped and had CIMT assessed; 44,963 were of European ancestry and 7,906 were of African ancestry.

222,701 (93.5\%) of the 238,065 variants were polymorphic in the CAC meta-analysis; of polymorphic variants, 193,373 (97.1\%) were annotated as nonsynonymous or splice-site variants. Similarly, 227,344 (95.5\%) of array variants were polymorphic in the CIMT meta-analysis and, of these, 217,235 (95.6\%) were nonsynonymous or splice-site variants.

Coronary Artery Calcification Association

Figure 1 plots the meta-analysis CAC association P-value by genomic locus for each variant. The top loci with lead variants associated with CAC among all participants are listed in Table 1.
No systematic association inflation was observed across the set of statistical tests performed (Figure S1).

We identified previously-described common non-coding variant associations at the 9p21 and 6p24 loci. A 9p21 haplotype marked by lead SNP rs10757278-G (MAF 43%), an intergenic variant, was replicated and associated with increased CAC quantity (23.4%; 95% CI: 18.6, 28.3%; $P = 2 \times 10^{-24}$). Similarly, rs9349379-G (MAF 34%), an intronic variant within PHACTR1, was associated with increased CAC quantity (20.9%; 95% CI: 16.3, 25.8%; $P = 5 \times 10^{-20}$). While these associations were robust for those of European ancestry, there was no apparent evidence for association in those of African ancestry (Figure S2, Figure S3). Both loci display locus heterogeneity, or multiple independent associations, for CAC in those of European ancestry (Table 1). We did not discover non-coding variants at other loci on the exome chip that met our stringent Bonferroni alpha threshold. Previously, rs3809346, an intronic variant of COL4A2, had a suggestive association with CAC, but now in our European ancestry sample size that is twice as large, genome-wide significant association was not observed ($P = 2 \times 10^{-3}$).

Among functional variants, a nonsynonymous APOB (rs5742904-T; MAF 0.2%; NM_000384.2:c.10580G>A; NP_000375.2:p.Arg3527Gln) variant was significantly associated with CAC quantity. Carriers of the rare APOB missense variant had markedly increased CAC (4.1-fold; 95% CI: 2.6-, 6.4-fold; $P = 3 \times 10^{-10}$). In our meta-analysis, the Old Order Amish cohort primarily accounted for the strong association, and the variant was extremely rarely observed within other cohorts. Furthermore, the variant was not seen among individuals of African ancestry (Figure S4). We also discovered a distinct rare APOB missense variant (rs1801696-T;
MAF 0.6% European ancestry; NM_000384.2:c.7696G>A; NP_000375.2:p.Glu2566Lys), detected in individuals of European ancestry in most cohorts, that was moderately associated with increased CAC (1.9-fold; 95% CI: 1.6-, 2.1-fold; \(P = 9 \times 10^{-6}\)). This variant was not observed in individuals of African ancestry.

Additionally, a missense 19q13 variant within the APOE gene (rs7412-T; MAF 7.4% European ancestry, 10.8% African ancestry; NM_000041.2:c.526C>T; NP_000032.1:p.Arg176Cys) was associated with diminished CAC quantity (-22.3%; 95% CI, -27.6- -16.7%; \(P = 1 \times 10^{-12}\)) (Figure S5). This association was consistent in those of both European ancestry (-17.3%; 95% CI: -23.7, -10.3%; \(P = 4 \times 10^{-6}\)) and African ancestry (-35.2%; 95% CI: -43.6, -25.7%; \(P = 5 \times 10^{-10}\)) without significant heterogeneity (\(P = 0.53\)) (Figure 2). Additionally, an independent variant (rs769449-A; MAF 11% European ancestry, 2.4% African ancestry) within an intron of APOE also had nominal evidence of association with increased CAC quantity only in individuals of European ancestry (+15.0%; 95% CI: 7.9, 22.6%; \(P = 2 \times 10^{-5}\)).

To improve power of discovery for rare protein-coding variants, we conducted gene-based analyses by aggregating such variants within a gene into a single statistical unit to increase the exposure rate. However, collapsing nonsynonymous variants on the exome chip within a gene did not yield genome-wide significant results (Figure S6).

Carotid Intima Media Thickness Association
There was no systematic inflation of CIMT associations with any variant (Figure S7). The top meta-analysis association findings are listed in Table 2 and a Manhattan plot of all associations is presented in Figure 3.

We noted that, in addition to diminished CAC, the rs7412-T APOE ε2 allele was associated with diminished CIMT (-1.4%: 95% CI: -1.8, -1.0%; P = 4x10^{-14}). There was consistency of association across European and African ancestry cohorts (Figure 4 and Figure S8). There was no significant heterogeneity among the cohorts for this association (P heterogeneity = 0.23).

There were two additional independent suggestive associations at 19q13 at non-coding variants. A variant 5kb upstream of LDLR (rs11668477) was associated with diminished CIMT (P = 5x10^{-7}) primarily among those of European ancestry. This variant has previously been associated with reduced LDL cholesterol.27 The nearby rs7188-G variant (MAF 33% European ancestry, 7.9% African ancestry) within the 3’UTR region of KANK2 was associated with CIMT in those of European ancestry (P = 1x10^{-6}). Additionally, a rare missense variant (rs143873045-A; MAF 0.5% African ancestry; NM_001136191.2:c.1274C>T; NP_001129663.1:p.Ser425Leu) in KANK2 only observed in individuals of African ancestry showed suggestive association with increased CIMT (P = 4x10^{-4}). Lastly, in gene-based analyses, collapsing nonsynonymous variants within a gene did not yield significant associations (Figure S9).

**APOE ε2’s Effect Conditional on LDL Cholesterol**
We sought to determine whether LDL cholesterol concentration accounted for the observed ε2 association with CAC. First, when restricting the original analysis only to participants with LDL cholesterol measurements (n = 20,527), ε2 remained significantly associated with reduced CAC quantity (-22.3%; 95% CI: -25.1, -19.3%; P = 2x10^{-11}) (Table S3). When further adjusting for medication-adjusted LDL cholesterol, the effect estimate was diminished yet the association remained genome-wide significant (-17.0%; 95% CI: -19.7, -14.2%; P = 2x10^{-8}).

APOE ε2’s Effect Conditional on ε3 and ε4

Given the absence of ε4 from the array, we sought to determine whether ε2’s apparent effect on reduced CAC quantity was due to a referent that includes a previously described risk allele (ε3 + ε4). 5,872 participants had CAC and the major APOE genotypes assessed by PCR. Each APOE genotype’s association with CAC (to the ε3/ε3 referent) was performed by cohort and ethnicity and subsequently meta-analyzed with fixed effects. ε2/ε3 was associated with 10.8% reduced CAC (95% CI: -19.6, -0.01%; P = 0.03) and ε2/ε2 with 27.4% reduced CAC (95% CI: -45.2, -0.04%; P = 0.03) (Figure S10).

Concordance of CHD Variants with Subclinical Atherosclerosis Associations

Of the 57 loci previously associated with CHD mainly in individuals of European or South Asian descent, 40 published variants were on the array and available for analysis. 32 of the 40 variants have the same effect direction for CAC and CHD (P = 1.8x10^{-4}) whereas only 23 variants were concordant for CIMT (P = 0.43) in European ancestry participants (Table S4). When restricting
the analysis to variants with at least nominal association \((P < 0.05)\) with CAC, all 17 had concordant effect directions \((P = 4.8 \times 10^{-7})\). A similar analysis with variants at least nominally associated with CIMT showed that 6 of 11 had concordant effect directions for CHD \((P = 0.56)\).

**Replication of Convergent Subclinical Atherosclerosis Finding with CHD**

21,182 individuals of European ancestry, independent of the sample for subclinical atherosclerosis investigations, were genotyped by the Illumina HumanExome BeadChip array, of whom 9,472 had CHD. In cross-sectional analyses, meta-analysis of rs7412-T confirmed a significantly lower odds of CHD (odds ratio 0.77; 95% CI: 0.71, 0.84; \(P = 1.47 \times 10^{-10}\)).

**Discussion**

In our exome-wide association analysis for subclinical atherosclerosis in two distinct ethnicities, we find that protein-coding mutations in *APOB* and *APOE* are associated with subclinical atherosclerosis. While the association for *APOB* was driven by a founder mutation in the Amish, a missense mutation in *APOE* (ε2) was associated with both reduced CAC and CIMT in individuals of European ancestry and African ancestry, even when adjusting for LDL cholesterol concentration. Furthermore, carriers of the ε2 allele had a reduced risk of coronary heart disease. Here, we provide evidence for the first exome-wide association across multiple subclinical atherosclerosis traits and multiple ethnicities for APOE ε2.
Both CAC and CIMT have been proposed as proximal clinical phenotypes of atherosclerosis that may identify individuals at high risk for developing clinical CHD. However, we see that alleles that associate with increased CHD risk also appear to largely result in increased CAC, which is less consistently observed with CIMT. This is concordant with the prior observation that CAC outperforms CIMT in predicting cardiovascular events. Recently, post hoc analyses in statin trials to prevent cardiovascular disease observed that those with a higher burden of CHD-predisposing alleles are more likely to derive clinical benefit from preventive statin therapy. The APOB p.Arg3527Gln (also known as p.Arg3500Gln) has been previously been shown to lead to increased concentrations of LDL cholesterol and premature CHD. Our association signal for this variant was nearly exclusively driven by the Old Order Amish, where it is known to be a founder mutation (MAF 12%) predisposing to increased LDL cholesterol concentrations and CAC quantity through disruption of the LDL receptor binding domain. We also observed a distinct APOB missense mutation, p.Glu2566Lys, with borderline association with increased CAC quantity. Unlike p.Arg3527Gln, p.Glu2566Lys does not occur within the LDL receptor binding domain but occurs within a conserved amphipathic motif of the β2 domain predicted to influence the conversion of VLDL to LDL. Furthermore, we demonstrated that APOE p.Arg176Cys (ε2 allele) was associated with reduced CAC and reduced CIMT in both individuals of European and African ancestry. APOE is an essential mediator of the catabolism and clearance of triglyceride-rich and cholesterol-rich lipoproteins. The major alleles, ε2, ε3, and ε4, have been previously linked to cardiovascular disease, from the candidate gene era, and ε2 is the least common allele. Previously, CHD
risk predisposition from ε4 was primarily thought to be mediated by LDL cholesterol raising effects but observations with ε2 have been mixed. Similarly, ε4, unlike ε2, has been generally linked to ischemic stroke risk. Major reasons for the lack of association of the major APOE alleles with cardiovascular traits in prior genome-wide association studies include the notable absence of rs7412 and rs429358 on population-based genotyping arrays as well as poor imputation of these variants. Similarly, rs429358 is not included on the array used for this study.

ApoE is a major ligand of LDL receptor and a key mediator of remnant lipoprotein particle clearance. The ε2 allele is believed to result in less efficient LDL receptor binding by altering the positive potential. Using publicly available data, ε2 does not impact expression of nearby genes in GTEx nor does it demonstrate enhancer or promoter chromatin marks in ENCODE HepG2 liver cells supporting ε2’s direct impact on ApoE itself. ApoE ε2 can alternatively clear lipoproteins via cell-surface heparan sulfate proteoglycan and LDL receptor-related protein. ApoE ε2 transgenic mice crossbred with ApoB transgenic mice have lower LDL cholesterol. Furthermore, ApoE ε2 transgenic mice lacking LDL receptor still had lower LDL cholesterol suggesting that hypocholesterolemia appears independent of ε2’s effects on LDL receptor. ApoE ε2 impairs lipoprotein lipase-mediated metabolism of VLDL to LDL potentially through the displacement of ApoCII, an activator of lipoprotein lipase. The consequent diminished hepatic cholesterol may subsequently increase LDL receptors for ApoB-containing lipoproteins like LDL.

Interestingly, despite accounting for LDL cholesterol or serum triglycerides, we observe that ε2 still is highly associated with reduced CAC quantity. It is likely that single cross-sectional
measure of lipoproteins, while correlates with, does not fully account for lifelong lipoprotein exposures. ApoE ε2 homozygotes who develop type III hyperlipoproteinemia have a marked increase in remnant lipoprotein particles unlike heterozygotes. Analogously, ApoE ε2-overexpressing mice have increased hepatic VLDL production.\(^4\) Thus, while ApoE ε2 heterozygotes may have an increase in VLDL production and decreased triglyceride catabolism via lipoprotein lipase, the observation of similar triglyceride levels compared to non-carriers suggests preservation of, or enhanced, clearance of remnant lipoprotein particles. We hypothesize that ApoE ε2’s association with reduced subclinical atherosclerosis may be due to increased clearance of both atherogenic LDL and remnant lipoprotein particles through LDL receptor-dependent and -independent pathways. Further work is needed to test this hypothesis.

Our study has several strengths. First, we perform a genetic association meta-analysis across the largest set of individuals to-date for subclinical atherosclerosis in two distinct ancestries. Second, we characterize the association of protein-coding genomic variation, which has not been well studied at the population level, with subclinical atherosclerosis. Third, we explore mechanisms of association through lipoprotein-mediation analyses. Fourth, we provide novel insights with both cross-ethnicity and cross-atherosclerosis trait observations. Fifth, we relate the associations of these subclinical atherosclerosis genetic variants on risk for CHD.

While our study has several strengths, we note some key limitations. First, not all protein-coding variation is catalogued on the exome chip. Due to purifying selection, disruptive protein-coding variation is rare.\(^4\) By potentially not accounting for the totality of disruptive variation not on the array, variance is increased and power is not optimized for gene-based analyses. Whole exome
sequencing can better address this limitation as such technologies continue to become more cost-effective for large-scale experiments. Second, our analyses of prior associations at non-coding sites are restricted to sites on the exome chip. We were able to robustly replicate prior non-coding association analyses for CAC at 9p21 and 6p24. A prior meta-analysis for CIMT genome-wide association discovered one genome-wide association, an intergenic common variant (rs11781551-A) 385kb from ZHX2 at 8q24. No variant with modest linkage disequilibrium with this variant was present on the exome chip thereby limiting ability for replication. An intronic variant in PINX1 at 8q23 and intergenic variant 2.3kb from APOC1 at 19q13 previously had suggestive association but no suitable proxies to replicate association were available on the exome chip. Third, our analysis still demonstrates a paucity of genome-wide associations for these quantitative atherosclerotic traits and highlights an important challenge to ongoing CAC association analyses.

Genetic determinants of CHD have been characterized among individuals of European ancestry but the strongest association signals have not replicated in those of African ancestry which may be due to smaller sample sizes hindering statistical power or different key genetic drivers. But now we demonstrate a cardioprotective genetic mechanism in those of European ancestry and African ancestry through the reduction of subclinical atherosclerosis. We propose potential mechanisms and call for renewed attention to APOE ε2 in the genesis of atherosclerosis underlying clinical cardiovascular disease. Lastly, given the strong concordance of subclinical atherosclerosis measures and clinical CHD, our findings support a future study of genotypes, subclinical atherosclerosis, and incident CHD.
**Sources of Funding**

P.N. is supported by the John S. LaDue Memorial Fellowship in Cardiology, Harvard Medical School. M.K. is supported by the NWO VENI grant (VENI, 91616079). Infrastructure for the CHARGE Consortium is supported in part by the National Heart, Lung, and Blood Institute grant R01HL105756. Funding support for the CHARGE Consortium Exome Chip analyses is provided in part by the National Heart, Lung, and Blood Institute grant R01HL120393. Please refer to the Supplementary Materials regarding additional sources of funding.
Disclosures Statement

P.N. reports grant support from Amarin Corporation. I.B.B. is the Executive Director of Analytical Genetics at Regeneron Pharmaceuticals Inc. I.I. reports grant support from Pie Medical Imaging PV. O.H.F. works in ErasmusAGE, which is funded by Nestlé Nutrition (Nestec Ltd.), Metagenics Inc., and AXA.
References


Table 1. Top meta-analysis variant associations for coronary artery calcification quantity

<table>
<thead>
<tr>
<th>Variant</th>
<th>Consequence</th>
<th>Nearest Gene*</th>
<th>Chrom:Pos†</th>
<th>Minor Allele</th>
<th>MAF</th>
<th>Beta‡</th>
<th>SE</th>
<th>P</th>
<th>MAF</th>
<th>P</th>
<th>MAF</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs10757278§</td>
<td>intergenic (CDKN2B)</td>
<td>9:22124477 G</td>
<td>0.43</td>
<td>0.21</td>
<td>0.020</td>
<td>3.1x10^-24</td>
<td>0.48</td>
<td>2.9x10^-25</td>
<td>0.21</td>
<td>0.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs9349379</td>
<td></td>
<td></td>
<td>intronic PHACTR1</td>
<td>6:12903957 G</td>
<td>0.34</td>
<td>0.19</td>
<td>0.020</td>
<td>4.9x10^-20</td>
<td>0.39</td>
<td>1.28x10^-19</td>
<td>0.094</td>
<td>0.088</td>
</tr>
<tr>
<td>rs7412#</td>
<td>missense APOE</td>
<td>19:45412079 T</td>
<td>0.081</td>
<td>-0.25</td>
<td>0.036</td>
<td>1.19x10^-12</td>
<td>0.074</td>
<td>4.43x10^-6</td>
<td>0.11</td>
<td>5.36x10^-10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs1412829§</td>
<td>intronic CDKN2B</td>
<td>9:22043926 C</td>
<td>0.34</td>
<td>-0.14</td>
<td>0.021</td>
<td>1.56x10^-11</td>
<td>0.41</td>
<td>5.58x10^-12</td>
<td>0.072</td>
<td>0.84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs5742904**</td>
<td>missense APOB</td>
<td>2:21229160 T</td>
<td>2.1x10^-3</td>
<td>1.41</td>
<td>0.22</td>
<td>2.93x10^-10</td>
<td>2.7x10^-3</td>
<td>2.93x10^-10</td>
<td>0</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs9369640</td>
<td></td>
<td></td>
<td>intronic PHACTR1</td>
<td>6:12901441 A</td>
<td>0.43</td>
<td>-0.11</td>
<td>0.019</td>
<td>4.91x10^-8</td>
<td>0.38</td>
<td>5.04x10^-9</td>
<td>0.36</td>
<td>0.71</td>
</tr>
<tr>
<td>rs769449#</td>
<td>intronic APOE</td>
<td>19:45410002 A</td>
<td>0.10</td>
<td>0.14</td>
<td>0.032</td>
<td>7.93x10^-6</td>
<td>0.11</td>
<td>1.86x10^-6</td>
<td>0.024</td>
<td>0.19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs1801696**</td>
<td>missense APOB</td>
<td>2:21232044 T</td>
<td>4.6x10^-3</td>
<td>0.63</td>
<td>0.14</td>
<td>1.44x10^-5</td>
<td>5.7x10^-3</td>
<td>9.77x10^-6</td>
<td>0</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Genes for SNPs that are outside the transcript boundary of the protein-coding gene are shown in parentheses [eg, (CDKN2B)].

† Genomic positions correspond to GRCh37.p13 reference, forward strand.

‡ β-Coefficients are estimated for natural log transformation of total Agatston CAC score+1.

§ CDKN2B lead variants show modest correlation among EA (r²=0.24) and no correlation among AA.
PHACTRI lead variants show modest correlation among EA ($r^2=0.36$) and AA ($r^2=0.05$).

# APOE lead variants show minimal correlation among EA ($r^2=0.01$) and no correlation among AA.

** APOB lead variants are not observed to be correlated.

Abbreviations: AA=African ancestry; AF=minor allele frequency; Chrom:Pos=hg19 build chromosome:position; EA=European ancestry; SE=standard error
# Table 2. Top meta-analysis variant associations for carotid intima media thickness

<table>
<thead>
<tr>
<th>Variant</th>
<th>Consequence</th>
<th>Nearest Gene*</th>
<th>Chrom:Pos†</th>
<th>Minor Allele</th>
<th>All MAF</th>
<th>Beta‡</th>
<th>SE</th>
<th>P</th>
<th>EA MAF</th>
<th>P</th>
<th>AA MAF</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs7412</td>
<td>missense</td>
<td>APOE</td>
<td>19:45412079</td>
<td>T</td>
<td>0.083</td>
<td>-0.014</td>
<td>0.0022</td>
<td>3.79x10^{-14}</td>
<td>0.079</td>
<td>1.97x10^{-10}</td>
<td>0.11</td>
<td>1.43x10^{-5}</td>
</tr>
<tr>
<td>rs11668477</td>
<td>intergenic</td>
<td>(LDLR)</td>
<td>19:11195030</td>
<td>G</td>
<td>0.27</td>
<td>-0.0064</td>
<td>0.0016</td>
<td>4.69x10^{-7}</td>
<td>0.20</td>
<td>5.26x10^{-6}</td>
<td>0.34</td>
<td>0.030</td>
</tr>
<tr>
<td>rs7188</td>
<td>3’UTR</td>
<td>(KANK2)</td>
<td>19:11275139</td>
<td>G</td>
<td>0.29</td>
<td>0.0054</td>
<td>0.0011</td>
<td>2.23x10^{-6}</td>
<td>0.33</td>
<td>1.36x10^{-6}</td>
<td>0.079</td>
<td>0.98</td>
</tr>
<tr>
<td>rs1712790</td>
<td>intergenic</td>
<td>(FAM55B)</td>
<td>11:114621469</td>
<td>C</td>
<td>0.47</td>
<td>-0.0048</td>
<td>0.0011</td>
<td>5.93x10^{-6}</td>
<td>0.48</td>
<td>1.89x10^{-6}</td>
<td>0.21</td>
<td>0.89</td>
</tr>
<tr>
<td>rs2298375</td>
<td>missense</td>
<td>C22orf15</td>
<td>22:24106448</td>
<td>A</td>
<td>0.086</td>
<td>0.0082</td>
<td>0.0019</td>
<td>9.51x10^{-6}</td>
<td>0.085</td>
<td>5.64x10^{-6}</td>
<td>0.091</td>
<td>0.061</td>
</tr>
<tr>
<td>rs174547</td>
<td>intronic</td>
<td>(FADS1)</td>
<td>11:61570783</td>
<td>C</td>
<td>0.30</td>
<td>-0.0049</td>
<td>0.0011</td>
<td>1.07x10^{-5}</td>
<td>0.34</td>
<td>3.84x10^{-5}</td>
<td>0.082</td>
<td>0.062</td>
</tr>
</tbody>
</table>

* Genes for SNPs that are outside the transcript boundary of the protein-coding gene are shown in parentheses [eg, (LDLR)].

† Genomic positions correspond to GRCh37.p13 reference, forward strand.

‡ β-Coefficients are estimated for natural log transformation of CIMT.

Abbreviations: AA=African ancestry; AF=minor allele frequency; Chrom:Pos=hg19 build chromosome:position; EA=European ancestry; SE=standard error; UTR=untranslated region
Figures

Figure 1. Association of each genotyped variant with CAC quantity

Plot of $-\log_{10}(P)$ for association of genotyped variants by chromosomal position for all autosomal polymorphisms analyzed in the age-, sex-, and principal components- adjusted model of coronary artery calcification quantity in the meta-analysis. The genes associated with the top associated variants are displayed.
CAC quantity for *APOE ε2* carriers relative to non-carriers is displayed for all cohorts stratified by European and African ancestries to demonstrate consistency across diverse cohorts and ethnicities.
Figure 3. Association of each genotyped variant with CIMT

Plot of $-\log_{10}(P)$ for association of genotyped variants by chromosomal position for all autosomal polymorphisms analyzed in the age-, sex-, and principal components- adjusted model of carotid intima media thickness in the meta-analysis. The genes associated with the top associated variants are displayed.
Figure 4. Forest plot of relative CIMT for $APOE \, \varepsilon2$ carriers

<table>
<thead>
<tr>
<th>Ethnicity</th>
<th>Cohort</th>
<th>Relative CIMT</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>European</td>
<td>AGES-Reykjavik</td>
<td>0.987</td>
<td>(0.973,1.001)</td>
</tr>
<tr>
<td></td>
<td>ARIC</td>
<td>0.982</td>
<td>(0.971,0.993)</td>
</tr>
<tr>
<td></td>
<td>Biolmage</td>
<td>0.987</td>
<td>(0.973,1.000)</td>
</tr>
<tr>
<td></td>
<td>Cardiovascular Health</td>
<td>0.981</td>
<td>(0.967,0.994)</td>
</tr>
<tr>
<td></td>
<td>Diabetes Heart</td>
<td>0.997</td>
<td>(0.985,1.010)</td>
</tr>
<tr>
<td></td>
<td>Erasmus Ruchphen Family</td>
<td>0.997</td>
<td>(0.965,1.031)</td>
</tr>
<tr>
<td></td>
<td>Framingham Heart</td>
<td>0.984</td>
<td>(0.967,1.001)</td>
</tr>
<tr>
<td></td>
<td>GeneSTAR</td>
<td>0.997</td>
<td>(0.966,1.029)</td>
</tr>
<tr>
<td></td>
<td>LBC1936</td>
<td>0.976</td>
<td>(0.937,1.017)</td>
</tr>
<tr>
<td></td>
<td>MESA</td>
<td>1.003</td>
<td>(0.985,1.023)</td>
</tr>
<tr>
<td></td>
<td>NEO</td>
<td>0.997</td>
<td>(0.986,1.008)</td>
</tr>
<tr>
<td></td>
<td>Rotterdam</td>
<td>0.975</td>
<td>(0.958,0.992)</td>
</tr>
<tr>
<td></td>
<td>SHIP</td>
<td>0.977</td>
<td>(0.963,0.990)</td>
</tr>
<tr>
<td></td>
<td>SHIP Trend</td>
<td>0.981</td>
<td>(0.968,0.995)</td>
</tr>
<tr>
<td></td>
<td>European</td>
<td>0.987</td>
<td>(0.983,0.991)</td>
</tr>
<tr>
<td>African</td>
<td>ARIC</td>
<td>0.980</td>
<td>(0.963,0.997)</td>
</tr>
<tr>
<td></td>
<td>Biolmage</td>
<td>0.995</td>
<td>(0.965,1.025)</td>
</tr>
<tr>
<td></td>
<td>Cardiovascular Health</td>
<td>0.969</td>
<td>(0.942,0.996)</td>
</tr>
<tr>
<td></td>
<td>GeneSTAR</td>
<td>0.981</td>
<td>(0.929,1.036)</td>
</tr>
<tr>
<td></td>
<td>Jackson Heart</td>
<td>0.985</td>
<td>(0.965,1.006)</td>
</tr>
<tr>
<td></td>
<td>MESA</td>
<td>0.968</td>
<td>(0.948,0.989)</td>
</tr>
<tr>
<td></td>
<td>African</td>
<td>0.979</td>
<td>(0.969,0.988)</td>
</tr>
<tr>
<td>ALL</td>
<td></td>
<td>0.986</td>
<td>(0.982,0.989)</td>
</tr>
</tbody>
</table>

CIMT for $APOE \, \varepsilon2$ carriers relative to non-carriers is displayed for all cohorts stratified by European and African ancestries to demonstrate consistency across diverse cohorts and ethnicities.