Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Qi Q, Chu AY, Kang JH, et al. Sugar-sweetened beverages and genetic risk of obesity. N Engl J Med 2012;367:1387-96. DOI: 10.1056/NEJMoa1203039

Supplementary Appendix

Table of Contents:

Methods for online publication

Study population

The NHS is a prospective cohort study of 121,700 female registered nurses who were 30 to 55 years old at study inception in 1976.[1] The HPFS is a prospective cohort study of 51,529 U.S. male health professionals who were 40 to 75 years old at study inception in 1986.[2] In both cohorts, information about medical history, lifestyle factors and health conditions has been collected biennially by self-administered questionnaires every 2 years since inception. Participants completed food-frequency questionnaires (FFQs) every 4 years. For this analysis, we used 1980 as baseline for the NHS and 1986 as baseline for HPFS when the first dietary data were collected. Study samples included participants of European ancestry with GWAS data that were initially designed to address various chronic diseases such as type 2 diabetes, coronary heart disease, kidney stone disease, primary open-angle glaucoma, and breast cancer (NHS only). Details regarding the study design, genotyping quality control and assurance of these GWASs have been reported elsewhere.[3-7] After excluding participants with these chronic diseases at baseline, a total of 6934 initially healthy women and 4423 initially healthy men were included in the current analysis.

The WGHS is a prospective cohort of initially healthy U.S. women.[8] Study participants were health professionals who were age 45 years and older and free of major chronic disease including cancer and cardiovascular disease at study entry (1992–1995). A total of 21,740 women who had confirmed self-reported European ancestry, had genotyping and FFQ data available, and were free of diabetes at baseline, were included in the current analysis. The present study was approved by the institutional review boards of Brigham and Women's Hospital and Harvard School of Public Health (Boston, MA).

Assessment of beverage intake

We used similar semi-quantitative FFQs to assess food and beverage intakes in the three cohorts.[9, 10] Participants were asked to report their usual intake (never to ≥6 times/day) of a standard portion size of foods and beverages. Total energy and nutrient intake was calculated by summing up from all foods and beverages. Participants with an implausible energy intake (<900 or >3500 kcal/day), or with more than nine blank responses on the FFQs were not included in the analysis. For SSBs, we included caffeinated colas, caffeine-free colas, other carbonated non-colas, and non-carbonated SSBs (fruit punches, lemonades, or other fruit drinks). Artificially sweetened beverages were defined as caffeinated, caffeine-free, and non-carbonated low-calorie beverages. We summed the intake of individual items to create total consumption of SSBs and artificially sweetened beverages. In the current analysis, 4 categories of frequency of intake were coded consistently across FFQs from all cohorts (<1 serving/month, 1-4 servings/month, 2-6 servings/week, and ≥1 servings/day). The reproducibility and validity of the FFQs have been described elsewhere.[9-11] The deattenuated correlations between FFQs and 7-day diet records were 0.84 for colas, 0.75 for low-calories colas, and 0.55 for other carbonated SSBs.[11]

Assessment of body mass index and covariates

In the NHS and HPFS, height and body weight were assessed by questionnaire at baseline, and weight was requested on each follow-up questionnaires. Self-reported weights were highly correlated with measured weight (r=0.97 in men and women).[12] BMI was calculated as body weight (kg)/height^2 (m^2). Participants with a BMI ≥ 30 kg/m^2 were defined as obese. Information about lifestyle factors was derived from the biennial questionnaires.[1, 2] The validity of the self-reported height, weight and physical activity data has been described previously.[12-14] Diet quality was assessed by using the alternative Healthy Eating Index[15] which comprises 9 components of

dietary factors: vegetables, fruit, nuts and soy protein, ratio of white to red meat, cereal fiber, trans fat, ratio of polyunsaturated to saturated fatty acids, duration of multivitamin use, and alcohol. In the WGHS, weight and physical activity were assessed by the baseline and follow-up questionnaires. Information about medical history and other dietary and lifestyle factors was collected from questionnaires at baseline. Details regarding the assessment of these variables have been reported previously.[16, 17]

Genotyping

We selected 32 single nucleotide polymorphisms (SNPs) that represent all 32 known BMI associated loci.[18] SNP genotyping and imputation have been described in detail elsewhere.[3-8, 17] Characteristics of the 32 SNPs are listed in Table S1. Most of the SNPs were genotyped or had a high imputation quality score (MACH $r^2 \geq 0.8$).

Genetic Predisposition Score Computation

Genetic predisposition score was calculated on the basis of the 32 established BMI associated variants by using a previously reported weighted method.[18] Briefly, each SNP was weighted by its relative effect size (β-coefficient). To obtain more accurate and precise population effect size of each SNP on BMI, we used β-coefficients derived from the previously reported meta-analysis of ~126,000 individuals.[18] The genetic predisposition score was calculated by multiplying each β-coefficient by the number of corresponding risk alleles and then summing the products. Because this produced a score out of 8.78 (twice the sum of the reported β-coefficients), all values were divided by 8.78 and multiplied by 64 to make the genetic predisposition score easier to interpret: each point of the genetic predisposition score corresponded to one risk allele.

Statistical analysis

In the NHS and HPFS, to minimize reverse causality, we analyzed the data prospectively with the assessment of beverage intake 4 years prior to the assessment of BMI. For example, the 1980 SSB intake data were related to the 1984 BMI in the NHS. Because of possible confounding due to aging-related weight change in the elderly population, we only used follow-up data up to 1998 as the mean age of our study samples was greater than 65 years old after 1998. Sensitivity analyses were conducted by using follow-up data up to 2006. Generalized linear models with repeated measures analysis were applied to estimate differences in BMI per increment of 10 risk alleles stratified by 4 categories of beverage intake (<1 serving/month, 1-4 servings/month, 2-6 servings/ week, and ≥1 servings/day). There were 5 repeated measures during 1980-1998 in the NHS and 3 repeated measures during 1986-1998 in the HPFS. Cox proportional hazards models were used to estimate relative risks (RRs) (95% CI) per increment of 10 risk alleles of developing obesity during follow-up by 4-year intervals in the NHS and HPFS, stratified by 4 categories of SSB intake. Obese participants at baseline were excluded from this analysis, leaving 6402 non-obese women and 3889 non-obese men in the analysis. Participants who became obese in a 4-year period were excluded from the analysis of the subsequent 4-year intervals. Interactions between the genetic predisposition score and beverage intake on BMI or risk of obesity were tested by including the respective interaction terms in the models (e.g., interaction term = SSB intake × genetic predisposition score). Multivariable models were used to adjust for age, genotyping data source, physical activity, TV watching, smoking, alcohol intake, alternative Healthy Eating Index, and total energy intake. For each covariate, we included an indicator for missing data in the models, if necessary.

Similar analyses were repeated in the WGHS to replicate the results in the NHS and HPFS. General linear models were applied to examine the interaction between beverage intake and the

genetic predisposition score on BMI, using data on beverage intake assessed at baseline and BMI assessed 3 years later. Cox proportional hazards models were used to examine interaction between SSB intake and the genetic predisposition score on the risk of developing obesity among the 18,127 non-obese women at baseline. Multivariable models were used to adjust for age, region, eigenvectors, physical activity, smoking, alcohol intake and total energy intake. In addition, we also used general linear models to estimate differences in BMI associated with one serving/day of SSB intake according to the quartiles of the genetic predisposition score in the NHS, HPFS and WGHS separately. Findings across cohorts were pooled by inverse-variance–weighted, fixed-effects meta-analyses. All reported P-values are nominal and two-sided. Statistical analyses were performed in SAS 9.1 (SAS Institute, Inc., Cary, NC, USA) or R 2.13.0 (R Foundation, Vienna, Austria).

Author contributions

Qibin Qi and Lu Qi designed the study. Jae H. Kang, Majken K. Jensen, Gary C. Curhan, Louis R. Pasquale, Paul M. Ridker, David J. Hunter, Walter C. Willett, Eric B. Rimm, Daniel I. Chasman, Frank B. Hu and Lu Qi gathered the data. Qibin Qi and Audrey Y. Chu analyzed the data. Lu Qi vouches for the data and the analysis. Qibin Qi and Lu Qi wrote the first draft; and all the co-authors contributed to writing the paper. Lu Qi provided funding support to the analysis and decided to publish the paper.

Table S1 Characteristics of 32 established SNPs for BMI

SNP	Nearest Gene	Chr	Allele*		Beta (SE)†	NHS		HPFS		WGHS	
			Effect	Other		EAF	r^2‡	EAF	r^2‡	EAF	r^2‡
rs543874	*SEC16B*	1	G	A	0.22 (0.03)	0.19	1.00	0.19	1.00	0.20	0.99
rs1514175	*TNNI3K*	1	A	G	0.07 (0.02)	0.42	1.00	0.40	1.00	0.42	1.00
rs1555543	*PTBP2*	1	C	A	0.06 (0.02)	0.59	1.00	0.58	1.00	0.59	1.00
rs2815752	*NEGR1*	1	A	G	0.13 (0.02)	0.63	1.00	0.63	1.00	0.62	1.00
rs2890652	*LRP1B*	2	C	T	0.09 (0.03)	0.17	0.99	0.17	0.99	0.17	0.97
rs887912	*FANCL*	2	T	C	0.10 (0.02)	0.30	0.99	0.30	0.99	0.28	0.99
rs713586	*RBJ*	2	C	T	0.14 (0.02)	0.47	1.00	0.48	1.00	0.46	0.95
rs2867125	*TMEM18*	2	C	T	0.31 (0.03)	0.82	1.00	0.81	1.00	0.82	1.00
rs13078807	*CADM2*	3	G	A	0.10 (0.02)	0.21	0.99	0.22	0.99	0.21	1.00
rs9816226	*ETV5*	3	T	A	0.14 (0.03)	0.82	0.97	0.82	0.97	0.82	0.95
rs13107325	*SLC39A8*	4	T	C	0.19 (0.04)	0.08	0.86	0.09	0.83	0.07	1.00
rs10938397	*GNPDA2*	4	G	A	0.18 (0.02)	0.44	0.98	0.44	0.99	0.43	0.85
rs4836133	*ZNF608*	5	A	C	0.07 (0.02)	0.48	0.94	0.52	0.93	0.48	0.90
rs2112347	*FLJ35779*	5	T	G	0.10 (0.02)	0.64	0.97	0.63	0.97	0.64	0.98
rs987237	*TFAP2B*	6	G	A	0.13 (0.03)	0.18	1.00	0.18	1.00	0.18	1.00
rs206936	*NUDT3*	6	G	A	0.06 (0.02)	0.20	1.00	0.20	1.00	0.20	0.98
rs10968576	*LRRN6C*	9	G	A	0.11 (0.02)	0.31	1.00	0.31	1.00	0.32	1.00
rs3817334	*MTCH2*	11	T	C	0.06 (0.02)	0.42	1.00	0.41	1.00	0.42	1.00
rs4929949	*RPL27A*	11	C	T	0.06 (0.02)	0.51	0.97	0.49	0.97	0.51	0.95
rs10767664	*BDNF*	11	A	T	0.19 (0.03)	0.78	1.00	0.78	1.00	0.79	1.00
rs7138803	*FAIM2*	12	A	G	0.12 (0.02)	0.38	1.00	0.39	1.00	0.38	1.00
rs4771122	*MTIF3*	13	G	A	0.09 (0.03)	0.22	0.95	0.21	0.95	0.23	0.94
rs11847697	*PRKD1*	14	T	C	0.17 (0.05)	0.05	0.85	0.05	0.80	0.04	0.95
rs10150332	*NRXN3*	14	C	T	0.13 (0.03)	0.22	1.00	0.20	1.00	0.22	0.99
rs2241423	*MAP2K5*	15	G	A	0.13 (0.02)	0.77	1.00	0.74	1.00	0.78	1.00
rs7359397	*SH2B1*	16	T	C	0.15 (0.02)	0.39	0.98	0.37	0.97	0.39	1.00
rs1558902	*FTO*	16	A	T	0.39 (0.02)	0.42	1.00	0.44	1.00	0.40	0.89
rs12444979	*GPRC5B*	16	C	T	0.17 (0.03)	0.86	0.99	0.86	0.98	0.86	0.99
rs571312	*MC4R*	18	A	C	0.23 (0.03)	0.24	1.00	0.24	1.00	0.24	0.96
rs29941	*KCTD15*	19	G	A	0.06 (0.02)	0.68	1.00	0.68	1.00	0.69	1.00
rs3810291	*TMEM160*	19	A	G	0.09 (0.02)	0.65	0.70	0.64	0.71	0.65	0.77
rs2287019	*QPCTL*	19	C	T	0.15 (0.03)	0.81	0.75	0.80	0.67	0.81	1.00

Chr: chromosome; EAF: effect allele frequency.

*Allele coding based on the forward strand.

†Effect sizes in kg/m^2 of BMI obtained from GWAS.

‡r^2 refers to the measurement of SNPs imputation quality.

Table S2 P-values for interactions between 32 SNPs and SSB intake on BMI in the NHS, HPFS and WGHS

SNP	Nearest Gene	NHS*	HPFS*	WGHS†	Pooled‡
rs543874	*SEC16B*	0.02	0.19	0.17	0.003
rs1514175	*TNNI3K*	0.24	0.81	0.07	0.21
rs1555543	*PTBP2*	0.32	0.09	0.09	0.26
rs2815752	*NEGR1*	0.99	0.93	0.38	0.68
rs2890652	*LRP1B*	0.68	0.87	0.08	0.81
rs887912	*FANCL*	0.81	0.98	0.47	0.91
rs713586	*RBJ*	0.09	0.38	0.75	0.05
rs2867125	*TMEM18*	0.75	0.91	0.21	0.88
rs13078807	*CADM2*	0.40	0.61	0.99	0.89
rs9816226	*ETV5*	0.30	0.95	0.95	0.56
rs13107325	*SLC39A8*	0.65	0.63	0.31	0.78
rs10938397	*GNPDA2*	0.76	0.30	0.51	0.51
rs4836133	*ZNF608*	0.12	0.31	0.88	0.08
rs2112347	*FLJ35779*	0.01	0.86	0.31	0.04
rs987237	*TFAP2B*	0.16	0.35	0.01	0.38
rs206936	*NUDT3*	0.19	0.45	0.25	0.28
rs10968576	*LRRN6C*	0.71	0.008	0.67	0.02
rs3817334	*MTCH2*	0.33	0.09	0.51	0.37
rs4929949	*RPL27A*	0.75	0.26	0.18	0.32
rs10767664	*BDNF*	0.43	0.33	0.66	0.26
rs7138803	*FAIM2*	0.32	0.10	0.09	0.03
rs4771122	*MTIF3*	0.23	0.65	0.51	0.43
rs11847697	*PRKD1*	0.26	0.14	0.59	0.86
rs10150332	*NRXN3*	0.38	0.6	0.50	0.98
rs2241423	*MAP2K5*	0.06	0.19	0.83	0.49
rs7359397	*SH2B1*	0.11	0.39	0.21	0.20
rs1558902	*FTO*	0.27	0.65	0.0003	0.03
rs12444979	*GPRC5B*	0.68	0.47	0.25	0.24
rs571312	*MC4R*	0.06	0.50	0.94	0.10
rs29941	*KCTD15*	0.70	0.87	0.91	0.67
rs3810291	*TMEM160*	0.41	0.47	0.90	0.89
rs2287019	*QPCTL*	0.73	0.31	0.45	0.51

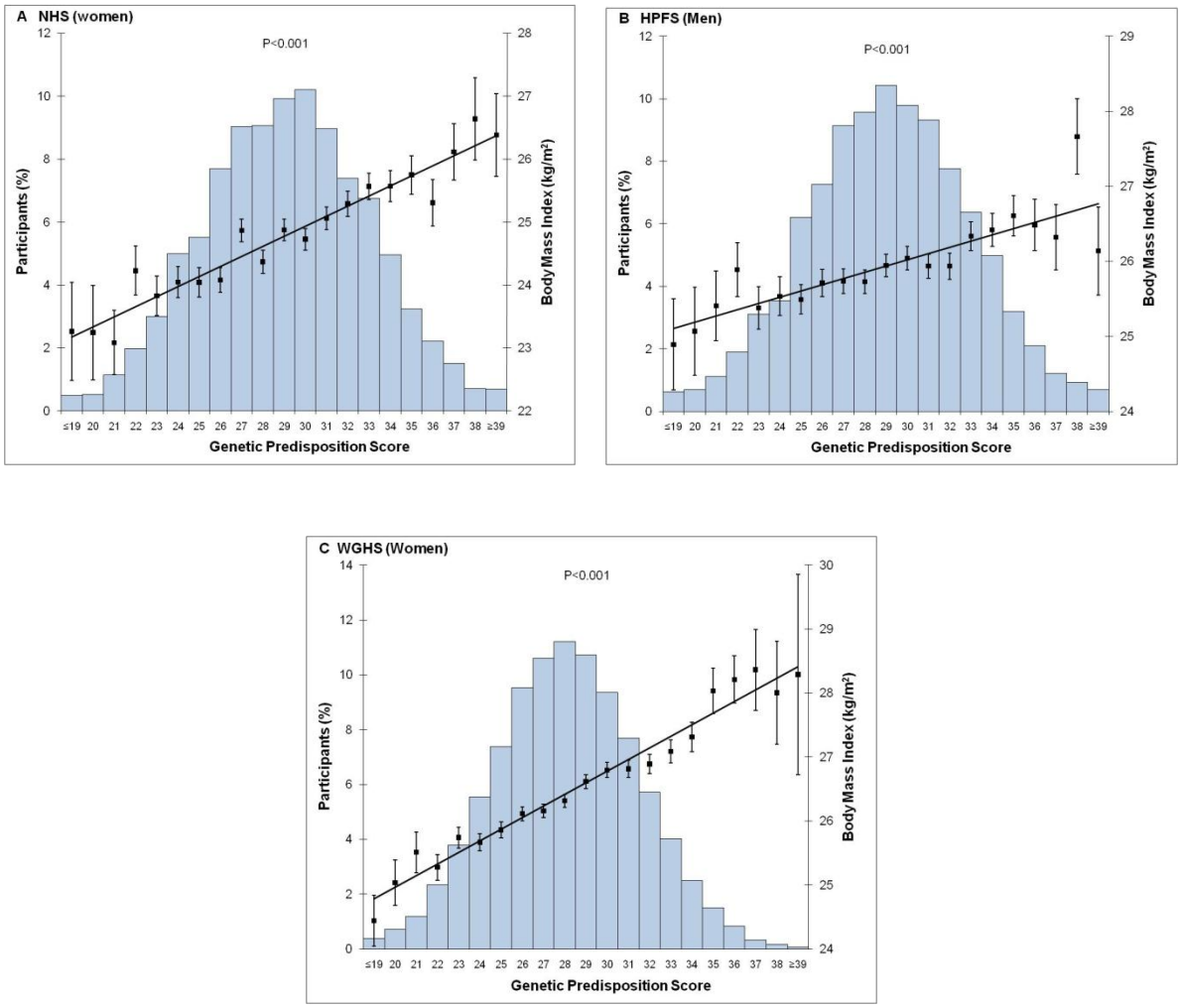
*Adjusted for age and genotyping data source.

†Adjusted for age, region and eigenvectors.

‡Results were pooled among the three cohorts by inverse-variance–weighted, fixed-effects meta-analyses

Figure S1 Genetic predisposition score and body mass index in three cohorts

The histograms represent the percentage of participants; and the means (±SE) of BMI are plotted, with the trend lines across the genetic predisposition score.



References

1. Colditz GA, Manson JE, Hankinson SE. The Nurses' Health Study: 20-year contribution to the understanding of health among women. J Womens Health 1997;6:49-62.

2. Rimm EB, Giovannucci EL, Willett WC, et al. Prospective study of alcohol consumption and risk of coronary disease in men. Lancet 1991;338:464-468.

3. Hunter DJ, Kraft P, Jacobs KB, et al. A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. Nat Genet 2007;39:870-874.

4. Qi L, Cornelis MC, Kraft P, et al. Genetic variants at 2q24 are associated with susceptibility to type 2 diabetes. Hum Mol Genet 2010;19:2706-2715.

5. Cornelis MC, Monda KL, Yu K, et al. Genome-wide meta-analysis identifies regions on 7p21 AHR and 15q24 CYP1A2 as determinants of habitual caffeine consumption. PLoS Genet 2011;7:e1002033.

6. Wiggs JL, Hee Kang J, Yaspan BL, et al. Common variants near CAV1 and CAV2 are associated with primary open-angle glaucoma in Caucasians from the USA. Hum Mol Genet 2011;20:4707-4713.

7. Jensen MK, Pers TH, Dworzynski P, Girman CJ, Brunak Sr, Rimm EB. Protein interaction-based genome-wide analysis of incident coronary heart disease. Circ Cardiovasc Genet 2011;4:549-556.

8. Ridker PM, Chasman DI, Zee RYL, et al. 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.

9. 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.

10. Willett WC, Sampson L, Stampfer MJ, et al. Reproducibility and validity of a semiquantitative food frequency questionnaire. Am J Epidemiol 1985;122:51-65.

11. Feskanich D, Rimm EB, Giovannucci EL, et al. Reproducibility and validity of food intake measurements from a semiquantitative food frequency questionnaire. J Am Diet Assoc 1993;93:790-796.

12. Rimm EB, Stampfer MJ, Colditz GA, Chute CG, Litin LB, Willett WC. Validity of self-reported waist and hip circumferences in men and women. Epidemiology 1990;1:466-1473.

13. Pirie P, Jacobs D, Jeffery R, Hannan P. Distortion in self-reported height and weight data. J Am Diet Assoc 1981;78:601-606.

14. Wolf AM, Hunter DJ, Colditz GA, et al. Reproducibility and validity of a self-administered physical activity questionnaire. Int J Epidemiol 1994;23:991-999.

15. McCullough ML, Feskanich D, Stampfer MJ, et al. Diet quality and major chronic disease risk in men and women: moving toward improved dietary guidance. Am J Clin Nutr 2002;76:1261-1271.

16. Weinstein AR, Sesso HD, Lee IM, et al. Relationship of physical activity vs body mass index with type 2 diabetes in women. JAMA 2004;292:1188-1194.

17. Ahmad T, Lee IM, Paré G, et al. Lifestyle interaction with fat mass and obesity-associated (FTO) genotype and risk of obesity in apparently healthy U.S. women. Diabetes Care 2011;34:675-680.

18. Speliotes EK, Willer CJ, Berndt SI, et al. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. Nat Genet 2010;42:937-948.