HHS Public Access

Author manuscript

Genet Epidemiol. Author manuscript; available in PMC 2022 September 01.

Published in final edited form as:

Genet Epidemiol. 2021 September ; 45(6): 651–663. doi:10.1002/gepi.22390.

Exome Sequence Association Study of Levels and Longitudinal Change of Cardiovascular Risk Factor Phenotypes in European-Americans and African-Americans from the Atherosclerosis Risk in Communities Study

Elena V. Feofanova[1], **Elise Lim**[2], **Han Chen**[1,3], **MinJae Lee**[4], **Ching-Ti Liu**[2], **L. Adrienne Cupples**[2,5], **Eric Boerwinkle**[1,6]

[1]Human Genetics Center, Department of Epidemiology, Human Genetics, and Environmental Sciences, School of Public Health, The University of Texas Health Science Center at Houston, Houston, TX, USA.

[2]Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA.

[3]Center for Precision Health, School of Public Health & School of Biomedical Informatics, The University of Texas Health Science Center at Houston, Houston, TX, USA.

[4]Division of Biostatistics, Department of Population & Data Sciences, UT Southwestern Medical Center, Dallas, Texas, USA.

[5]Framingham Heart Study, Framingham, MA, USA.

[6]Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX, USA.

Abstract

Cardiovascular disease (CVD) is responsible for 31% of all deaths worldwide. Among CVD risk factors are age, race, increased systolic blood pressure (BP), and dyslipidemia. Both BP and blood lipids levels change with age, with a dose-dependent relationship between the cumulative exposure to hyperlipidemia and the risk of CVD.

We performed an exome sequence association study using longitudinal data with up to 7,805 European Americans (EA) and 3,171 African Americans (AA) from the Atherosclerosis Risk in Communities (ARIC) study. We assessed associations of common (Minor Allele Frequency>5%) nonsynonymous and splice-site variants and gene-based sets of rare variants with levels and with longitudinal change of seven CVD risk factor phenotypes (BP traits: systolic BP, diastolic BP, pulse pressure; lipids traits: triglycerides, total cholesterol, high-density lipoprotein cholesterol [HDL-C], low-density lipoprotein cholesterol [LDL-C]). Further, we investigated the relationship of the identified variants and genes with select CVD endpoints.

Address for Correspondence: Elena V. Feofanova, MD, MS, PhD, 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., Houston, TX 77030, Phone: +1 (713) 500-9827, elena.v.feofanova@uth.tmc.edu.

Disclosures

The authors declare no conflict of interest.

We identified two novel genes: *DCLK3* associated with the change of HDL-C levels in AAs, and *RAB7L1* associated with the change of LDL-C levels in EAs. *RAB7L1* is further associated with increased risk of heart failure in ARIC EAs.

Investigation of the contribution of genetic factors to the longitudinal change of CVD risk factor phenotypes promotes our understanding of the etiology of CVD outcomes, stressing the importance of incorporating the longitudinal structure of the cohort data in future analyses.

Introduction

Cardiovascular disease (CVD) is the leading cause of morbidity and mortality in the world and is responsible for 31% of all deaths worldwide ("Cardiovascular diseases," 2017). CVD risk factors include but are not limited to age, race, increased systolic blood pressure (SBP), and dyslipidemia (increased total cholesterol [TC], increased low-density lipoprotein cholesterol [LDL-C], decreased high-density lipoprotein cholesterol [HDL-C]) (Goff et al., 2014). Prevention and control of dyslipidemia and hypertension are known to reduce the risk of CVD (Xi et al., 2013).

Both blood pressure (BP) and lipids levels change with age (Abbott et al., 1997; Anagnostis, Stevenson, Crook, Johnston, & Godsland, 2015; Park et al., 2015; Petruski-Ivleva et al., 2016; Sun, 2015) and there is a dose-dependent relationship between the cumulative exposure to hyperlipidemia (such as increased non-HDL-C (Navar-Boggan et al., 2015) and triglycerides [TG] (Glynn, Rosner, & Silbert, 1982)) and the risk of coronary heart disease (CHD). Both the levels and the change of CVD risk factor phenotypes over time are heritable (Azizi et al., 2009; Chen, Li, Srinivasan, Boerwinkle, & Berenson, 2005, 2007; Friedlander et al., 1997; Goode, Cherny, Christian, Jarvik, & de Andrade, 2012; Lin et al., 2014; Tarnoki et al., 2013; X. Wang & Snieder, 2017; Woo, Morrison, Stroop, Aronson Friedman, & Martin, 2014; Zhang et al., 2010; Zhang et al., 2009), and a number of genetic polymorphisms have been identified to have age-dependent influences on blood lipid (Fornage, Papanicolaou, Lewis, Boerwinkle, & Siscovick, 2010; Shirts, Hasstedt, Hopkins, & Hunt, 2011; Srinivasan, Li, Chen, Boerwinkle, & Berenson, 2003) and BP levels (He et al., 2015; Joubert, Diao, Lin, North, & Franceschini, 2009; Simino et al., 2014). Nonetheless, only a few studies have investigated the impact of genetic factors on longitudinal change of CVD risk factors.

Both the levels and the pattern of longitudinal change of the known CVD risk factors throughout life are important for an individual's health. A better understanding of the longitudinal change of risk factors may improve CVD risk prediction. In this study, we investigated whether common (Minor Allele Frequency [MAF]>5%) nonsynonymous and splice-site variants and gene-based sets of rare variants (MAF≤5%) are associated with the levels and longitudinal change of 7 CVD risk factor phenotypes (SBP, diastolic BP [DBP], pulse pressure [PP], TG, TC, HDL-C, LDL-C) using whole exome sequencing (WES) data from the biracial Atherosclerosis Risk In Communities (ARIC) study.

Materials and Methods

Study population

The data used in this research was obtained from the longitudinal ARIC study, a detailed description of which can be found elsewhere ("The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators," 1989). Briefly, a total of 15,792 individuals, mostly of European and African ancestry, 45 to 64 years of age, participated in the baseline examination in 1987–1989, with six follow-up visits in 1990–92 (Visit 2), 1993–95 (Visit 3), 1996–98 (Visit 4), 2011–13 (Visit 5), 2016–2017 (Visit 6), 2018–2019 (Visit 7), and an eighth visit ongoing, which began in 2020, from four communities (Forsyth County, North Carolina; Jackson, Mississippi; Minneapolis, Minnesota; and Washington County, Maryland). In this analysis, we used data from the first five visits. Institutional IRBs and an external advisory board approved the protocols of the study. Informed consent was obtained from each participant.

Measurements

Sitting SBP and DBP were measured by a certified trained technician in accordance with ARIC protocols, 3 times at visits 1–3 and twice at visit 4 using a random zero sphygmomanometer, and 3 times at visit 5 using an automatic sphygmomanometer (OMRON HEM-907 XL) ("ARIC Manual 2 Home and Field Center Procedures ARIC Visit 5 and NCS Study Protocol Version 1.0," 2013; "Atherosclerosis Risk in Communities Study Protocol Manual 2 Cohort Component Procedures Version 2.0," 1988). The average of the second and third measurements (first and second at visit 4) was used here. To account for the use of antihypertensive medications, 15 mm Hg was added to SBP and 10 mm Hg to DBP measurements (Morrison et al., 2014; Wu et al., 2005). PP was calculated as the difference between the SBP and the DBP.

TG, TC and HDL-C (mg/dl) were measured in the fasting state. During visits 1–4, TC and TG were measured by enzymatic procedure (reagents from Boehringer Manneheim Biochemical, Indianapolis, catalog numbers 236691 and 701912, respectively), on a Cobas-Bio analyzer (visits 1–2) and Cobas-Fara II Analyzer (visits 3–4) (*Lipid and Lipoprotein Determinations*, 1987) (*Lipid and Lipoprotein Determinations*, 1991) (*Lipid and Lipoprotein Determinations*, 1994). On visit 5, TC and TG were measured using system reagents on an OLYMPUS analyzer (*Laboratory Methods*, 2012). HDL-C was measured using standard enzymatic and lipoprotein particle precipitation methods. LDL-C was obtained using the Friedewald formula (TC-[TG/5+HDL-C]) (McNamara, Cohn, Wilson, & Schaefer, 1990). Standardized anthropomorphic measurements of weight (in kilograms) and height (in centimeters) were obtained, and body mass index (BMI, kg/m^2) was calculated. Statin medication use was defined as statin use in past 4 weeks, and was self-reported by participants and confirmed by staff from medications brought to the visit (Ballew et al., 2017).

End points of CVD, including hospitalized heart failure (HF), CHD and ischemic stroke (IS), were obtained using yearly hospital medical record review and telephone interviews until 2012; telephone contacts became semi-annual since 2012 (*Heart Failure Cohort*

Surveillance Procedures, 2020; *Stroke Cohort Surveillance Procedures*, 2020; *Surveillance Component Procedures*, 2021; *Surveillance of Heart Failure Manual of Operations*, 2011). Individuals were followed for events from the time of the baseline examination (1987–1989) to 31 December 2017, and those who were lost to follow-up were censored at the date of last contact. To identify incident HF, CHD and IS, surveillance and medical record abstraction of discharge lists from local hospitals and death certificates from state vital statistics offices was performed. A trained abstractor screened eligible hospitalization records (Rosamond et al., 1999) for ICD-9/ICD-10 discharge diagnoses of interest, and for related diagnostic and therapeutic information (Jones, Gottesman, Shahar, Wruck, & Rosamond, 2014; Liao et al., 1997; Rosamond et al., 1999). Death certificates, an interview with one or more next of kin, and a questionnaire completed by the patient's physician were examined in case of out-of-hospital deaths. If available, coroner reports and autopsy reports were also investigated (Liao et al., 1997). Collected evidence was independently reviewed by two physicians on the ARIC morbidity and mortality classification panel (for HF and CHD) or by a physician reviewer and a computer algorithm (for stroke), with discrepancies in the diagnostic classification adjudicated by a third panel member (Jones et al., 2014; Liao et al., 1997; Rosamond et al., 2012).

The diagnosis of HF was based on the *International Classification of Diseases, Ninth Revision* (ICD-9) code 428, or ICD-10, code I50 (Yamagishi, Folsom, Rosamond, Boerwinkle, & Investigators, 2009). A CHD event was defined as a validated definite or probable hospitalized myocardial infarction (MI), a definite CHD death, an unrecognized MI defined by ARIC electrocardiogram readings, or coronary revascularization (Barbalic et al., 2011; White et al., 1996). Diagnosis of stroke was based on ICD-9-CM (clinical modification) codes 430 to 438 (1987–1996) or 430 to 436 (after 1997), with each event classified according to criteria adapted from the National Survey of Stroke (Jones et al., 2014).

Whole Exome Sequencing

All sequencing was performed at the Baylor College of Medicine Human Genome Sequencing Center (HGSC). Exomes were captured using the HGSC VCRome 2.1 reagent (Bainbridge et al., 2011) (42Mb, NimbleGen) and all samples were paired-end sequenced using Illumina GAII or HiSeq instruments. Variant calling was completed using the Atlas2 (Challis et al., 2012) suite. Quality control procedures have been reported elsewhere (Li et al., 2015). Briefly, variants with a single-nucleotide polymorphism (SNP) posterior probability <0.95, depth of coverage <6×, <3 variant reads, an allelic fraction of <0.1, 99% reads in a single direction and homozygous reference alleles with <6× coverage were excluded (Li et al., 2015).

Whole exome variants were annotated using ANNOVAR (K. Wang, Li, & Hakonarson, 2010) and dbNSFP v2.0 (X. Liu, Jian, & Boerwinkle, 2013) according to the reference genome GRCh37 and National Center for Biotechnology Information RefSeq. Coding variants were annotated to a unique gene as well as splicing and nonsynonymous variants categories. Overall, up to 18,822 common (MAF>5%; up to 16,111 in African Americans [AAs]; up to 11,711 in European Americans [EAs]), and up to 949,283 rare (MAF≤5%;

up to 419,613 in AAs; up to 636,570 in EAs) nonsynonymous and splice-site variants were taken forward for analyses.

Statistical Analyses

Each of the selected 7 complex traits was treated as continuous outcome variable to be used in "linear" longitudinal regression models using generalized estimating equations (GEE) (https://cran.r-project.org/web/packages/geeM/geeM.pdf), as well as in Tests for Genetic Association/Gene-Environment Interaction in Longitudinal Studies (LGEWIS) in R (https://cran.r-project.org/web/packages/LGEWIS/LGEWIS.pdf) (Supplemental Methods). All models (including longitudinal regression models and survival analyses) were adjusted for sex, centered age at the first visit, centered BMI, centered BMI squared, study center, and the first three ancestry principal components to account for population stratification (obtained using EIGENSTRAT (Price et al., 2006)). Models for TG were additionally adjusted for TC and statin medication use. Longitudinal regression models also included "Year" (in both GEE and LGEWIS) and squared "Year" (in GEE only, to handle non-linear relationship of time with dependent variables) as structural variables that determine repeated observations per subject.

Single variant analysis

To assess associations with levels of 7 risk factor phenotypes (SBP, DBP, PP, TG, TC, HDL-C, LDL-C), for each of up to 18,822 common (MAF>5%) nonsynonymous and splice-site variants, test was performed using the GEE and the LGEWIS models (Supplemental Methods).

For testing main effect, associations with p-values$<5\times10^{-8}$ (genome-wide significance threshold) by both analytical methods were considered statistically significant. Statistically significant variant-CVD risk factor association pairs with consistent direction of effect and a Bonferroni-corrected p-value in the other race (Supplemental Methods) were considered to have generalized effect.

For each variant with nominal evidence of a main effect (p-value≤0.05) using both analytical methods, association with longitudinal change of the corresponding CVD risk factor phenotype was tested by inclusion of an interaction variable with "Year". To be considered significant, the effect of a variant on longitudinal change of CVD risk factor phenotype had to reach the Bonferroni-adjusted p-value accounting for the total number of variants with nominally statistically significant effect, using both analytical methods. Furthermore, variants reaching the Bonferroni-adjusted p-value by one method, and p-value$<5\times10^{-4}$ by another method were considered suggestively associated with the change of CVD risk factor phenotypes. For the subset of variants with the genome-wide significant main effect specifically, to be considered significant, the effect of a variant on longitudinal change of CVD risk factor phenotype had to reach the Bonferroni-adjusted p-value accounting for the total number of variants with genome-wide significant main effect.

Gene-based test

To assess associations of gene-based aggregation of rare (MAF≤5%) nonsynonymous and splice-site variants with levels of CVD risk factor phenotypes, a set-based test was performed using the naïve burden test collapsing the variants within each gene, and using LGEWIS. For the main effect, associations with p-value<2.5×10^{-6} (accounting for 20,000 genes), using both analytical methods, were considered statistically significant. Statistically significant gene-CVD risk factor association pairs with consistent direction of effect and a Bonferroni-corrected p-value in the other race (Supplemental Methods) were considered to have generalized effect.

For each gene with nominal evidence of a main effect (p-value≤0.05) using both methods, association with longitudinal change in the corresponding risk factor phenotype was assessed by inclusion of an interaction variable with “Year”. To be considered significant, the effect of gene on longitudinal change of CVD risk factor phenotype had to reach the Bonferroni-adjusted p-value accounting for the number of genes reaching nominal main effect significance, using both analytical methods. Furthermore, genes reaching the Bonferroni-adjusted p-value by one method, and p-value<5×10^{-4} by another method were considered suggestively associated with the change of CVD risk factor phenotypes. For the subset of genes reaching the main effect significance threshold (p-value<2.5×10^{-6}), to be considered significant, the effect of gene on longitudinal change of CVD risk factor phenotype had to reach the Bonferroni-adjusted p-value accounting for the number of genes reaching the main effect significance threshold.

Conditional analyses

To identify whether an observed gene-based test association is due to a single rare or low frequency variant, conditional set-based test was performed using the LGEWIS test, conditioning on the variant with the lowest p-value within the gene. For each gene, the variant with the lowest p-value within the gene was identified using LGEWIS single variant test, since GEE method for single rare variants resulted in unstable or unrealistic parameter estimates. Genes with $p\text{-value}_{conditional}<1\times10^{-4}$ were considered to be driven by more than one variant.

Associations with HF, CHD and IS

To investigate the relationship of the identified statistically significant variants and genes and clinical CVD endpoints (HF, CHD, IS), survival analysis was conducted via the ‘survival’ R package, using single variant and naïve burden tests, respectively (https://cran.r-project.org/web/packages/survival/survival.pdf). Continuous time-dependent coefficients were estimated to account for the non-proportional hazards (since we are assuming that the effect of variant or a gene-based aggregation of variants on the clinical CVD endpoint changes over time), to evaluate the association of variants as well as gene-based aggregation of variants with the longitudinal change of the risk of the clinical CVD endpoints (https://cran.r-project.org/web/packages/survival/vignettes/timedep.pdf). For each incident disease analysis, prevalent cases were excluded.

Replication

For novel associations, replication was performed in Exome Chip data from the Framingham Offspring Study (FOS). FOS is a prospective, longitudinal community-based cohort study of the offspring of the participants of the original Framingham Heart Study and their spouses. FOS started with enrollment of 5,124 EA in 1971 to expand knowledge on risk factors and outcomes associated with CVD (Feinleib, Kannel, Garrison, McNamara, & Castelli, 1975; Kannel, Feinleib, McNamara, Garrison, & Castelli, 1979), with participants examined every 4–8 years. Data from up to 9 visits were used (exams 6–8 were used for HF analyses).

In each exam, anthropometric measurements were obtained, and blood was drawn after a ≥8-hour fast. At the first three visits, TC was measured by the method of Abel et al. (Abel, Levy, Brodie, & Kendall, 1952), TG was assessed by the method of Kessler and Lederer (Kessler G, 1966), and HDL-C was determined after plasma precipitation with heparin-manganese (*Lipid and lipoprotein analysis*, 1974). At the following visits the TC and TG levels were measured using the standard enzymatic methods (Wilson, Christiansen, Anderson, & Kannel, 1989), HDL-C was determined after plasma treatment with dextran-magnesium (Warnick, Benderson, & Albers, 1982). LDL-C levels were estimated according to the Friedwald formula.

Genotyping of ~250,000 variants was performed using Illumina HumanExome BeadChip v1.0 (Illumina Inc., San Diego, CA). Genotype calling was completed using Genome Studio Software v1.9.4 (Illumina Inc.). Incident HF events were defined according to Framingham criteria published previously (McKee, Castelli, McNamara, & Kannel, 1971). Analyses were performed in up to 1,617 EAs (909 individuals were included in analyses of HF, 1,617 – for all other analyses). Associations with Bonferroni-adjusted p-value<0.05 were considered statistically significant.

Results

Longitudinal analyses were performed in up to 7,805 EA and 3,171 AA participants from the ARIC study. The average age at baseline was 54.3 and 53.2 for EAs and AAs, respectively, with higher proportion of females in both races (Table S1). Mean (SD) number of examinations was 2.7 (1.3) and 2.6 (1.3) for EAs and AAs, respectively (Table S2). Trajectories of SBP and PP means across subjects at each visit tended to gradually increase with age in both EAs and AAs, with the means being higher in AAs. DBP means tended to decrease with time in both races. No clear trend was seen for HDL-C means in either race. TC and LDL-C means were comparable between races, and tended to decrease with age. TG means were higher in EAs, and tended to increase until visit 3 (in AAs) or 4 (in EAs), and to subsequently decrease. Prevalence of hypertension was higher, while statin use was lower, among AAs compared to EAs throughout the observed study period.

Single variant analysis

Main Effect—Overall, three SNPs were significantly associated with HDL-C in EAs (one in AAs), four – with LDL-C in EAs (one in AAs), three – with TC in EAs (none in AAs), and nine – with TG in EAs (one of which was also detected in AAs), using both

analytical methods (Table S3). Four variant-CVD risk factor associations had generalized effect across EAs and AAs (p-value<3.13×10^{-3}, Supplemental Methods). One additional variant (rs3135506) was associated with HDL-C in EAs according to GEE only. All of the single variants with statistically significant main effect on the blood lipids levels were located in previously implicated genes.

Longitudinal Effect—On average, 694 variants per CVD risk factor phenotype were nominally significant (p-value≤0.05, Table S4) for the main effect association, and were taken forward for the analyses of longitudinal change. Although we observed no statistically significant associations, two variants were suggestively (p-value<5×10^{-4}) associated with the change of levels of CVD risk factors in ARIC participants. One loss-of-function variant (rs58542926, located in *TM6SF2*) was detected by GEE to be associated with the longitudinal change of LDL-C in EAs (Bonferroni-corrected p-value<0.05, Table S3), but was only suggestively associated in LGEWIS. Another missense variant (rs3218721 belonging to *CDT1*) was detected only by GEE to be associated with the longitudinal change of SBP in AAs (Bonferroni-corrected p-value<0.05, Table S3), and was also suggestively associated in LGEWIS.

Among the single variants with genome-wide statistically significant main effect on the blood lipids levels, we didn't find significant (p-value<2.27×10^{-3}, Supplemental Methods) effect of these variants on the change in the corresponding risk factors over time in the corresponding race.

Single variant analysis tended to have a very strong correlation between the p-values from the two methods (r>0.8) both for the main effect and for the longitudinal effect over all seven CVD risk factors in either race (Figure S1) (Akoglu, 2018; Chan, 2003).

Gene-based analysis

Main Effect—In the main effect analysis, four genes (*LPL*, *CETP*, *FAM65A*, *LIPG*) were associated with HDL-C in EAs, one (*PCSK9*) - with LDL-C and TC in AAs, two (*LPL*, *APOC3*) – with TG in EAs (p-value<2.5×10^{-6} using both methods, Table S5). Five of the above genes showed a generalized effect in the other race group (p-value<6.25×10^{-3} by both methods, Supplemental Methods). All of the above genes were previously reported to be associated with the blood lipids traits; all except *LIPG*-HDL-C pair were driven by one leading variant (Table S5). Additionally, three previously reported genes (*MYBPHL*, *APOB*, *BCAM*) were associated with LDL, and one gene (*CNTN2*) – with SBP according only to the LGEWIS approach in EAs. The naïve burden test identified 9 additional genes associated with DBP, 7 – with PP, 7 – with SBP, 8 – with HDL-C, 4 - with LDL, 6 – with TC, and 6 – with TG in either EAs or AAs (Table S5).

Longitudinal Effect—Across all seven CVD risk factors, genes with nominally significant main effects (p-value≤0.05 using both methods, Table S4) were tested for the longitudinal effect of the corresponding CVD risk phenotypes. Two genes were detected by both methods to have a significant effect (Bonferroni-corrected p-value<0.05, Methods, Table 1) on the change of lipids levels. Each rare variant minor allele copy in *RAB7L1* in EAs is expected to decrease LDL-C levels, but with time this effect decreases (until at some point it changes

the direction, after which each copy of the rare variant minor allele carried is expected to increase LDL-C levels; Table S6, Figure 1). A similar trend is seen for the effect of *DCLK3* on HDL-C levels in AAs (Table S6, Figure 1). Both associations can be explained by the leading variant within the genes (rs41302139 for *RAB7L1*, MAF_{EA}=0.013; rs75314929 for *DCLK3*, MAF_{AA}=0.035), given that gene-based analyses for longitudinal effect conditional on the leading variant showed no association ($p\text{-value}_{conditional} > 1\times10^{-4}$, Table S6).

Additionally, one gene was detected only by LGEWIS to be associated with the longitudinal change of LDL-C in AAs. The GEE-only approach detected 48 additional gene-longitudinal change associations. Among the latter, *SLC8A1* was detected by GEE to be associated with the longitudinal change of TG (p-value$<4.87\times10^{-5}$, Tables S3-S4), while LGEWIS showed suggestive association (p-value$<5\times10^{-4}$) in AAs.

None of the genes with statistically significant main effects (p-value$<2.5\times10^{-6}$) had statistically significant (p-value$<6.25\times10^{-3}$) association with the longitudinal change of CVD risk factor phenotypes.

For the main effect analyses, set-based analysis p-values obtained using the two methods tended to have moderate correlation ($0.6<r<0.8$), while the correlation of the p-values for the longitudinal effect analyses was fair ($0.3<r<0.6$) for all seven CVD risk factor phenotypes (Figure S2) (Akoglu, 2018; Chan, 2003). In both races, GEE-only associations were enriched with genes having cMAC<25 in both the main effect (94%) analysis, and in longitudinal effect (83%) analysis.

Survival analysis

None of the single variants detected by both methods with statistically significant main effect were significantly associated with any of the risk of incident CHD, HF or IS events (p-value$<10.4\times10^{-3}$, Supplemental Methods; Table S7). Likewise, for genes with statistically significant main effect via both burden test and LGEWIS test, no statistically significant effect on CVD endpoints were detected (p-value$<1.39\times10^{-3}$, Supplemental Methods; Table S8).

RAB7L1 variants in aggregate were associated with increased HF risk in EAs (HR=1.60, 95%CI:1.24–2.07; Table S9, Table 2). Conditional on the variant with the lowest p-value within the gene (rs41302139), the above association was no longer significant (Table S10, Figure 2).

Replication

The *DCLK3* leading variant was monomorphic in FOS EAs. The *DCLK3* gene-based test performed in the absence of the variant leading in ARIC AAs was not associated with the levels or the change of HDL-C in FOS EAs. A race-specific replication for *DCLK3* may be warranted. Neither *RAB7L1* variants in aggregate, nor rs41302139 in particular were associated with the levels or the change of LDL-C in FOS (Table S11). In contrast to the results in ARIC, in FOS, both *RAB7L1* in aggregate, and rs41302139 in particular, were significantly associated with decreased risk of HF (p-value<0.00001).

Discussion

We identified two loci that are significantly associated with the longitudinal change of blood lipids levels in ARIC participants. Gene-based aggregation tests showed that *DCLK3* is associated with an increase in HDL-C levels with time in AAs, while *RAB7L1* is associated with an increase in LDL-C levels with time in EAs. Additionally, *RAB7L1* is associated with increased risk of HF in ARIC EAs. Investigation of the contribution of genetic factors to the longitudinal change of CVD risk factor phenotypes increases our understanding of CVD outcomes etiology, stressing the importance of incorporating the longitudinal nature of the prospective cohort data in future analyses. The focus of this discussion is primarily the single-variant and gene-based associations identified by both analytical approaches applied – naïve GEE and LGEWIS.

Main effect single variant analysis using coding WES variants identified 16 SNPs associated with at least one blood lipids trait, all of which were located in previously implicated loci. Lack of novel discoveries for the main effect is not surprising, given the limited power of the present study, which was specifically designed with the longitudinal change hypotheses in mind.

Six genes were associated with blood lipids levels using both burden (GEE) and score type (LGEWIS) test approaches, all located in previously implicated loci, but none of them were significantly associated with the change in CVD risk factors (Table S5). Two novel genes were associated with the longitudinal change of lipids levels, with both associations driven by one low-frequency missense variant (Table S6). Both variants are rare in the other race (rs41302139 $MAF_{AA}=2.5\times10^{-3}$; rs75314929 $MAF_{EA}=6.5\times10^{-5}$).

RAB7L1 was associated with increased LDL-C levels with time in ARIC EAs. Importantly, slope of the LDL-C accrual over time increases the risk of CVD events, including HF (Domanski et al., 2020). *RAB7L1* is involved in lysosomal trafficking and maintenance, with *RAB7L1* knock-out mice having an age-associated lysosomal defects characterized by accumulation of lipids in kidney proximal tubule cells (Kuwahara et al., 2016). Lipotoxicity in proximal tubule cells may lead to chronic kidney dysfunction (Nishi, Higashihara, & Inagi, 2019) - a risk factor for incident HF (Kottgen et al., 2007). In ARIC EAs, *RAB7L1* was also associated with increased incidence of HF, primarily due to the effect of one lead variant - rs41302139 (Table S9). Rs41302139 is considered to be among the 3% most deleterious variants of the genome, according to Combined Annotation Dependent Depletion scores (Rentzsch, Witten, Cooper, Shendure, & Kircher, 2019), and its minor allele C results in substitution of a non-polar glutamine for a negatively charged glutamic acid. Amino-acid charge may affect protein properties, such as structure, folding and binding, while modulation of the charges on glutamic acid in particular may lead to protein denaturation (Zhou & Pang, 2018). Lack of replication of the above *RAB7L1* associations in FOS, as well as the opposite direction of association for rs41302139 with HF in FOS, may be explained by differences in LD patterns between the studies, or differences in population ages and observation period lengths.

DCLK3 was associated with increased HDL-C levels over time in ARIC AAs. *DCLK3* is an epithelial-to-mesenchymal transition (EMT)-related genes regulator, possibly promoting EMT events (de Vries et al., 2017; N. Q. Liu et al., 2017). EMT contributes to liver fibrosis (Zhao, Zhu, & Sun, 2016), and more severe liver fibrosis is associated with higher serum HDL-C level (Klisic et al., 2019).

Our suggestive findings may also provide some insight into genetic contribution to CVD risk factor phenotypes change over time. For example, nonsynonymous rs58542926 (belonging to *TM6SF2*) was associated with increased LDL-C levels with time in EAs. Rs58542926 is a known missense variant previously associated with several blood lipids, including decreased LDL-C levels (Gamaldo, Ferrucci, Rifkind, & Zonderman, 2011). *TM6SF2* is involved in VLDL-C metabolism (Ehrhardt et al., 2017), lipid synthesis (Luukkonen et al., 2017) and regulation of the number of secreted lipoprotein particles (Prill et al., 2019), with loss of its function leading to steatohepatosis and lower serum LDL-C and TG levels (Ehrhardt et al., 2017; Fan et al., 2016). Our findings suggest a potential time-dependent nature of the effect of rs58542926 on LDL-C levels.

CVD is a late onset disease and, generally speaking, the levels of its risk factors increase over time, as does risk itself. The trajectory of this increase varies among individuals, and studies of siblings have demonstrated that the longitudinal change of CVD risk factors is also affected by genetic factors (Bonati et al., 2014; Chen et al., 2005, 2007; Friedlander et al., 1997). In addition, there are well-known precedent-setting examples of variants having age dependent effects on CVD risk factor levels (e.g. Familial Hypercholesterolemia and LDL-C) (Khera et al., 2016). It is therefore important to carry out a more comprehensive analysis of change in CVD risk factor levels over time. To help facilitate these analyses, it is noteworthy that most large epidemiologic cohort studies have these risk factors measured over time – often decades. It is, therefore, somewhat surprising that genetic analyses of change are not more common.

Conclusion

In this study, we used two methods, naïve GEE and LGEWIS, to analyze genotype-phenotype associations in the longitudinal change in 7 risk factor traits (SBP, DBP, PP, TC, HDL-C, LDL-C and TG) in the large biracial ARIC study. In the data presented in our study, using the gene-based tests, we discovered two novel genes associated with the change of the CVD risk factors. Both genes possess a time-related biological function, and one gene is further associated with a CVD end-point, thus making both genes good candidates for future research and replication, and emphasizes the benefits of investigating CVD risk factor levels in a longitudinal setting.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The authors thank the staff and participants of the ARIC and FOS study for their important contributions.

Sources of Funding

The Atherosclerosis Risk in Communities study has been funded in whole or in part with Federal funds from the National Heart, Lung, and Blood Institute, National Institutes of Health, Department of Health and Human Services (contract numbers HHSN268201700001I, HHSN268201700002I, HHSN268201700003I, HHSN268201700004I and HHSN268201700005I). 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). Sequencing was carried out at the Baylor College of Medicine Human Genome Sequencing Center (U54HG003273 and R01HL086694).

The Framingham Offspring Study is supported by NIH NHLBI Framingham Heart Study contracts N01-HC-25195, HHSN268201500001I and 75N92019D00031. Genotyping, quality control and calling of the Illumina HumanExome BeadChip in the Framingham Heart Study was supported by funding from the National Heart, Lung and Blood Institute Division of Intramural Research (Daniel Levy and Christopher J. O'Donnell, Principle Investigators). Support for the centralized genotype calling was provided by Building on GWAS for NHLBI-diseases: the U.S. CHARGE consortium through the National Institutes of Health (NIH) American Recovery and Reinvestment Act of 2009 (5RC2HL102419).

Data Availability Statement

The primary data for the ARIC cohort is available via dbGaP Study Accession phs000280, through the established application procedure.

References

Abbott RD, Sharp DS, Burchfiel CM, Curb JD, Rodriguez BL, Hakim AA, & Yano K (1997). Cross-sectional and longitudinal changes in total and high-density-lipoprotein cholesterol levels over a 20-year period in elderly men: The Honolulu heart program. Annals of Epidemiology, 7(6), 417–424. doi:10.1016/s1047-2797(97)00043-4 [PubMed: 9279451]

Abel LL, Levy BB, Brodie BB, & Kendall FE (1952). A simplified method for the estimation of total cholesterol in serum and demonstration of its specificity. J Biol Chem, 195(1), 357–366. [PubMed: 14938387]

Akoglu H (2018). User's guide to correlation coefficients. Turk J Emerg Med, 18(3), 91–93. doi:10.1016/j.tjem.2018.08.001 [PubMed: 30191186]

Anagnostis P, Stevenson JC, Crook D, Johnston DG, & Godsland IF (2015). Effects of menopause, gender and age on lipids and high-density lipoprotein cholesterol subfractions. Maturitas, 81(1), 62–68. doi:10.1016/j.maturitas.2015.02.262 [PubMed: 25804951]

ARIC Manual 2 Home and Field Center Procedures ARIC Visit 5 and NCS Study Protocol Version 1.0 (2013).

The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. (1989). Am J Epidemiol, 129(4), 687–702. [PubMed: 2646917]

Atherosclerosis Risk in Communities Study Protocol Manual 2 Cohort Component Procedures Version 2.0 (1988).

Azizi F, Farahani ZK, Ghanbarian A, Sheikholeslami F, Mirmiran P, Momenan AA, . . . Eskandari F (2009). Familial aggregation of the metabolic syndrome: Tehran Lipid and Glucose Study. Ann Nutr Metab, 54(3), 189–196. doi:10.1159/000217816 [PubMed: 19420912]

Bainbridge MN, Wang M, Wu Y, Newsham I, Muzny DM, Jefferies JL, . . . Gibbs RA (2011). Targeted enrichment beyond the consensus coding DNA sequence exome reveals exons with higher variant densities. Genome Biol, 12(7), R68. doi:10.1186/gb-2011-12-7-r68 [PubMed: 21787409]

Ballew SH, Chen Y, Daya NR, Godino JG, Windham BG, McAdams-DeMarco M, . . . Grams ME (2017). Frailty, Kidney Function, and Polypharmacy: The Atherosclerosis Risk in Communities (ARIC) Study. American Journal of Kidney Diseases, 69(2), 228–236. doi:10.1053/j.ajkd.2016.08.034 [PubMed: 27884475]

Barbalic M, Reiner AP, Wu C, Hixson JE, Franceschini N, Eaton CB, . . . Boerwinkle E (2011). Genome-wide association analysis of incident coronary heart disease (CHD) in African Americans: a short report. PLoS Genet, 7(8), e1002199. doi:10.1371/journal.pgen.1002199 [PubMed: 21829389]

Bonati MT, Graziano F, Monti MC, Crocamo C, Terradura-Vagnarelli O, Cirillo M, . . . Zanchetti A (2014). Heritability of blood pressure through latent curve trajectories in families from the Gubbio population study. J Hypertens, 32(11), 2179–2187. doi:10.1097/HJH.0000000000000311 [PubMed: 25275247]

Cardiovascular diseases (2017). Retrieved from https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)

Challis D, Yu J, Evani US, Jackson AR, Paithankar S, Coarfa C, . . . Yu F (2012). An integrative variant analysis suite for whole exome next-generation sequencing data. BMC Bioinformatics, 13, 8. doi:10.1186/1471-2105-13-8 [PubMed: 22239737]

Chan YH (2003). Biostatistics 104: correlational analysis. Singapore Med J, 44(12), 614–619. [PubMed: 14770254]

Chen W, Li S, Srinivasan SR, Boerwinkle E, & Berenson GS (2005). Autosomal genome scan for loci linked to blood pressure levels and trends since childhood: the Bogalusa Heart Study. Hypertension, 45(5), 954–959. doi:10.1161/01.HYP.0000161881.02361.11 [PubMed: 15809362]

Chen W, Li S, Srinivasan SR, Boerwinkle E, & Berenson GS (2007). A genome scan for loci influencing levels and trends of lipoprotein lipid-related traits since childhood: The Bogalusa Heart Study. Atherosclerosis, 190(2), 248–255. doi:10.1016/j.atherosclerosis.2006.03.006 [PubMed: 16678832]

de Vries PS, Yu B, Feofanova EV, Metcalf GA, Brown MR, Zeighami AL, . . . Morrison AC (2017). Whole-genome sequencing study of serum peptide levels: the Atherosclerosis Risk in Communities study. Hum Mol Genet, 26(17), 3442–3450. doi:10.1093/hmg/ddx266 [PubMed: 28854705]

Domanski MJ, Tian X, Wu CO, Reis JP, Dey AK, Gu Y, . . . Fuster V (2020). Time Course of LDL Cholesterol Exposure and Cardiovascular Disease Event Risk. J Am Coll Cardiol, 76(13), 1507–1516. doi:10.1016/j.jacc.2020.07.059 [PubMed: 32972526]

Ehrhardt N, Doche ME, Chen S, Mao HZ, Walsh MT, Bedoya C, . . . Peterfy M (2017). Hepatic Tm6sf2 overexpression affects cellular ApoB-trafficking, plasma lipid levels, hepatic steatosis and atherosclerosis. Hum Mol Genet, 26(14), 2719–2731. doi:10.1093/hmg/ddx159 [PubMed: 28449094]

Fan Y, Lu H, Guo Y, Zhu T, Garcia-Barrio MT, Jiang Z, . . . Chen YE (2016). Hepatic Transmembrane 6 Superfamily Member 2 Regulates Cholesterol Metabolism in Mice. Gastroenterology, 150(5), 1208–1218. doi:10.1053/j.gastro.2016.01.005 [PubMed: 26774178]

Feinleib M, Kannel WB, Garrison RJ, McNamara PM, & Castelli WP (1975). The framingham offspring study. Design and preliminary data. Preventive Medicine, 4(4), 518–525. doi:10.1016/0091-7435(75)90037-7 [PubMed: 1208363]

Fornage M, Papanicolaou G, Lewis CE, Boerwinkle E, & Siscovick DS (2010). Common INSIG2 polymorphisms are associated with age-related changes in body size and high-density lipoprotein cholesterol from young adulthood to middle age. Metabolism, 59(8), 1084–1091. doi:10.1016/j.metabol.2009.11.005 [PubMed: 20045156]

Friedlander Y, Austin MA, Newman B, Edwards K, Mayer-Davis EJ, & King M-C (1997). Heritability of Longitudinal Changes in Coronary-Heart-Disease Risk Factors in Women Twins. The American Journal of Human Genetics, 60(6), 1502–1512. doi:10.1086/515462 [PubMed: 9199573]

Gamaldo AA, Ferrucci L, Rifkind JM, & Zonderman AB (2011). Age-related changes in mean corpuscular volume in adult whites and African Americans. J Am Geriatr Soc, 59(9), 1763–1764. doi:10.1111/j.1532-5415.2011.03583.x [PubMed: 22136544]

Glynn RJ, Rosner B, & Silbert JE (1982). Changes in cholesterol and triglyceride as predictors of ischemic heart disease in men. Circulation, 66(4), 724–731. doi:10.1161/01.cir.66.4.724 [PubMed: 7116589]

Goff DC Jr., Lloyd-Jones DM, Bennett G, Coady S, D'Agostino RB Sr., Gibbons R, . . . American College of Cardiology/American Heart Association Task Force on Practice, G. (2014). 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol, 63(25 Pt B), 2935–2959. doi:10.1016/j.jacc.2013.11.005 [PubMed: 24239921]

Goode EL, Cherny SS, Christian JC, Jarvik GP, & de Andrade M (2012). Heritability of Longitudinal Measures of Body Mass Index and Lipid and Lipoprotein Levels in Aging Twins. Twin Research and Human Genetics, 10(05), 703–711. doi:10.1375/twin.10.5.703

He WJ, Li C, Rao DC, Hixson JE, Huang J, Cao J, . . . Kelly TN (2015). Associations of Renin-Angiotensin-Aldosterone System Genes With Blood Pressure Changes and Hypertension Incidence. Am J Hypertens, 28(11), 1310–1315. doi:10.1093/ajh/hpv033 [PubMed: 25820244]

Heart Failure Cohort Surveillance Procedures (2020). Retrieved from https://sites.cscc.unc.edu/aric/sites/default/files/public/manuals/Manual%203a.1.%20Cohort%20HF%20Procedures_0.pdf

Jones SA, Gottesman RF, Shahar E, Wruck L, & Rosamond WD (2014). Validity of hospital discharge diagnosis codes for stroke: the Atherosclerosis Risk in Communities Study. Stroke, 45(11), 3219–3225. doi:10.1161/STROKEAHA.114.006316 [PubMed: 25190443]

Joubert BR, Diao G, Lin D, North KE, & Franceschini N (2009). Longitudinal age-dependent effect on systolic blood pressure. BMC Proceedings, 3(Suppl 7), S87. doi:10.1186/1753-6561-3-s7-s87 [PubMed: 20018083]

Kannel WB, Feinleib M, McNamara PM, Garrison RJ, & Castelli WP (1979). An investigation of coronary heart disease in families. The Framingham offspring study. Am J Epidemiol, 110(3), 281–290. doi:10.1093/oxfordjournals.aje.a112813 [PubMed: 474565]

Kessler G LH (1966). Fluorometric measurements of triglycerides. In Skeggs JLT, et al. (Ed.), Automation in analytical chemistry, Technicon Symposia 1965 (pp. p. 341). New York: Eds. Mediad.

Khera AV, Won H-H, Peloso GM, Lawson KS, Bartz TM, Deng X, . . . Kathiresan S (2016). Diagnostic Yield and Clinical Utility of Sequencing Familial Hypercholesterolemia Genes in Patients With Severe Hypercholesterolemia. Journal of the American College of Cardiology, 67(22), 2578–2589. doi:10.1016/j.jacc.2016.03.520 [PubMed: 27050191]

Klisic A, Abenavoli L, Fagoonee S, Kavaric N, Kocic G, & Ninic A (2019). Older age and HDL-cholesterol as independent predictors of liver fibrosis assessed by BARD score. Minerva Med, 110(3), 191–198. doi:10.23736/S0026-4806.19.05978-0 [PubMed: 30784251]

Kottgen A, Russell SD, Loehr LR, Crainiceanu CM, Rosamond WD, Chang PP, . . . Coresh J (2007). Reduced Kidney Function as a Risk Factor for Incident Heart Failure: The Atherosclerosis Risk in Communities (ARIC) Study. Journal of the American Society of Nephrology, 18(4), 1307–1315. doi:10.1681/asn.2006101159 [PubMed: 17344421]

Kuwahara T, Inoue K, D'Agati VD, Fujimoto T, Eguchi T, Saha S, . . . Abeliovich A (2016). LRRK2 and RAB7L1 coordinately regulate axonal morphology and lysosome integrity in diverse cellular contexts. Sci Rep, 6, 29945. doi:10.1038/srep29945 [PubMed: 27424887]

Laboratory Methods (2012). Retrieved from https://sites.cscc.unc.edu/aric/sites/default/files/public/manuals/Manual%208%20Laboratory%20Determinations%20revised%2009242012.pdf

Li AH, Morrison AC, Kovar C, Cupples LA, Brody JA, Polfus LM, . . . Boerwinkle E (2015). Analysis of loss-of-function variants and 20 risk factor phenotypes in 8,554 individuals identifies loci influencing chronic disease. Nature Genetics, 47(6), 640–642. doi:10.1038/ng.3270 [PubMed: 25915599]

Liao D, Cai J, Rosamond WD, Barnes RW, Hutchinson RG, Whitsel EA, . . . Heiss G (1997). Cardiac autonomic function and incident coronary heart disease: a population-based case-cohort study. The ARIC Study. Atherosclerosis Risk in Communities Study. Am J Epidemiol, 145(8), 696–706. [PubMed: 9125996]

Lin CC, Peyser PA, Kardia SL, Li CI, Liu CS, Chu JS, . . . Li TC (2014). Heritability of cardiovascular risk factors in a Chinese population--Taichung Community Health Study and Family Cohort. Atherosclerosis, 235(2), 488–495. doi:10.1016/j.atherosclerosis.2014.05.939 [PubMed: 24953488]

Lipid and lipoprotein analysis (1974). Retrieved from https://books.google.com/books?id=CpIGjwEACAAJ

Lipid and Lipoprotein Determinations (1987). Retrieved from https://sites.cscc.unc.edu/aric/sites/default/files/public/manuals/Lipid_and_Lipoprotein_Determinations.1_8.pdf

Lipid and Lipoprotein Determinations (1991). Retrieved from https://sites.cscc.unc.edu/aric/sites/default/files/public/manuals/Lipid_and_Lipoprotein_Determinations.2_8.pdf

Lipid and Lipoprotein Determinations (1994). Retrieved from https://sites.cscc.unc.edu/aric/sites/default/files/public/manuals/Lipid_and_Lipoprotein_Determinations.3_8.pdf

Liu NQ, Ter Huurne M, Nguyen LN, Peng T, Wang SY, Studd JB, . . . Stunnenberg HG (2017). The non-coding variant rs1800734 enhances DCLK3 expression through long-range interaction and promotes colorectal cancer progression. Nat Commun, 8, 14418. doi:10.1038/ncomms14418 [PubMed: 28195176]

Liu X, Jian X, & Boerwinkle E (2013). dbNSFP v2.0: a database of human non-synonymous SNVs and their functional predictions and annotations. Hum Mutat, 34(9), E2393–2402. doi:10.1002/humu.22376 [PubMed: 23843252]

Luukkonen PK, Zhou Y, Nidhina Haridas PA, Dwivedi OP, Hyotylainen T, Ali A, . . . Yki-Jarvinen H (2017). Impaired hepatic lipid synthesis from polyunsaturated fatty acids in TM6SF2 E167K variant carriers with NAFLD. J Hepatol, 67(1), 128–136. doi:10.1016/j.jhep.2017.02.014 [PubMed: 28235613]

McKee PA, Castelli WP, McNamara PM, & Kannel WB (1971). The Natural History of Congestive Heart Failure: The Framingham Study. New England Journal of Medicine, 285(26), 1441–1446. doi:10.1056/nejm197112232852601

McNamara JR, Cohn JS, Wilson PW, & Schaefer EJ (1990). Calculated values for low-density lipoprotein cholesterol in the assessment of lipid abnormalities and coronary disease risk. Clin Chem, 36(1), 36–42. [PubMed: 2297935]

Morrison AC, Bis JC, Hwang SJ, Ehret GB, Lumley T, Rice K, . . . Levy D (2014). Sequence analysis of six blood pressure candidate regions in 4,178 individuals: the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) targeted sequencing study. PLoS One, 9(10), e109155. doi:10.1371/journal.pone.0109155 [PubMed: 25275628]

Navar-Boggan AM, Peterson ED, D'Agostino RB Sr., Neely B, Sniderman AD, & Pencina MJ (2015). Hyperlipidemia in early adulthood increases long-term risk of coronary heart disease. Circulation, 131(5), 451–458. doi:10.1161/CIRCULATIONAHA.114.012477 [PubMed: 25623155]

Nishi H, Higashihara T, & Inagi R (2019). Lipotoxicity in Kidney, Heart, and Skeletal Muscle Dysfunction. Nutrients, 11(7). doi:10.3390/nu11071664

Park YM, Sui X, Liu J, Zhou H, Kokkinos PF, Lavie CJ, . . . Blair SN (2015). The effect of cardiorespiratory fitness on age-related lipids and lipoproteins. J Am Coll Cardiol, 65(19), 2091–2100. doi:10.1016/j.jacc.2015.03.517 [PubMed: 25975472]

Petruski-Ivleva N, Viera AJ, Shimbo D, Muntner P, Avery CL, Schneider ALC, . . . Kucharska-Newton A (2016). Longitudinal Patterns of Change in Systolic Blood Pressure and Incidence of Cardiovascular DiseaseNovelty and Significance. Hypertension, 67(6), 1150–1156. doi:10.1161/hypertensionaha.115.06769 [PubMed: 27045024]

Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, & Reich D (2006). Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet, 38(8), 904–909. doi:10.1038/ng1847 [PubMed: 16862161]

Prill S, Caddeo A, Baselli G, Jamialahmadi O, Dongiovanni P, Rametta R, . . . Romeo S (2019). The TM6SF2 E167K genetic variant induces lipid biosynthesis and reduces apolipoprotein B secretion in human hepatic 3D spheroids. Sci Rep, 9(1), 11585. doi:10.1038/s41598-019-47737-w [PubMed: 31406127]

Rentzsch P, Witten D, Cooper GM, Shendure J, & Kircher M (2019). CADD: predicting the deleteriousness of variants throughout the human genome. Nucleic Acids Res, 47(D1), D886–D894. doi:10.1093/nar/gky1016 [PubMed: 30371827]

Rosamond WD, Chang PP, Baggett C, Johnson A, Bertoni AG, Shahar E, . . . Chambless LE (2012). Classification of heart failure in the atherosclerosis risk in communities (ARIC) study: a comparison of diagnostic criteria. Circ Heart Fail, 5(2), 152–159. doi:10.1161/CIRCHEARTFAILURE.111.963199 [PubMed: 22271752]

Rosamond WD, Folsom AR, Chambless LE, Wang CH, McGovern PG, Howard G, . . . Shahar E (1999). Stroke Incidence and Survival Among Middle-Aged Adults : 9-Year Follow-Up of the Atherosclerosis Risk in Communities (ARIC) Cohort. Stroke, 30(4), 736–743. doi:10.1161/01.str.30.4.736 [PubMed: 10187871]

Shirts BH, Hasstedt SJ, Hopkins PN, & Hunt SC (2011). Evaluation of the gene-age interactions in HDL cholesterol, LDL cholesterol, and triglyceride levels: the impact of the SORT1 polymorphism on LDL cholesterol levels is age dependent. Atherosclerosis, 217(1), 139–141. doi:10.1016/j.atherosclerosis.2011.03.008 [PubMed: 21466885]

Simino J, Shi G, Bis JC, Chasman DI, Ehret GB, Gu X, . . . Rao DC (2014). Gene-age interactions in blood pressure regulation: a large-scale investigation with the CHARGE, Global BPgen, and ICBP Consortia. Am J Hum Genet, 95(1), 24–38. doi:10.1016/j.ajhg.2014.05.010 [PubMed: 24954895]

Srinivasan SR, Li S, Chen W, Boerwinkle E, & Berenson GS (2003). R219K polymorphism of the ABCA1 gene and its modulation of the variations in serum high-density lipoprotein cholesterol and triglycerides related to age and adiposity in white versus black young adults. The bogalusa heart study. Metabolism, 52(7), 930–934. doi:10.1016/s0026-0495(03)00076-3 [PubMed: 12870173]

Stroke Cohort Surveillance Procedures (2020). Retrieved from https://sites.cscc.unc.edu/aric/sites/default/files/public/manuals/Manual%203c%20Cohort%20Stroke%20Procedures.pdf

Sun Z (2015). Aging, arterial stiffness, and hypertension. Hypertension, 65(2), 252–256. doi:10.1161/HYPERTENSIONAHA.114.03617 [PubMed: 25368028]

Surveillance Component Procedures (2021). Retrieved from https://sites.cscc.unc.edu/aric/sites/default/files/public/manuals/Manual%203%20Version%206.7_0.pdf

Surveillance of Heart Failure Manual of Operations (2011). Retrieved from https://sites.cscc.unc.edu/aric/sites/default/files/public/manuals/Man3a_updated_6172011_0.pdf

Tarnoki AD, Tarnoki DL, Bogl LH, Medda E, Fagnani C, Nistico L, . . . Pietilainen KH (2013). Association of body mass index with arterial stiffness and blood pressure components: a twin study. Atherosclerosis, 229(2), 388–395. doi:10.1016/j.atherosclerosis.2013.05.001 [PubMed: 23880193]

Wang K, Li M, & Hakonarson H (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res, 38(16), e164. doi:10.1093/nar/gkq603 [PubMed: 20601685]

Wang X, & Snieder H (2017). Heritability and Familial Aggregation of Blood Pressure 1–18. doi:10.1007/978-3-319-31420-4_14-1

Warnick GR, Benderson J, & Albers JJ (1982). Dextran sulfate-Mg2+ precipitation procedure for quantitation of high-density-lipoprotein cholesterol. Clin Chem, 28(6), 1379–1388. [PubMed: 7074948]

White AD, Folsom AR, Chambless LE, Sharret AR, Yang K, Conwill D, . . . Tyroler HA (1996). Community surveillance of coronary heart disease in the Atherosclerosis Risk in Communities (ARIC) Study: methods and initial two years' experience. J Clin Epidemiol, 49(2), 223–233. [PubMed: 8606324]

Wilson PW, Christiansen JC, Anderson KM, & Kannel WB (1989). Impact of national guidelines for cholesterol risk factor screening. The Framingham Offspring Study. JAMA, 262(1), 41–44. [PubMed: 2733123]

Woo JG, Morrison JA, Stroop DM, Aronson Friedman L, & Martin LJ (2014). Genetic architecture of lipid traits changes over time and differs by race: Princeton Lipid Follow-up Study. J Lipid Res, 55(7), 1515–1524. doi:10.1194/jlr.M049932 [PubMed: 24859784]

Wu J, Kraja AT, Oberman A, Lewis CE, Ellison RC, Arnett DK, . . . Rao DC (2005). A summary of the effects of antihypertensive medications on measured blood pressure. Am J Hypertens, 18(7), 935–942. doi:10.1016/j.amjhyper.2005.01.011 [PubMed: 16053990]

Xi B, Zhao X, Chandak GR, Shen Y, Cheng H, Hou D, . . . Mi J (2013). Influence of Obesity on Association Between Genetic Variants Identified by Genome-Wide Association Studies and Hypertension Risk in Chinese Children. American Journal of Hypertension, 26(8), 990–996. doi:10.1093/ajh/hpt046 [PubMed: 23591986]

Yamagishi K, Folsom AR, Rosamond WD, Boerwinkle E, & Investigators A (2009). A genetic variant on chromosome 9p21 and incident heart failure in the ARIC study. Eur Heart J, 30(10), 1222–1228. doi:10.1093/eurheartj/ehp087 [PubMed: 19329499]

Zhang S, Liu X, Necheles J, Tsai HJ, Wang G, Wang B, . . . Wang X (2010). Genetic and environmental influences on serum lipid tracking: a population-based, longitudinal Chinese

twin study. Pediatr Res, 68(4), 316–322. doi:10.1203/00006450-201011001-00618 10.1203/PDR.0b013e3181eeded6 [PubMed: 20606601]

Zhang S, Liu X, Yu Y, Hong X, Christoffel KK, Wang B, . . . Wang X (2009). Genetic and environmental contributions to phenotypic components of metabolic syndrome: a population-based twin study. Obesity (Silver Spring), 17(8), 1581–1587. doi:10.1038/oby.2009.125 [PubMed: 19407809]

Zhao YL, Zhu RT, & Sun YL (2016). Epithelial-mesenchymal transition in liver fibrosis. Biomed Rep, 4(3), 269–274. doi:10.3892/br.2016.578 [PubMed: 26998262]

Zhou HX, & Pang X (2018). Electrostatic Interactions in Protein Structure, Folding, Binding, and Condensation. Chem Rev, 118(4), 1691–1741. doi:10.1021/acs.chemrev.7b00305 [PubMed: 29319301]

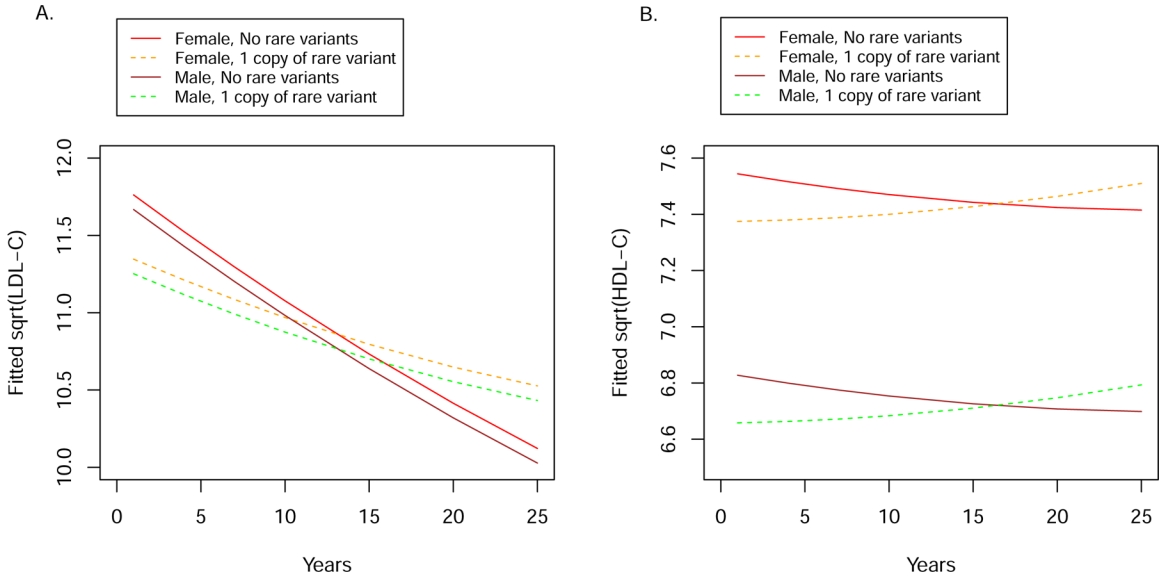

Figure 1.

Estimated effect of rare variants in A. *RAB7L1* (EAs), B. *DCLK3* (AAs) genes on longitudinal change of CVD risk factor phenotypes by gender.

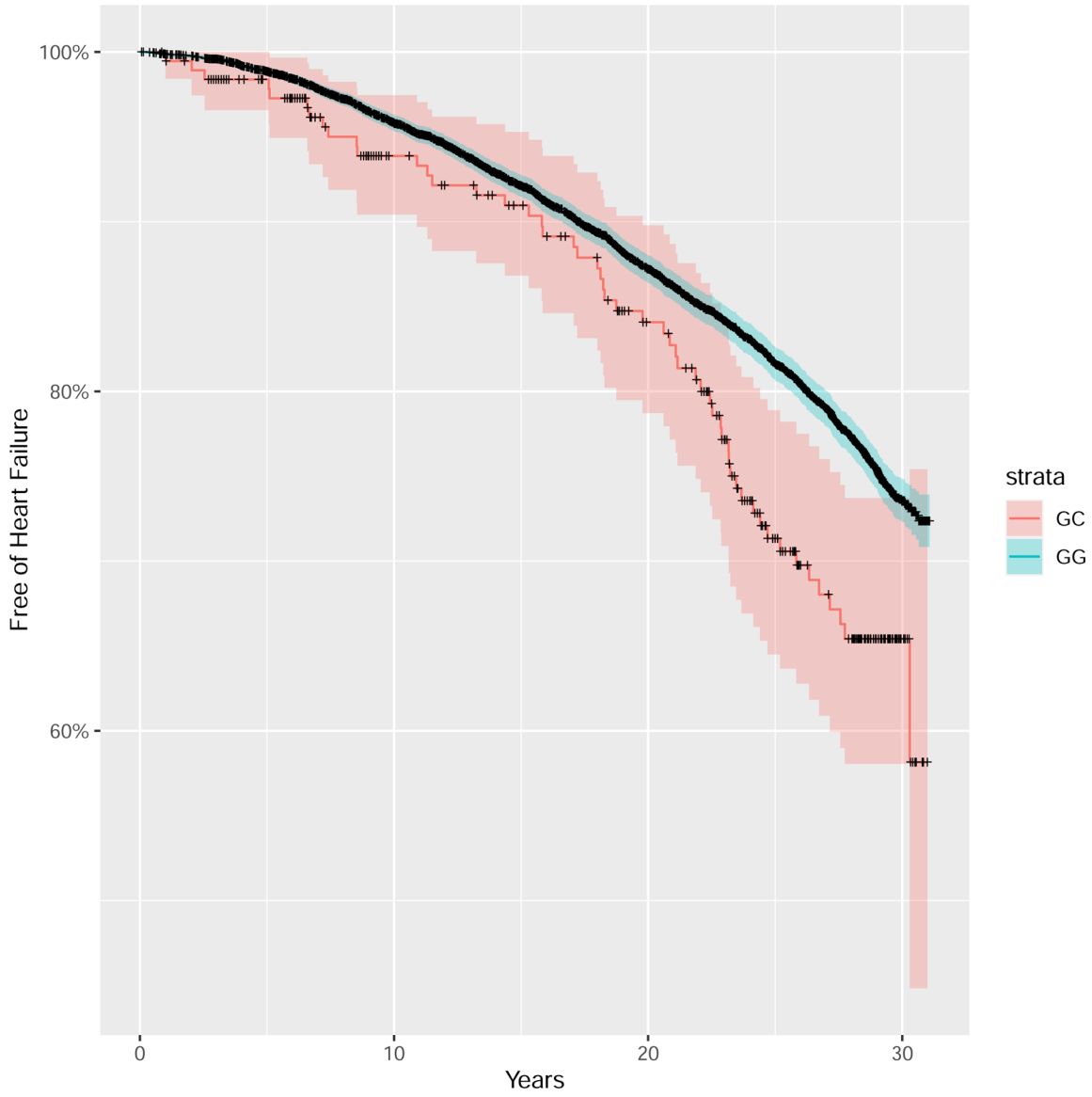

Figure 2.
Survival analysis for rs41302139: effect on Heart Failure.

Table 1.

Gene-based Analysis Results for longitudinal change

Information							Main Effect				Longitudinal Change				N
Pheno-type	Gene	Race	Chr	Location	# SNPs	cMAC	Beta	SE	*P*-value (Burden)	*P*-value (LGEWIS)	Beta	SE	*P*-value (Burden)	*P*-value (LGEWIS)	
HDL	*DCLK3*	AA	3	36756821:36780115	30	329	−0.11	0.05	3.55×10^{-2}	3.54×10^{-2}	0.01	0.002	4.65×10^{-6}	1.14×10^{-6}	3170
LDL	*RAB7L1*	EA	1	205739490:205741653	9	220	−0.22	0.08	7.35×10^{-3}	8.32×10^{-3}	0.03	0.008	1.61×10^{-5}	2.05×10^{-5}	7788

Table 2.

Gene-based Analysis Results for incident Heart Failure

Gene Information						Main Effect Incident Heart Failure			
Gene	**Race**	**Chr**	**Location**	**# SNPs**	**cMAC**	**Beta**	**SE**	***P*-value (Burden)**	**N**
RAB7L1	EA	1	205739490:205741653	9	208	0.47	0.132	3.64×10^{-4}	7384