Plasma Vitamin C and Type 2 Diabetes: Genome-Wide Association Study and Mendelian Randomization Analysis in European Populations

Diabetes Care 2021;44:98–106 | https://doi.org/10.2337/dc20-1328

Ju-Sheng Zheng,[1,2] Jian'an Luan,[1] Eleni Sofianopoulou,[3] Fumiaki Imamura,[1] Isobel D. Stewart,[1] Felix R. Day,[1] Maik Pietzner,[1] Eleanor Wheeler,[1] Luca A. Lotta,[1] Thomas E. Gundersen,[4] Pilar Amiano,[5] Eva Ardanaz,[6,7,8] María-Dolores Chirlaque,[8,9] Guy Fagherazzi,[10,11] Paul W. Franks,[12] Rudolf Kaaks,[13] Nasser Laouali,[11] Francesca Romana Mancini,[11] Peter M. Nilsson,[12] N. Charlotte Onland-Moret,[14] Anja Olsen,[15,16] Kim Overvad,[16,17] Salvatore Panico,[18] Domenico Palli,[19] Fulvio Ricceri,[20,21] Olov Rolandsson,[22] Annemieke M.W. Spijkerman,[23] María-José Sánchez,[24] Matthias B. Schulze,[25,26,27] Núria Sala,[28] Sabina Sieri,[29] Anne Tjønneland,[16,30] Rosario Tumino,[31] Yvonne T. van der Schouw,[14] Elisabete Weiderpass,[32] Elio Riboli,[33] John Danesh,[3,34,35,36,37,38] Adam S. Butterworth,[3,34,36,37,38] Stephen J. Sharp,[1] Claudia Langenberg,[1] Nita G. Forouhi,[1] and Nicholas J. Wareham[1]

OBJECTIVE

Higher plasma vitamin C levels are associated with lower type 2 diabetes risk, but whether this association is causal is uncertain. To investigate this, we studied the association of genetically predicted plasma vitamin C with type 2 diabetes.

RESEARCH DESIGN AND METHODS

We conducted genome-wide association studies of plasma vitamin C among 52,018 individuals of European ancestry to discover novel genetic variants. We performed Mendelian randomization analyses to estimate the association of genetically predicted differences in plasma vitamin C with type 2 diabetes in up to 80,983 case participants and 842,909 noncase participants. We compared this estimate with the observational association between plasma vitamin C and incident type 2 diabetes, including 8,133 case participants and 11,073 noncase participants.

RESULTS

We identified 11 genomic regions associated with plasma vitamin C ($P < 5 \times 10^{-8}$), with the strongest signal at *SLC23A1*, and 10 novel genetic loci including *SLC23A3*, *CHPT1*, *BCAS3*, *SNRPF*, *RER1*, *MAF*, *GSTA5*, *RGS14*, *AKT1*, and *FADS1*. Plasma vitamin C was inversely associated with type 2 diabetes (hazard ratio per SD 0.88; 95% CI 0.82, 0.94), but there was no association between genetically predicted plasma vitamin C (excluding *FADS1* variant due to its apparent pleiotropic effect) and type 2 diabetes (1.03; 95% CI 0.96, 1.10).

CONCLUSIONS

These findings indicate discordance between biochemically measured and genetically predicted plasma vitamin C levels in the association with type 2 diabetes among European populations. The null Mendelian randomization findings provide no strong evidence to suggest the use of vitamin C supplementation for type 2 diabetes prevention.

Observational evidence suggests that plasma vitamin C (ascorbic acid) is inversely associated with the incidence of type 2 diabetes (1). However, evidence from randomized controlled trials (RCTs) of supplementation with vitamin C suggests no effect of supplementation on glycemia in individuals without diabetes (2) or on type 2 diabetes incidence among women at high risk of cardiovascular disease (CVD) (3).

[1]*Medical Research Council Epidemiology Unit, University of Cambridge, Cambridge, U.K.*
[2]*Key Laboratory of Growth Regulation and Translational Research of Zhejiang Province, School of Life Sciences, Westlake University, Hangzhou, China*
[3]*British Heart Foundation Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge, U.K.*
[4]*Vitas Ltd, Oslo, Norway*
[5]*Ministry of Health of the Basque Government, Public Health Division of Gipuzkoa, Biodonostia Health Research Institute, Donostia-San Sebastian, Spain*
[6]*Navarra Public Health Institute, Pamplona, Spain*
[7]*IdiSNA, Navarra Institute for Health Research, Pamplona, Spain*
[8]*CIBER in Epidemiology and Public Health (CIBERESP), Madrid, Spain*
[9]*Department of Epidemiology, Regional Health Council, Instituto Murciano de Investigatión Biosanitaria (IMIB)-Arrixaca, Murcia University, Murcia, Spain*
[10]*Digital Epidemiology and e-Health Research Hub, Department of Population Health, Luxembourg Institute of Health, Strassen, Luxembourg, France*

Important limitations of the RCT evidence have included small sample size of some individual trials, uncertainty of the optimal dose, baseline status, the inability to separate out the effects of vitamin C when these were used in combination with other vitamins/antioxidants, and declining compliance with the intervention over time. Conducting further RCTs to assess an impact of vitamin C on type 2 diabetes is challenging, requiring large sample size, long follow-up duration, and high cost. Meanwhile, it remains unclear whether there is a causal role of vitamin C in the prevention of type 2 diabetes.

Mendelian randomization (MR) is an approach that can help to assess the potential causal relevance of a risk exposure for a disease end point, taking advantage of the random assortment of alleles at conception (4). This means that the association between a disease and a genetically determined phenotype is unlikely to be affected by reverse causation or by confounding. A limitation of MR is weak instrument bias if the genetic variants explain a small fraction of the variation in the risk factor of interest, but this could be partially addressed through the use of multiple independent genetic variants identified in genome-wide association studies (GWAS) (5). Thus far, only one genetic variant (rs33972313) at *SLC23A1* has been identified to be associated with plasma vitamin C levels (6), based on a candidate gene approach, and no GWAS has been published so far.

The objective of this study was to study the relationship between genetic variants and vitamin C levels and to develop a genetic instrument to be used in an MR analysis of the association of genetically predicted levels of plasma vitamin C with type 2 diabetes risk.

RESEARCH DESIGN AND METHODS

Participating Studies and Genotyping

Figure 1 provides an overview of the participating cohorts, and our overall approach to data inclusion and analysis. The three steps in the process were as follows: *1*) we conducted GWAS to identify genetic variants associated with plasma vitamin C levels; *2*) using the GWAS-discovered genetic variants, we examined the association between genetically predicted plasma vitamin C and type 2 diabetes risk by MR analysis; and *3*) for comparison with MR results, we performed analyses to evaluate the observational association between plasma vitamin C and incident type 2 diabetes. All of the studies included in the present analyses were approved by local ethics committees, and participants provided written informed consent.

We performed meta-analysis of GWAS of plasma vitamin C levels in up to 52,018 individuals from the Fenland study (n = 10,771) (2), European Prospective Investigation into Cancer and Nutrition (EPIC)-InterAct study (n = 16,841) (7), EPIC Norfolk study (8) (n = 16,756, excluding duplicated samples with EPIC-InterAct), and the EPIC-CVD study (9) (n = 7,650, excluding duplicated samples with EPIC-InterAct or EPIC-Norfolk).

We then estimated the association of lead genetic variants from the vitamin C GWAS meta-analysis with the risk of type 2 diabetes in a meta-analysis including participants (80,983 case participants with type 2 diabetes, and 842,909 noncase participants) from UK Biobank (10), DIAbetes Meta-ANalysis of Trans-Ethnic association studies (DIAMANTE) (European) (11), and EPIC-Norfolk study (additional case participants with type 2 diabetes not included in the DIAMANTE study: 1,220 case participants with diabetes and 18,026 noncase participants) (8).

The details of each participating study and genotyping strategy and quality control procedures are described in Supplementary Table 1 and Supplementary

[11]*Center of Epidemiology and Population Health UMR 1018, INSERM, Paris South - Paris Saclay University, Gustave Roussy Institute, Villejuif, France*
[12]*Department of Clinical Sciences, Lund University, Malmö, Sweden*
[13]*Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany*
[14]*Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands*
[15]*Danish Cancer Society Research Center, Copenhagen, Denmark*
[16]*Department of Public Health, Aarhus University, Aarhus, Denmark*
[17]*Department of Cardiology, Aalborg University Hospital, Aarhus, Denmark*
[18]*Dipartimento di Medicina Clinica e Chirurgia, Federico II University, Naples, Italy*
[19]*Cancer Risk Factors and Life-Style Epidemiology Unit, Institute for Cancer Research, Prevention and Clinical Network - ISPRO, Florence, Italy*
[20]*Department of Clinical and Biological Sciences, University of Turin, Orbassano, Turin, Italy*
[21]*Unit of Epidemiology, Regional Health Service ASL TO3, Grugliasco, Turin, Italy*
[22]*Department of Public Health and Clinical Medicine, Family Medicine, Umeå University, Umeå, Sweden*
[23]*National Institute for Public Health and the Environment (RIVM), Bilthoven, the Netherlands*
[24]*Andalusian School of Public Health, Granada, Spain*
[25]*Department of Molecular Epidemiology, German Institute of Human Nutrition Potsdam-Rehbruecke, Nuthetal, Germany*
[26]*German Center for Diabetes Research (DZD), München-Neuherberg, Germany*
[27]*Institute of Nutrition Science, University of Potsdam, Nuthetal, Germany*
[28]*Unit of Nutrition and Cancer, Cancer Epidemiology Research Program and Translational Research Laboratory; Catalan Institute of Oncology - ICO, Group of Research on Nutrition and Cancer, Bellvitge Biomedical Research Institute (IDIBELL), L'Hospitalet of Llobregat, Barcelona, Spain*
[29]*Epidemiology and Prevention Unit, Fondazione Istituto di Ricovero e Cura a Carattere Scientifico (IRCCS) Istituto Nazionale dei Tumori di Milano Via Venezian, Milan, Italy*
[30]*Institute of Public Health, University of Copenhagen, Copenhagen, Denmark*
[31]*Cancer Registry and Histopathology Department, Azienda Sanitaria Provinciale (ASP), Ragusa, Italy*
[32]*International Agency for Research on Cancer, Lyon, France*
[33]*School of Public Health, Imperial College, London, U.K.*
[34]*National Institute for Health Research Blood and Transplant Research Unit in Donor Health and Genomics, University of Cambridge, Cambridge, U.K.*
[35]*British Heart Foundation Center of Research Excellence, University of Cambridge, Cambridge, U.K.*
[36]*Department of Human Genetics, Wellcome Sanger Institute, Hinxton, U.K.*
[37]*National Institute for Health Research Cambridge Biomedical Research Center, University of Cambridge and Cambridge University Hospitals, Cambridge, U.K.*
[38]*Health Data Research UK Cambridge, Wellcome Genome Campus and University of Cambridge, Cambridge, U.K.*

Corresponding authors: Nita G. Forouhi, nita.forouhi@mrc-epid.cam.ac.uk, and Nicholas J. Wareham, nick.wareham@mrc-epid.cam.ac.uk

Received 2 June 2020 and accepted 15 October 2020

This article contains supplementary material online at https://doi.org/10.2337/figshare.13105559.

C.L., N.G.F., and N.J.W. contributed equally.

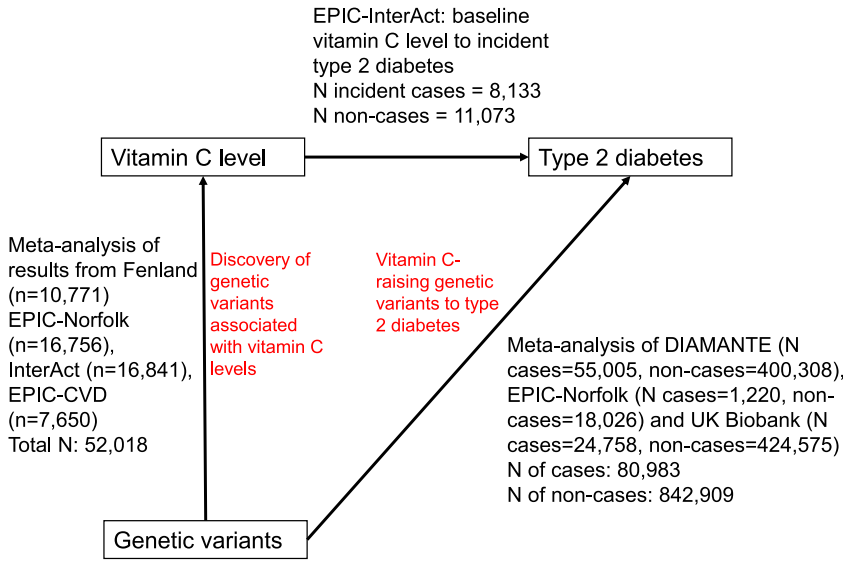


Figure 1—Design of the study for MR analysis of plasma vitamin C with type 2 diabetes.

Text. Briefly, the Fenland study is an ongoing, population-based cohort study including 12,435 adults at baseline in Cambridgeshire, U.K. (2). The GWAS included 10,771 Fenland participants with both genotype and vitamin C data. Genotype imputation was performed to the Haplotype Reference Consortium reference panel using IMPUTE v4 or the Sanger imputation server.

EPIC-InterAct is a case-cohort study including 12,403 participants with incident type 2 diabetes from among 340,234 participants across eight European countries (France, Italy, Spain, U.K., Netherlands, Germany, Sweden, and Denmark) and 16,835 subcohort participants randomly selected from the full cohort, of whom 16,154 remained after exclusion of participants with prevalent diabetes (7); 16,841 participants with both genotype and vitamin C data contributed to the present GWAS. EPIC-Norfolk is an ongoing U.K.-based prospective cohort study with 25,639 men and women aged 40–79 at baseline (8), among whom 16,756 participants had the required data to be included in the vitamin C GWAS. EPIC-CVD is also a case-cohort study nested within the EPIC study, with a random subcohort of 18,249 participants and 24,557 adults who later developed CVD during the follow-up. EPIC-CVD shared the random subcohort with the EPIC-InterAct for eight participating countries (9), and a total of 7,650 EPIC-CVD participants contributed to the present vitamin C GWAS. Genotype imputation of the EPIC-InterAct, EPIC-Norfolk, and EPIC-CVD was performed to the Haplotype Reference Consortium reference panel using IMPUTE v4 software.

The UK Biobank study is a population-based cohort of ∼0.5 million U.K. individuals aged 40–69 years recruited between 2006 and 2010 across the U.K. (10). Genotype data and prevalent type 2 diabetes information were both available among a total of 449,333 individuals in the UK Biobank data set (24,758 case participants and 424,575 noncase participants). DIAMANTE is a consortium that published a large-scale meta-analysis of GWAS of type 2 diabetes in individuals of different ethnicities. Summary results from the DIAMANTE meta-analyses among Europeans were made publicly available (11).

Plasma Vitamin C Measurement

Plasma vitamin C was measured using high-performance liquid chromatography with ultraviolet detection in EPIC-InterAct and EPIC-CVD using samples that had been stored at −196°C (−150°C for Danish samples) (12). In EPIC-Norfolk and Fenland, plasma vitamin C was measured with a fluorometric assay using samples stored at −70°C. For each of these studies, plasma was stabilized in a standardized volume of metaphosphoric acid. The distribution of plasma vitamin C levels in each participating cohort is shown in Supplementary Fig. 1.

Statistical Analysis

Meta-analysis of GWAS

For the GWAS of plasma vitamin C among each participating cohort, standardized residuals (in SD units) for plasma vitamin C were calculated, adjusting for age, sex, and study center (where appropriate). Moreover, in each cohort, GWAS was performed by running linear regression with SNPTEST (v2.5.4) assuming an additive effect, adjusting for the first 10 genetic principal components of ancestry within each cohort. Meta-analysis of the GWAS results was conducted by combining β coefficients and SEs using inverse variance-weighted fixed-effect meta-analysis across the participating studies using METAL (13). Associated loci were identified using the conventional threshold for genome-wide statistical significance ($P < 5 \times 10^{-8}$). At each locus, a lead single-nucleotide polymorphism (SNP) was identified as the SNP with the lowest *P* value within a 1 megabase-pair window. Functional annotation and pathway analysis of the GWAS results were performed using MAGENTA and DEPICT (14,15) (Supplementary Text).

Linear regression models were used to estimate the variance in plasma vitamin C explained by the identified lead SNPs in each of Fenland, EPIC-InterAct, EPIC-Norfolk, and EPIC-CVD studies. We also calculated the *F* statistic in the EPIC-Norfolk study to evaluate the strength of the instrument. We estimated the genetic correlations of vitamin C levels with type 2 diabetes and related glycemic markers (fasting glucose, insulin, 2-h glucose, HOMA of insulin resistance [HOMA-IR], HOMA of β-cell function [HOMA-β], and hemoglobin A_{1c} [HbA_{1c}]) using the meta-analyzed GWAS results and linkage disequilibrium score regression on the LD Hub platform (16), and the results were corrected for multiple testing.

Association of Genetically Predicted Levels of Vitamin C With Type 2 Diabetes: MR Analysis

We used two-sample MR analysis to test for an association between a genetically predicted difference in plasma vitamin C levels and type 2 diabetes risk. To this end, we used estimates of each of the "SNP to vitamin C" and "SNP to diabetes" associations to examine the association between a genetically predicted 1-SD difference in plasma vitamin C and type 2 diabetes risk. Estimates from multiple SNPs were pooled using an inverse variance–weighted method (5). MR assumptions are violated if there is horizontal

pleiotropy. Therefore, in the EPIC-InterAct study, we constructed a vitamin C-raising genetic score based on the identified genetic variants and examined its association with anthropometric, lifestyle, dietary factors, and lipid biomarkers to examine the potential pleiotropic effect of the genetic score. The potential pleiotropic effect of the identified SNPs was further examined by investigating their associations with 174 blood metabolites using GWAS summary statistics from genome-wide meta-analysis of these metabolites across Fenland, EPIC-Norfolk, and Efficiency and Safety of Varying the Frequency of Whole Blood Donation (INTERVAL) studies, as previously reported (17).

In addition, directional pleiotropy for the used SNPs was tested using MR-Egger intercept (18). We used MR-Egger regression to detect and adjust for potential unbalanced pleiotropy in the MR analysis (18), and the Cochran *Q* statistic to examine the heterogeneity of the association between different genetic variants, because different genetic variants may represent different biological pathways. We also used a weighted median MR method (19) as a sensitivity analysis and highlighted the weighted median results if significant heterogeneity of the associations among different genetic variants was observed. We conducted a reverse MR analysis to examine the potential association of genetically predicted risk of type 2 diabetes on plasma vitamin C concentrations, using a score composed of 231 type 2 diabetes SNPs identified in a recent GWAS (11).

We calculated the estimates of a genetically predicted 1-SD difference in plasma vitamin C with several glycemic traits using data from Meta-Analysis of Glucose and Insulin-related traits Consortium (MAGIC): fasting glucose (n = 133,010), 2-h glucose (n = 42,854), fasting insulin (n = 108,557), HOMA-IR (n = 46,186), HOMA-β (n = 46,186), and HbA_{1c} (n = 123,665).

Observational Association of Plasma Vitamin C With Type 2 Diabetes

For the observational association of plasma vitamin C and incident type 2 diabetes in the EPIC-InterAct study, we used Prentice-weighted Cox regression, accounting for the design of a case-cohort study, to estimate the country-specific hazard ratios and 95% CIs for associations per 1-SD difference (SD calculated in the subcohort) of plasma vitamin C with incident type 2 diabetes. The estimates were adjusted for potential confounding factors of age as the underlying time scale, sex, center, physical activity, smoking status, employment status, marital status, education level, alcohol intake, total energy intake, individual plasma carotenoids, BMI, and waist circumference. The results across countries were pooled by random-effects meta-analysis.

Statistical analyses were performed in Stata 14 (StataCorp, College Station, TX), R 3.2.2 (R Foundation for Statistical Computing), and METAL 2011-03-25.

Data and Resource Availability

GWAS summary statistics for the plasma vitamin C levels can be accessed at figshare.com (https://doi.org/10.6084/m9.figshare.13227443.v1).

RESULTS

Population Characteristics

Plasma vitamin C was associated with a variety of anthropometric, lifestyle behavioral factors (education, employment status, physical activity, and smoking status), dietary factors, and lipid biomarkers (Supplementary Table 2). In contrast, the vitamin C-raising genetic score calculated from the vitamin C-raising alleles (generated from the GWAS results) was positively associated only with plasma vitamin C levels, and it was not associated with other factors including anthropometric, lifestyle, diet factors, or lipid biomarkers (Supplementary Table 3).

Meta-analysis of GWAS of Plasma Vitamin C

Genome-wide meta-analysis identified 11 independent genomic loci associated with plasma vitamin C levels (Table 1), of which 10 were novel, while *SLC23A1* (6) was a known plasma vitamin C locus (Table 1 and Supplementary Figs. 2 and 3). The regional plots of the identified loci are presented in Supplementary Figs. 4–12. The heterogeneity for the association of the identified lead SNPs with plasma vitamin C across individual cohorts was low (Supplementary Fig. 13). The strongest signal was rs33972313 within the *SLC23A1* gene ($P = 4.61 \times 10^{-90}$) on chromosome 5, a missense mutation that was reported to be associated with circulating vitamin C levels in a previous candidate gene study (6). *SLC23A1* encodes solute carrier family 23 member 1, also known as sodium-dependent vitamin C transporter 1 (SVCT1), which is responsible for the uptake of vitamin C into target

Table 1—Lead SNPs identified in the genome-wide meta-analysis of plasma vitamin C across four cohort studies (Fenland, EPIC-Norfolk, EPIC-InterAct, and EPIC-CVD)*

Chromosome	Position (GRCh37)	Lead SNPs	Effect allele	Other allele	EAF	β	SE	*P* value	Nearest gene
1	2330190	rs6693447	T	G	0.551	0.039	0.006	6.25×10^{-10}	RER1
2	220031255	rs13028225	T	C	0.857	0.102	0.009	2.38×10^{-30}	SLC23A3
5	138715502	rs33972313	C	T	0.968	0.360	0.018	4.61×10^{-90}	SLC23A1
5	176799992	rs10051765	C	T	0.342	0.039	0.007	3.64×10^{-9}	RGS14
6	52725787	rs7740812	G	A	0.594	0.038	0.006	1.88×10^{-9}	GSTA5
11	61570783	rs174547	C	T	0.328	0.036	0.007	3.84×10^{-8}	FADS1
12	96249111	rs117885456	A	G	0.087	0.078	0.012	1.70×10^{-11}	SNRPF
12	102093459	rs2559850	A	G	0.598	0.058	0.006	6.30×10^{-20}	CHPT1
14	105253581	rs10136000	A	G	0.283	0.040	0.007	1.33×10^{-8}	AKT1
16	79740541	rs56738967	C	G	0.321	0.041	0.007	7.62×10^{-10}	MAF
17	59456589	rs9895661	T	C	0.817	0.063	0.008	1.05×10^{-14}	BCAS3

EAF, effect allele frequency. *The β coefficients are in SD unit per allele; effect allele is the vitamin C-raising allele. The pooled sample size for the genome-wide meta-analyses is 52,018.

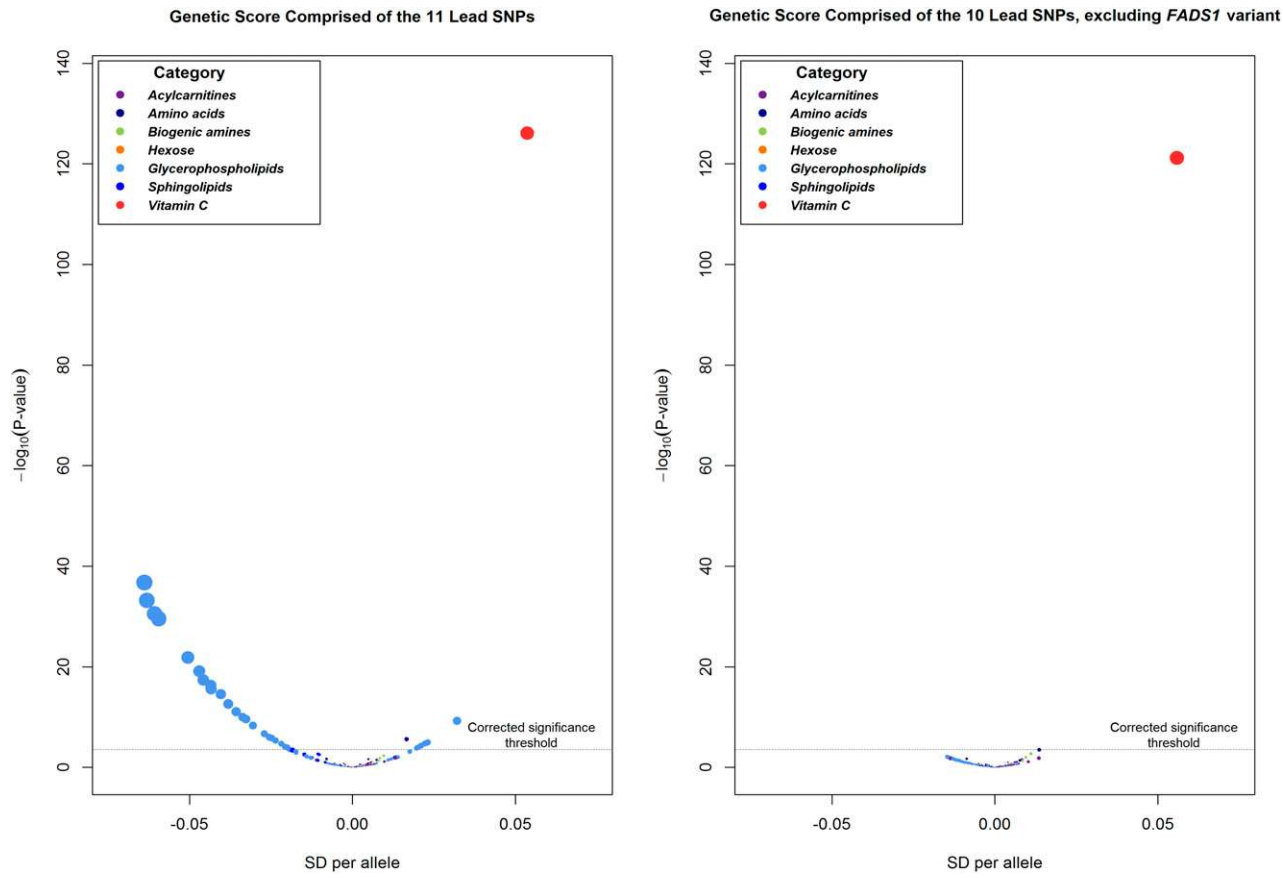


Figure 2—Volcano plot for the association of the plasma vitamin C genetic score with blood metabolites. Associations of the 11 variant scores and the score excluding the *FADS1* variant were estimated using fixed-effects meta-analyses of individual variant-metabolite associations aligned to the plasma vitamin C-raising alleles. The horizontal dashed line indicates a Bonferroni-corrected significance threshold (corrected for 175 tests), and diameters of points are proportional to abs(effect size)*50, where "abs" is the function to calculate the absolute value of the effect size.

tissues (20). The second strongest signal rs13028225 ($P = 2.38 \times 10^{-30}$) was within the *SLC23A3* gene on chromosome 2, which encodes sodium-dependent vitamin C transporter 3 (SVCT3), belonging to the same family as *SLC23A3*.

We identified further genomic loci with less obvious linkage to the metabolism of vitamin C, such as *RER1* (rs6693447, $P = 6.25 \times 10^{-10}$, encoding retention in endoplasmic reticulum sorting receptor 1), *RGS14* (rs10051765, $P = 3.64 \times 10^{-9}$, encoding regulator of G protein signaling 14), *GSTA5* (rs7740812, $P = 1.88 \times 10^{-9}$, encoding glutathione *S*-transferase α 5), *FADS1* (rs174547, $P = 3.84 \times 10^{-8}$, encoding fatty acid desaturase 1), *SNRPF* (rs117885456, $P = 1.7 \times 10^{-11}$, encoding small nuclear ribonucleoprotein polypeptide F), *CHPT1* (rs2559850, $P = 6.3 \times 10^{-20}$, encoding choline phosphotransferase 1), *AKT1* (rs10136000, $P = 1.33 \times 10^{-8}$, encoding serine-threonine protein kinase), *MAF* (rs56738967, $P = 7.62 \times 10^{-10}$, encoding MAF bZIP transcription factor), and *BCAS3* (rs9895661, $P = 1.05 \times 10^{-14}$, encoding microtubule associated cell migration factor).

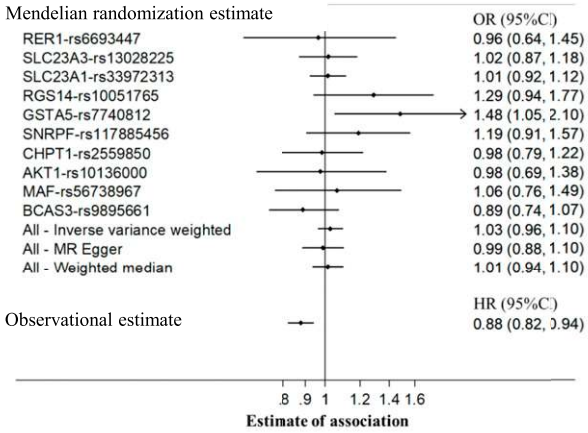
Results of pathway analysis and tissue and gene set enrichment analysis did not reveal a specific pathway or enriched tissues/gene sets (Supplementary Tables 4 and 5). The variance of plasma vitamin C explained by the 11 lead SNPs was 1.87% on average, with 2.51% in Fenland, 1.74% in EPIC-Norfolk, 1.61% in the InterAct subcohort, 1.63% in the InterAct nonsubcohort, 1.07% in the EPIC-CVD subcohort, and 0.7% in the EPIC-CVD nonsubcohort. The *F* statistic of the vitamin C genetic score for MR analysis was 30.5. The estimates of genome-wide genetic correlations of vitamin C levels with type 2 diabetes and glycemic traits were significant only for fasting insulin ($r_{genetics} = -0.22$, $P = 0.005$) after controlling for multiple testing (Supplementary Table 6).

Association of Plasma Vitamin C–Raising Alleles With Type 2 Diabetes

From among 11 SNPs that were significant at the genome-wide level, there was a pleiotropic effect of the variant at *FADS1*, which was associated with a large number of glycerophospholipids or sphingolipids (Fig. 2). Thus, we excluded this SNP from our MR analysis, using the other 10 SNPs in a genetic instrument to examine the causal association of plasma vitamin C with type 2 diabetes (Supplementary Table 7).

Genetic predisposition to a higher level of plasma vitamin C was not associated with odds of type 2 diabetes, with an estimated odds ratio of 1.03 (95% CI 0.96, 1.10) per 1-SD difference in plasma vitamin C level (Fig. 3). There was no evidence of directional horizontal pleiotropy ($P = 0.41$ for the test of MR-Egger intercept) and heterogeneity for the genetic instrument used (Supplementary Table 8), which confirmed the validity of our genetic instrument. Sensitivity analysis using different methods (MR-Egger, weighted median) of MR did not find a substantial difference compared with the inverse variance–weighted method. The reverse MR analysis suggested that genetically higher risk of type 2 diabetes

A Estimate of the Mendelian randomization and observational analysis for the plasma vitamin C and risk of type 2 diabetes



B Scatter plot of genetic association with type 2 diabetes against genetic association with the plasma vitamin C levels

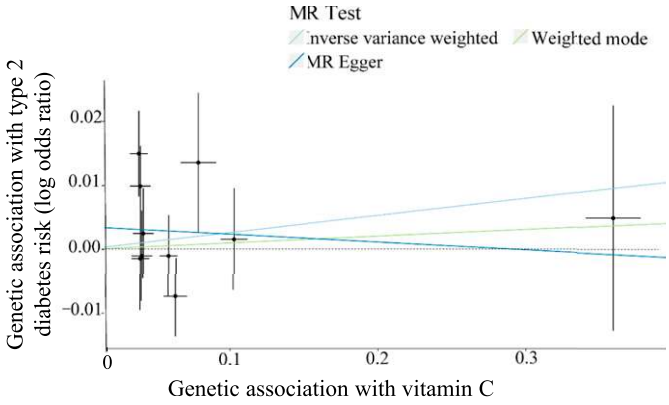


Figure 3—Genetically predicted and observational associations of plasma vitamin C with type 2 diabetes. *A*: MR and observational estimate of plasma vitamin C with type 2 diabetes. MR estimate of the association between a genetically predicted difference of 1 SD in plasma vitamin C and type 2 diabetes (odds ratio [OR]) was determined using different methods: inverse-variance weighted, MR Egger, and weighted median. Ten lead genetic variants (excluding *FADS1* variant due to pleiotropic effects) identified in the present genome-wide meta-analysis were used for the MR analysis. Observational estimate of the association of plasma vitamin C with type 2 diabetes (hazard ratio [HR]) was estimated in EPIC-InterAct using Prentice-weighted Cox regression, adjusting for potential confounders. *B*: Scatter plot of genetic association with type 2 diabetes and genetic association with the plasma vitamin C levels.

was not significantly associated with plasma vitamin C levels (Supplementary Fig. 14). In the observational analysis, a 1-SD difference in plasma vitamin C level was associated with a lower hazard of type 2 diabetes (hazard ratio 0.88; 95% CI 0.82, 0.94) (Fig. 3).

Association of Vitamin C-Raising Genetic Risk Score With Glycemic Traits

In the MR analysis, a genetically predicted higher plasma vitamin C level was not significantly associated with any of the examined glycemic markers, including fasting insulin, glucose, 2-h glucose, HOMA-IR, HOMA-β, and HbA_{1c} (Supplementary Fig. 15). This is consistent with the lack of association seen for type 2 diabetes. The associations of individual vitamin C-raising alleles with these traits are presented in Supplementary Fig. 16.

CONCLUSIONS

This study represents the first GWAS to be conducted for plasma vitamin C, identifying 11 genomic regions associated with plasma vitamin C, including 10 that were novel loci and confirming the previously identified candidate locus at *SLC23A1* (rs33972313). The results from using these newly identified genome-wide significant lead SNPs at corresponding loci in a genetic instrument suggested that a genetically higher level of plasma vitamin C was not significantly associated with type 2 diabetes risk or with glycemic traits.

Findings in Context of Other Evidence: MR Findings

There is no prior published MR analysis of plasma vitamin C and type 2 diabetes or glycemic traits to compare with our current findings. However, our finding of no evidence for an association of genetic markers of plasma vitamin C with type 2 diabetes or with any of the glycemic traits is consistent with prior null findings from RCTs of vitamin C supplementation (2,3). An obvious explanation for the discrepancy between the observational and the MR or RCT evidence is likely to be confounding by several factors, including socioeconomic, dietary, and lifestyle behavioral factors, as previously discussed (21), and as we found in the current analyses.

Findings in Context of Other Evidence: GWAS Findings

Our novel GWAS findings stimulate greater understanding of the biology of plasma vitamin C. Prior evidence suggests that vitamin C homeostasis is mainly regulated by two categories of genes. Some genes are involved in direct transport and regulation of vitamin C concentrations, including vitamin C absorption from food, renal reabsorption, and cellular uptake, particularly the sodium-dependent vitamin C transporters of the SLC23 family (e.g., *SLC23A1* and *SLC23A2*), whereas others are related to antioxidant and oxidative stress (22). *SLC23A1* encodes a well-known protein SVCT1 affecting the circulating vitamin C levels (6,22), and the lead SNP rs33972313 identified in the current study also showed consistent, reproducible evidence in a previous study (6).

Unlike *SLC23A1*, we did not identify SNPs of genome-wide significance at *SLC23A2*, although both genes encode SVCTs, which are specific for vitamin C transportation. It may be that SVCT2 encoded by *SLC23A2* mainly regulates tissue levels of vitamin C and that its impact on circulating vitamin C is minimal, as indicated in a recent review (22). Notably, we identified GWAS hits at *SLC23A3* associated with plasma vitamin

C levels. *SLC23A3* encodes SVCT3, which is an orphan transporter, and its functional role is largely unknown (23). There is evidence showing that SVCT3 is mainly expressed in the kidney (23) and that the *SLC23A3* mRNA expression level is highly correlated with the estimated glomerular filtration rate, a marker of renal function (24). Therefore, genetic variation at *SLC23A3* may affect circulating vitamin C levels by regulating renal function and related vitamin C reabsorption by the kidney. In addition, we identified several other genomic regions that might affect vitamin C levels potentially through regulating renal function and renal reabsorption of vitamin C, including *RGS14*, *MAF*, and *BCAS3* (25).

Other genetic loci identified in the present GWAS may be related to antioxidant and oxidative stress. The *GSTA5* gene encodes glutathione *S*-transferase (GST)-α5, a detoxifying enzyme belonging to the GST families, which contributes to the glutathione–vitamin C antioxidant cycle (22). Genetic variants at *FADS1* were associated with lipid peroxides, as well as other oxidative stress markers (26), and might influence vitamin C levels through the oxidative stress-related pathway. The *RER1* may be related to endoplasmic reticulum stress, which is closely related to oxidative stress (27). Genetic variation at *CHPT1* may lead to disruption of phosphatidylcholine synthesis, which is closely related to the antioxidant system (28). Finally, we identified genetic variants at *SNRPF* gene associated with plasma vitamin C, but a potentially functional link between this gene and vitamin C metabolism is unclear.

Strengths and Limitations

This study has several strengths. We conducted the first GWAS for plasma vitamin C using a large sample from multiple cohorts and identified 10 novel genetic loci for plasma vitamin C in addition to confirming the single variant described previously from candidate gene analysis (6). The current work expands our understanding about vitamin C biology, and it enabled the derivation of a genetic instrument to use in MR analysis. Our study is the first large MR analysis for circulating vitamin C and type 2 diabetes, with a large number of case participants in the genotype–type 2 diabetes association. We examined the association of genetically predicted vitamin C levels with a variety of diabetes-related traits, which also confirmed the observed null results for type 2 diabetes.

Limitations of our work include that despite the large GWAS meta-analysis we conducted, the variation of vitamin C explained by the SNPs used in the MR was still relatively small. The low variability of plasma vitamin C explained by these SNPs may limit the statistical power and precision for the MR analysis. Our analyses included White European populations; hence, our findings may not be generalizable to other populations. It is possible that pleiotropic effects of genetic variants used in MR may potentially exist, although we did not find evidence of pleiotropic effects (i.e., none of the SNPs used in the MR were associated with tested confounders) or heterogeneity across the genetic variants. Therefore, confounding may not be a substantial concern, and our instrument is specific for vitamin C, which fulfils an important criterion to identify causality. In addition, the MR estimate may be biased by gene-environment interaction or canalization, which should be investigated in future research.

Implications of the Current Findings

Clarification of the potential role of vitamin C in type 2 diabetes development is important as the prevalence of type 2 diabetes continues to expand globally and identification of modifiable risk factors is critical to curb the type 2 diabetes pandemic. Despite interest in vitamin C supplement use, the current findings indicate that there is unlikely to be a causal role of vitamin C as an isolated micronutrient on type 2 diabetes and its related traits in general populations. This should be distinguished from the use of plasma vitamin C as a biomarker to objectively estimate or be a proxy indicator for the consumption of fruit and vegetables (29). The latter provides support for the previous interpretation that the inverse observational prospective associations between plasma vitamin C and cardiometabolic diseases, including type 2 diabetes (1,30–32), reflect the potential benefits of promoting higher dietary consumption of fruit and vegetables. It is important to clarify the two distinct phenomena: the current findings suggest no evidence for an association with type 2 diabetes of genetically predicted plasma vitamin C level as a nutrient per se, but this does not rule out the utility of plasma vitamin C as a valid biomarker of fruit and vegetable intake, reducing measurement error and recall bias of subjective reporting. It is likely that a synergetic effect of various constituents in fruits and vegetables (including vitamins, minerals, numerous phytochemicals, and dietary fiber) rather than vitamin C alone may be responsible for observed inverse association of fruit and vegetable consumption on type 2 diabetes risk.

In summary, using a genome-wide meta-analysis approach, we identified 10 novel genetic loci and confirmed a known locus at *SLC23A1* for plasma vitamin C levels. With these newly identified genetic variants as a genetic instrument, our MR analysis found that there was no evidence to support a causal association between plasma vitamin C and type 2 diabetes or with related metabolic traits among European populations. These findings suggest that previous observations on higher vitamin C levels and lower type 2 diabetes risk may be attributed to a dietary pattern high in fruit and vegetable intake rather than to vitamin C level as an isolated nutrient. The implication of the current study is that vitamin C supplementation aiming to increase circulating vitamin C levels is unlikely to be helpful for the prevention of type 2 diabetes among individuals of European ancestries.

Acknowledgments. The authors thank all EPIC participants and staff for their contribution to the study. The authors thank Nicola Kerrison (MRC Epidemiology Unit, Cambridge) for managing the data for the InterAct Project. The authors thank the technical and functional operational teams of the Medical Research Council Epidemiology Unit and laboratory team at VITAS AS, Norway, for the measurements of plasma vitamin C. The authors thank staff from the EPIC-CVD and EPIC-InterAct Coordinating Centers for performing sample preparation and data-handling work, particularly Sarah Spackman (EPIC-CVD Data Manager), and Cambridge Genomic Services for genotyping.

Funding. The InterAct project was funded by the European Union Sixth Framework Programme (grant number LSHM_CT_2006_037197). In addition, InterAct investigators acknowledge funding from the following agencies: Medical Research Council Epidemiology Unit MC_UU_12015/1 and MC_UU_12015/5 (J.-S.Z., F.I., N.G.F., N.J.W.), and National Institute for Health Research Biomedical Research Center Cambridge: Nutrition, Diet, and Lifestyle Research Theme (IS-BRC-1215-20014; N.G.F., N.J.W.). The coordination of EPIC is financially supported by the European Commission (DG-SANCO) and the International

Agency for Research on Cancer. The Spanish National cohort is supported by Health Research Fund (FIS) - Instituto de Salud Carlos III (ISCIII), Regional Governments of Andalucía, Asturias, Basque Country, Murcia, and Navarra, and the Catalan Institute of Oncology – ICO (Spain). Biomarker measurements for vitamin C were funded by the Medical Research Council Cambridge Initiative (RG71466, SJAH/004) and the EPIC-CVD project, which is supported by the European Union Framework 7 (HEALTH-F2-2012-279233), European Research Council (268834), UK Medical Research Council (G0800270 and MR/L003120/1), British Heart Foundation (SP/09/002, RG/08/014, and RG13/13/30194) and the UK National Institute for Health Research (Cambridge Biomedical Research Center at the Cambridge University Hospitals National Health Service Foundation Trust). This work was supported by Health Data Research UK, which is funded by the UK Medical Research Council, Engineering and Physical Sciences Research Council, Economic and Social Research Council, Department of Health and Social Care (England), Chief Scientist Office of the Scottish Government Health and Social Care Directorates, Health and Social Care Research and Development Division (Welsh Government), Public Health Agency (Northern Ireland), British Heart Foundation, and Wellcome. This research has been conducted using the UK Biobank Resource (application number 44448). J.-S.Z. was additionally supported by the European Union's Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie Actions grant agreement No 701708, the National Natural Science Foundation of China (81903316), and Zhejiang Ten-thousand Talents Program (2019R52039).

The views expressed are those of the authors and not necessarily those of the National Health Service, the National Institute for Health Research, or the Department of Health and Social Care.

Where authors are identified as personnel of the International Agency for Research on Cancer/World Health Organization, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy, or views of the International Agency for Research on Cancer/World Health Organization.

Duality of Interest. J.D. is a member of the International Cardiovascular and Metabolic Advisory Board for Novartis (since 2010), the Steering Committee of UK Biobank (since 2011), the Medical Research Council International Advisory Group (ING) member, London (since 2013), the Medical Research Council High Throughput Science "Omics" Panel Member, London (since 2013), the Scientific Advisory Committee for Sanofi (since 2013), the International Cardiovascular and Metabolism Research and Development Portfolio Committee for Novartis, and the AstraZeneca Genomics Advisory Board (2018). No other potential conflicts of interest relevant to this article were reported.

Author Contributions. J.-S.Z., J.L., and E.S. performed the statistical analyses. J.-S.Z. and N.G.F. drafted the manuscript. The Working Group (J.-S.Z., J.L., E.S., F.I., I.D.S., F.R.D., M.P., E.W., L.A.L., T.E.G., A.S.B., C.L., and N.G.F.) and P.A., E.A., M.-D.C., G.F., P.W.F., R.K., N.L., F.R.M., P.M.N., N.C.O.-M., A.O., K.O., S.P., D.P., F.R., O.R., A.M.W.S., M.-J.S., M.B.S., N.S., S.S., A.T., R.T., Y.T.v.d.S., E.W., E.R., and J.D. contributed to interpretation of data and revised the article critically for important intellectual content. All authors approved the final version of the manuscript. S.J.S., C.L., N.G.F., and N.J.W. coordinated the InterAct project with N.J.W. as chief investigator. J.-S.Z. and N.G.F. are guarantors of the work and, as such, had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

References

1. Harding AH, Wareham NJ, Bingham SA, et al. Plasma vitamin C level, fruit and vegetable consumption, and the risk of new-onset type 2 diabetes mellitus: the European prospective investigation of cancer–Norfolk prospective study. Arch Intern Med 2008;168:1493–1499
2. Ashor AW, Werner AD, Lara J, Willis ND, Mathers JC, Siervo M. Effects of vitamin C supplementation on glycaemic control: a systematic review and meta-analysis of randomised controlled trials. Eur J Clin Nutr 2017;71:1371–1380
3. Song Y, Cook NR, Albert CM, Van Denburgh M, Manson JE. Effects of vitamins C and E and beta-carotene on the risk of type 2 diabetes in women at high risk of cardiovascular disease: a randomized controlled trial. Am J Clin Nutr 2009;90:429–437
4. Didelez V, Sheehan N. Mendelian randomization as an instrumental variable approach to causal inference. Stat Methods Med Res 2007;16:309–330
5. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. Genet Epidemiol 2013;37:658–665
6. Timpson NJ, Forouhi NG, Brion MJ, et al. Genetic variation at the SLC23A1 locus is associated with circulating concentrations of L-ascorbic acid (vitamin C): evidence from 5 independent studies with >15,000 participants. Am J Clin Nutr 2010;92:375–382
7. Langenberg C, Sharp S, Forouhi NG, et al.; InterAct Consortium. Design and cohort description of the InterAct Project: an examination of the interaction of genetic and lifestyle factors on the incidence of type 2 diabetes in the EPIC Study. Diabetologia 2011;54:2272–2282
8. Day N, Oakes S, Luben R, et al. EPIC-Norfolk: study design and characteristics of the cohort. European Prospective Investigation of Cancer. Br J Cancer 1999;80(Suppl. 1):95–103
9. Danesh J, Saracci R, Berglund G, et al.; EPIC-Heart. EPIC-Heart: the cardiovascular component of a prospective study of nutritional, lifestyle and biological factors in 520,000 middle-aged participants from 10 European countries. Eur J Epidemiol 2007;22:129–141
10. Collins R. What makes UK Biobank special? Lancet 2012;379:1173–1174
11. Mahajan A, Taliun D, Thurner M, et al. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. Nat Genet 2018;50:1505–1513
12. Karlsen A, Blomhoff R, Gundersen TE. High-throughput analysis of vitamin C in human plasma with the use of HPLC with monolithic column and UV-detection. J Chromatogr B Analyt Technol Biomed Life Sci 2005;824:132–138
13. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 2010;26:2190–2191
14. Pers TH, Karjalainen JM, Chan Y, et al.; Genetic Investigation of ANthropometric Traits (GIANT) Consortium. Biological interpretation of genome-wide association studies using predicted gene functions. Nat Commun 2015;6:5890
15. Segrè AV, Groop L, Mootha VK, Daly MJ, Altshuler D; DIAGRAM Consortium; MAGIC investigators. Common inherited variation in mitochondrial genes is not enriched for associations with type 2 diabetes or related glycemic traits. PLoS Genet 2010;6:e1001058
16. Zheng J, Erzurumluoglu AM, Elsworth BL, et al.; Early Genetics and Lifecourse Epidemiology (EAGLE) Eczema Consortium. LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. Bioinformatics 2017;33:272–279
17. Lotta AL, Pietzner M, Stewart DI, et al. Cross-platform genetic discovery of small molecule products of metabolism and application to clinical outcomes. 4 February 2020 [preprint]. BioRxiv 2020.02.03.932541
18. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol 2015;44:512–525
19. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. Genet Epidemiol 2016;40:304–314
20. Shaghaghi MA, Kloss O, Eck P. Genetic variation in human vitamin C transporter genes in common complex diseases. Adv Nutr 2016;7:287–298
21. Lawlor DA, Davey Smith G, Kundu D, Bruckdorfer KR, Ebrahim S. Those confounded vitamins: what can we learn from the differences between observational versus randomised trial evidence? Lancet 2004;363:1724–1727
22. Michels AJ, Hagen TM, Frei B. Human genetic variation influences vitamin C homeostasis by altering vitamin C transport and antioxidant enzyme function. Annu Rev Nutr 2013;33:45–70
23. Bürzle M, Suzuki Y, Ackermann D, et al. The sodium-dependent ascorbic acid transporter family SLC23. Mol Aspects Med 2013;34:436–454
24. Bunnag S, Einecke G, Reeve J, et al. Molecular correlates of renal function in kidney transplant biopsies. J Am Soc Nephrol 2009;20:1149–1160
25. Okada Y, Sim X, Go MJ, et al.; KidneyGen Consortium; CKDGen Consortium; GUGC consortium. Meta-analysis identifies multiple loci associated with kidney function-related traits in east Asian populations. Nat Genet 2012;44:904–909
26. Hong SH, Kwak JH, Paik JK, Chae JS, Lee JH. Association of polymorphisms in FADS gene with age-related changes in serum phospholipid polyunsaturated fatty acids and oxidative stress markers in middle-aged nonobese men. Clin Interv Aging 2013;8:585–596

27. Hara T, Hashimoto Y, Akuzawa T, Hirai R, Kobayashi H, Sato K. Rer1 and calnexin regulate endoplasmic reticulum retention of a peripheral myelin protein 22 mutant that causes type 1A Charcot-Marie-Tooth disease. Sci Rep 2014;4:6992
28. Lee HS, Kim BK, Nam Y, et al. Protective role of phosphatidylcholine against cisplatin-induced renal toxicity and oxidative stress in rats. Food Chem Toxicol 2013;58:388–393
29. Pennant M, Steur M, Moore C, Butterworth A, Johnson L. Comparative validity of vitamin C and carotenoids as indicators of fruit and vegetable intake: a systematic review and meta-analysis of randomised controlled trials. Br J Nutr 2015;114:1331–1340
30. Cooper AJ, Sharp SJ, Luben RN, Khaw KT, Wareham NJ, Forouhi NG. The association between a biomarker score for fruit and vegetable intake and incident type 2 diabetes: the EPIC-Norfolk study. Eur J Clin Nutr 2015;69:449–454
31. Khaw KT, Bingham S, Welch A, et al. Relation between plasma ascorbic acid and mortality in men and women in EPIC-Norfolk prospective study: a prospective population study. European Prospective Investigation into Cancer and Nutrition. Lancet 2001;357:657–663
32. Zheng JS, Sharp SJ, Imamura F, et al. Association of plasma biomarkers of fruit and vegetable intake with incident type 2 diabetes: EPIC-InterAct case-cohort study in eight European countries. BMJ 2020;370:m2194