Role of rare and low-frequency variants in gene-alcohol interactions on plasma lipid levels

Wang, Z.; Chen, H.; Bartz, T.M.; Bielak, L.F.; Chasman, D.I.; Feitosa, M.F.; ... ; CHARGE Gene Lifestyle Interactions

Citation
Wang, Z., Chen, H., Bartz, T. M., Bielak, L. F., Chasman, D. I., Feitosa, M. F., … Morrison, A. C. (2020). Role of rare and low-frequency variants in gene-alcohol interactions on plasma lipid levels. *Circulation: Genomic And Precision Medicine*, *13*(4), 307-313. doi:10.1161/CIRCGEN.119.002772

Version:	Publisher's Version
License:	Creative Commons CC BY-NC-ND 4.0 license
Downloaded from:	https://hdl.handle.net/1887/3182379

Note: To cite this publication please use the final published version (if applicable).

ORIGINAL ARTICLE

Role of Rare and Low-Frequency Variants in Gene-Alcohol Interactions on Plasma Lipid Levels

Zhe Wang, PhD; Han Chen, PhD; Traci M. Bartz, MS; Lawrence F. Bielak, PhD; Daniel I. Chasman, PhD; Mary F. Feitosa, PhD; Nora Franceschini, MD, MPH; Xiuqing Guo, PhD; Elise Lim, MS; Raymond Noordam, PhD; Melissa A. Richard, PhD; Heming Wang, PhD; Brian Cade, PhD; L. Adrienne Cupples, PhD; Paul S. de Vries, PhD; Franco Giulanini, PhD; Jiwon Lee, MS; Rozenn N. Lemaitre, PhD; Lisa W. Martin, MD; Alex P. Reiner, MD; Stephen S. Rich, PhD; Pamela J. Schreiner, PhD; Stephen Sidney, MD, MPH; Colleen M. Sitlani, PhD; Jennifer A. Smith, PhD; Ko Willems van Dijk, PhD; Jie Yao, MS; Wei Zhao, PhD; Myriam Fornage, PhD; Sharon L.R. Kardia, PhD; Charles Kooperberg, PhD; Ching-Ti Liu, PhD; Dennis O. Mook-Kanamori, PhD; Michael A. Province, PhD; Bruce M. Psaty, MD, PhD; Susan Redline, MD, MPH; Paul M. Ridker, MD, MPH; Jerome I. Rotter, MD; Eric Boerwinkle, PhD; Alanna C. Morrison, PhD; on behalf of the CHARGE Gene-Lifestyle Interactions Working Group*

BACKGROUND: Alcohol intake influences plasma lipid levels, and such effects may be moderated by genetic variants. We aimed to characterize the role of aggregated rare and low-frequency protein-coding variants in gene by alcohol consumption interactions associated with fasting plasma lipid levels.

METHODS: In the Cohorts for Heart and Aging Research in Genomic Epidemiology consortium, fasting plasma triglycerides and high- and low-density lipoprotein cholesterol were measured in 34 153 individuals with European ancestry from 5 discovery studies and 32 277 individuals from 6 replication studies. Rare and low-frequency functional protein-coding variants (minor allele frequency, ≤5%) measured by an exome array were aggregated by genes and evaluated by a gene-environment interaction test and a joint test of genetic main and gene-environment interaction effects. Two dichotomous self-reported alcohol consumption variables, current drinker, defined as any recurrent drinking behavior, and regular drinker, defined as the subset of current drinkers who consume at least 2 drinks per week, were considered.

RESULTS: We discovered and replicated 21 gene-lipid associations at 13 known lipid loci through the joint test. Eight loci (*PCSK9*, *LPA*, *LPL*, *LIPG*, *ANGPTL4*, *APOB*, *APOC3*, and *CD300LG*) remained significant after conditioning on the common index single-nucleotide polymorphism identified by previous genome-wide association studies, suggesting an independent role for rare and low-frequency variants at these loci. One significant gene-alcohol interaction on triglycerides in a novel locus was significantly discovered (P=6.65×10^{-6} for the interaction test) and replicated at nominal significance level (P=0.013) in *SMC5*.

CONCLUSIONS: In conclusion, this study applied new gene-based statistical approaches and suggested that rare and low-frequency genetic variants interacted with alcohol consumption on lipid levels.

Key Words: exome ■ gene-environment interaction ■ genome-wide association study ■ lipids ■ self-report

Downloaded from http://ahajournals.org by on September 28, 2022

Plasma lipid profiles, including high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C), and triglyceride (TG) levels, have been well characterized for their roles in the development and prevention of cardiovascular disease.[1,2] Genome-wide association studies (GWAS) and advanced DNA sequence technology have uncovered >200 genetic loci influencing lipid levels,[3–8] and these common (minor allele frequency [MAF], >5%) single-nucleotide polymorphisms (SNPs) often reside in noncoding regions of the genome. In addition to the evidence that genetic factors affect plasma lipid profiles, environmental factors

Correspondence to: Alanna C. Morrison, PhD, UTHealth School of Public Health, RAS E447, 1200 Pressler St, Houston, TX 77030. Email alanna.c.morrison@uth.tmc.edu

*A list of the CHARGE Gene-Lifestyle Interactions Working Group is given in the Data Supplement.

The Data Supplement is available at https://www.ahajournals.org/doi/suppl/10.1161/CIRCGEN.119.002772.

For Sources of Funding and Disclosures, see page 312.

© 2020 American Heart Association, Inc.

Circulation: Genomic and Precision Medicine is available at www.ahajournals.org/journal/circgen

Nonstandard Abbreviations and Acronyms

HDL-C	high-density lipoprotein cholesterol
LDL-C	low-density lipoprotein cholesterol
TG	triglyceride
CVD	cardiovascular disease
GWAS	genome-wide association study
SNP	single nucleotide polymorphism
MAF	minor allele frequency
G×E	gene-by-environment
MAC	minor allele count

influence lipid levels as well. Epidemiological studies have demonstrated an association between moderate alcohol consumption and improved lipid profile, including higher HDL-C levels, HDL particle concentration, and HDL-C subfractions.[9,10] However, the association evidence between alcohol use and LDL-C or TG levels is inconsistent. Some studies reported positive associations, whereas others reported negative associations.[11–15]

Studying gene-by-environment (G×E) interactions is important, as it extends our knowledge of the genetic architecture of complex traits and improves our understanding of the underlying mechanisms of common diseases for novel and known loci.[16–18] Several large-scale genome-wide G×E studies have successfully identified novel common variants accounting for the environmental effects such as alcohol consumption and smoking status on lipid levels and other cardiovascular disease–related traits.[19–22] These studies have successfully identified common variant loci that were not detected in main effects GWAS. However, unlike well-established G×E interaction tests for common variants,[23,24] methods for detecting rare variant G×E interactions are emerging. Recently developed novel approaches for testing rare variant G×E interaction effects include a joint test that allows for simultaneous testing of the genetic main effect and interaction effect, as well as the ability to assess gene-based G×E interactions for both related and unrelated individuals.[25]

Accounting for the effect of alcohol consumption in defining the genetic architecture of lipid levels may not only provide valuable insights into relationship between alcohol consumption and lipids but also may help refine association signals at previously identified GWAS loci or identify new loci. This study is the first to incorporate G×E interaction in modeling rare and low-frequency variant genetic and alcohol effects on plasma lipid levels.

METHODS

This study includes 66 430 men and women between 18 and 80 years of age from 11 European ancestry population studies that are part of the Cohorts for Heart and Aging Research in Genomic Epidemiology Gene-Lifestyle Interactions Working Group[18] (Figure I in the Data Supplement). Each study obtained informed consent from participants and approval from the appropriate institutional review boards. Additional detail for these studies and full Methods are available in the Data Supplement of the article. Data from consortia were accessed subject to the applicable data-sharing agreements. Summary data are available to other researchers on reasonable request to the corresponding authors.

RESULTS

Descriptive statistics for up to 34 153 participants of the 5 discovery and 32 277 participants of the 6 replication studies are summarized in Table 1 and Table I in the Data Supplement. On average, two-thirds of the study participants were current drinkers and 39.5% were regular drinkers. The proportion of current and regular drinkers was greater in the discovery studies as compared with the replication studies.

We performed gene-based analyses for each lipid/alcohol consumption combination using (1) a G×E test that considers the genetic main effects as random effects and (2) a joint analysis of the genetic main and the G×E interaction effects in each study participating in the discovery phase. Significant genes from meta-analysis of the discovery studies were pursued for replication. Overall, meta-analyses showed highly consistent results across current drinker and regular drinker (Table II in the Data Supplement). Distributions of QQ plots for meta-analyzing discovery studies are shown in Figure II in the Data Supplement. In the discovery phase, we observed 31 gene-lipid associations ($P<5\times10^{-5}$) in the joint analysis and 5 gene-lipid associations ($P<5\times10^{-5}$) in the interaction test, with 3 genes (*INDK*, *REM2*, and *SMC5*) overlapping between the 2 approaches (Table II in the Data Supplement). These gene-lipid pairs were taken forward for replication, one of which (*IDNK*) was only available in 1 replication study (the CHS [Cardiovascular Health Study]). Therefore, we evaluated 30 gene-lipid associations for replication using the joint test and 4 using the gene-alcohol interaction test (Table II in the Data Supplement). Thirteen known lipid loci (21 gene-lipid associations) were replicated, and 1 novel interaction at a novel locus was replicated at the borderline for Bonferroni-corrected significant level (P_{int}=0.013) for the *SMC5*-by-current drinker interaction on TG levels (Table 2). The average TG levels for SMC5 carriers and noncarriers by current drinker status among discovery studies were showed in Figure 1. Among the replicated genes, 4 were shared between TG and HDL-C but none were shared between LDL-C and TG or HDL-C, as shown in a Venn diagram (Figure 2).

For the 13 known lipid loci that were replicated through the joint test, additional analyses were conducted following the flowchart shown in Figure 3. First, we performed conditional analyses to examine whether the gene-based

Downloaded from http://ahajournals.org by on September 28, 2022

Table 1. Descriptive Characteristics for Discovery and Replication Studies

	Study	Design	n	CurDrinker, %	RegDrinker, %
Discovery	ARIC	Unrelated	10 989	64.9	36.8
	FHS	Family	7258	83.6	65.5
	NEO	Unrelated	5718	86.8	69.0
	WHI	Unrelated	8021	76.5	32.6
	CARDIA	Unrelated	2167	68.7	59.6
	Total/average		34 153	75.5	48.7
Replication	WGHS	Unrelated	22 478	56.7	29.3
	CFS	Family	253	50.2	25.1
	CHS	Unrelated	3690	53.8	25.0
	FamHS	Family	1735	50.7	28.3
	GENOA	Family	1543	53.1	29.2
	MESA	Unrelated	2578	71.9	43
	Total/average		32 277	57.2	29.8
Overall			66 430	66.6	39.5

ARIC indicates Atherosclerosis Risk in Communities study; CARDIA, Coronary Artery Risk Development in Young Adults study; CFS, Cleveland Family Study; CHS, Cardiovascular Health Study; CurDrinker, current drinker; FamHS, Family Heart Study; FHS, Framingham Heart Study; GENOA, Genetic Epidemiology Network of Arteriopathy study; MESA, Multi-Ethnic Study of Atherosclerosis; NEO, Netherlands Epidemiology of Obesity study; RegDrinker, regular drinker; WGHS, Women's Genome Health Study; and WHI, Women's Health Initiative study.

rare variant associations are independent of the common index SNP identified by previous GWAS. In total, 8 loci (*PCSK9*, *LPA*, *LPL*, *LIPG*, *ANGPTL4*, *APOB*, *APOC3*, and *CD300LG*; 10 gene-lipid associations) remained significant after conditioning on a common index SNP. However, the genes that were not reported to be associated with lipids themselves but in known lipid loci, such as *BCAM* and *CBLC* on LDL-C, were strongly attenuated after adjusting for rs7412, the index SNP of *APOE* identified by previous GWAS and, in part, defining the *APOE2/3/4* alleles (Table III in the Data Supplement).

Second, single-variant analyses were performed for the 5 gene-lipid associations that were not evaluated in the conditional analyses because they did not have previously reported common SNPs and for the 10 gene-lipid pairs that remained significant following conditional analyses (Figure 3; Table III in the Data Supplement). Single-variant tests at these genes confirmed previous known low-frequency lipid variants. For example, rs11591147 in *PCSK9* was associated with LDL-C, and rs77960347 in *LIPG* and rs116843064 in *ANGPTL4* were associated with HDL-C. Additionally, we provide evidence that 2 of the driving variants underlying the joint test results are novel rare variants associated with LDL-C (Table IV in the Data Supplement). One of them is rs41267813, a variant in the *LPA* gene (P=6.55×10^{-29} discovery, P=1.83×10^{-03} replication), and the other is rs41288783 of *APOB* gene (P=5.40×10^{-08} discovery, P=7.92×10^{-07} replication). In the joint model, the genetic main effect per A allele of rs41267813 was associated with a 31.6-mg/dL decrease in LDL-C levels (β_{main} [se_{main}], −31.55 [2.78]), while the estimated interaction effect indicated a positive interaction with regular drinker status (β_{int} [se_{int}], 27.07 [5.66]). In contrast, the genetic main effect of rs41288783 was associated with an increase in LDL-C levels among regular drinkers and nondrinkers (β_{main} [se_{main}], 16.18 [5.34]; β_{int} [se_{int}], 11.03 [7.68]). For the novel interaction between *SMC5* and current drinker on TG levels, we identified the driving variant as rs142488686, a missense mutation (minor allele count, 5–7 discovery [ARIC (Atherosclerosis Risk in Communities) and CARDIA (Coronary Artery Risk Development in Young Adults)]; minor allele count, 7–17 replication [WGHS (Women's Genome Health Study), CHS, and MESA (Multi-Ethnic Study of Atherosclerosis)]), with a replicating interaction effect (P_{int}=0.016 discovery, P_{int}=0.008 replication), while the genetic main effect was modest (P<0.1 discovery and replication, respectively).

DISCUSSION

This is the first large-scale study to evaluate the role of rare and low-frequency variants in lipids by incorporating gene-alcohol consumption interactions. We tested for gene-alcohol interaction effects on lipid levels, as well as the joint effects of genetic main effects and gene-alcohol interactions. We replicated 13 gene-lipid associations at known lipid loci, among which 2 leading rare variants in *APOB* and *LPA* genes associated with LDL-C were novel. Only 1 novel gene-alcohol interaction was identified as significant and replicated at nominal significance level (the interaction between rare and low-frequency variants in *SMC5* and current drinker on TG levels).

Using a single-variant test, we confirmed numerous previously identified rare and low-frequency missense lipid variants. For example, rs11591147 (MAF, ≈1.5%) of *PCSK9* has been associated with LDL-C levels,[24,26] rs77960347 (MAF, ≈1.2%) of *LIPG* and rs116843064 (MAF, ≈2.0%) of *ANGPTL4* have been associated with

Downloaded from http://ahajournals.org by on September 28, 2022

Table 2. Genes Discovered and Replicated by the Joint Test or Interaction-Only Test

Trait	Gene	CHR	Alcohol*	Test	Discovery, No. of Studies	cMAF Range†	Discovery, *P* Value	Replication, No. of Studies	Replication, *P* Value
HDL-C	*LPL*	8	Both	Joint	5	0.036–0.040	8.76×10^{-22}	5	4.25×10^{-21}
	APOC3	11	Both	Joint	3	0.001–0.001	2.82×10^{-6}	2	4.62×10^{-6}
	CD300LG	17	Both	Joint	5	0.031–0.055	2.64×10^{-12}	6	5.94×10^{-10}
	LIPG	18	Both	Joint	5	0.014–0.019	7.65×10^{-17}	5	4.09×10^{-11}
	ANGPTL4	19	Both	Joint	5	0.024–0.031	2.34×10^{-20}	5	5.53×10^{-9}
	HNF4A	20	Both	Joint	5	0.031–0.034	3.37×10^{-10}	5	3.20×10^{-7}
	CELSR2	1	Both	Joint	5	0.079–0.093	1.63×10^{-10}	6	3.21×10^{-8}
	MYBPHL	1	Both	Joint	5	0.044–0.051	7.26×10^{-9}	6	6.49×10^{-6}
	PCSK9	1	Both	Joint	5	0.050–0.055	3.16×10^{-62}	6	9.06×10^{-11}
LDL-C	*APOB*	2	Both	Joint	5	0.174–0.226	5.33×10^{-18}	6	1.20×10^{-15}
	LPA	6	RegDrink	Joint	5	0.096–0.147	2.28×10^{-5}	6	3.7×10^{-4}
	APOH	17	Both	Joint	5	0.074–0.081	1.11×10^{-5}	6	1.18×10^{-5}
	BCAM	19	Both	Joint	5	0.120–0.166	1.49×10^{-18}	6	1.77×10^{-37}
	CBLC	19	Both	Joint	5	0.084–0.104	$7.48\times10^{-2}2$	6	1.64×10^{-35}
	LPL	8	Both	Joint	5	0.036–0.040	8.55×10^{-19}	5	7.30×10^{-16}
	APOA4	11	Both	Joint	5	0.019–0.024	8.83×10^{-9}	6	3.77×10^{-9}
	APOA5	11	Both	Joint	5	0.025–0.033	8.93×10^{-7}	5	2.3×10^{-4}
TG	*APOC3*	11	Both	Joint	3	0.001–0.001	2.09×10^{-10}	3	7.92×10^{-8}
	MAP1A	15	Both	Joint	5	0.129–0.166	1.70×10^{-6}	6	4.30×10^{-5}
	CD300LG	17	Both	Joint	5	0.031–0.055	1.39×10^{-9}	6	5.26×10^{-8}
	ANGPTL4	19	Both	Joint	5	0.024–0.031	1.33×10^{-24}	5	3.56×10^{-15}
	SMC5	9	CurDrink	Interaction	4	0.001–0.002	6.65×10^{-6}	4	0.013‡

Significant threshold for replication was set as $P<0.0017$ for joint test and $P<0.0125$ for interaction test using Bonferroni correction. CHR indicates chromosome; cMAF, cumulative minor allele frequency; CurDrinker, current drinker; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; RegDrinker, regular drinker; and TG, triglycerides.

*Both indicates the gene-lipid pair was identified through using both current and regular drinker as the alcohol consumption variable.

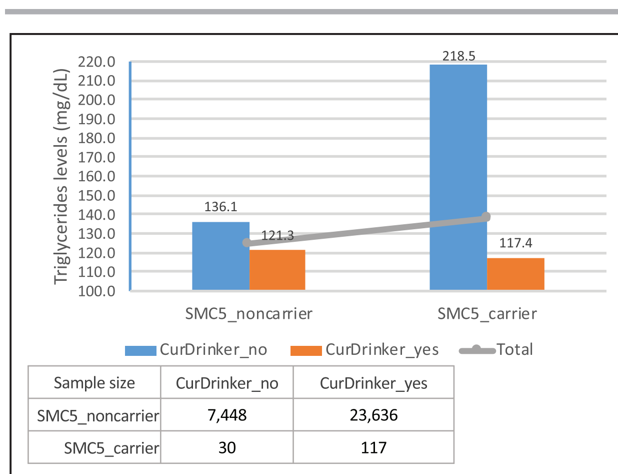
†cMAF for variants aggregated in the genes across studies involved in discovery phase for that gene.

‡*SMC5*, current drinking interaction on TG levels just missed the Bonferroni-corrected threshold of significance (P=0.0125) for replication but reached nominal significance.

Downloaded from http://ahajournals.org by on September 28, 2022

HDL-C levels.[27,28] A missense mutation in the *APOC3* gene, rs147210663(MAF, ≈0.07%), has been associated with a >40% lower average TG level in individuals carrying 1 A allele.[29,30] In the present study, we observed a novel relationship between increased HDL-C levels in individuals carrying rs147210663 (A) allele as rs147210663 was previously reported as a founder mutation in a Pennsylvania Amish population associated with TG.[31]

Between the 2 novel rare driving variants that were identified and replicated, rs41267813 (*LPA*, missense variant; MAF, ≈0.16%) is located 28 kb away from a stop/gain variant rs41267811 (*LPA*; MAF, ≈0.02%) that was also significantly associated with LDL-C levels in the discovery phase. However, we were unable to replicate the association with rs41267811 as it was only available in 1 replication study (WGHS) and, therefore, did not meet our criteria to be included in replication. *LPA* encoded protein constitutes a substantial portion of lipoprotein(a) and associated with inherited conditions including type III hyperlipoproteinemia and familial hyperlipidemia.[32] A stop/gain mutation in this gene would be associated with lower LDL-C levels in carriers, which is true among non-drinkers. However, such association may be modified by alcohol consumption as we observed the carriers of this variant with a higher LDL-C levels compared with noncarriers in a population who had at least 2 drinks per week in the ARIC study. A previous study of gene-alcohol interaction on lipids focusing on common variants identified rs5014650 (MAF, 15%; intergenic), at the *LPA* locus that was associated with LDL-C levels in a joint test,[21] suggesting that this locus affects LDL-C levels through both main effects and an interaction with alcohol consumption. Previous studies have reported a relationship between moderate alcohol consumption and lower Lp(a) lipoprotein concentrations[33,34] yet no published evidence of an association between genetic variation at the *LPA* locus and alcohol consumption. It is possible that that alcohol modifies the *LPA* expression for carriers of rs41267813, changing the Lp(a) lipoprotein concentrations and thereby influencing LDL-C levels. Unfortunately, the LDL-C measurement we used did not distinguish Lp(a) concentrations from the LDL-C levels; further experimental study is warranted to test such hypothesis. However, the observed modification effect should be interpreted with caution as LDL-C levels in ARIC was determined by the Friedewald formula, and this does not distinguish between cholesterol



Sample size	CurDrinker_no	CurDrinker_yes
SMC5_noncarrier	7,448	23,636
SMC5_carrier	30	117

Figure 1. SMC5, current alcohol consumption and average triglyceride levels across 4 discovery studies: the Atherosclerosis Risk in Communities study, the Framingham Heart Study, the Netherlands Epidemiology of Obesity study, and the Women's Health Initiative study.

derived from LDL and lipoprotein(a) and, therefore, represent the sum of cholesterol from both. It is possible that the observed association represents a relationship with lipoprotein(a) levels.

In addition to the variant described above, the other driving rare variant had not been previously associated with a lipid trait; rs41288783 (p.Pro994Leu), is a deleterious variant in *APOB* gene (missense variant; MAF, ≈0.10%). A previous study reported its existence in a patient who was clinically diagnosed as familial hypercholesterolemia without a detectable mutation.[35] Familial hypercholesterolemia is characterized by very high levels of LDL-C, and we observed an association with higher LDL-C levels though jointly testing the effects of rs41288783 and its interaction with alcohol consumption. Nevertheless, the exact biological function of rs41288783 remains unknown. We note that a Mendelian randomization study has suggested a causal role of alcohol consumption in reducing plasma apo B (Apolipoprotein B) and LDL-C levels in a general population.[36] Considering this, alcohol consumption may have contributed to the observed significant joint effect of *APOB* and alcohol consumption on LDL-C levels. It is also worth noting that these 2 novel rare driving variants showed 4 to 16× larger main effect sizes on LDL-C levels as compared with the effect sizes of previously identified common variants (rs1367117 and rs1564348).[3] Such observations supported the hypothesis of rare alleles of large effect[37] and pinpointed the importance of analyzing rare variant G×E interactions.

For the significant gene-alcohol interaction effect we observed on TG levels, the driving variant was identified as rs142488686—a missense mutation in *SMC5* (Structural Maintenance of Chromosomes 5). *SMC5* encodes a core component involved in repair of DNA double-strand breaks and required for telomere maintenance.[38–40] Variants in *SMC5* have been previously reported to be associated with body mass index in a Japanese population[41]

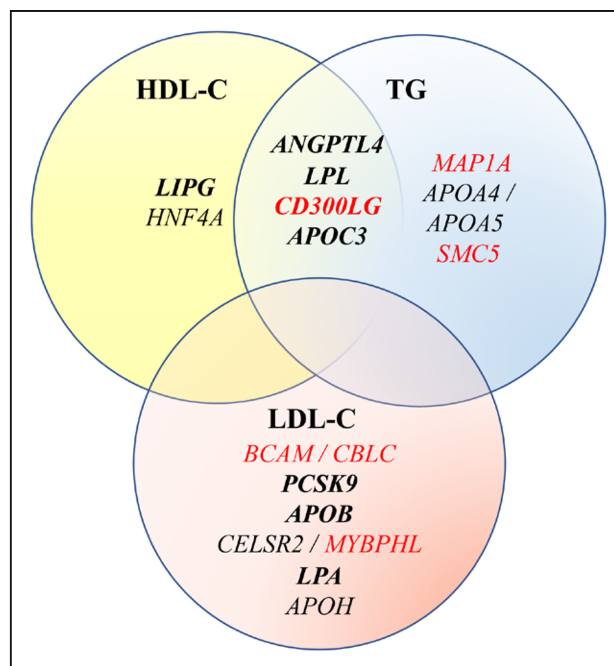


Figure 2. Genes as revealed by gene-environment (G×E) interaction test or jointly testing the gene and G×E interaction effects in association with plasma lipid levels. Bolded genes were genes that remained significant after conditioning on common index single-nucleotide polymorphisms. Genes in red were not previously reported to be associated with ≥1 lipid trait, but they are in known lipid loci. HDL-C indicates high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; and TG, triglycerides.

but not with lipid levels or alcohol consumption, and it is unknown whether the interaction we observed between *SMC5* locus and current drinking behavior on TG levels has a biological aspect. As the gene level interaction test results just missed Bonferroni-corrected significance level for replication, further studies are warranted to validate such findings.

A limitation of this study is the imbalance in percentage of alcohol consumers between discovery (on average 48.7% regular drinkers and 78.5% current drinkers) and replication studies (on average 29.8% regular drinkers and 57.2% current drinkers), which may have impacted our ability to identify and replicate additional loci beyond what is reported here. All participating studies used similar questionnaires (either interviewer-administered or self-reported; Table I in the Data Supplement) to collect alcohol consumption information. The variation in percentage of current and regular drinkers may represent the heterogeneity of drinking behaviors across populations and, therefore, may contribute to the generalizability of our results. Additionally, as self-reported alcohol consumption was used and may very likely be underreported, this study may suffer from loss of statistical power due to potential misclassification.[42] Similarly, dichotomizing alcohol consumption into regular drinkers and current drinkers may also reduce power as compared with treating it as a continuous variable.[43] It is possible that a more

Downloaded from http://ahajournals.org by on September 28, 2022

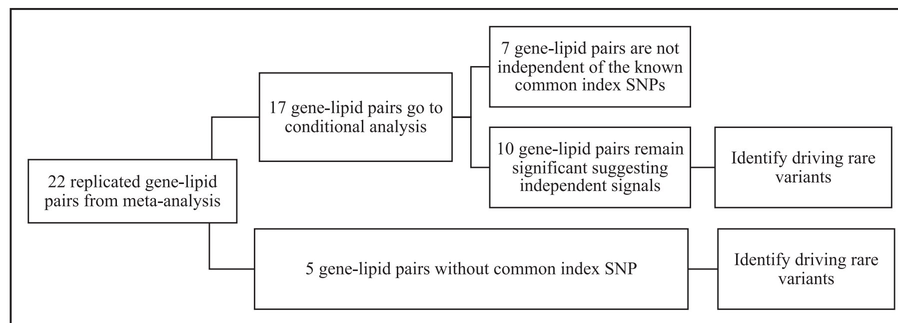


Figure 3. Flowchart of follow-up analyses, including conditional analysis and single-variant test to identify driving rare variants. For conditional analysis, significant results were defined as $P<5\times10^{-5}$ in meta-analysis of discovery studies and $P<0.05/10$ (Bonferroni correction for 10 gene-lipid pairs with $P<5\times10^{-5}$ in discovery phase) in meta-analysis of replication studies. For single-variant test to identify driving rare variants, we included variants with minor allele count at least 5 and present in at least 2 studies. Bonferroni correction for number of single-nucleotide polymorphisms (SNPs) tested in discovery phase and number of SNPs taken forward to replication was applied separately for joint test and interaction test for each lipid trait.

comprehensive characterization of alcohol consumption could reveal associations that were missed in the present study. In addition, although the sample size of 66 428 may provide sufficient power for a traditional GWAS, on the identification of rare variants and G×E interactions, may require even larger sample sizes or bigger effect sizes.[17,44]

In conclusion, this study applied emerging statistical approaches to investigate the role of rare and low-frequency variants in gene-alcohol consumption interaction effects on lipid levels, and identified 2 novel rare variants at known lipid loci for LDL-C levels, with larger effect sizes than those of the previously known common variants, and suggested 1 novel locus for gene-alcohol interaction on TG levels. Our results show promise for other larger scale studies analyzing rare variant G×E interactions to refine association signals at previously identified loci to reveal novel biology.

ARTICLE INFORMATION

Received September 20, 2019; accepted April 9, 2020.

Affiliations

Department of Epidemiology, Human Genetics Center, Human Genetics and Environmental Sciences, School of Public Health (Z.W., H.C., P.S.d.V., M.F., E.B., A.C.M.), Center for Precision Health, School of Public Health and School of Biomedical Informatics (H.C.), Brown Foundation Institute of Molecular Medicine (M.A.R., M.F.), University of Texas Health Science Center at Houston, Houston, TX. Cardiovascular Health Research Unit, Department of Biostatistics and Medicine (T.M.B.), Cardiovascular Health Research Unit, Department of Medicine (R.N.L., C.M.S.), and Cardiovascular Health Research Unit, Department of Medicine, Epidemiology and Health Services (B.M.P.), University of Washington, Seattle, WA. Department of Epidemiology, School of Public Health (L.F.B., J.A.S., W.Z., S.L.R.K.), Institute for Social Research, Survey Research Center (J.A.S.), University of Michigan, Ann Arbor, MI. Division of Preventive Medicine (D.I.C., F.G.), Division of Sleep and Circadian Disorders, Department of Medicine (H.W., B.C., J.L., S.R., P.M.R.), Brigham and Women's Hospital, Boston, MA. Harvard Medical School, Boston, MA (D.I.C., H.W., B.C., J.L., S.R., P.M.R.). Division of Statistical Genomics, Department of Genetics, Washington University School of Medicine, St. Louis, MO (M.F.F., M.A.P.). Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina, Chapel Hill (N.F.). Department of Pediatrics, Institute for Translational Genomics and Population Sciences, The Lundquist Institute at Harbor–UCLA Medical Center, Torrance, CA (X.G., J.Y., J.I.R.). Biostatistics Department, Boston University School of Public Health, MA (E.L., L.A.C., C.-T.L.). Section of Gerontology and Geriatrics, Department of Internal Medicine (R.N.), Department of Human Genetics (K.W.v.D.), Division of Endocrinology, Department of Internal Medicine (K.W.v.D.), Department of Clinical Epidemiology (D.O.M.-K.), and Department of Public Health and Primary Care (D.O.M.-K.), Leiden University Medical Center, the Netherlands. NHLBI Framingham Heart Study, MA (L.A.C.). George Washington University School of Medicine and Health Sciences, DC (L.W.M.). Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA (A.P.R., C.K.). Department of Public Health Sciences, Center for Public Health Genomics, University of Virginia, Charlottesville (S.S.R.). Epidemiology and Community Health, School of Public Health, University of Minnesota, Minneapolis (P.J.S.). Division of Research, Kaiser Permanente Northern California, Oakland (S.S.). Kaiser Permanente Washington Health Research Institute, Seattle, WA (B.M.P.). Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX (E.B.).

Sources of Funding

The various Gene-Lifestyle Interaction projects, including this one, are largely supported by a grant from the US National Heart, Lung, and Blood Institute, the National Institutes of Health, R01HL118305. Dr de Vries is supported by American Heart Association grant 18CDA34110116. Dr Franceschini is supported by the National Institutes of Health grants R01-DK117445, R01-MD012765, and R21-HL140385. Full set of study-specific funding sources appears in the Data Supplement.

Disclosures

Dr Mook-Kanamori is a part-time research consultant with Metabolon, Inc, and Dr Psaty serves on the Steering Committee of the Yale Open Access Project funded by Johnson & Johnson. The other authors report no conflicts.

REFERENCES

1. Scandinavian Simvastatin Survival Study Group. Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: the scandinavian simvastatin survival study (4s). *Lancet*. 1994;344:1383–1389.
2. American College of Physicians. Clinical guideline, part 1: guidelines for using serum cholesterol, high-density lipoprotein cholesterol, and triglyceride levels as screening tests for preventing coronary heart disease in adults. *Ann Intern Med*. 1996;124:515–517.
3. Teslovich TM, Musunuru K, Smith AV, Edmondson AC, Stylianou IM, Koseki M, Pirruccello JP, Ripatti S, Chasman DI, Willer CJ, et al. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature*. 2010;466:707–713. doi: 10.1038/nature09270
4. Asselbergs FW, Guo Y, van Iperen EP, Sivapalaratnam S, Tragante V, Lanktree MB, Lange LA, Almoguera B, Appelman YE, Barnard J, et al; LifeLines Cohort Study. Large-scale gene-centric meta-analysis across 32 studies identifies multiple lipid loci. *Am J Hum Genet*. 2012;91:823–838. doi: 10.1016/j.ajhg.2012.08.032
5. Willer CJ, Schmidt EM, Sengupta S, Peloso GM, Gustafsson S, Kanoni S, Ganna A, Chen J, Buchkovich ML, Mora S, et al; Global Lipids Genetics Consortium. Discovery and refinement of loci associated with lipid levels. *Nat Genet*. 2013;45:1274–1283. doi: 10.1038/ng.2797

Downloaded from http://ahajournals.org by on September 28, 2022

6. Peloso GM, Auer PL, Bis JC, Voorman A, Morrison AC, Stitziel NO, Brody JA, Khetarpal SA, Crosby JR, Fornage M, et al; NHLBI GO Exome Sequencing Project. Association of low-frequency and rare coding-sequence variants with blood lipids and coronary heart disease in 56,000 whites and blacks. *Am J Hum Genet.* 2014;94:223–232. doi: 10.1016/j.ajhg.2014.01.009
7. Surakka I, Horikoshi M, Mägi R, Sarin AP, Mahajan A, Lagou V, Marullo L, Ferreira T, Miraglio B, Timonen S, et al; ENGAGE Consortium. The impact of low-frequency and rare variants on lipid levels. *Nat Genet.* 2015;47:589–597. doi: 10.1038/ng.3300
8. Tang CS, Zhang H, Cheung CY, Xu M, Ho JC, Zhou W, Cherny SS, Zhang Y, Holmen O, Au KW, et al. Exome-wide association analysis reveals novel coding sequence variants associated with lipid traits in Chinese. *Nat Commun.* 2015;6:10206. doi: 10.1038/ncomms10206
9. Muth ND, Laughlin GA, von Mühlen D, Smith SC, Barrett-Connor E. High-density lipoprotein subclasses are a potential intermediary between alcohol intake and reduced risk of cardiovascular disease: the Rancho Bernardo Study. *Br J Nutr.* 2010;104:1034–1042. doi: 10.1017/S0007114510001595
10. Gardner CD, Tribble DL, Young DR, Ahn D, Fortmann SP. Associations of HDL, HDL(2), and HDL(3) cholesterol and apolipoproteins A-I and B with lifestyle factors in healthy women and men: the Stanford Five City Project. *Prev Med.* 2000;31:346–356. doi: 10.1006/pmed.2000.0716
11. Wakabayashi I. Relationship between alcohol intake and lipid accumulation product in middle-aged men. *Alcohol Alcohol.* 2013;48:535–542. doi: 10.1093/alcalc/agt032
12. Whitfield JB, Heath AC, Madden PA, Pergadia ML, Montgomery GW, Martin NG. Metabolic and biochemical effects of low-to-moderate alcohol consumption. *Alcohol Clin Exp Res.* 2013;37:575–586. doi: 10.1111/acer.12015
13. Rakic V, Puddey IB, Dimmitt SB, Burke V, Beilin LJ. A controlled trial of the effects of pattern of alcohol intake on serum lipid levels in regular drinkers. *Atherosclerosis.* 1998;137:243–252. doi: 10.1016/s0021-9150(97)00269-4
14. Brien SE, Ronksley PE, Turner BJ, Mukamal KJ, Ghali WA. Effect of alcohol consumption on biological markers associated with risk of coronary heart disease: systematic review and meta-analysis of interventional studies. *BMJ.* 2011;342:d636. doi: 10.1136/bmj.d636
15. Rimm EB, Williams P, Fosher K, Criqui M, Stampfer MJ. Moderate alcohol intake and lower risk of coronary heart disease: meta-analysis of effects on lipids and haemostatic factors. *BMJ.* 1999;319:1523–1528. doi: 10.1136/bmj.319.7224.1523
16. Thomas D. Gene–environment-wide association studies: emerging approaches. *Nat Rev Genet.* 2010;11:259–272. doi: 10.1038/nrg2764
17. Hunter DJ. Gene-environment interactions in human diseases. *Nat Rev Genet.* 2005;6:287–298. doi: 10.1038/nrg1578
18. Rao DC, Sung YJ, Winkler TW, Schwander K, Borecki I, Cupples LA, Gauderman WJ, Rice K, Munroe PB, Psaty BM, et al. Multiancestry study of gene-lifestyle interactions for cardiovascular traits in 610 475 individuals from 124 cohorts: design and rationale. *Circ Cardiovasc Genet.* 2017;10:e001649. doi: 10.1161/CIRCGENETICS.116.001649
19. Feitosa MF, Kraja AT, Chasman DI, Sung YJ, Winkler TW, Ntalla I, Guo X, Franceschini N, Cheng CY, Sim X, et al; InterAct Consortium. Novel genetic associations for blood pressure identified via gene-alcohol interaction in up to 570K individuals across multiple ancestries. *PLoS One.* 2018;13:e0198166. doi: 10.1371/journal.pone.0198166
20. Kilpeläinen TO, Bentley AR, Noordam R, Sung YJ, Schwander K, Winkler TW, Jakupović H, Chasman DI, Manning A, Ntalla I, et al; Lifelines Cohort Study. Multi-ancestry study of blood lipid levels identifies four loci interacting with physical activity. *Nat Commun.* 2019;10:376. doi: 10.1038/s41467-018-08008-w
21. de Vries PS, Brown MR, Bentley AR, Sung YJ, Winkler TW, Ntalla I, Schwander K, Kraja AT, Guo X, Franceschini N, et al; InterAct Consortium; Lifelines Cohort, Groningen, The Netherlands (Lifelines Cohort Study). Multiancestry genome-wide association study of lipid levels incorporating gene-alcohol interactions. *Am J Epidemiol.* 2019;188:1033–1054. doi: 10.1093/aje/kwz005
22. Sung YJ, Winkler TW, de Las Fuentes L, Bentley AR, Brown MR, Kraja AT, Schwander K, Ntalla I, Guo X, Franceschini N, et al; CHARGE Neurology Working Group; COGENT-Kidney Consortium; GIANT Consortium; Lifelines Cohort Study. A large-scale multi-ancestry genome-wide study accounting for smoking behavior identifies multiple significant loci for blood pressure. *Am J Hum Genet.* 2018;102:375–400. doi: 10.1016/j.ajhg.2018.01.015
23. Kraft P, Yen YC, Stram DO, Morrison J, Gauderman WJ. Exploiting gene-environment interaction to detect genetic associations. *Hum Hered.* 2007;63:111–119. doi: 10.1159/000099183
24. Manning AK, LaValley M, Liu CT, Rice K, An P, Liu Y, Miljkovic I, Rasmussen-Torvik L, Harris TB, Province MA, et al. Meta-analysis of gene-environment interaction: joint estimation of SNP and SNP × environment regression coefficients. *Genet Epidemiol.* 2011;35:11–18. doi: 10.1002/gepi.20546
25. Chen H, Meigs JB, Dupuis J. Incorporating gene-environment interaction in testing for association with rare genetic variants. *Hum Hered.* 2014;78:81–90. doi: 10.1159/000363347
26. Cohen JC, Boerwinkle E, Mosley TH Jr, Hobbs HH. Sequence variations in PCSK9, low LDL, and protection against coronary heart disease. *N Engl J Med.* 2006;354:1264–1272. doi: 10.1056/NEJMoa054013
27. Romeo S, Pennacchio LA, Fu Y, Boerwinkle E, Tybjaerg-Hansen A, Hobbs HH, Cohen JC. Population-based resequencing of ANGPTL4 uncovers variations that reduce triglycerides and increase HDL. *Nat Genet.* 2007;39:513–516. doi: 10.1038/ng1984
28. Kanoni S, Masca NG, Stirrups KE, Varga TV, Warren HR, Scott RA, Southam L, Zhang W, Yaghootkar H, Müller-Nurasyid M, et al; Wellcome Trust Case Control Consortium. Analysis with the exome array identifies multiple new independent variants in lipid loci. *Hum Mol Genet.* 2016;25:4094–4106. doi: 10.1093/hmg/ddw227
29. Jørgensen AB, Frikke-Schmidt R, Nordestgaard BG, Tybjærg-Hansen A. Loss-of-function mutations in APOC3 and risk of ischemic vascular disease. *N Engl J Med.* 2014;371:32–41. doi: 10.1056/NEJMoa1308027
30. Crosby J, Peloso GM, Auer PL, Crosslin DR, Stitziel NO, Lange LA, Lu Y, Tang ZZ, Zhang H, Hindy G, et al. Loss-of-function mutations in apoc3, triglycerides, and coronary disease. *N Engl J Med.* 2014;371:22–31.
31. Crawford DC, Dumitrescu L, Goodloe R, Brown-Gentry K, Boston J, McClellan B Jr, Sutcliffe C, Wiseman R, Baker P, Pericak-Vance MA, et al. Rare variant APOC3 R19X is associated with cardio-protective profiles in a diverse population-based survey as part of the Epidemiologic Architecture for Genes Linked to Environment Study. *Circ Cardiovasc Genet.* 2014;7:848–853. doi: 10.1161/CIRCGENETICS.113.000369
32. Langsted A, Kamstrup PR, Benn M, Tybjærg-Hansen A, Nordestgaard BG. High lipoprotein(a) as a possible cause of clinical familial hypercholesterolaemia: a prospective cohort study. *Lancet Diabetes Endocrinol.* 2016;4:577–587. doi: 10.1016/S2213-8587(16)30042-0
33. Sharpe PC, Young IS, Evans AE. Effect of moderate alcohol consumption on lp(a) lipoprotein concentrations. Reduction is supported by other studies. *BMJ.* 1998;316:1675.
34. Paassilta M, Kervinen K, Rantala AO, Savolainen MJ, Lilja M, Reunanen A, Kesäniemi YA. Social alcohol consumption and low Lp(a) lipoprotein concentrations in middle aged Finnish men: population based study. *BMJ.* 1998;316:594–595. doi: 10.1136/bmj.316.7131.594
35. Alves AC, Etxebarria A, Soutar AK, Martin C, Bourbon M. Novel functional APOB mutations outside LDL-binding region causing familial hypercholesterolaemia. *Hum Mol Genet.* 2014;23:1817–1828. doi: 10.1093/hmg/ddt573
36. Vu KN, Ballantyne CM, Hoogeveen RC, Nambi V, Volcik KA, Boerwinkle E, Morrison AC. Causal role of alcohol consumption in an improved lipid profile: the Atherosclerosis Risk in Communities (ARIC) study. *PLoS One.* 2016;11:e0148765. doi: 10.1371/journal.pone.0148765
37. Gorlov IP, Gorlova OY, Sunyaev SR, Spitz MR, Amos CI. Shifting paradigm of association studies: value of rare single-nucleotide polymorphisms. *Am J Hum Genet.* 2008;82:100–112. doi: 10.1016/j.ajhg.2007.09.006
38. Potts PR, Yu H. The SMC5/6 complex maintains telomere length in ALT cancer cells through SUMOylation of telomere-binding proteins. *Nat Struct Mol Biol.* 2007;14:581–590. doi: 10.1038/nsmb1259
39. Behlke-Steinert S, Touat-Todeschini L, Skoufias DA, Margolis RL. SMC5 and MMS21 are required for chromosome cohesion and mitotic progression. *Cell Cycle.* 2009;8:2211–2218. doi: 10.4161/cc.8.14.8979
40. Jeppsson K, Kanno T, Shirahige K, Sjögren C. The maintenance of chromosome structure: positioning and functioning of SMC complexes. *Nat Rev Mol Cell Biol.* 2014;15:601–614. doi: 10.1038/nrm3857
41. Okada Y, Kubo M, Ohmiya H, Takahashi A, Kumasaka N, Hosono N, Maeda S, Wen W, Dorajoo R, Go MJ, et al; GIANT Consortium. Common variants at CDKAL1 and KLF9 are associated with body mass index in East Asian populations. *Nat Genet.* 2012;44:302–306. doi: 10.1038/ng.1086
42. Livingston M, Callinan S. Underreporting in alcohol surveys: whose drinking is underestimated? *J Stud Alcohol Drugs.* 2015;76:158–164.
43. Altman DG, Royston P. The cost of dichotomising continuous variables. *BMJ.* 2006;332:1080. doi: 10.1136/bmj.332.7549.1080
44. VanderWeele TJ. Sample size and power calculations for additive interactions. *Epidemiol Methods.* 2012;1:159–188. doi: 10.1515/2161-962X.1010

Downloaded from http://ahajournals.org by on September 28, 2022