DOI: 10.1161/CIRCGEN.120.003288

Sugar-sweetened Beverage Consumption May Modify Associations between Genetic Variants in the CHREBP Locus and HDL-C and TG Concentrations

Running title: *Haslam et al.; SSB-CHREBP interactions on lipid traits*

Danielle E. Haslam, PhD[1-3]; Gina M. Peloso, PhD[4]; Melanie Guirette, MS[1]; Fumiaki Imamura, PhD[5]; Traci M. Bartz, MS[6,17]; Achilleas N. Pitsillides, PhD[4]; Carol A. Wang, PhD[7]; Ruifang Li-Gao, PhD[8]; Jason M. Westra, MS[10]; Niina Pitkänen, PhD[11,12]; Kristin L. Young, PhD[13]; Mariaelisa Graff, PhD[13,15]; Alexis C. Wood, PhD[14]; Kim V.E. Braun, PhD[15]; Jian'an Luan, PhD[5]; Mika Kähönen, PhD[16]; Jessica C. Kiefte-de Jong, RD, PhD[9,15]; Mohsen Ghanbari, PhD[13,15]; Nathan Tintle, PhD[10]; Rozenn N. Lemaitre, PhD[17]; Dennis O. Mook-Kanamori, PhD[8,9]; Kari North, PhD[13,18]; Mika Helminen, PhD[19]; Yasmin Mossavar-Rahmani, PhD[20]; Linda Snetselaar, RD, PhD[21]; Lisa W. Martin, MD[22]; Jorma S. Viikari, MD, PhD[23]; Wendy H. Oddy, PhD[24]; Craig E. Pennell, PhD[7]; Frits R. Rosendall, MD, PhD[8]; M Arfan Ikram, MD, PhD[15]; Andre G Uitterlinden, PhD[25]; Bruce M. Psaty, MD, PhD[6,17,26]; Dariush Mozaffarian, MD, DrPH[27]; Jerome I. Rotter, MD[28]; Kent D. Taylor, PhD[28]; Terho Lehtimäki, PhD[29]; Olli T. Raitakari, MD, PhD[12,30]; Kara A. Livingston, MPH[1]; Trudy Voortman, PhD[15]; Nita G. Forouhi, PhD[5]; Nick J. Wareham, MB, PhD[5]; Renée de Mutsert, PhD[8]; Steven S. Rich, PhD[31]; JoAnn E. Manson, DrPH, MD[2,32,33]; Samia Mora, MD[33,34]; Paul M Ridker, MD[34]; Jordi Merino, PhD[35-38]; James B. Meigs, MD[35,36,39]; Hassan S. Dashti, PhD, RD[35,38,40]; Daniel I. Chasman, PhD[33]; Alice H. Lichtenstein, DSc[41]; Caren E. Smith, MS, DVM[42]; Josée Dupuis, PhD[4]; Mark A Herman, MD[43]; Nicola M. McKeown, PhD[1]

[1]Nutritional Epidemiology Program, [41]Cardiovascular Nutrition Laboratory, [42]Nutrition & Genomics Laboratory, Jean Mayer U.S. Dept of Agriculture Human Nutrition Rsrch Ctr on Aging, [27]Friedman School of Nutrition Science & Policy, Tufts Univ; [2]Channing Division of Network Med, [33]Division of Preventive Med, [34]Cardiovascular Division of Med & Ctr for Lipid Metabolomics, Brigham and Women's Hospital & Harvard Medical School; [3]Dept of Nutrition, [32]Dept of Epidemiology, Harvard T.H. Chan School of Public Hlth; [4]Dept of Biostatistics, Boston Univ School of Public Hlth, Boston, MA; [5]Medical Rsrch Council Epidemiology Unit, Univ of Cambridge, Cambridge, UK; [6]Cardiovascular Hlth Research Unit, Depts of Biostatistics, [17]Dept of Med, Univ of Washington, Seattle, WA; [7]School of Med & Public Hlth, Faculty of Med & Hlth, The Univ of Newcastle, NSW, Australia; [8]Dept of Clinical Epidemiology, Leiden Univ Medical Ctr; [9]Dept of Public Hlth & Primary Care, Medical Ctr, Leiden Univ, Leiden, the Netherlands; [10]Dordt Univ, Sioux Ctr, IA; [11]Auria Biobank, [12]Rsrch Ctr of Applied & Preventive Cardiovascular Med, [23]Division of Med, Turku Univ Hospital & Dept of Medicine, [30]Dept of Clinical Physiology & Nuclear Med, Turku Univ Hospital & Ctr for Population Health Rsrch, Univ of Turku, Turku, Finland; [13]Dept of Epidemiology, Gillings School of Global Public Hlth, [18]Carolina Ctr for Genome Sciences, Univ of North Carolina; Chapel Hill, NC; [14]USDA/ARS Children's Nutrition Rsrch Ctr, Dept of Pediatrics, Baylor College of Med, Houston, TX; [15]Dept of Epidemiology, [25]Dept of Internal Med, Erasmus MC Univ Medical Ctr Rotterdam, the Netherlands; [16]Dept of Clinical Physiology, Erasmus MC Univ Medical Ctr Rotterdam, the Netherlands & Dept of Clinical Physiology, Tampere Univ & Dept of Clinical Physiology, Finnish Cardiovascular Rsrch Ctr - Tampere, Faculty of Med & Hlth Technology; [19]Research, Development & Innovation Ctr, Tampere Univ Hospital & Faculty of Social Sciences, Hlth Sciences, Tampere Univ, Tampere, Finland; [20]Dept of Epidemiology & Population Hlth, Albert Einstein College of Med, Bronx, NY; [21]Dept of Epidemiology, Univ of Iowa, Iowa City, IA; [22]George Washington Univ School of Med & Hlth Sciences, Washington, DC; [24]Menzies Inst for Medical Rsrch, Univ of Tasmania, HOB, Australia; [26]Dept of Health Services, Univ of Washington & Kaiser Permanente Washington Hlth Rsrch Inst, Seattle, WA; [28]The Inst for Translational Genomics & Population Sciences, Dept of Pediatrics, The Lundquist Inst for Biomedical Innovation at Harbor-UCLA Medical Ctr, Torrance, CA; [29]Dept of Clinical Chemistry Tampere Univ & Dept of Clinical Chemistry, Fimlab Laboratories, Tampere, Finland; [31]Ctr for Public Hlth Genomics & Dept of Public Hlth Sciences, Univ of Virginia School of Med, Charlottesville, VA; [35]Programs in Medical & Population Genetics, Broad Inst of MIT & Harvard, Cambridge; [36]Program in Metabolism, Broad Inst of MIT & Harvard, Cambridge & Dept of Medicine, Harvard Medical School, Boston, MA; [37]Institut d'Investigació Sanitària Pere Virgili, Universitat Rovira i Virgili, Reus, Spain; [38]Diabetes Unit & Ctr for Genomic Med, [39]Division of General Internal Med, [40]Dept of Anesthesia, Critical Care & Pain Med, Mass Gen Hospital & Harvard Medical School, Boston, MA; [43]Division Of Endocrinology, Metabolism & Nutrition, Dept of Med & Duke Molecular Physiology Inst, Duke Univ School of Med, Durham, NC

Correspondence:
Nicola McKeown, PhD
Nutrition Epidemiology Department
Jean Mayer U.S. Department of Agriculture
Human Nutrition Research Center on Aging
Tufts University
711 Washington Street
Boston MA 02111
Tel: (617) 556-3008
Email: nicola.mckeown@tufts.edu

Journal Subject Terms: Diet and Nutrition; Epidemiology; Genetic, Association Studies; Lipids and Cholesterol

DOI: 10.1161/CIRCGEN.120.003288

Abstract:

Background - Carbohydrate responsive element binding protein (ChREBP) is a transcription factor that responds to sugar consumption. Sugar-sweetened beverage (SSB) consumption and genetic variants in the *CHREBP* locus have separately been linked to high-density lipoprotein cholesterol (HDL-C) and triglyceride (TG) concentrations. We hypothesized SSB consumption would modify the association between genetic variants in the *CHREBP* locus and dyslipidemia.

Methods - Data from 11 cohorts from the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium (N=63,599) and the UK Biobank (UKB) (N=59,220) were used to quantify associations of SSB consumption, genetic variants, and their interaction on HDL-C and TG concentrations using linear regression models. A total of 1,606 single-nucleotide polymorphisms (SNPs) within or near *CHREBP* were considered. SSB consumption was estimated from validated questionnaires and participants were grouped by their estimated intake.

Results - In a meta-analysis, rs71556729 was significantly associated with higher HDL-C concentrations only among the highest SSB consumers [β (95% CI) = 2.12 (1.16, 3.07) mg/dl; $p<0.0002$], but not significantly among the lowest SSB consumers ($p=0.81$; $p_{Diff}<0.0001$). Similar results were observed for two additional variants (rs35709627 and rs71556736). For TG, rs55673514 was positively associated with TG concentrations only among the highest SSB consumers [β (95% CI): 0.06 (0.02, 0.09) per allele count for log(mg/dl), $p=0.001$], but not the lowest SSB consumers ($p=0.84$; $p_{Diff}=0.0005$).

Conclusions - Our results identified genetic variants in the *CHREBP* locus that may protect against SSB-associated reductions in HDL-C and other variants that may exacerbate SSB-associated increases in TG concentrations.

Clinical Trial Registration - Some participating cohorts were registered at URL: https://www.clinicaltrials.gov/ with unique identifiers: NCT00005131 (Atherosclerosis Risk in Communities), NCT00005133 (Cardiovascular Health Study), NCT00005121 (Framingham Offspring Study), NCT00005487 (Multi-Ethnic Study of Atherosclerosis), and NCT00000479 (Women's Health Study: parent study of the Women's Genome Health Study).

Key words: genetics, association studies; lipids and lipoproteins; epidemiology; nutrigenomics genetics; nutrition; sugar-sweetened beverages

DOI: 10.1161/CIRCGEN.120.003288

Nonstandard Abbreviations and Acronyms

HDL-C: high-density lipoprotein cholesterol

TG: triglyceride

T2D: type 2 diabetes

CVD: cardiovascular disease

GWAS: genome-wide association studies

ChREBP: Carbohydrate Responsive Element Binding Protein

SSB: sugar-sweetened beverages

SNPs: single nucleotide polymorphisms

CHARGE: Cohorts for Heart and Aging Research in Genetic Epidemiology

UKB: UK Biobank

Introduction

Low circulating high-density lipoprotein cholesterol (HDL-C) and elevated fasting triglyceride (TG) concentrations are positively associated with risk of type 2 diabetes (T2D) and cardiovascular disease (CVD).[1–5] Both genetic and environmental factors, including diet, are important determinants of HDL-C and TG concentrations.[5–7] Genetic determinants of HDL-C and TG concentrations have been identified in genome-wide association studies (GWAS),[8–12] but the extent to which genetic variants interact with environmental exposures is unknown. It is plausible that unrecognized genetic variants or genetic effects may be suppressed or exacerbated by environmental factors, such as diet.

Carbohydrate Responsive Element Binding Protein (ChREBP) is a transcription factor that regulates glucose and lipid metabolism in response to sugar consumption, including sugar from sugar sweetened beverages (SSB).[13,14] GWAS have consistently observed an association

DOI: 10.1161/CIRCGEN.120.003288

between single nucleotide polymorphisms (SNPs) in the *CHREBP* locus (also known as *MLXIPL*), and HDL-C and TG concentrations.[8,9,15,16] In animal studies, hepatic ChREBP is robustly activated by dietary fructose, a major constituent of SSB, and potentiates hepatic lipogenesis and TG secretion.[14,17–20] These findings are consistent with large population-based studies in which high SSB consumption has been associated with elevated fasting plasma TG and reduced HDL-C concentrations,[21–24] and increased T2D[25–27] and CVD[21] risk. Thus, SNPs within the *CHREBP* locus present promising candidates for gene-SSB interactions on circulating HDL-C and TG concentrations.

These pieces of biological, epidemiological and genetic evidence suggest that SSB consumption may modify how genetic variants within the *CHREBP* locus influence plasma lipid concentrations in some individuals. Although reduction of SSB consumption is increasingly being encouraged globally,[28] public health efforts to reduce SSB consumption have achieved limited success and SSB consumption remains a modifiable dietary exposure that contributes substantially to the burden of T2D and CVD worldwide.[29,30] A better understanding of the mechanisms underlying the SSB-ChREBP-lipid relationship may reveal novel mechanisms that contribute to the pathogenesis of T2D and CVD risk. Understanding these mechanisms may provide alternative strategies and approaches to reduce metabolic disease that may complement or facilitate dietary interventions.

The present study aimed to examine whether SSB consumption may modify the association of genetic variants within the *CHREBP* locus on HDL-C and TG concentrations in aggregated data from cohorts who are part of the Cohorts for Heart and Aging Research in Genetic Epidemiology (CHARGE) consortium.[31] Descriptions of the CHARGE cohorts are

DOI: 10.1161/CIRCGEN.120.003288

included in the supplemental material, Table I. We further used data from the UK Biobank (UKB) to assess the reproducibility of the finding in an independent cohort.[32]

Methods

Methods are available in the Supplemental Material. The data that support the findings of this study are available from the corresponding author upon reasonable request. All study participants provided written informed consent, and approval for all study protocols was granted by local institutional review boards and/or oversight committees.

Results

General characteristics and mean dietary intakes for the eleven CHARGE cohorts are shown in Table 1. Replication of previous findings on associations of SSB consumption and SNPs with lipid traits in the CHARGE cohorts are presented in the Supplemental Results in the Supplemental Material.

Difference Test Interactions between SSB Consumption and SNPs on HDL-C and TG in CHARGE Cohorts

We identified 55 SNPs that displayed a significant (p_{Diff} <0.0001) or suggestive (p_{Diff} <0.005) difference in estimated effect by category of SSB consumption on HDL-C concentrations in either of the two covariate models in the meta-analysis of the CHARGE cohorts. Among these 55 top SNPs, four were identified as distinct signals for HDL-C concentrations were observed when applying the difference test interaction. Two distinct SNPs in moderate LD with one another [rs35709627 and rs71556729; R^2 = 0.55 (Figure II in the Supplemental Material)] and in low LD with the top SNP identified in the overall analysis for HDL-C concentrations (R^2<0.3)

displayed a statistically significant difference in effect by category of SSB intake on HDL-C concentrations in fully adjusted models (Model 2; p_{Diff}<0.0001) (Table 2 and Figures III and IV in the Supplemental Material). In model 2, each additional minor allele at rs35709627 [β (SE): 2.72 (0.72), *p*=0.0002] and rs71556729 [β (SE): 3.89 (1.04), *p*=0.0002] was associated with higher mean concentrations of HDL-C concentrations among the highest SSB consumers (> 1 serving/day), but was not associated with mean HDL-C concentrations among the lowest SSB consumers (<1 serving/month; p>0.05). The effect sizes of these SNPs among the highest SSB consumers were consistent across all the cohorts. There was no heterogeneity (I^2 = 0%) observed the top four distinct signals (statistically significant and suggestive) among the highest SSB consumers (>1 serving/day), which could be due to low power to detect heterogeneity given the smaller sample size available among the highest SSB consumers (maximum n=4,033).

No statistically significant differences in effect by category of SSB intake on TG concentrations were observed when applying the difference test (p_{Diff} >0.0001 for all SNPs). One SNP (rs799157) in moderate LD with a top SNP identified in the overall analysis for TG concentrations (Table X in the Supplemental Material; R^2 with rs42124=0.44) displayed a suggestive difference in effect by category of SSB intake on TG concentrations in minimally adjusted models (Model 1; p_{Diff} =0.005) (Table 2). Each additional minor allele at rs799157 was associated with higher mean TG concentrations among the highest SSB consumers (> 1 serving/day) [β (SE): 0.11 (0.03) ln-mg/dl, *p*=0.002], but this association was attenuated among the lowest SSB consumers [β (SE): 0.01 (0.01) ln-mg/dl, *p*=0.11] (Figure V in the Supplemental Material). The direction of the effect size of this SNP among the highest SSB consumers was consistent across all the cohorts in which these SNPs were available, and heterogeneity was low among the highest SSB consumers (I^2 = 0%).

DOI: 10.1161/CIRCGEN.120.003288

Cross-Product Interactions between SSB Consumption and SNPs on HDL-C and TG in CHARGE Cohorts

No statistically significant cross-product interactions between SNPs and SSB consumption on HDL-C or TG concentrations were observed ($p_{interaction}$>0.0001), while some tests were suggestive ($p_{interaction}$<0.005) (Table 2). Three SNPs displayed a suggestive interaction with SSB consumption on HDL-C concentrations in either covariate model, and the clumping identified two distinct signals (rs71556729 and rs79578725). One SNP (rs55673514) displayed a suggestive interaction with SSB on TG concentrations in Model 2. Forest plots for top distinct signals in SSBxSNP interaction analyses on lipid traits are presented in Figures VI and VII in the Supplemental Material.

Interactions between SSB Consumption and SNPs on Lipid Traits in the UKB and Meta-Analysis with CHARGE Cohort Results

General characteristics and mean dietary intakes for the 59,220 UKB participants are shown in Table VI in the Supplemental Material. Two out of five top signals for HDL-C (rs35709627 and rs71556729) and one out of two top signals for TG in the CHARGE consortium were replicated among the UKB participants (Table VII in the Supplemental Material). In a meta-analysis of the top results from the CHARGE consortium and data from the UKB, three out of the five top SNPs for HDL-C and one out of the two top SNPs for TG concentrations displayed statistically significant interactions (Table 3). The top SNP for HDL-C concentrations was located at rs71556729 (Figure 1A). In fully adjusted models, the association between the minor allele at rs71556729 with HDL-C concentrations was observed only among the highest SSB consumers [β (95% CI): 2.12 (1.16, 3.07) mg/dl, p<0.0001], and not the lowest SSB consumers (p=0.81; p_{Diff} <0.0001). Similarly, two SNPs in low to moderate LD with rs71556729 (*TBL2*-rs35709627:

DOI: 10.1161/CIRCGEN.120.003288

R^2 with rs71556729=0.55; rs71556736: R^2 with rs71556729=0.19) displayed similar statistically significant differences in effect by category of SSB intake (p_{Diff} <0.0001). The SNP at rs55673514 displayed a suggestive interaction with TG concentrations in the CHARGE meta-analysis and was statistically significant after including data from the UKB (Figure 1B, p_{Diff} <0.0005). The association of the minor allele at rs55673514 with TG concentrations was observed only among the highest SSB consumers [β (95% CI): 0.06 (0.02, 0.09) ln-mg/dl, p=0.001], and not the lowest SSB consumers (p=0.84). The SNP at rs55673514 is not in appreciable LD with any of the top SNPs in the overall analysis for TG concentrations (R^2<0.1). A heatmap of LD among top SNPs in overall and interaction analyses is provided in Figure II in the Supplemental Material. Sensitivity analyses examining the influence of adjustment for other dietary factors and fasting hours among UKB participants yielded similar results for the top SNPs identified in the meta-analysis (Supplemental Results in the Supplemental Material).

Discussion

In this study, including up to 86,241 participants for whom genetic and SSB consumption data were available, we identified novel interactions between genetic variants at the CHREBP locus and SSB consumption on HDL-C and TG concentrations. Our data suggest that the magnitude of the inverse association between SSB consumption and HDL-C concentrations is lower among individuals harboring genetic variants at rs71556729, rs35709627, and/or rs71556736 and the positive association between SSB consumption and TG concentrations is exacerbated among individuals harboring genetic variants at rs55673514. In the CHARGE cohorts, we also observed a consistent inverse association between SSB consumption on fasting HDL-C and positive

association on TG concentrations. We also replicated previously observed main associations between SNPs in the *CHREBP* locus and HDL-C and TG concentrations.

Our study provides evidence that SSB consumption may modify the association of genetic variants in the *CHREBP* locus with HDL-C and TG concentrations. Participants with the minor allele at rs71556729, rs35709627, and/or rs71556736 and high SSB consumption had higher mean HDL-C concentrations than those with the major allele who also had high SSB consumption. This suggests that participants with the minor allele at rs71556729 (MAF = 0.05), rs35709627 (MAF = 0.05), and/or rs71556736 (MAF = 0.13) may be protected against SSB-induced reductions in HDL-C concentrations. The region containing these SNPs is enriched for enhancer histone marks and these SNPs lie within putative regulatory motifs for transcription factors that could potentially regulate ChREBP expression and function in an SSB dependent manner.[33] Similarly, rs55673514, which associates with TG only among the highest SSB consumers, lies within a region enriched for enhancer histone marks in several tissues, including liver.[33] Given the strong inverse relationship between HDL-C and TG concentrations, additional investigation into how these SNPs may independently influence HDL-C or TG concentrations could provide new insights into the distinct mechanisms contributing to plasma HDL-C and TG concentrations. Additional discussion of main associations between SNPs and SSB on TG and HDL-C in the CHARGE cohorts is provided in the Supplemental Discussion in the Supplemental Material.

The rs71556729 interaction was a top signal when testing for interactions using the difference test and the cross-product interaction test on HDL-C concentrations in the CHARGE cohorts. However, when applying the cross-product interaction test, the interaction appeared less significant than the result from the difference test. This may be due to heterogeneity in the

association between rs71556729 and HDL-C concentrations resulting from increased misclassification of SSB consumption among those reporting low (1-4 servings/month) to moderate (1-2 and 3-7 servings/week) SSB consumption (Figure IV in the Supplemental Material). These results suggest that the difference test may be a useful method for identifying gene-diet interactions in observational studies, and this could be due to a reduction in misclassification of SSB intake and the potential to detect non-linear interaction effects. However, we do not comprehensively compare the difference test to the cross-product interaction test. Future methodological studies comparing the usefulness of these two methods with varying degrees of misclassification and types of exposures may be useful to inform future gene-diet interaction studies.

There is a limited body of evidence describing how genes implicated in various diseases may interact with SSB consumption to modify cardiometabolic health and noncommunicable disease risk.[34] One large prospective cohort study among Swedish adults examined whether genetic risk for dyslipidemia (using a weighted genetic risk score) interacted with SSB consumption to influence plasma lipid concentrations, but no significant interactions were observed.[35] Although genetic risk scores can be useful for translation, as previously shown for the interaction between SSB consumption and genetic risk for obesity,[36] a weakness of genetic risk scores is that aggregation of multiple SNPs from across the genome does not allow inclusion of potential interacting SNPs that may not be associated with the outcome in overall analyses. In addition, interaction effects of SNPs may be mitigated by the null interaction effects of other SNPs included in the genetic risk score. The candidate gene approach in the current study allows for the potential to generate hypotheses of the mechanisms underlying the interaction that could be tested using animal and human models in future studies.

DOI: 10.1161/CIRCGEN.120.003288

No previous studies have examined the interaction between SNPs in the *CHREBP* region and SSB consumption on lipid concentrations. We previously investigated how selected SNPs in the ChREBP-FGF21 pathway interacted with SSB consumption to influence fasting insulin and glucose measures among 34,748 adults from CHARGE cohorts, but we did not identify a significant cross-product interaction that was consistent among the discovery and replication phases of that study.[37] In the current study, we applied a comprehensive approach that tested a wide range of SNPs in the *CHREBP* region that were not necessarily identified in GWAS. Given that our suggestive interaction results do not include any SNPs that were statistically significant in the overall SNP analyses, our data indicate that there may be additional SNPs not identified in GWAS contributing to the heritability of HDL-C and TG concentrations, but their contribution is influenced by SSB consumption. Similar to previous GWAS for body mass index that have identified new loci when adjusting for environmental factors[38,39], we provide an additional example of how missing genetic heritability may be revealed when accounting for environmental factors, such as SSB consumption in the current study.

The strengths of our study include the large sample size attained through meta-analysis of multiple independent cohorts, the ability to standardize the analyses conducted in all cohorts through a collaborative approach, the use of an external cohort to validate findings, and the use of multiple methods to screen for potential interactions between SSB consumption and over 1,606 SNPs in the *CHREBP* region on HDL-C and TG concentrations. The analytic approach revealed novel SNPs that may contribute to unexplained heritability of HDL-C and TG concentrations. Limitations of this study include its observational design that constrain our ability to infer causality, the sample of European-descent adults that limits generalizability, the use of self-reported dietary data from food frequency questionnaires and 24-hour recall that may

lead to misclassification of food and nutrient intakes, and the possibility of residual confounding, even after controlling for potential dietary and lifestyle factors that co-vary with SSB intake. Our focus on the comparison of the highest SSB consumers to the lowest SSB consumers helps minimize this potential misclassification by focusing on extreme consumption patterns. Misclassification in the UKB is likely given that a snapshot of intake on a single day cannot provide a reliable estimate of usual SSB consumption. However, this misclassification is likely non-differential by genotype, which would only result in attenuation of our results. Additionally, while our definition of SSB did consider a range of SSB, it was not comprehensive. For example, it did not include commonly consumed beverages, such as sweetened tea or coffee, and we included several types of SSB in the same exposure definition (colas and fruit drinks). The blood collection among UKB participants was conducted after less than the recommended 8 hours of fasting prior to measurement of lipids. We adjusted for fasting hours to help account for this variability and conducted a sensitivity analysis to examine the top interactions observed by fasting hours. The LD-based method used to estimate the number of independent tests in the region may be overly conservative, which could potentially lead to inflation of type II error rate. Thus, we additionally present suggestive results that did not reach statistical significance. Given these weaknesses, results from this study should be used to inform future studies with larger samples sizes or detailed experimental studies. Minority populations are disproportionality burdened by dyslipidemia and have higher SSB intake,[40,41] and thus more studies in these populations may help reduce health inequality and disparity.

In conclusion, our findings suggest that the minor alleles of three SNPs in the *CHREBP* region (rs71556729, rs35709627, and rs71556736) may be protective against SSB-induced low HDL-C concentrations and the minor allele at rs55673514 may exacerbate positive associations

between SSB consumption and TG concentrations. Several of the top SNPs identified in the interaction analyses were not top SNPs identified in the overall analyses, providing evidence that some genetic associations may be revealed only when conditioned on environmental factors, such as the range of SSB consumption in the current study. As larger datasets with genetics, diet, and lipids data become available, additional suggestive interactions between SSB consumption and SNPs within the *CHREBP* region on HDL-C and TG concentrations observed here may warrant further investigation.

Acknowledgments: Preliminary results were presented as abstracts at the annual meeting for the American Society for Nutrition 2020. Please see Table I in the Supplemental Material for cohort-specific acknowledgements.

The authors' responsibilities were as follows: DEH, GMP, MG, HSD, AHL, CES, JD, MAH, and NMM designed the study. DEH, GMP, MG, RNL, NT, DOM-K, KN, VM, JSV, LS, YM-R, LWM, WHO, CEP, FRR, MAI, AGU, TV, BMP, DM, JIR, KDT, TL, OTR, KAL, NGF, NJW, JL, RM, SSR, JEM, SM, PMR, JBM, DIC, AHL, CES, JD, MAH, and NMM played a role in acquisition of the data and critical editing of the manuscript; DEH, GMP, MG, FI, TMB, ANP, CAW, RLG, JMW, NP, KLY, MG, ACW, KVEB, JL, MK, JCK-dJ, MG, and NT conducted statistical analyses; DEH, GMP, MG, HSD, JM, AHL, CES, JD, MAH, and NMM interpreted the data; DEH, GMP, MG, HSD, JM, KAL, AHL, CES, JD, MAH, and NMM contributed to writing of the manuscript; all authors read and approved the final version of the manuscript. DEH and NMM are the guarantors of this work and, as such, had full access to all the data in the study and take responsibility for the integrity of the data and accuracy of the data analysis.

Sources of Funding: This work is supported by National Institutes of Health (NIH) 5T32HL069772-15 and NIH 2T32CA009001-39 (Haslam), American Heart Association 16CSA28590003 (Haslam, McKeown, and Herman), NIH R01 DK100425 (Herman), R01 DK121710 (McKeown, Herman, Smith, and Dupuis), K08 HL112845 (Smith), USDA ARS agreement No. 58-1950-4-003 (McKeown) and 588-1950-9-001 (Lichtenstein). Infrastructure for the CHARGE Consortium is supported in part by the National Heart, Lung, and Blood Institute

grant HL105756. Please see Table I in the Supplemental Material for funding sources associated with investigators and infrastructure of individual CHARGE cohorts.

Disclosures: SM received institutional research grant support from Atherotech Diagnostics for research outside the current work, served as a consultant (modest) to Quest Diagnostics and Pfizer outside the current work.

Supplemental Materials
Supplemental Methods
Supplemental Results
Supplemental Discussion
Supplemental Tables I-XI
Supplemental Figures I-XXII
Appendix I

References:

1. Wu L, Parhofer KG. Diabetic dyslipidemia. *Metabolism*. 2014;63:1469–1479.

2. Rader DJ, Hovingh GK. HDL and cardiovascular disease. *The Lancet*. 2014;384:618–625.

3. Navar AM. The Evolving Story of Triglycerides and Coronary Heart Disease Risk. *JAMA*. 2019;321:347–349.

4. Goff David C., Lloyd-Jones Donald M., Bennett Glen, Coady Sean, D'Agostino Ralph B., Gibbons Raymond, Greenland Philip, Lackland Daniel T., Levy Daniel, O'Donnell Christopher J., et al. 2013 ACC/AHA Guideline on the Assessment of Cardiovascular Risk. *Circulation*. 2014;129:S49–S73.

5. Grundy Scott M., Stone Neil J., Bailey Alison L., Beam Craig, Birtcher Kim K., Blumenthal Roger S., Braun Lynne T., de Ferranti Sarah, Faiella-Tommasino Joseph, Forman Daniel E., et al. 2018 AHA/ACC/AACVPR/AAPA/ABC/ACPM/ADA/AGS/APhA/ASPC/NLA/PCNA Guideline on the Management of Blood Cholesterol. *Circulation*. 0:CIR.0000000000000625.

6. Heller DA, de Faire U, Pedersen NL, Dahlén G, McClearn GE. Genetic and environmental influences on serum lipid levels in twins. *N Engl J Med*. 1993;328:1150–1156.

7. Abney M, McPeek MS, Ober C. Broad and Narrow Heritabilities of Quantitative Traits in a Founder Population. *Am J Hum Genet*. 2001;68:1302–1307.

8. 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.

9. Kathiresan S, Melander O, Guiducci C, Surti A, Burtt NP, Rieder MJ, Cooper GM, Roos C, Voight BF, Havulinna AS, et al. Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. *Nat Genet*. 2008;40:189–197.

10. Peloso GM, Demissie S, Collins D, Mirel DB, Gabriel SB, Cupples LA, Robins SJ, Schaefer EJ, Brousseau ME. Common genetic variation in multiple metabolic pathways influences susceptibility to low HDL-cholesterol and coronary heart disease. *J Lipid Res*. 2010;51:3524–3532.

11. Do R, Willer CJ, Schmidt EM, Sengupta S, Gao C, Peloso GM, Gustafsson S, Kanoni S, Ganna A, Chen J, et al. Common variants associated with plasma triglycerides and risk for coronary artery disease. *Nat Genet*. 2013;45:1345–1352.

12. Klarin D, Damrauer SM, Cho K, Sun YV, Teslovich TM, Honerlaw J, Gagnon DR, DuVall SL, Li J, Peloso GM, et al. Genetics of Blood Lipids Among ~300,000 Multi-Ethnic Participants of the Million Veteran Program. *Nat Genet*. 2018;50:1514–1523.

13. Uyeda K, Repa JJ. Carbohydrate response element binding protein, ChREBP, a transcription factor coupling hepatic glucose utilization and lipid synthesis. *Cell Metab*. 2006;4:107–110.

14. Fisher ffolliott M, Kim M, Doridot L, Cunniff JC, Parker TS, Levine DM, Hellerstein MK, Hudgins LC, Maratos-Flier E, Herman MA. A critical role for ChREBP-mediated FGF21 secretion in hepatic fructose metabolism. *Mol Metab*. 2017;6:14–21.

15. Kooner JS, Chambers JC, Aguilar-Salinas CA, Hinds DA, Hyde CL, Warnes GR, Gómez Pérez FJ, Frazer KA, Elliott P, Scott J, et al. Genome-wide scan identifies variation in MLXIPL associated with plasma triglycerides. *Nat Genet*. 2008;40:149–151.

16. Chasman DI, Paré G, Mora S, Hopewell JC, Peloso G, Clarke R, Cupples LA, Hamsten A, Kathiresan S, Mälarstig A, et al. Forty-Three Loci Associated with Plasma Lipoprotein Size, Concentration, and Cholesterol Content in Genome-Wide Analysis. *PLOS Genet*. 2009;5:e1000730.

17. Postic C, Dentin R, Denechaud P-D, Girard J. ChREBP, a transcriptional regulator of glucose and lipid metabolism. *Annu Rev Nutr*. 2007;27:179–192.

18. Erion DM, Popov V, Hsiao JJ, Vatner D, Mitchell K, Yonemitsu S, Nagai Y, Kahn M, Gillum MP, Dong J, et al. The Role of the Carbohydrate Response Element-Binding Protein in Male Fructose-Fed Rats. *Endocrinology*. 2013;154:36–44.

19. Kim M, Astapova II, Flier SN, Hannou SA, Doridot L, Sargsyan A, Kou HH, Fowler AJ, Liang G, Herman MA. Intestinal, but not hepatic, ChREBP is required for fructose tolerance. *JCI Insight* [Internet]. [cited 2018 Jun 8];2. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5752301/

20. Linden AG, Li S, Choi HY, Fang F, Fukasawa M, Uyeda K, Hammer RE, Horton JD, Engelking LJ, Liang G. Interplay between ChREBP and SREBP-1c coordinates postprandial glycolysis and lipogenesis in livers of mice. *J Lipid Res*. 2018;59:475–487.

21. de Koning L, Malik VS, Kellogg MD, Rimm EB, Willett WC, Hu FB. Sweetened beverage consumption, incident coronary heart disease, and biomarkers of risk in men. *Circulation*. 2012;125:1735–1741, S1.

22. Welsh JA, Sharma A, Abramson JL, Vaccarino V, Gillespie C, Vos MB. Caloric Sweetener Consumption and Dyslipidemia Among US Adults. *JAMA*. 2010;303:1490–1497.

23. Stanhope KL, Schwarz JM, Keim NL, Griffen SC, Bremer AA, Graham JL, Hatcher B, Cox CL, Dyachenko A, Zhang W, et al. Consuming fructose-sweetened, not glucose-sweetened, beverages increases visceral adiposity and lipids and decreases insulin sensitivity in overweight/obese humans. *J Clin Invest*. 2009;119:1322–1334.

24. Haslam Danielle E., Peloso Gina M., Herman Mark A., Dupuis Josée, Lichtenstein Alice H., Smith Caren E., McKeown Nicola M. Beverage Consumption and Longitudinal Changes in Lipoprotein Concentrations and Incident Dyslipidemia in US Adults: The Framingham Heart Study. *J Am Heart Assoc*. 2020;9:e014083.

25. de Koning L, Malik VS, Rimm EB, Willett WC, Hu FB. Sugar-sweetened and artificially sweetened beverage consumption and risk of type 2 diabetes in men123. *Am J Clin Nutr*. 2011;93:1321–1327.

26. Hu FB, Malik VS. Sugar-sweetened beverages and risk of obesity and type 2 diabetes: Epidemiologic evidence. *Physiol Behav*. 2010;100:47–54.

27. Imamura F, O'Connor L, Ye Z, Mursu J, Hayashino Y, Bhupathiraju SN, Forouhi NG. Consumption of sugar sweetened beverages, artificially sweetened beverages, and fruit juice and incidence of type 2 diabetes: systematic review, meta-analysis, and estimation of population attributable fraction. *The BMJ*. 2015;351:h2376.

28. Popkin BM, Hawkes C. Sweetening of the global diet, particularly beverages: patterns, trends, and policy responses. *Lancet Diabetes Endocrinol*. 2016;4:174–186.

29. Singh GM, Micha R, Khatibzadeh S, Lim S, Ezzati M, Mozaffarian D. Estimated Global, Regional, and National Disease Burdens Related to Sugar-Sweetened Beverage Consumption in 2010. *Circulation*. 2015;132:639–666.

30. Dai H, Much AA, Maor E, Asher E, Younis A, Xu Y, Lu Y, Liu X, Shu J, Bragazzi NL. Global, regional, and national burden of ischaemic heart disease and its attributable risk factors, 1990–2017: results from the Global Burden of Disease Study 2017. *Eur Heart J - Qual Care Clin Outcomes* [Internet]. [cited 2020 Dec 1];Available from: http://academic.oup.com/ehjqcco/advance-article/doi/10.1093/ehjqcco/qcaa076/5918025

31. Psaty BM, O'Donnell CJ, Gudnason V, Lunetta KL, Folsom AR, Rotter JI, Uitterlinden AG, Harris TB, Witteman JCM, Boerwinkle E. Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium: Design of prospective meta-analyses of genome-wide association studies from five cohorts. *Circ Cardiovasc Genet*. 2009;2:73–80.

32. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, Downey P, Elliott P, Green J, Landray M, Liu B, et al. UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLOS Med*. 2015;12:e1001779.

33. Kheradpour P, Kellis M. Systematic discovery and characterization of regulatory motifs in ENCODE TF binding experiments. *Nucleic Acids Res*. 2014;42:2976–2987.

34. Haslam DE, McKeown NM, Herman MA, Lichtenstein AH, Dashti HS. Interactions between Genetics and Sugar-Sweetened Beverage Consumption on Health Outcomes: A Review of Gene–Diet Interaction Studies. *Front Endocrinol* [Internet]. 2018 [cited 2018 Jan 8];8. Available from: https://www.frontiersin.org/articles/10.3389/fendo.2017.00368/full?&utm_source=Email_to_authors_&utm_medium=Email&utm_content=T1_11.5e1_author&utm_campaign=Email_publication&field=&journalName=Frontiers_in_Endocrinology&id=320441

35. Sonestedt E, Hellstrand S, Drake I, Schulz C-A, Ericson U, Hlebowicz J, Persson MM, Gullberg B, Hedblad B, Engström G, et al. Diet Quality and Change in Blood Lipids during 16 Years of Follow-up and Their Interaction with Genetic Risk for Dyslipidemia. *Nutrients* [Internet]. 2016 [cited 2019 Feb 1];8. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4882687/

36. Qi Q, Chu AY, Kang JH, Jensen MK, Curhan GC, Pasquale LR, Ridker PM, Hunter DJ, Willett WC, Rimm EB, et al. Sugar-Sweetened Beverages and Genetic Risk of Obesity. *N Engl J Med*. 2012;367:1387–1396.

37. McKeown NM, Dashti HS, Ma J, Haslam DE, Jong JCK, Smith CE, Tanaka T, Graff M, Lemaitre RN, Rybin D, et al. Sugar-sweetened beverage intake associations with fasting glucose and insulin concentrations are not modified by selected genetic variants in a ChREBP-FGF21 pathway: a meta-analysis. *Diabetologia*. 2018;61:317–330.

38. Justice AE, Winkler TW, Feitosa MF, Graff M, Fisher VA, Young K, Barata L, Deng X, Czajkowski J, Hadley D, et al. Genome-wide meta-analysis of 241,258 adults accounting for smoking behaviour identifies novel loci for obesity traits. *Nat Commun* [Internet]. 2017 [cited 2021 Feb 8];8. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5414044/

39. Graff M, Scott RA, Justice AE, Young KL, Feitosa MF, Barata L, Winkler TW, Chu AY, Mahajan A, Hadley D, et al. Genome-wide physical activity interactions in adiposity — A meta-analysis of 200,452 adults. *PLoS Genet* [Internet]. 2017 [cited 2021 Feb 8];13. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5407576/

40. Pu J, Romanelli R, Zhao B, Azar KMJ, Hastings KG, Nimbal V, Fortmann SP, Palaniappan LP. Dyslipidemia in Special Ethnic Populations. *Cardiol Clin*. 2015;33:325–333.

41. Han E, Powell LM. CONSUMPTION PATTERNS OF SUGAR SWEETENED BEVERAGES IN THE UNITED STATES. *J Acad Nutr Diet*. 2013;113:43–53.

42. Becker BJ, Wu M-J. The Synthesis of Regression Slopes in Meta-Analysis. *Stat Sci*. 2007;22:414–429.

43. Cochran WG. The combination of estimates from different experiments. *Biometrics*. 1954;10:101–129.

44. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med*. 2002;21:1539–1558.

45. Yang J, Ferreira T, Morris AP, Medland SE, Genetic Investigation of ANthropometric Traits (GIANT) Consortium, DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium, Madden PAF, Heath AC, Martin NG, Montgomery GW, et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat Genet*. 2012;44:369–375.

46. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*. 2010;26:2190–2191.

47. Winkler TW, Justice AE, Cupples LA, Kronenberg F, Kutalik Z, Heid IM, Consortium the G. Approaches to detect genetic effects that differ between two strata in genome-wide meta-analyses: Recommendations based on a systematic evaluation. *PLOS ONE*. 2017;12:e0181038.

48. Voorman A, Lumley T, McKnight B, Rice K. Behavior of QQ-Plots and Genomic Control in Studies of Gene-Environment Interaction. *PLoS ONE* [Internet]. 2011 [cited 2017 Oct 23];6. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3093379/

49. Gao X. Multiple testing corrections for imputed SNPs. *Genet Epidemiol*. 2011;35:154–158.

50. 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature*. 2015;526:68.

51. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, et al. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am J Hum Genet*. 2007;81:559–575.

52. Zhou W, Nielsen JB, Fritsche LG, Dey R, Gabrielsen ME, Wolford BN, LeFaive J, VandeHaar P, Gagliano SA, Gifford A, et al. Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. *Nat Genet*. 2018;50:1335–1341.

53. Hert KA, Fisk II PS, Rhee YS, Brunt AR. Decreased consumption of sugar-sweetened beverages improved selected biomarkers of chronic disease risk among US adults: 1999 to 2010. *Nutr Res*. 2014;34:58–65.

54. Duffey KJ, Gordon-Larsen P, Steffen LM, Jacobs DR, Popkin BM. Drinking caloric beverages increases the risk of adverse cardiometabolic outcomes in the Coronary Artery Risk Development in Young Adults (CARDIA) Study. *Am J Clin Nutr*. 2010;92:954–959.

55. Dhingra R, Sullivan L, Jacques PF, Wang TJ, Fox CS, Meigs JB, D'Agostino RB, Gaziano JM, Vasan RS. Soft drink consumption and risk of developing cardiometabolic risk factors and the metabolic syndrome in middle-aged adults in the community. *Circulation*. 2007;116:480–488.

56. Yu Z, Ley SH, Sun Q, Hu FB, Malik VS. Cross-sectional association between sugar-sweetened beverage intake and cardiometabolic biomarkers in US women. *Br J Nutr*. 2018;119:570–580.

57. Morenga LAT, Howatson AJ, Jones RM, Mann J. Dietary sugars and cardiometabolic risk: systematic review and meta-analyses of randomized controlled trials of the effects on blood pressure and lipids. *Am J Clin Nutr*. 2014;100:65–79.

58. Talmud PJ, Drenos F, Shah S, Shah T, Palmen J, Verzilli C, Gaunt TR, Pallas J, Lovering R, Li K, et al. Gene-centric Association Signals for Lipids and Apolipoproteins Identified via the HumanCVD BeadChip. *Am J Hum Genet*. 2009;85:628–642.

59. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. *Am J Epidemiol*. 1989;129:687–702.

60. The cardiovascular health study: Design and rationale. *Ann Epidemiol*. 1991;1:263–276.

61. Dawber TR, Kannel WB, Lyell LP. An Approach to Longitudinal Studies in a Community: The Framingham Study. *Ann N Y Acad Sci*. 1963;107:539–556.

62. Feinleib M, Kannel WB, Garrison RJ, McNamara PM, Castelli WP. The Framingham Offspring Study. Design and preliminary data. *Prev Med*. 1975;4:518–525.

63. Splansky GL, Corey D, Yang Q, Atwood LD, Cupples LA, Benjamin EJ, D'Agostino RB, Fox CS, Larson MG, Murabito JM, et al. The Third Generation Cohort of the National Heart, Lung, and Blood Institute's Framingham Heart Study: Design, Recruitment, and Initial Examination. *Am J Epidemiol*. 2007;165:1328–1335.

64. Bild DE, Bluemke DA, Burke GL, Detrano R, Diez Roux AV, Folsom AR, Greenland P, Jacob DR, Kronmal R, Liu K, et al. Multi-Ethnic Study of Atherosclerosis: objectives and design. *Am J Epidemiol*. 2002;156:871–881.

65. Mutsert R de, Heijer M den, Rabelink TJ, Smit JWA, Romijn JA, Jukema JW, Roos A de, Cobbaert CM, Kloppenburg M, Cessie S le, et al. The Netherlands Epidemiology of Obesity (NEO) study: study design and data collection. *Eur J Epidemiol*. 2013;28:513–523.

66. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, Downey P, Elliott P, Green J, Landray M, et al. UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLOS Med*. 2015;12:e1001779.

67. Newnham JP, Evans SF, Michael CA, Stanley FJ, Landau LI. Effects of frequent ultrasound during pregnancy: a randomised controlled trial. *Lancet Lond Engl*. 1993;342:887–891.

68. Ikram MA, Brusselle G, Ghanbari M, Goedegebure A, Ikram MK, Kavousi M, Kieboom BCT, Klaver CCW, de Knegt RJ, Luik AI, et al. Objectives, design and main findings until 2020 from the Rotterdam Study. *Eur J Epidemiol*. 2020;35:483–517.

69. Ridker PM, Chasman DI, Zee RYL, Parker A, Rose L, Cook NR, Buring JE. Rationale, Design, and Methodology of the Women's Genome Health Study: A Genome-Wide Association Study of More Than 25 000 Initially Healthy American Women. *Clin Chem*. 2008;54:249–255.

70. Design of the Women's Health Initiative clinical trial and observational study. The Women's Health Initiative Study Group. *Control Clin Trials*. 1998;19:61–109.

71. Raitakari OT, Juonala M, Rönnemaa T, Keltikangas-Järvinen L, Räsänen L, Pietikäinen M, Hutri-Kähönen N, Taittonen L, Jokinen E, Marniemi J, et al. Cohort profile: the cardiovascular risk in Young Finns Study. *Int J Epidemiol*. 2008;37:1220–1226.

72. Stevens J, Metcalf PA, Dennis BH, Tell GS, Shimakawa T, Folsom AR. Reliability of a food frequency questionnaire by ethnicity, gender, age and education. *Nutr Res*. 1996;16:735–745.

73. Willett WC, Sampson L, Stampfer MJ, Rosner B, Bain C, Witschi J, Hennekens CH, Speizer FE. Reproducibility and validity of a semiquantitative food frequency questionnaire. *Am J Epidemiol*. 1985;122:51–65.

74. Kumanyika S, Tell GS, Fried L, Martel JK, Chinchilli VM. Picture-Sort Method for Administering a Food Frequency Questionnaire to Older Adults. *J Am Diet Assoc*. 1996;96:137–144.

75. Bingham SA, Gill C, Welch A, Day K, Cassidy A, Khaw KT, Sneyd MJ, Key TJ, Roe L, Day NE. Comparison of dietary assessment methods in nutritional epidemiology: weighed records v. 24 h recalls, food-frequency questionnaires and estimated-diet records. *Br J Nutr*. 1994;72:619–643.

76. Bingham SA, Gill C, Welch A, Cassidy A, Runswick SA, Oakes S, Lubin R, Thurnham DI, Key TJ, Roe L, et al. Validation of dietary assessment methods in the UK arm of EPIC using weighed records, and 24-hour urinary nitrogen and potassium and serum vitamin C and carotenoids as biomarkers. *Int J Epidemiol*. 1997;26 Suppl 1:S137-151.

77. Rimm EB, Giovannucci EL, Stampfer MJ, Colditz GA, Litin LB, Willett WC. Reproducibility and validity of an expanded self-administered semiquantitative food frequency questionnaire among male health professionals. *Am J Epidemiol*. 1992;135:1114–1126; discussion 1127-1136.

78. Mayer-Davis EJ, Vitolins MZ, Carmichael SL, Hemphill S, Tsaroucha G, Rushing J, Levin S. Validity and reproducibility of a food frequency interview in a Multi-Cultural Epidemiology Study. *Ann Epidemiol*. 1999;9:314–324.

79. Verkleij-Hagoort AC, de Vries JHM, Stegers MPG, Lindemans J, Ursem NTC, Steegers-Theunissen RPM. Validation of the assessment of folate and vitamin B 12 intake in women of reproductive age: the method of triads. *Eur J Clin Nutr*. 2007;61:610–615.

80. Voortman T, Kiefte-de Jong JC, Ikram MA, Stricker BH, van Rooij FJA, Lahousse L, Tiemeier H, Brusselle GG, Franco OH, Schoufour JD. Adherence to the 2015 Dutch dietary guidelines and risk of non-communicable diseases and mortality in the Rotterdam Study. *Eur J Epidemiol*. 2017;32:993–1005.

81. Klipstein-Grobusch K, den Breeijen J, Goldbohm RA, Geleijnse JM, Hofman A, Grobbee DE, Witteman JCM. Dietary assessment in the elderly: validation of a semiquantitative food frequency questionnaire. *Eur J Clin Nutr*. 1998;52:588–596.

82. Goldbohm RA, van den Brandt PA, Brants HA, van't Veer P, Al M, Sturmans F, Hermus RJ. Validation of a dietary questionnaire used in a large-scale prospective cohort study on diet and cancer. *Eur J Clin Nutr*. 1994;48:253–265.

83. Feunekes GI, Van Staveren WA, De Vries JH, Burema J, Hautvast JG. Relative and biomarker-based validity of a food-frequency questionnaire estimating intake of fats and cholesterol. *Am J Clin Nutr*. 1993;58:489–496.

84. Paalanen L, Männistö S, Virtanen MJ, Knekt P, Räsänen L, Montonen J, Pietinen P. Validity of a food frequency questionnaire varied by age and body mass index. *J Clin Epidemiol*. 2006;59:994–1001.

85. Liu B, Young H, Crowe FL, Benson VS, Spencer EA, Key TJ, Appleby PN, Beral V. Development and evaluation of the Oxford WebQ, a low-cost, web-based method for assessment of previous 24 h dietary intakes in large-scale prospective studies. *Public Health Nutr*. 2011;14:1998–2005.

86. McCance and Widdowson's The Composition of Foods: Edition 6. 2002.

87. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, Motyer A, Vukcevic D, Delaneau O, O'Connell J, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature*. 2018;562:203–209.

88. Brage S, Westgate K, Franks PW, Stegle O, Wright A, Ekelund U, Wareham NJ. Estimation of Free-Living Energy Expenditure by Heart Rate and Movement Sensing: A Doubly-Labelled Water Study. *PLoS ONE* [Internet]. 2015 [cited 2020 Oct 28];10. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4562631/

89. Kannel WB, Sorlie P. Some Health Benefits of Physical Activity: The Framingham Study. *Arch Intern Med*. 1979;139:857–861.

90. Machiela MJ, Chanock SJ. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics*. 2015;31:3555–3557.

DOI: 10.1161/CIRCGEN.120.003288

Table 1. General characteristics of participating CHARGE consortium cohorts*

	Raine Study	ARIC	FHS	NEO	Fenland	YFS	WGHS	WHI	MESA	CHS	RS
Characteristics											
Country	Australia	USA	USA	Netherlands	UK	Finland	USA	USA	USA	USA	Netherlands
n	617	10,924	6,382	5,694	10,022	1,782	16,284	1,109	1,805	3,196	5,784
Age (years)	20 (1)	55 (6)	49 (14)	56 (6)	49 (7)	38 (5)	55 (7)	65 (7)	70 (10)	72 (5)	66 (8)
Sex (% women)	52.4	52.7	54.3	52.0	53.3	55.9	100	100	51.2	61.0	57.8
Body Mass Index (kg/m^2)	24.5 (5.2)	27.0 (4.8)	27.4 (5.5)	30.0 (4.8)	26.9 (4.8)	25.9 (4.6)	25.9 (4.9)	28.6 (5.7)	28.0 (5.3)	26.3 (4.4)	26.5 (3.7)
Current Smoker (%)	13.5	24.2	13.4	16.0	12.0	27.6	11.7	10.1	7.0	11.4	23.4
Completed High School (%)	81.5	84.9	98.0	93.0	81.8	75.4	100	94.7	96.5	75.1	60.8
Fasting HDL-C (mg/dl)	51 (13)	51 (17)	54 (17)	55 (16)	59 (16)	52 (13)	54 (15)	58 (15)	57 (18)	55 (16)	53 (14)
Fasting TG (mg/dl)	85 (2)	137 (90)	117 (87)	130 (85)	106 (81)	122 (82)	119 (89)	156 (92)	107 (59)	140 (76)	137 (71.0)
Dietary Intakes											
SSB intake (servings/d)	0.7 (1.0)	0.5 (0.9)	0.4 (0.8)	0.4 (0.8)	0.3 (0.6)	0.3 (0.5)	0.3 (0.6)	0.2 (0.6)	0.1 (0.5)	0.1 (0.3)	0.1 (0.2)
<1 serving/month (%)	13.6	35.7	33.9	49.4	35.8	23.6	44.8	58.0	70.0	63.4	71.9
1-4 serving/month (%)	14.4	16.3	24.3	13.8	24.6	31.9	22.0	19.3	12.4	16.9	13.5
1-2 serving/week (%)	23.8	12.1	9.76	14.1	14.0	17.1	13.1	3.5	2.2	0.06	6.4
3-7 serving/week (%)	29.2	25.7	21.3	11.7	15.2	21.0	15.1	15.3	8.6	18.7	7.5
>1 serving/day (%)	19.0	10.3	10.8	11.0	10.4	6.3	5.0	3.9	2.3	0.9	0.8
Energy Intake (kcal/d)	1,857 (850)	1,644 (599)	1,956 (645)	2,291 (763)	1,935 (578)	2,381 (762)	1,732 (524)	1,698 (670)	1708 (734)	2,024 (654)	2,046 (1,409)
Saturated Fat Intake (% total energy)	16.1 (3.1)	12.2 (3.1)	11.1 (2.9)	12.4 (2.9)	12.5 (3.0)	11.8 (2.4)	10.2 (2.5)	11.6 (3.3)	11.3 (3.3)	10.4 (2.2)	14.4 (3.1)
Fruit intake (servings/d)	1.9 (1.3)	1.5 (1.3)	1.1 (1.0)	1.1 (0.9)	2.7 (2.2)	3.4 (3.1)	1.9 (1.2)	1.8 (1.2)	2.1 (1.7)	2.7 (1.5)	1.2 (1.0)
Vegetable Intake (servings/d)	1.7 (0.9)	1.7 (1.2)	2.0 (1.1)	2.8 (1.5)	5.0 (2.5)	1.4 (1.8)	3.9 (2.3)	2.2 (1.3)	2.4 (1.5)	2.8 (1.5)	2.8 (2.1)
Whole Grain Intake (servings/d)	0.8 (1.0)	1.1 (1.1)	1.2 (1.2)	NA	1.8 (1.4)	3.2 (1.9)	1.5 (1.2)	1.2 (0.8)	1.0 (0.8)	1.0 (0.7)	3.4 (2.9)
Fish Intake (servings/d)	0.4 (0.6)	0.3 (0.3)	0.4 (0.4)	0.2 (0.2)	0.4 (0.3)	1.3 (0.9)	0.3 (0.2)	0.2 (0.2)	0.3 (0.3)	0.3 (0.3)	0.1 (0.2)
Nuts/Seeds Intake (servings/d)	0.1 (0.2)	0.4 (0.6)	0.6 (0.9)	0.8 (1.0)	0.2 (0.3)	0.1 (0.1)	0.3 (0.4)	0.2 (0.3)	0.5 (0.6)	0.2 (0.3)	0.2 (2.1)
Alcohol Intake (g/d)	7.8 (8.9)[b]	6.7 (13.5)	10.5 (14.8)	15.5 (17.4)	9.5 (12.7)	8.6 (13.4)	4.3 (8.5)	5.0 (10.2)	8.8 (15.5)	5.5 (12.9)	11.1 (15.5)

*Means (standard deviation) or percentage for maximum observations available for analysis. Sample sizes may vary depending on availability of genotype and covariate information. Cohorts are ordered by estimate of sugar-sweetened beverage intake. Cohort study abbreviations: The Raine Study (Raine Study), Atherosclerosis Risk in Communities Study (ARIC), Framingham Heart Study (FHS), Netherlands Epidemiology in Obesity Study (NEO), The Fenland Study (Fenland), Young Finns Study (YFS), Women's Genome Health Study (WGHS), Women's Health Initiative (WHI), Multi-Ethnic Study of Atherosclerosis (MESA), Cardiovascular Health Study (CHS), and the Rotterdam Study (RS).

CHARGE, Cohorts for Heart and Aging Research in Genomic Epidemiology; HDL-C, high-density lipoprotein cholesterol concentrations; *n*, total sample size; SSB, sugar-sweetened beverages; TG, triglyceride concentrations.

DOI: 10.1161/CIRCGEN.120.003288

Table 2. Top SNPs in meta-analysis of difference test (p_{Diff}<0.005) and cross-product ($p_{interact}$<0.005) interactions between SSB consumption and SNPs on HDL-C and TG concentrations in CHARGE consortium cohorts*

SNP	Location (Hg19)	Alleles (E/A)†	Minor Allele Frequency	Model‡	SSB Intake Category	*n*	Effect Size (SE)§	*P*-value	Direction‖	I^2	*p* #
HDL-C (mg/dl)											
Difference Test											p_{Diff}
rs35709627††	72999171	A/G	0.05	Model 1	<1 serving/month	24,389	-0.01 (0.04)	0.86	+-++--+++??	23%	1.98E-05**
					>1 serving/day	4,033	3.23 (0.77)	2.94E-05	+?+?+?+?+??	0%	
				Model 2	<1 serving/month	23,801	0.006 (0.04)	0.86	+-++--+++??	30%	0.0001
					>1 serving/day	3,955	2.72 (0.72)	0.0002	+?+?+?+?+??	0%	
rs71556729††	72989516	T/C	0.05	Model 1	<1 serving/month	23,974	0.02 (0.06)	0.77	+?++-+++-??	0%	4.78E-05**
					>1 serving/day	3,359	4.47 (1.10)	5.02E-05	??+?+?+?+??	0%	
				Model 2	<1 serving/month	22,835	0.01 (0.05)	0.83	+?++-+-?-??	0%	0.0001
					>1 serving/day	3,299	3.89 (1.04)	0.0002	??+?+?+?+??	0%	
rs71556736	73034929	T/C	0.13	Model 1	<1 serving/month	24,389	-0.0005 (0.02)	0.98	+-+++--++?-	60%	0.0003
					>1 serving/day	4,033	1.65 (0.47)	0.0004	+?+?+?+?+??	0%	
				Model 2	<1 serving/month	23,801	0.007 (0.02)	0.69	+-++++-++??	67%	0.002
					>1 serving/day	3,955	1.34 (0.43)	0.002	+?+?+?+?+??	0%	
rs73137017	72974413	G/A	0.04	Model 1	<1 serving/month	24,020	-0.05 (0.06)	0.46	-+--+-++-??	0%	0.002
					>1 serving/day	3,933	-3.13 (0.99)	0.002	-?-?-?-?-??	0%	
				Model 2	<1 serving/month	23,437	-0.008 (0.05)	0.88	++--+-++-??	0%	
					>1 serving/day	3,855	-2.64 (0.91)	0.004	-?-?-?-?-??	0%	0.003
Cross-Product Interaction Test											$p_{interact}$
rs71556729	72989516	T/C	0.03	Model 1	-	55,418	0.66 (0.21)	-	+++++?+-+--	0%	0.001
				Model 2	-	53,394	0.68 (0.20)	-	++-++?+++?-	26%	0.0007
rs79578725	73002455	A/G	0.05	Model 1	-	53,662	-0.51 (0.18)	-	+?-+-?-----	0%	0.005
				Model 2	-	52,328	-0.18 (0.17)	-	+?++-?-----	0%	0.28
TG (ln-mg/dl)											
Difference Test											p_{Diff}
rs799157	73020301	T/C	0.05	Model 1	<1 serving/month	23,974	0.01 (0.01)	0.11	+?++++-++??	59%	0.005
					>1 serving/day	4,033	0.11 (0.03)	0.002	+?+?+?+?+??	0%	
				Model 2	<1 serving/month	23,403	0.02 (0.01)	0.17	+?++--?+?	68%	0.008
					>1 serving/day	3,955	0.09 (0.03)	0.004	+?+?++?+?	0%	
Cross-Product Interaction Test											$p_{interact}$
rs55673514	73021456	G/A	0.04	Model 1	-	57,977	0.02 (0.01)	-	-+++++++?++	17%	0.04
				Model 2	-	56,578	0.02 (0.01)	-	-++++++++??	0%	0.005

*Top signals represent suggestive interactions p_{Diff}<0.005 or $p_{interact}$<0.005[c]

†Alleles presented as effect (E)/alternative (A) alleles

‡Model 1 adjusted for age (years), sex (male/female), total energy intake (kcal/day) field center (CHS, FHS, YFS, Fenland, RS, MESA), and accounted for family or population structure where applicable (FHS, YFS, Fenland, NEO, MESA, WGHS, Raine Study, MESA); Model 2 adjusted for Model 1 covariates plus cohort-specific definition of education, smoking, physical activity, alcohol intake, and body mass index (kg/m^2).

§For the difference test, β (SE) represents the direction and magnitude of the difference in the outcome trait with each additional effect allele among categories of SSB consumption. For the cross-product interaction test, β (SE) represents the direction and magnitude of the difference in the outcome trait with each additional effect allele, per each increase in category of SSB intake (<1 serving/month, 1-4 servings/month, 1-2 servings/week, 3-7 servings/week, >1 serving/day).

‖Order of cohorts for regression coefficient directions: Framingham Heart Study, Young Finns Study, Fenland Study, Cardiovascular Health Study, Netherlands Epidemiology in Obesity Study, Rotterdam Study, Women's Genome Health Study, Women's Health Initiative, Atherosclerosis Risk in Communities Sutyd, The Raine Study, Multi-Ethnic Study of Atherosclerosis (+, positive effect size; -, negative effect size; ?, SNP not available in cohort).

#*P* represents p_{Diff} for the difference test for the highest and lowest category of SSB intake (<1 serving/month vs. >1 serving/day). *P* represents $p_{interact}$ for the cross-product interaction regression coefficient of additive SSBxSNP categories.

††Linkage disequilibrium (R^2) between rs13240662 and rs71556729=0.55 in European ancestry groups of Phase 3 (Version 5) of the 1000 Genomes Project.

**Indicates a statistically significant interaction based on Bonferonni-corrected p_{Diff} or $p_{interact}$<0.0001

CHARGE, Cohorts for Heart and Aging Research in Genetic Epidemiology; HDL-C, high-density lipoprotein cholesterol concentrations; SE, standard error; SNP, single nucleotide polymorphism; SSB, sugar-sweetened beverages; TG, triglyceride concentrations.

DOI: 10.1161/CIRCGEN.120.003288

Table 3. Meta-analysis of top candidate SNPs for difference test interactions between SSB consumption and SNPs on HDL-C and TG concentrations in CHARGE consortium cohorts and UKB*

SNP	Location (Hg19)	Alleles (E/A)†	MAF	SSB Intake Category	*n*	Effect Size (SE)	*P*-value	Direction‡	I^2	p_{Diff}
HDL-C (mg/dl)										
rs71556729§	72989516	T/C	0.05	Low	68,701	0.01 (0.05)	0.81	++	0 %	1.5E-06‖
				High	15,227	2.06 (0.44)	3.48E-06	++	74 %	
rs35709627§	72999171	A/G	0.05	Low	69,667	0.01 (0.04)	0.74	++	0 %	1.0E-05‖
				High	15,883	1.37 (0.32)	2.15E-05	++	87 %	
rs71556736	73034929	T/C	0.13	Low	69.667	0.02 (0.02)	0.33	++	93 %	2.5E-05‖
				High	15,882	0.84 (0.20)	3.27E-05	++	42 %	
rs73137017	72974413	G/A	0.04	Low	69,303	0.01 (0.05)	0.82	++	0 %	0.04
				High	15,783	0.73 (0.37)	0.05	++	81 %	
rs79578725	73002455	A/G	0.05	Low	68,929	-0.02 (0.04)	0.64	--	21 %	0.55
				High	15,783	-0.22 (0.36)	0.53	--	0 %	
TG (ln-mg/dl)										
rs55673514	73021456	G/A	0.04	Low	69,096	-0.002 (0.01)	0.84	+-	29 %	0.0005‖
				High	15,395	-0.06 (0.02)	0.001	--	0 %	
rs799157	73020301	T/C	0.05	Low	70,235	0.03 (0.01)	2.55E-07	++	59 %	0.05
				High	16,006	0.06 (0.02)	0.0002	++	19 %	

*Top candidates represent statistically significant or suggestive interactions (p_{Diff}<0.005 or $p_{interact}$<0.005) in CHARGE cohort meta-analysis. Models adjusted for age, sex, total energy intake, field center and accounted for family or population structure where applicable plus education, smoking, physical activity, alcohol intake, and body mass index (kg/m^2). For the difference test, interaction coefficients are shown as β (SE), where β represents the direction and magnitude of change in the outcome trait with each additional effect allele among participants with low (CHARGE:<1 serving/month; UKB: non-consumers) or high (CHARGE: >1 serving/day; UKB: consumers) SSB consumption.

†Alleles presented as effect (E)/alternative (A) alleles

‡Order of cohorts for regression coefficient directions: CHARGE cohorts, UKB (+, positive effect size; -, negative effect size).

§Linkage disequilibrium (R^2) between rs13240662 and rs71556729=0.55 in European ancestry groups of Phase 3 (Version 5) of the 1000 Genomes Project.

‖Indicates a statistically significant interaction based on Bonferroni-corrected p_{Diff}<0.01 (0.05/5 top signals) for HDL-C and p_{Diff}<0.025 (0.05/2 top signals) for TG concentrations

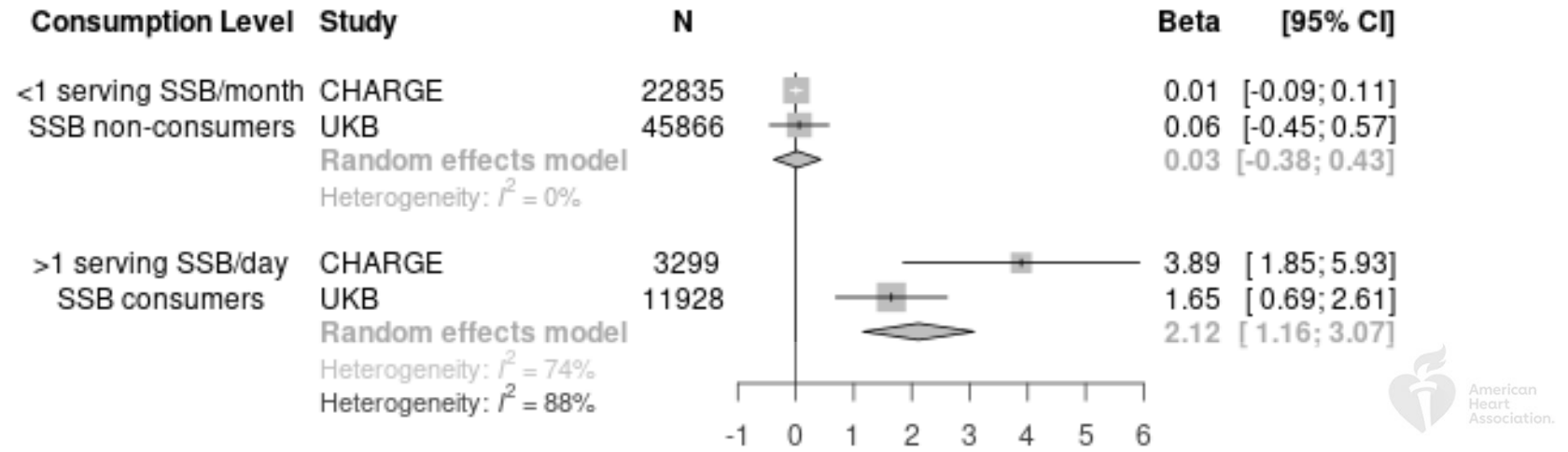
CHARGE, Cohorts for Heart and Aging Research in Genetic Epidemiology; HDL-C, high-density lipoprotein cholesterol concentrations; SNP, single nucleotide polymorphism; SSB, sugar-sweetened beverages; TG, triglyceride concentrations; UKB, UK Biobank.

DOI: 10.1161/CIRCGEN.120.003288

Figure Legends:

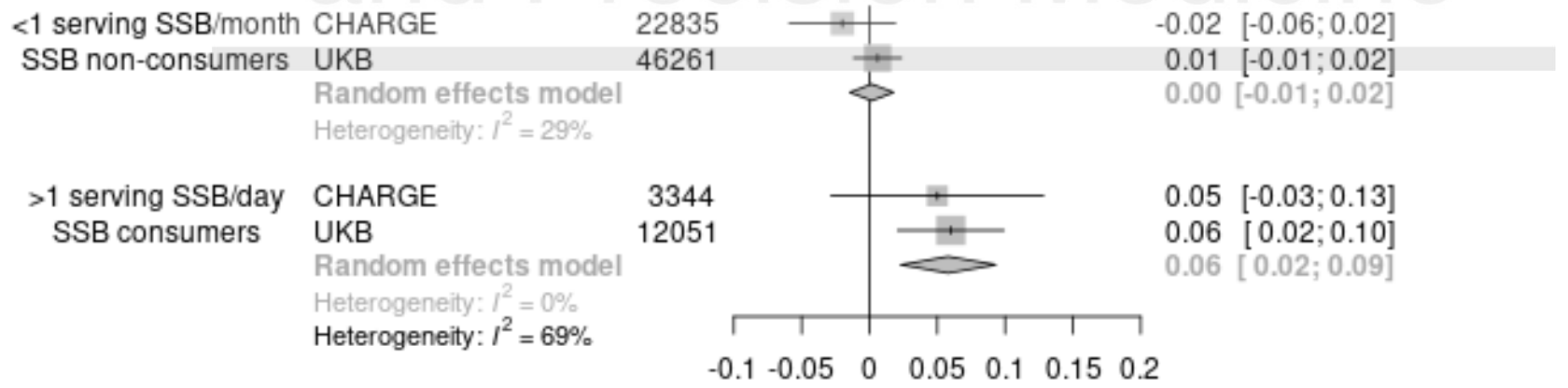
Figure 1. Associations between top candidate SNPs and HDL-C and TG concentrations stratified by category of SSB intake in a meta-analysis of the CHARGE cohorts and the UKB. **A)** In a meta-analysis of the CHARGE cohorts and the UKB, the association of the minor allele at *TBL2*-rs71556729 with HDL-C concentrations was observed only among the highest SSB consumers [β (95% CI): 2.12 (1.16, 3.07) mg/dl, $p<0.0001$], and not the lowest SSB consumers ($p=0.81$; $p_{Diff}<0.0001$); **B)** In a meta-analysis of the CHARGE cohorts and the UKB, the association of the minor allele at *CHREBP*-rs55673514 with TG concentrations was observed only among the highest SSB consumers [β (95% CI): 0.06 (0.02, 0.09) ln-mg/dl, $p=0.001$], and not the lowest SSB consumers ($p=0.84$; $p_{Diff}<0.0005$); Linear regression models represent associations between each additional effect allele and HDL-C (mg/dl) or TG (ln-mg/dl) concentrations among SSB consumption categories accounting for family, population structure, and/or field center (where applicable) and adjusting for age, sex, total energy intake, education, smoking, physical activity, alcohol intake, and body mass index. Intake categories are different for the highest SSB consumers (CHARGE: >1 serving/day; UKB: SSB consumers) and lowest SSB consumers (CHARGE: <1 serving/month; UKB: SSB non-consumers) in the two samples. CI, confidence interval; CHARGE, Cohorts for Heart and Aging Research in Genetic Epidemiology; HDL-C, high-density lipoprotein cholesterol concentrations; I^2, percentage of variance in a meta-analysis that is attributable to study heterogeneity; SSB, sugar-sweetened beverage consumption; TG, triglyceride concentrations; UKB, UK Biobank.

A



Mean difference in HDL-C concentrations (mg/dl) with each additional T allele at rs71556729 by category of SSB consumption

B



Mean difference in TG concentrations (ln-mg/dl) with each additional G allele at rs55673514 by category of SSB consumption