Published in final edited form as:
Circ Cardiovasc Genet. 2014 June ; 7(3): 359–364. doi:10.1161/CIRCGENETICS.113.000116.

Sequencing of Two Subclinical Atherosclerosis Candidate Regions in 3,669 Individuals: the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Targeted Sequencing Study

Joshua C. Bis, PhD[1,*], **Charles C. White, PhD**[2,3,*], **Nora Franceschini, MD, MPH**[4,*], **Jennifer Brody, BA**[1], **Xiaoling Zhang, PhD**[2], **Donna Muzny, MS**[5], **Jireh Santibanez, MS**[5], **Richard Gibbs, PhD**[5], **Xiaoming Liu, PhD**[6], **Honghuang Lin, PhD**[2,7], **Eric Boerwinkle, PhD**[5,6], **Bruce M. Psaty, MD, PhD**[1,8,9,10], **Kari E. North, PhD**[4,11], **L. Adrienne Cupples, PhD**[2,3], and **Christopher J. O'Donnell, MD**[2,12,13,*] **on behalf of the CHARGE Subclinical Atherosclerosis Working**

[1]Cardiovascular Health Research Unit & Department of Medicine, University of Washington, Seattle, WA

[2]National Heart, Lung and Blood Institute's Framingham Heart Study, Framingham

[3]Department of Biostatistics, Boston University School of Public Health, Boston, MA

[4]Department of Epidemiology, University of North Carolina Chapel Hill, Chapel Hill, NC

[5]Human Genome Sequencing Center, Baylor College of Medicine, University of Texas Health Science Center at Houston, Houston, TX

[6]Human Genetics Center, University of Texas Health Science Center at Houston, Houston, TX

[7]Department of Medicine, Boston University School of Medicine, Boston, MA

[8]Department of Epidemiology, University of Washington, Seattle, WA

[9]Department of Health Services, University of Washington, Seattle, WA

[10]Group Health Research Institute, Group Health, Seattle, WA

[11]Carolina Center for Genome Sciences, University of North Carolina Chapel Hill, Chapel Hill, NC

[12]Division of Intramural Research, National Heart, Lung and Blood Institute, Bethesda, MD

[13]Cardiology Division, Department of Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, MA

Abstract

Correspondence: Joshua C. Bis, PhD, Cardiovascular Health Research Unit, University of Washington, 1730 Minor Avenue, Suite 1360, Seattle, WA 98101, Tel: (206) 287-2777, Fax: (206) 287-2777, joshbis@uw.edu.
*contributed equally

Conflict of Interest Disclosures: B.M.P. serves on the DSMB of a clinical trial of a device funded by Zoll LifeCor and on the Steering Committee of the Yale Open Data Access Project funded by Medtronic. Other authors declare no commercial conflicts of interest.

Background—Atherosclerosis, the precursor to coronary heart disease and stroke, is characterized by accumulation of fatty cells in the arterial intimal-medial layers. Common carotid intima media thickness (cIMT) and plaque are subclinical atherosclerosis measures that predict cardiovascular disease events. Previously, genome-wide association studies demonstrated evidence for association with cIMT (*SLC17A4*) and plaque (*PIK3CG*).

Methods and Results—We sequenced 120kb around *SLC17A4* (6p22.2) and 251kb around *PIK3CG* (7q22.3) among 3,669 European ancestry participants from the Atherosclerosis Risk in Communities Study, Cardiovascular Health Study, and Framingham Heart Study in the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium. Primary analyses focused on 438 common variants (minor allele frequency [MAF] ≥ 1%), which were independently meta-analyzed. A 3' UTR *CCDC71L* variant (rs2286149), upstream from *PIK3CG*, was the most significant finding in cIMT analyses (p= 0.00033) and plaque (p=0.0004). A *SLC17A4* intronic variant was also associated with cIMT (p=0.008). Both were in low LD with the GWAS SNPs. Gene-based tests including T1 count and SKAT for rare variants (MAF < 1%), did not yield statistically significant associations. However, we observed nominal associations for rare variants in the *CCDC71L* and *SLC17A3* with cIMT and of the entire *7q22* region with plaque (p=0.05).

Conclusions—Common and rare variants in the *PIK3CG* and *SLC17A4* regions demonstrated modest association with subclinical atherosclerosis traits. While not conclusive, these findings may help to understand the genetic architecture of regions previously implicated by GWAS and identify variants within these regions for further investigation in larger samples.

Keywords

subclinical atherosclerosis risk factor; common carotid artery; epidemiology; genetics

Atherosclerosis, a precursor to clinical coronary artery disease and some strokes, is characterized by an accumulation of fatty and inflammatory cells and fibrosis in the intimal-medial layers of the arteries. Common carotid intima media thickness (cIMT) and plaque, reflecting a thickening of the carotid artery wall or the presence of large irregular arterial wall deposits, respectively, are established measures of subclinical atherosclerotic disease that can be detected non-invasively and with reasonable precision in population samples using high resolution ultrasound techniques. Multiple independent studies have established consistent association of abnormal carotid phenotypes with coronary events and stroke in prospective studies of young, middle-aged, and older adults,[1, 2] and recent consensus prevention guidelines cite cIMT as a potentially useful measure for prediction of these outcomes.[3]

We have previously identified two regions that demonstrated evidence for association with cIMT (*SLC17A4*) and carotid plaque (*PIK3CG*) in a large-scale genome-wide association meta-analysis conducted among over 40,000 participants from population-based studies of European-ancestry adults.[4] Recent studies have shown that targeted genome sequencing can identify a significant excess burden of functional variants underlying GWAS signals. [5] To identify the causal variants accounting for these signals, we performed a targeted sequencing of these loci using next generation technology. The first region is a 120kb window on

chromosome 6p22.2 including the *SLC17A4*, *SLC17A1*, and *SLC17A3* genes that showed suggestive evidence for association with common cIMT in the GWAS (rs4712972, MAF=0.12, β=0.0099, $p = 7.8 \times 10^{-8}$).[4] The second targeted region is a 251kb stretch of chromosome 7q22.3 including the *PIK3CG* gene that was significantly associated with increased risk of carotid plaque in the GWAS (rs17398575, MAF=0.25, OR=1.18, $p=2\times10^{-12}$). [4] In addition to association with plaque, this region was also selected for sequencing on the basis of its association in recent GWAS with both platelet volume [6] and aggregation [7] as well as pulse pressure. [8].

Rather than identifying new susceptibility loci, our aim was to better characterize the landscape of common and rare variation in these previously associated regions and to determine whether novel or low frequency variation therein was associated with cIMT and carotid plaque. Such fine-mapping across the full spectrum of allele frequencies may help to explain previous genome-wide associations and provide new information on potential mechanisms of atherosclerosis that could contribute to subclinical cardiovascular disease.

Methods

Participating studies

Our analyses were performed within the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium's[9] Targeted Sequencing Study, which included 3,688 European ancestry participants from the Atherosclerosis Risk in Communities Study[10] (ARIC), the Cardiovascular Health Study[11] (CHS), and the Framingham Heart Study[12, 13] (FHS). This sample included a sex-stratified subset of approximately 200 individuals that were selected for elevated measures of age-adjusted cIMT (see Supplementary Note). The remainder of the participants in this study was selected as part of a large random sample or for extreme values of 13 other cardiovascular related phenotypes. Institutional review boards at all participating centers approved the study, and participants gave informed consent. Additional information about the design of these studies is included in the Online Data Supplement.

Carotid artery phenotypes

Each study evaluated the carotid arteries using high-resolution B-mode ultrasonography, using previously described reading protocols to define phenotypes, as per our previous report.[4] For these analyses, we used data from the baseline examination or the first examination in which carotid ultrasonography was obtained. Our primary analysis concerned the common carotid artery using the intima media thickness, typically summarized as the mean of the maximum of several carotid measurements. For most studies, this was an average of multiple measurements from both the left and right arteries. All studies measured the far wall, and several also included the near wall. We also examined the atherosclerotic thickening of the carotid artery wall, defined in two studies by the presence of plaque (CHS, ARIC) or the proxy measure of luminal stenosis greater than 25% (FHS). Specific details for each study's ultrasound, reading, and plaque definition protocols are provided in the Online Data Supplement.

Sequencing

The methods of the CHARGE Targeted Sequencing Study have been described fully in a separate manuscript (Lin H, Wang M, Brody JA, Bis JC, Dupuis J, Lumley T, et al., submitted to *Circ Cardiovasc Genet* along with this manuscript). Briefly, the Study sequenced a total of 77 target regions that harbor genetic variants associated with 14 phenotypes implicated by GWAS within the CHARGE Consortium. Two of the selected genes were particularly relevant to subclinical atherosclerosis. First, the Subclinical Phenotype Group used the University of California at Santa Cruz (UCSC) Genome Browser to select target sequence upstream and downstream of the *SLC17A4/3/1* genes. Secondly, the Pleiotropy Phenotype Group selected the *PIK3CG* region with the aim of capturing both the gene as well as the upstream *CCDC71L* gene and other untranslated regions that harbored SNPs identified in several GWAS.

Approximately 2Mb of target regions were captured by a customized NimbleGen Capture array and sequenced using the ABI SOLiD V4.0 platform. The raw short reads were aligned to the reference human genome (NCBI Genome Build 36, hg18) by BFAST.[14] Samtools[15] was used to pile up aligned reads and call variants with quality filters. The resulting data was then subjected to quality control procedures. Variants were categorized as known or novel by comparison with the dbSNP database and the 1000 Genomes Project. The functional impact of identified variants on the encoded proteins was predicted by the ANNOVAR software package.[16]

Statistical analysis within studies

Each study independently implemented a predefined and standardized analysis plan, as described below.

For the continuous measures of cIMT, we evaluated cross-sectional associations of natural log of cIMT [ln(cIMT)] and genetic variation using linear regression models (or linear mixed effects models in FHS to account for family relatedness). For the dichotomous outcome of plaque, each study used logistic regression models (or general estimating equations clustering on family, to account for familial correlations). In our primary analyses, all studies were adjusted for age, sex, and principal components to account for population substructure (as needed). Some studies made additional adjustments including study site (CHS, ARIC) or familial structure (FHS). For cIMT, we expressed the association of each SNP and ln(cIMT) as the regression slope (β), its standard error [SE(β)] and a corresponding p-value. For the presence of plaque, we calculated a log odds ratio (OR), 95% confidence interval, and p-value. In this case, the OR represents the increase or decrease in the odds of plaque for each additional copy of the variant ’s coded allele. Each study repeated these analyses weighted by each participant’s sampling probability to obtain valid estimates of effect size. (Lumley T, Dupuis J, Rice KM, Barbalic M, Bis JC, Cupples LA, et al. http://stattech.wordpress.fos.auckland.ac.nz/files/2012/05/design-paper.pdf)

We applied two methods of analysis within each region for each of these traits. For each variant with a MAF ≥ 1% in the combined population, each study fit additive genetic models, regressing trait on genotype dosage (0 to 2 copies of the variant allele). For rare

variants, our primary analyses aggregated rare alleles, i.e., those with MAF < 1% into a T1 count statistic, which was defined as the number of variant sites in the target at which a person has at least one rare allele with MAF < 1%. In secondary rare variant analyses of our target regions, we used two approaches: (1) First, we used a Madsen-Browning (MB) type test, which aggregates all variants with MAF< 1% in a genomic region, weighting each variant by a function of its MAF.[17] (2) Second, to explore the possibility that rare variants within a gene did not have the same direction or magnitude of association, we implemented the Sequence Kernel Association Test (SKAT)[18], which approximates the score test that would be obtained fitting a model that includes all the variants.[19]

We then conducted a meta-analysis of regression estimates and standard errors using an inverse-variance weighting approach, implemented in METAL[20] or, for SKAT, using customized R scripts. (Lumley T, Brody J, Dupuis J, Cupples LA http://stattech.wordpress.fos.auckland.ac.nz/files/2012/11/skat-meta-paper.pdf). For the meta-analysis of common variants we excluded SNPs observed in only one study and discarded results with heterogeneity p-values less than 1×10^{-5}. As described above, meta-analysis of unweighted regression coefficients was used to determine significance and meta-analysis of weighted regression coefficients were used to estimate effect size for single variants.

Our primary hypotheses focused on descriptive analyses of genetic sequence variants in the *SLC17A4* and *PIK3CG* regions. Given the prior evidence for these two regions, we used a p-value threshold of $p<0.01$ to identify common variants of potential interest. However, we also performed exploratory analyses across the remaining 75 sequenced targets. Given the hypothesis-free nature of these analyses, we considered only regions satisfying a more stringent "target-wide significance" threshold ($p<1\times10^{-5}$ for common SNPs: 0.05/4800 common variants across the entire sequencing project; or $p<0.00067$, 0.05/75 targets for burden tests).

Regulatory function of SNPs

We used HaploReg[21] to evaluate regulatory function for variants identified in the study (www.broadinstitute.org/mammals/haploreg/haploreg.php). HaploReg is a tool for exploring annotations of the noncoding genome at variants on haplotype blocks that uses LD information from the 1000 Genomes Project. Linked SNPs and small indels are visualized for conservation across mammals, and their effect on regulatory motifs from the ENCODE project, and their predicted chromatin state in nine cell types: embryonic stem cells (H1 ES); erythrocytic leukemia cells (K562); B-lymphoblastoid cells (GM12878); hepatocellular carcinoma cells (HepG2); umbilical vein endothelial cells (HUVEC); skeletal muscle myoblasts (HSMM); normal lung fibroblasts (NHLF); normal epidermal keratinocytes (NHEK); and mammary epithelial cells (HMEC).

Results

Our analysis included 3,669 participants who had carotid ultrasound measures and successful targeted sequencing; characteristics of these participants are shown in Table 1 according to whether they were sampled into the study for extreme values of cIMT, extreme values of another cardiovascular phenotype, or as part of the Cohort Random Sample. As

expected, participants sampled from the cIMT extremes had greater cIMT and were more likely to have plaque, while those sampled for other cardiovascular traits resembled the Cohort Random Sample in terms of these subclinical atherosclerosis phenotypes.

The overall CHARGE Targeted Sequencing Study identified 52,736 variants across all successfully sequenced participants – most were rare variants (only 4,800 had a minor allele frequency greater than 1%). Our primary analyses focused on 3,767 variants in two genomic regions – 6p22 (110 common, 656 rare) and 7q22 (328 common, 2,673 rare) – that had previously associated with subclinical atherosclerosis traits. A summary of the genomic annotation for these variants is shown in Supplementary Table 1.

Common Variants Results

Figure 1 displays the regional plots with results of the meta-analysis for common variants in the 6p22 and 7q22 regions; Supplementary Tables 2A and 2B present the results for individual common variants with p-values less than 0.01 for either cIMT or plaque. Supplementary Figures 1A and B shows the Manhattan plot of all common variant associations across the 77 target regions for cIMT and plaque, respectively.

For the 6p22 region spanning *SLC17A4*, the strongest finding was for a newly-discovered intronic variant with cIMT (rs141877104, MAF=0.015, p=0.008). This variant was in low LD (r^2=0.003) with the common GWAS signal SNP, rs4712972, which was only nominally associated with cIMT in our sequenced subsample (p=0.04). In the analysis of plaque another intronic *SLC17A4* variant (rs76788698, MAF=0.021) was modestly associated with the presence of plaque (p=0.04) and was in low LD with rs4712972 (r^2=0.1). While the correlation-based estimates of LD for these SNPs and the GWAS signal variant were low, there is little evidence of recombination in this region (D'>0.9, Figure 1) suggesting that the signal GWAS SNP could be tagging underlying rarer variant associations. Although neither of these variants were present in the Phase II HapMap CEU panel, they have been reported in the latest version of the 1000 Genomes data.

In the chromosome 7q22 region, we observed the smallest p-value for a common variant rs2286149 (MAF=0.12) for both cIMT (β= 0.018, p-value= 0.0003) and plaque (O.R.=1.4; p-value=0.0003). This variant falls in the 3' UTR of *CCDC71L*, upstream from the targeted *PIK3CG* gene. The plaque GWAS signal variant (rs17398575) is located between these two genes, but was not significantly associated with the presence of plaque (p=0.5) in this subset sample of the large GWAS and is in low LD (r^2 = 0.002, D'= 0.11) with rs2286149 in our sequenced sample.

Functional follow-up of the common variants

Applying HaploReg, we evaluated regulatory function for variants shown in Supplementary Table 2A. We found that rs2190093 affects the HP1-site-factor motif, which is necessary to generate a liver-specific promoter, and is in perfect LD with rs2286149, itself strongly associated with both cIMT (p=0.0007) and plaque (p=0.0004) in this sequencing study. This common variant is located in the 3'UTR region of gene *CCDC71L*, upstream from the targeted *PIK3CG* gene. The implicated promoter binds to a liver-specific transcription

factor, hepatocyte nuclear factor 1 (HNF1A) which is also known as HNF1 and HNF-1alpha. HNF1A is known to regulate the expression of several liver-specific genes. We queried TRANSFAC (http://www.biobase-international.com/product/transcription-factor-binding-sites), a unique knowledge-base of published data on eukaryotic transcription factors, their experimentally-proven binding sites, and regulated genes, we found that target genes of HNF1A include those related to coagulation, lipids, C-reactive protein, inflammation and toxic stress (Supplementary Table 3).

Rare variants

We conducted gene-based rare variant analyses for each of the two chromosomal regions as well as for the five sub-regions within these loci defined by the UCSC gene boundaries (Table 2). We did not observe any statistically significant findings for association between cIMT or plaque with any of these gene-based tests, although we noted suggestive associations of the *SLC17A3* sub-region with cIMT (MB p=0.04), the *CCDC71L* sub-region with cIMT (T1 p=0.05; SKAT p=0.05) and the overall 7q22 region with plaque (T1, p=0.05). Further, limiting the SKAT tests to a set of potentially functional variants (non-synonymous, splice sites, or non-coding SNPs with some evidence of regulatory function defined by RegulomeDB.org score < 6) with MAF < 0.05 did not improve the evidence for association for these genes.

For the remaining 75 regions identified by the CHARGE Targeted Sequencing Study, no common variants or T1 rare variant burdens showed statistically significant associations between cIMT or plaque.

Discussion

In the meta-analysis of association between common sequence variants and subclinical atherosclerosis traits, a variant in the 3' UTR of *CCDC71L*, rs2286149, showed the smallest p-value in both analyses of cIMT and plaque (p-value=0.0003). This variant is located upstream of the targeted *PIK3CG* gene. For the 6p22 region, the strongest finding was for anintronic *SLC17A4* variant (rs141877104) with cIMT (p=0.008), which provides modest evidence toward localizing the broad GWAS peak to one member of the three-gene cluster.

The 7q22 region targeted for sequencing in this study contains two genes, *PIK3CG* and *CCDC71L*. In addition to the association with plaque in our previous GWAS, variants in this region are also strongly associated with mean platelet volume [22, 23]. In particular, variants in the *CCDC71L*, (coiled-coil domain containing 71-like) have been previously associated with expression of the *PIK3CG* gene as a platelet eQTL. [6] *PIK3CG* encodes one of the pi3/pi4-kinase family of proteins, important modulators of extracellular signals, including those elicited by E-cadherin-mediated cell-cell adhesion, which plays an important role of endothelin in maintenance of the structural and functional integrity of epithelia. In our sequencing study a common variant in the *CCDC71L* was strongly associated with both cIMT and plaque. This variant, rs2286149, shows only modest evidence of regulatory function as it alters the Zfp105 regulatory motifs; is in an active promoter for chromatin states in several cell types (HMEC, Uuvec, K562, and NHLF), polycomb repression in HepG2, and falls within a DNAseI hypersensitivity peak in hepatocytes. In the 3'UTR

region of gene *CCDC71L*, another common variant, rs2190093, is in perfect LD with the top associated SNP rs2286149 (r^2=1.0), and is also found to be strongly associated with both cIMT and plaque in this sequencing study. rs2190093 affects the binding site of a liver-specific transcription factor, hepatocyte nuclear factor 1 (HNF1A). Reported targets of HNF1A include *CRP, APOB, APOA2* and coagulation factors including *F2, F8*, and β-fibrinogen (*FGB*).[2425]. *ITGA2B*, platelet glycoprotein IIb, is also a target of HNF1A. Although rs2190093 is in 3'UTR of *CCDC71L*, it is also located in 5' upstream regulatory region of *PIK3CG*, which harbors variants associated with both carotid phenotypes as well as platelet number and volume [26]and aggregation [27]in large GWAS. These results generate the hypothesis that *PIK3CG* may be an HNF1A target expressed in platelets. Rare variant tests in the *CCDC71L* gene showed borderline evidence for association with cIMT and rare variant tests across the entire 7q22 region showed borderline association with plaque.

The chromosome 6p22 region, targeted in this study on the basis of a suggestive association of rs4712972 with common carotid IMT in our previous GWAS, includes *SLC17A4*, *SLC17A1*, and *SLC17A3*; the products of these genes are involved in renal phosphate homeostasis. Serum phosphate has been associated with subclinical atherosclerosis: [2829]. *SLC17A1* is located in the proximal tubules in the kidney and is responsible for renal excretion of Ph; *SLC17A4* is located in the intestine. In addition, two GWAS have associated this region with uric acid levels (*SLC17A1* [30], *SLC17A3* [31]). The association peak was broad in our GWAS that used imputation to the Phase II HapMap CEU reference panel, and the strongest associations in this sequenced subset were for intronic variants in the *SLC17A4* gene that were not present in the HapMap. There is little evidence for regulatory potential for these two variants, though one (rs112544908), which modifies a Pax-3 binding site, was associated with cIMT (p=0.009) and modestly, with higher risk of plaque (p=0.09).

The GWAS that originally identified these loci as associated with subclinical atherosclerosis traits included a much larger sample size and relied on imputation to the Phase II HapMap reference panel. In the current targeted sequencing study, conducted among a subset of participants from three cohorts that comprised the large-scale discovery effort, neither of the prior GWAS top SNPs that identified these loci as targets was strongly associated with either of the subclinical traits. This effort was primarily a comprehensive, descriptive analysis targeted at better characterize variation in regions identified by GWAS; as such, we reported results for variants with p<0.01. However, we acknowledge that few of the findings we have described would satisfy a conservative Bonferroni correction for the total number of tests (p=0.0001, 0.05/~450 common variants) or for one expected false positive accounting for the correlation between the variants (p=0.007, or 1/~140 effectively independent tests). In part, this lack of association could reflect diminished power, particularly for the dichotomous analyses, of this smaller sample size. For instance, our power in the targeted sequencing analysis to discover common variants was excellent (99%) for a single SNP that explained 1% of the overall variance in the trait, but was only moderate (69%) for a variant that explained 0.5% of the trait variance. Nevertheless, these findings from fine-mapping may better localize associations. For instance, the original GWAS signal for plaque fell in an apparent recombination region between *PIK3CG* and

CCDC71L. The sequencing data showed no evidence for association of this GWAS variant with plaque, instead exposing *CCDC71L*, as the potentially relevant locus.

In summary, sequencing of the *PIK3CG* and *SLC17A4* regions within a limited number of participants from three CHARGE cohorts demonstrated evidence of association between plaque and the *PIK3CG* region for both common and rare variants. While not conclusive, these findings may help to better understand the genetic architecture of two regions previously implicated in subclinical atherosclerosis.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding Sources: Funding support for "Building on GWAS for NHLBI-diseases: the U.S. CHARGE consortium" was provided by the NIH through the American Recovery and Reinvestment Act of 2009 (ARRA) (5RC2HL102419). Data for "Building on GWAS for NHLBI-diseases: the U.S. CHARGE consortium" was provided by Eric Boerwinkle on behalf of the Atherosclerosis Risk in Communities (ARIC) Study, L. Adrienne Cupples, principal investigator for the Framingham Heart Study, and Bruce Psaty, principal investigator for the Cardiovascular Health Study. Sequencing was carried out at the Baylor Genome Center (U54 HG003273).

The ARIC Study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute (NHLBI) contracts (HHSN268201100005C, HHSN268201100006C, HHSN268201100007C, HHSN268201100008C, HHSN268201100009C, HHSN2682011000010C, HHSN2682011000011C, and HHSN2682011000012C), R01HL087641, R01HL59367 and R01HL086694. The authors thank the staff and participants of the ARIC study for their important contributions. The Framingham Heart Study is conducted and supported by the NHLBI in collaboration with Boston University (Contract No. N01-HC-25195), and its contract with Affymetrix, Inc., for genome-wide genotyping services (Contract No. N02-HL-6-4278), for quality control by Framingham Heart Study investigators using genotypes in the SNP Health Association Resource (SHARe) project. A portion of this research was conducted using the Linux Cluster for Genetic Analysis (LinGA) computing resources at Boston University Medical Campus. This CHS research was supported by NHLBI contracts N01HC85239, N01HC85079, N01HC85080, N01HC85081, N01HC85082, N01HC85083, N01HC85084, N01HC85085, N01HC-85086; N01HC35129, N01HC15103, N01HC-55222, N01HC75150, N01HC45133, HHSN268201200036C and NHLBI grants HL080295, HL087652, HL105756 with additional contribution from NINDS. Additional support was provided through AG023629, AG15928, AG20098, and AG027058 from the NIA. See also http://www.chs-nhlbi.org/pi.htm.

References

1. Lorenz MW, Markus HS, Bots ML, Rosvall M, Sitzer M. Prediction of clinical cardiovascular events with carotid intima-media thickness: a systematic review and meta-analysis. Circulation. 2007; 115:459–467. [PubMed: 17242284]
2. Stein JH, Korcarz CE, Hurst RT, Lonn E, Kendall CB, Mohler ER, et al. Use of carotid ultrasound to identify subclinical vascular disease and evaluate cardiovascular disease risk: a consensus statement from the American Society of Echocardiography Carotid Intima-Media Thickness Task Force. Endorsed by the Society for Vascular Medicine. J Am Soc Echocardiogr. 2008; 21:93–111. quiz 189-190. [PubMed: 18261694]
3. Greenland P, Alpert JS, Beller GA, Benjamin EJ, Budoff MJ, Fayad ZA, et al. 2010 ACCF/AHA guideline for assessment of cardiovascular risk in asymptomatic adults: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. Circulation. 2010; 122:e584–636. [PubMed: 21098428]
4. Bis JC, Kavousi M, Franceschini N, Isaacs A, Abecasis GR, Schminke U, et al. Meta-analysis of genome-wide association studies from the CHARGE consortium identifies common variants associated with carotid intima media thickness and plaque. Nat Genet. 2011; 43:940–947. [PubMed: 21909108]

5. Johansen CT, Wang J, Lanktree MB, Cao H, McIntyre AD, Ban MR, et al. Excess of rare variants in genes identified by genome-wide association study of hypertriglyceridemia. Nat Genet. 2010; 42:684–687. [PubMed: 20657596]
6. Soranzo N, Rendon A, Gieger C, Jones CI, Watkins NA, Menzel S, et al. A novel variant on chromosome 7q22.3 associated with mean platelet volume, counts, and function. Blood. 2009; 113:3831–3837. [PubMed: 19221038]
7. Johnson AD, Yanek LR, Chen MH, Faraday N, Larson MG, Tofler G, et al. Genome-wide meta-analyses identifies seven loci associated with platelet aggregation in response to agonists. Nat Genet. 2010; 42:608–613. [PubMed: 20526338]
8. Wain LV, Verwoert GC, O'Reilly PF, Shi G, Johnson T, Johnson AD, et al. Genome-wide association study identifies six new loci influencing pulse pressure and mean arterial pressure. Nat Genet. 2011; 43:1005–1011. [PubMed: 21909110]
9. Psaty BM, O'Donnell CJ, Gudnason V, Lunetta KL, Folsom AR, Rotter JI, 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; 2:73–80. [PubMed: 20031568]
10. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. Am J Epidemiol. 1989; 129:687–702. [PubMed: 2646917]
11. Fried LP, Borhani NO, Enright P, Furberg CD, Gardin JM, Kronmal RA, et al. The Cardiovascular Health Study: design and rationale. Ann Epidemiol. 1991; 1:263–276. [PubMed: 1669507]
12. Dawber TR, Kannel WB. The Framingham study. An epidemiological approach to coronary heart disease. Circulation. 1966; 34:553–555. [PubMed: 5921755]
13. Splansky GL, Corey D, Yang Q, Atwood LD, Cupples LA, Benjamin EJ, et al. The Third Generation Cohort of the National Heart, Lung, and Blood Institute's Framingham Heart Study: design, recruitment, and initial examination. Am J Epidemiol. 2007; 165:1328–1335. [PubMed: 17372189]
14. Homer N, Merriman B, Nelson SF. BFAST: an alignment tool for large scale genome resequencing. PLoS ONE. 2009; 4:e7767. [PubMed: 19907642]
15. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009; 25:2078–2079. [PubMed: 19505943]
16. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010; 38:e164. [PubMed: 20601685]
17. Madsen BE, Browning SR. A groupwise association test for rare mutations using a weighted sum statistic. PLoS Genet. 2009; 5:e1000384. [PubMed: 19214210]
18. Wu MC, Lee S, Cai T, Li Y, Boehnke M, Lin X. Rare-variant association testing for sequencing data with the sequence kernel association test. American journal of human genetics. 2011; 89:82–93. [PubMed: 21737059]
19. Chen H, Meigs JB, Dupuis J. Sequence kernel association test for quantitative traits in family samples. Genet Epidemiol. 2013; 37:196–204. [PubMed: 23280576]
20. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics. 2010; 26:2190–2191. [PubMed: 20616382]
21. 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:D930–934. [PubMed: 22064851]
22. Gieger C, Radhakrishnan A, Cvejic A, Tang W, Porcu E, Pistis G, et al. New gene functions in megakaryopoiesis and platelet formation. Nature. 2011; 480:201–208. [PubMed: 22139419]
23. Soranzo N, Spector TD, Mangino M, Kuhnel B, Rendon A, Teumer A, et al. A genome-wide meta-analysis identifies 22 loci associated with eight hematological parameters in the HaemGen consortium. Nat Genet. 2009; 41:1182–1190. [PubMed: 19820697]
24. Courtois G, Morgan JG, Campbell LA, Fourel G, Crabtree GR. Interaction of a liver-specific nuclear factor with the fibrinogen and alpha 1-antitrypsin promoters. Science. 1987; 238:688–692. [PubMed: 3499668]
25. Kugler W, Wagner U, Ryffel GU. Tissue-specificity of liver gene expression: a common liver-specific promoter element. Nucleic Acids Res. 1988; 16:3165–3174. [PubMed: 2453838]

26. Gieger C, Radhakrishnan A, Cvejic A, Tang W, Porcu E, Pistis G, et al. New gene functions in megakaryopoiesis and platelet formation. Nature. 2011; 480:201–208. [PubMed: 22139419]
27. Johnson AD, Yanek LR, Chen MH, Faraday N, Larson MG, Tofler G, et al. Genome-wide meta-analyses identifies seven loci associated with platelet aggregation in response to agonists. Nat Genet. 2010; 42:608–613. [PubMed: 20526338]
28. Onufrak SJ, Bellasi A, Shaw LJ, Herzog CA, Cardarelli F, Wilson PW, et al. Phosphorus levels are associated with subclinical atherosclerosis in the general population. Atherosclerosis. 2008; 199:424–431. [PubMed: 18093595]
29. Ruan L, Chen W, Srinivasan SR, Xu J, Toprak A, Berenson GS. Relation of serum phosphorus levels to carotid intima-media thickness in asymptomatic young adults (from the Bogalusa Heart Study). Am J Cardiol. 2010; 106:793–797. [PubMed: 20816119]
30. Kolz M, Johnson T, Sanna S, Teumer A, Vitart V, Perola M, et al. Meta-analysis of 28,141 individuals identifies common variants within five new loci that influence uric acid concentrations. PLoS Genet. 2009; 5:e1000504. [PubMed: 19503597]
31. Dehghan A, Kottgen A, Yang Q, Hwang SJ, Kao WL, Rivadeneira F, et al. Association of three genetic loci with uric acid concentration and risk of gout: a genome-wide association study. Lancet. 2008; 372:1953–1961. [PubMed: 18834626]

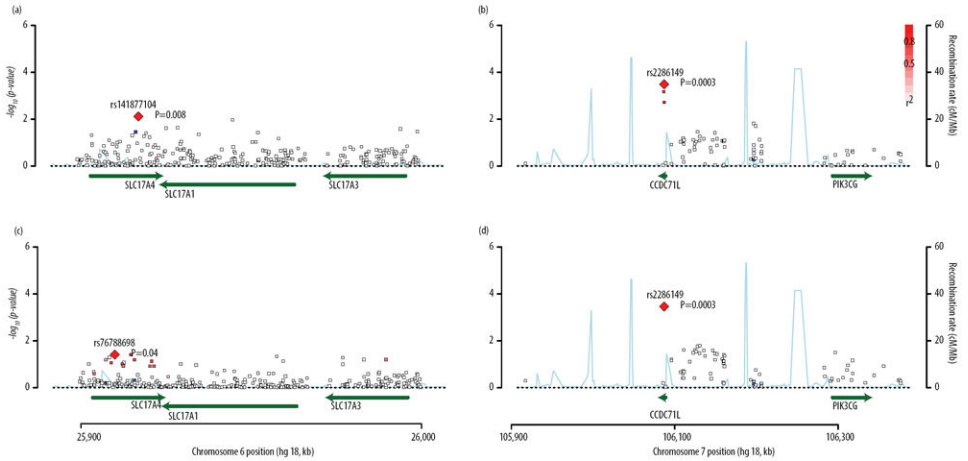

Figure 1.

Regional association plots for common sequence variants and subclinical atherosclerosis phenotypes. Associations of sequenced common variants for the 6p22 (*SLC17A4*, a & c) and 7q22 (*PIK3CG*, b & d) regions and subclinical atherosclerosis phenotypes are plotted according to their chromosomal position on the x-axis (NCBI Genome Build 37, 2009). The y-axis provided the -$\log_{10}$ *P* value of each variant's association with cIMT (a & b) or plaque (b & d). Variants are colored based on the linkage disequilibrium (r^2) to the most significant SNP (red diamond) in each region. Genes are indicated by green arrows and the recombination rate is shown with blue lines.

Table 1

Characteristics of participants

Cohort	Subgroup	N	Age Mean (SD)	Female (%)	cca-IMT median	% with Plaque*
ARIC	Carotid Extremes	69	55 (5)	55	1.24	71
	Random Cohort	803	55 (6)	52	0.73	15
	Other	815	55 (6)	48	0.76	22
CHS	Carotid Extremes	67	74 (6)	46	1.67	94
	Random Cohort	356	73 (5)	51	1.01	66
	Other	702	72 (5)	55	1.01	68
FHS	Carotid Extremes	60	62 (10)	50	1.22	52
	Random Cohort	435	60 (9)	48	0.73	20
	Other	362	60 (10)	53	0.72	20
	Total N	3669				

*Indicates percent with the presence of carotid plaque (ARIC, CHS) or stenosis greater than 25% (FHS)

Table 2

Rare variant analyses, stratified by sub-regions/genes (cIMT and plaque)

		cIMT			Plaque		
Gene	**#SNPs**	**T1**	**MB**	**SKAT**	**T1**	**MB**	**SKAT**
6p22 Region	2673	0.3	0.06	0.7	0.6	1.0	1.0
SLC17A4	561	0.7	1.0	0.6	0.5	0.6	1.0
SLC17A1	1147	0.1	0.1	0.8	0.9	0.4	1.0
SLC17A3	561	0.3	0.04	0.6	0.9	0.8	1.0
7p22	656	0.2	0.3	0.08	0.05	1.0	1.0
CCDC71L	30	0.05	0.4	0.05	0.8	0.4	1.0
PIK3CG	167	0.1	0.8	0.07	0.4	0.9	0.9

Numbers in table represent p-values from gene-based tests of rare variants: T1 indicates T1 count; MB, weighted sum statistic; SKAT, sequence kernel association test.