Online Supplemental Files

Title: Plasma vitamin C and type 2 diabetes: genome-wide association study and Mendelian randomization analysis in European populations

Supplemental Tables

Supplemental Figures

Online Supplemental Files

Supplemental Figure 7 Regional plots for SNPs on chromosome 6 identified in the genome-wide meta-analysis

Supplemental Figure 8 Regional plots for SNPs on chromosome 11 identified in the genome-wide meta-analysis

Supplemental Figure 9 Regional plots for SNPs on chromosome 12 identified in the genome-wide meta-analysis

Supplemental Figure 10 Regional plots for SNPs on chromosome 14 identified in the genome-wide meta-analysis

Supplemental Figure 11 Regional plots for SNPs on chromosome 16 identified in the genome-wide meta-analysis

Supplemental Figure 12 Regional plots for SNPs on chromosome 17 identified in the genome-wide meta-analysis

Supplemental Figure 13 Association of GWAS-identified lead SNPs with plasma vitamin C by individual cohorts

Supplemental Figure 14 Mendelian randomization estimate for the causal effect of type 2 diabetes on vitamin C. In the Mendelian randomization (MR) analysis of the association of genetically predicted type 2 diabetes risk with vitamin C, we used three methods: inverse variance weighted, MR-Egger regression and weighted median, and presented the results below. Significant heterogeneity ($P<0.001$) was observed in the MR analysis and results from the weighted median method were highlighted in the manuscript. A total of 231 independent SNPs of type 2 diabetes were used as genetic instrument (Nat Genet. 2018 Nov;50(11):1505-1513. doi: 10.1038/s41588-018-0241-6).

Supplemental Figure 15 Mendelian randomization for effect of a 1-SD higher plasma vitamin C concentration on glycemic traits. Estimates and 95% CIs are shown for beta, i.e. the difference in SD units of the outcome per 1-SD genetic predicted increase of plasma vitamin C concentration. Fasting glucose, 2-hour glucose, ln-transformed fasting insulin, ln-transformed HOMA-IR, ln-transformed HOMA-B and HbA1c are from the Meta-Analysis of Glucose and Insulin-related traits Consortium [MAGIC]; SD, standard deviation; HOMA-IR, homeostatic model assessment of insulin resistance; HOMA-B, homeostatic model assessment of beta-cell function; HbA1c, hemoglobin A1c.

Supplemental Figure 16 Association of vitamin C-raising alleles with continuous glycemic traits. Associations are in standardized units (in SD unit) per vitamin C-raising allele. Fasting glucose (n=133,010), 2-hour glucose (n=42,854), ln-transformed fasting insulin (n=108,557), ln-transformed HOMA-IR (n=46,186), ln-transformed HOMA-B (n=46,186) and HbA1c (n=46,368) are from the Meta-Analysis of Glucose and Insulin-related traits Consortium (MAGIC). 10 SNPs used as genetic instrument in the Mendelian randomization were assessed for their associations with glycemic traits. For SNRPF-rs117885456, we used the estimates of this SNP derived from Fenland, as the information of this SNP was not available in the MAGIC and no proxy SNP was found.

Online Supplemental Files

Supplemental text - Supplemental Materials

Online Supplemental Files

Supplemental Table 1 Characteristics of the cohorts included in the genome-wide meta-analysis of plasma vitamin C (n=52,018)

Variable	Fenland GWAS array	Fenland UKBB array	Fenland Core-Exome array	InterAct subcohort GWAS array	InterAct subcohort core-exome array	InterAct non-subcohort GWAS array	InterAct non-subcohort core-exome array	EPIC-Norfolk GWAS array	EPIC-CVD subcohort	EPIC-CVD non-subcohort
Participants included in the GWAS, N	1,349	8,391	1,031	3,521	6,504	2,944	3,872	16,756	885	6,765
Age, mean years (SD)	45 (7)	49 (7)	51 (7)	51 (9)	53 (9)	55 (8)	56 (7)	59 (9)	53 (12)	57 (8)
Female sex, N (%)	754 (56)	4,413 (53)	566 (55)	2,293 (65)	4,000 (62)	1,543 (52)	1,856 (48)	8,947 (53)	522 (59)	3,061 (45)
Plasma vitamin C, mean umol/l (SD)	66.2 (21.3)	68.6 (21.5)	68.3 (21.8)	42.8 (19)	42.9 (19.1)	36.4 (17.6)	36.5 (18.9)	53.8 (20.2)	41.0 (21.0)	37.9 (20.8)
Genotyping array	Affymetrix genome-Wide Human SNP Array 5.0	Affymetrix UK Biobank Axiom Array	Illumina HumanCore Exome chip	Illumina 660W-Quad BeadChip	Illumina HumanCoreExome array	Illumina 660W-Quad BeadChip	Illumina HumanCoreExome array	Affymetrix UK Biobank Axiom Array	Illumina HumanCoreExome array	Illumina HumanCoreExome array
Imputation panel	HRC	HRC	HRC	HRC	HRC	HRC	HRC	HRC	HRC	HRC
Number of GWAS SNPs*	7,622,992	7,727,349	7,933,258	7,737,656	7,693,434	7,739,029	7,686,676	7,716,054	7,705,977	7,692,630

*Number of GWAS SNPs indicates number of SNPs with minor allele frequency≥1% within each cohort, beta coefficient and standard error≤5, imputation quality (info score)≥4 and P-Hardy-Weinberg equilibrium≥10^{-6}.

Online Supplemental Files

Supplemental Table 2 Baseline characteristics by quintiles of plasma vitamin C levels in in the EPIC-InterAct study (n=16,841)*

	EPIC-InterAct: subcohort (n=10,025)						EPIC-InterAct: non-subcohort (n=6,816)					
	Q1 (n=1,901)	Q2 (n=2,002)	Q3 (n=2,039)	Q4 (n=2,008)	Q5 (n=2,075)	p-trend	Q1 (n=2,014)	Q2 (n=1,740)	Q3 (n=1,200)	Q4 (n=1,065)	Q5 (n=797)	p-trend
Plasma vitamin C, umol/l	16.3 (7.6)	32.8 (3.2)	42.4 (2.4)	51.3 (2.9)	69.2 (13.3)		15.7 (7.6)	32.6 (3.1)	42.1 (2.4)	51.2 (2.9)	69.1 (13.5)	
Age, y	52.8 (8.7)	52 (8.6)	51.9 (8.6)	51.4 (9)	52.5 (9.1)	0.313	55.6 (7.5)	55.1 (7.3)	55.4 (7.2)	55.7 (7.3)	56.4 (7.1)	0.002
BMI, kg/m2	26.9 (4.4)	26.6 (4.1)	26.5 (4.2)	25.9 (4.2)	25.2 (4.1)	<0.001	30.2 (4.8)	30.4 (4.8)	29.8 (4.8)	29.5 (4.7)	29.2 (4.8)	<0.001
Systolic blood pressure, mmHg	135.9 (20.1)	132.7 (19.3)	131.5 (18.6)	129.8 (18.8)	128.4 (18.3)	<0.001	145.9 (20.3)	143.2 (20.4)	142.8 (20.6)	142.4 (20.1)	139.4 (20.3)	<0.001
Diastolic blood pressure, mmHg	83.2 (11.1)	82.2 (11)	81.4 (10.7)	80.8 (10.7)	79.6 (10.3)	<0.001	88.3 (11.4)	87.2 (11)	86.6 (10.9)	86.4 (11.1)	84.5 (11.1)	<0.001
Total cholesterol, mmol/l	6 (1.1)	5.9 (1.1)	5.9 (1.1)	5.9 (1.1)	5.9 (1.1)	0.266	6.2 (1.2)	6.1 (1.1)	6.1 (1.1)	6.1 (1.2)	6.2 (1.1)	0.981
Triglycerides, mmol/l	1.6 (1.1)	1.4 (0.9)	1.3 (0.9)	1.2 (0.8)	1.2 (0.8)	<0.001	2.3 (1.5)	2.1 (1.5)	2 (1.3)	1.9 (1.2)	1.7 (1)	<0.001
LDL-C, mmol/l	3.9 (1)	3.8 (1)	3.8 (1)	3.8 (1)	3.8 (1)	<0.001	4 (1.1)	4 (1)	4 (1)	4 (1)	4 (1)	0.594
HDL-C, mmol/l	1.4 (0.4)	1.4 (0.4)	1.5 (0.4)	1.6 (0.4)	1.6 (0.4)	<0.001	1.2 (0.4)	1.2 (0.3)	1.3 (0.4)	1.3 (0.4)	1.4 (0.4)	<0.001
Vegetable intake, g/d	172 (112.1)	195.1 (118.5)	202.7 (122.4)	206.5 (127.4)	194.6 (118.5)	<0.001	168.1 (114.1)	198 (127.5)	200.5 (127.8)	203.8 (129.1)	201.2 (130.6)	<0.001
Fruit intake, g/d	176.6 (162)	230.3 (182.1)	253.5 (194.4)	273.2 (204.6)	262.6 (191.3)	<0.001	185.8 (180.9)	232.1 (184.9)	257.7 (199.2)	271.1 (208.6)	270.2 (208.3)	<0.001
Sex						<0.001						<0.001
Men	1045 (55)	936 (46.8)	760 (37.3)	554 (27.6)	437 (21.1)		1261 (62.6)	952 (54.7)	559 (46.6)	419 (39.3)	226 (28.4)	
Women	856 (45)	1066 (53.2)	1279 (62.7)	1454 (72.4)	1638 (78.9)		753 (37.4)	788 (45.3)	641 (53.4)	646 (60.7)	571 (71.6)	
Education						<0.001						0.001
None	130 (6.8)	172 (8.6)	212 (10.4)	202 (10.1)	149 (7.2)		174 (8.6)	203 (11.7)	160 (13.3)	135 (12.7)	83 (10.4)	
Primary school	720 (37.9)	639 (31.9)	620 (30.4)	579 (28.8)	557 (26.8)		877 (43.5)	660 (37.9)	462 (38.5)	389 (36.5)	292 (36.6)	
Technical/professional school	445 (23.4)	463 (23.1)	497 (24.4)	461 (23)	508 (24.5)		473 (23.5)	418 (24)	283 (23.6)	258 (24.2)	204 (25.6)	
Secondary school	239 (12.6)	282 (14.1)	289 (14.2)	288 (14.3)	304 (14.7)		203 (10.1)	170 (9.8)	112 (9.3)	114 (10.7)	89 (11.2)	
Longer education	321 (16.9)	399 (19.9)	373 (18.3)	435 (21.7)	523 (25.2)		236 (11.7)	236 (13.6)	155 (12.9)	135 (12.7)	105 (13.2)	

Online Supplemental Files

	EPIC-InterAct: subcohort (n=10,025)						EPIC-InterAct: non-subcohort (n=6,816)					
Employment						0.008						<0.001
No	514 (27)	447 (22.3)	431 (21.1)	480 (23.9)	581 (28)		664 (33)	489 (28.1)	363 (30.3)	339 (31.8)	307 (38.5)	
Yes	931 (49)	981 (49)	984 (48.3)	936 (46.6)	1075 (51.8)		927 (46)	801 (46)	502 (41.8)	443 (41.6)	316 (39.6)	
Physical activity						<0.001						0.032
Inactive	529 (27.8)	436 (21.8)	439 (21.5)	447 (22.3)	387 (18.7)		657 (32.6)	493 (28.3)	332 (27.7)	277 (26)	217 (27.2)	
Moderately inactive	571 (30)	679 (33.9)	689 (33.8)	677 (33.7)	690 (33.3)		633 (31.4)	545 (31.3)	386 (32.2)	356 (33.4)	267 (33.5)	
Moderately active	403 (21.2)	438 (21.9)	437 (21.4)	451 (22.5)	484 (23.3)		359 (17.8)	338 (19.4)	247 (20.6)	208 (19.5)	151 (18.9)	
Active	365 (19.2)	405 (20.2)	437 (21.4)	403 (20.1)	486 (23.4)		340 (16.9)	329 (18.9)	215 (17.9)	206 (19.3)	150 (18.8)	
Smoke						<0.001						<0.001
Never	605 (31.8)	858 (42.9)	1008 (49.4)	1043 (51.9)	1068 (51.5)		588 (29.2)	678 (39)	545 (45.4)	510 (47.9)	387 (48.6)	
Former	480 (25.2)	559 (27.9)	539 (26.4)	556 (27.7)	559 (26.9)		600 (29.8)	582 (33.4)	349 (29.1)	316 (29.7)	217 (27.2)	
Current	791 (41.6)	554 (27.7)	461 (22.6)	388 (19.3)	426 (20.5)		800 (39.7)	458 (26.3)	292 (24.3)	219 (20.6)	183 (23)	
Alcohol						<0.001						0.123
None	305 (16)	302 (15.1)	315 (15.4)	364 (18.1)	266 (12.8)		318 (15.8)	282 (16.2)	225 (18.8)	197 (18.5)	131 (16.4)	
Current	1586 (83.4)	1694 (84.6)	1714 (84.1)	1635 (81.4)	1803 (86.9)		1681 (83.5)	1445 (83)	971 (80.9)	857 (80.5)	661 (82.9)	
Vitamin supplement						<0.001						<0.001
No	1162 (61.1)	1167 (58.3)	1128 (55.3)	1051 (52.3)	956 (46.1)		1196 (59.4)	957 (55)	582 (48.5)	510 (47.9)	341 (42.8)	
Yes	568 (29.9)	687 (34.3)	732 (35.9)	789 (39.3)	884 (42.6)		556 (27.6)	598 (34.4)	470 (39.2)	440 (41.3)	369 (46.3)	

*Values are expressed as mean (SD) or No. of participants (%). p-trend is calculated with nonparametric test for trend across the quintiles for continuous variables and calculated with chi-square test for categorical variables.

Online Supplemental Files

Supplemental Table 3 Baseline characteristics by quintiles of plasma vitamin C-raising genetic risk score in EPIC-InterAct (n=16,841)*

	InterAct: subcohort (10,025)						InterAct: non-subcohort (n=6,816)					
	Q1 (n=2,002)	Q2 (n=2,008)	Q3 (n=2,002)	Q4 (n=2,008)	Q5 (n=2,005)	p-trend	Q1 (n=1,369)	Q2 (n=1,271)	Q3 (n=1,369)	Q4 (n=1,398)	Q5 (n=1,409)	p-trend
Genetic risk score	8.26 (0.85)	9.84 (0.28)	10.8 (0.26)	11.8 (0.30)	13.4 (0.82)		8.30 (0.82)	9.83 (0.29)	10.8 (0.25)	11.8 (0.31)	13.4 (0.84)	
Plasma vitamin C, umol/l	39.8 (18.2)	41.7 (18.4)	43.5 (18.8)	43.8 (20.2)	45.5 (19.2)	0.000	33.8 (16.5)	35.9 (18.1)	35.7 (18.6)	38 (18.1)	38.8 (20)	0.000
Age, y	52.1 (8.7)	52.3 (8.9)	51.9 (8.9)	51.8 (9)	52.4 (8.4)	0.679	55.5 (7.3)	55.6 (7.5)	55.7 (7.1)	55.6 (7.3)	55.3 (7.3)	0.399
BMI, kg/m2	26.3 (4.2)	26.4 (4.4)	26.1 (4.1)	26.1 (4.1)	26.2 (4.2)	0.159	29.9 (4.7)	29.9 (4.7)	30.1 (4.8)	29.9 (4.8)	29.9 (4.9)	0.486
Systolic blood pressure, mmHg	132.3 (19.6)	131.7 (19.3)	131.7 (18.8)	130.9 (19.5)	131.4 (18.6)	0.142	142.9 (19.5)	143.4 (19.8)	143.7 (20.2)	143.8 (21.7)	143.1 (20.8)	0.785
Diastolic blood pressure, mmHg	81.4 (10.8)	81.5 (10.8)	81.3 (10.7)	81.2 (10.9)	81.5 (10.9)	0.788	86.9 (11)	86.8 (11.3)	86.9 (11.2)	86.9 (11.2)	87.4 (11.3)	0.296
Total cholesterol, mmol/l	5.96 (1.07)	5.92 (1.1)	5.92 (1.1)	5.9 (1.1)	5.92 (1.09)	0.206	6.15 (1.16)	6.1 (1.15)	6.14 (1.14)	6.13 (1.17)	6.18 (1.16)	0.380
Triglycerides, mmol/l	1.35 (0.86)	1.37 (0.87)	1.35 (0.97)	1.35 (0.91)	1.36 (0.94)	0.534	2.03 (1.35)	2.09 (1.41)	2.04 (1.36)	1.97 (1.33)	2.07 (1.43)	0.677
LDL-C, mmol/l	3.85 (0.98)	3.81 (0.98)	3.8 (0.99)	3.79 (0.99)	3.8 (1)	0.241	4.01 (1.02)	3.93 (0.98)	4 (0.99)	4 (1.03)	4.01 (0.99)	0.487
HDL-C, mmol/l	1.51 (0.43)	1.49 (0.41)	1.51 (0.42)	1.49 (0.42)	1.51 (0.42)	0.546	1.25 (0.38)	1.23 (0.36)	1.24 (0.36)	1.26 (0.37)	1.25 (0.37)	0.141
Vegetable intake, g/d	189.8 (118.8)	193.5 (120.5)	199.8 (121.3)	195.1 (119.3)	194.3 (122.7)	0.191	192.1 (134)	190.1 (122.9)	191 (119.8)	195.2 (124)	185.8 (124.8)	0.826
Fruit intake, g/d	234.2 (186.1)	237 (191.7)	246.3 (196.5)	245.5 (194.6)	238.4 (184.3)	0.072	232.6 (202.6)	240 (201.6)	232.3 (194.9)	235.6 (193)	227.7 (188.1)	0.824
Sex						0.818						0.002
Men	744 (37.2)	770 (38.3)	745 (37.2)	738 (36.8)	735 (36.7)		724 (52.9)	663 (52.2)	697 (50.9)	642 (45.9)	691 (49)	
Women	1258 (62.8)	1238 (61.7)	1257 (62.8)	1270 (63.2)	1270 (63.3)		645 (47.1)	608 (47.8)	672 (49.1)	756 (54.1)	718 (51)	
Education						0.142						0.859
None	172 (8.6)	184 (9.2)	181 (9)	168 (8.4)	160 (8)		166 (12.1)	149 (11.7)	151 (11)	152 (10.9)	137 (9.7)	
Primary school	614 (30.7)	661 (32.9)	579 (28.9)	627 (31.2)	634 (31.6)		545 (39.8)	505 (39.7)	546 (39.9)	545 (39)	539 (38.3)	
Technical/professional school	511 (25.5)	456 (22.7)	448 (22.4)	463 (23.1)	496 (24.7)		331 (24.2)	305 (24)	330 (24.1)	328 (23.5)	342 (24.3)	
Secondary school	264 (13.2)	270 (13.4)	311 (15.5)	285 (14.2)	272 (13.6)		130 (9.5)	119 (9.4)	135 (9.9)	150 (10.7)	154 (10.9)	
Longer education	405 (20.2)	389 (19.4)	434 (21.7)	419 (20.9)	404 (20.1)		165 (12.1)	161 (12.7)	166 (12.1)	180 (12.9)	195 (13.8)	

Online Supplemental Files

	InterAct: subcohort (10,025)						InterAct: non-subcohort (n=6,816)					
Employment						0.712						0.851
No	494 (24.7)	490 (24.4)	458 (22.9)	502 (25)	509 (25.4)		402 (29.4)	404 (31.8)	435 (31.8)	461 (33)	460 (32.6)	
Yes	973 (48.6)	973 (48.5)	985 (49.2)	988 (49.2)	988 (49.3)		590 (43.1)	547 (43)	604 (44.1)	613 (43.8)	635 (45.1)	
Physical activity						0.904						0.894
Inactive	440 (22)	458 (22.8)	460 (23)	445 (22.2)	435 (21.7)		403 (29.4)	374 (29.4)	396 (28.9)	397 (28.4)	406 (28.8)	
Moderately inactive	658 (32.9)	667 (33.2)	674 (33.7)	667 (33.2)	640 (31.9)		433 (31.6)	400 (31.5)	434 (31.7)	477 (34.1)	443 (31.4)	
Moderately active	448 (22.4)	436 (21.7)	425 (21.2)	432 (21.5)	472 (23.5)		250 (18.3)	239 (18.8)	268 (19.6)	271 (19.4)	275 (19.5)	
Active	428 (21.4)	412 (20.5)	404 (20.2)	425 (21.2)	427 (21.3)		258 (18.8)	242 (19)	247 (18)	232 (16.6)	261 (18.5)	
Smoke						0.536						0.270
Never	898 (44.9)	922 (45.9)	940 (47)	899 (44.8)	923 (46)		540 (39.4)	471 (37.1)	536 (39.2)	579 (41.4)	582 (41.3)	
Former	539 (26.9)	540 (26.9)	539 (26.9)	526 (26.2)	549 (27.4)		410 (29.9)	422 (33.2)	412 (30.1)	411 (29.4)	409 (29)	
Current	544 (27.2)	517 (25.7)	494 (24.7)	557 (27.7)	508 (25.3)		399 (29.1)	363 (28.6)	403 (29.4)	387 (27.7)	400 (28.4)	
Alcohol						0.953						0.237
None	299 (14.9)	313 (15.6)	315 (15.7)	315 (15.7)	310 (15.5)		241 (17.6)	220 (17.3)	208 (15.2)	228 (16.3)	256 (18.2)	
Current	1698 (84.8)	1681 (83.7)	1679 (83.9)	1687 (84)	1687 (84.1)		1114 (81.4)	1045 (82.2)	1152 (84.1)	1161 (83)	1143 (81.1)	
Vitamin supplement						0.263						0.120
No	1078 (53.8)	1064 (53)	1128 (56.3)	1115 (55.5)	1079 (53.8)		698 (51)	663 (52.2)	760 (55.5)	713 (51)	752 (53.4)	
Yes	741 (37)	752 (37.5)	705 (35.2)	712 (35.5)	750 (37.4)		513 (37.5)	449 (35.3)	464 (33.9)	520 (37.2)	487 (34.6)	

*Values are expressed as mean (SD) or No. of participants (%). p-trend is calculated with nonparametric test for trend across the quintiles for continuous variables and calculated with chi-square test for categorical variables. The unweighted genetic risk score was generated by summing the number of vitamin C-raising alleles across the 10 genetic variants identified in the present genome-wide meta-analysis of plasma vitamin C, excluding FADS1 variant.

Online Supplemental Files

Supplemental Table 4 Pathway analysis using MAGENTA

Gene sets	ORIG_GS_SIZE	EFF_GS_SIZE	NOMINAL_GSEA_PVAL_95PERC_CUTOFF	FDR_95PERC_CUTOFF	EXP_#_GENES_ABOVE_95PERC_CUTOFF	OBS_#_GENES_ABOVE_95PERC_CUTOFF
ENSG00000139835	37	30	2.00E-04	0.06	2	8
ENSG00000057657	47	44	4.00E-05	0.07	2	10
GO:0042809 (VITAMIN_D_RECEPTOR_BINDING)	76	58	2.60E-05	0.07	3	12
ENSG00000166046	72	68	1.00E-04	0.08	3	12
MP:0008803	52	47	2.00E-04	0.09	2	10
ENSG00000142611	58	52	2.00E-04	0.09	3	10
MP:0000602	53	48	1.00E-04	0.09	2	10
GO:0046415 (URATE_METABOLIC_PROCESS)	134	119	2.00E-04	0.09	6	17
GO:0015293 (SYMPORTER_ACTIVITY)	193	180	2.10E-05	0.09	9	23
GO:0015103 (INORGANIC_ANION_TRANSMEMBRANE_TRANSPORTER_ACTIVITY)	165	149	4.70E-05	0.09	7	20
GO:0015101 (INORGANIC_CATION_TRANSMEMBRANE_TRANSPORTER_ACTIVITY)	168	152	1.00E-04	0.09	8	20
ENSG00000138041	37	29	3.00E-04	0.09	1	7
REACTOME_ORGANIC_CATIONANIONZWITTERION_TRANSPORT	181	166	1.00E-04	0.09	8	22
REACTOME_DOWNREGULATION_OF_ERRB2ERBB3_SIGNALING	36	34	1.00E-04	0.09	2	8
GO:0030332 (CYCLIN_BINDING)	50	46	2.00E-04	0.10	2	9
ENSG00000188620	55	49	4.00E-04	0.10	2	10

ORIG_GS_SIZE=Original number of genes per gene set. EFF_GS_SIZE=Effective number of genes per gene set analyzed by GSEA, after removing genes that were not assigned a gene score (e.g. no SNPs in their region), or after adjusting for physical clustering of genes in a given gene set (removing all but one gene from a subset of genes assigned the same best SNP, keeping the gene with the most significant gene score. NOMINAL_GSEA_PVAL_95PERC_CUTOFF = GSEA p-value using 95 percentile of all gene scores for the enrichment cut-off. FDR_95PERC_CUTOFF = Estimated false discovery rate (q-value) using 95 percentile cut-off.
EXP_#_GENES_ABOVE_95PERC_CUTOFF = Expected number of genes with a corrected gene p-value above the 95 percentile enrichment cut-off.
OBS_#_GENES_ABOVE_95PERC_CUTOFF = Observed number of genes with a corrected gene p-value above the 95 percentile enrichment cut-off.

Online Supplemental Files

Supplemental Table 5 DEPICT analysis*

Tissue enrichment (nominal p_val < 0.05)				
Name	**MeSH first level term**	**MeSH second level term**	**Nominal P value**	**False discovery rate**
Adrenal Cortex	Endocrine System	Endocrine Glands	5.67E-03	>=0.20
Parotid Gland	Digestive System	Gastrointestinal Tract	0.01	>=0.20
Salivary Glands	Digestive System	Gastrointestinal Tract	0.01	>=0.20
Adrenal Glands	Endocrine System	Endocrine Glands	0.01	>=0.20
Kidney Cortex	Urogenital System	Urinary Tract	0.02	>=0.20
Urinary Tract	Urogenital System	Urinary Tract	0.03	>=0.20
Kidney	Urogenital System	Urinary Tract	0.03	>=0.20
Geneset enrichment (nominal p_val < 1E-03)				
Original gene set description	**Nominal P value**	**False.discovery.rate**		
SYNGAP1 PPI subnetwork	1.69E-04	>=0.20		
pup cannibalization	1.69E-04	>=0.20		
regulation of sodium ion transport	1.94E-04	>=0.20		
myotube differentiation	3.35E-04	>=0.20		
sialyltransferase activity	3.95E-04	>=0.20		
ZBED1 PPI subnetwork	4.14E-04	>=0.20		
kidney cortex cysts	4.26E-04	>=0.20		
interaction with host	5.05E-04	>=0.20		
abnormal interventricular septum morphology	5.91E-04	>=0.20		
fatty-acyl-CoA binding	6.16E-04	>=0.20		
ZFPM1 PPI subnetwork	6.92E-04	>=0.20		
disorganized long bone epiphyseal plate	7.19E-04	>=0.20		
WWP2 PPI subnetwork	7.65E-04	>=0.20		
aging	8.75E-04	>=0.20		
increased spleen white pulp amount	8.80E-04	>=0.20		

*****Tissue enrichment and Geneset enrichment for meta-analysis of SNPs with *P*-values < 1×10^{-5}.

Online Supplemental Files

Supplemental Table 6 Genetic correlation of plasma vitamin C with type 2 diabetes and related glycemic traits, using Linkage Disequilibrium Score regression*

Trait	PMID	$r_{genetic}$	se	P-value
Type 2 Diabetes	22885922	-0.19	0.08	0.018
Fasting insulin	22581228	-0.22	0.08	0.005†
Fasting glucose	22581228	-0.10	0.07	0.168
2hr glucose	20081857	-0.01	0.13	0.946
HOMA-IR	20081858	-0.22	0.10	0.032
HOMA-B	20081858	-0.12	0.10	0.241
HbA1c	20858683	-0.14	0.09	0.114

*In the LD Score regression, the total observed scale heritability for vitamin C is 0.102 (se=0.012). PMID refers to the reference ID in the PubMed.
Abbreviation: r-genetics, genetic correlation of these traits with plasma vitamin C; se, standard error of genetic correlation;
Abbreviation: LD, linkage disequilibrium; HOMA-IR, homeostatic model assessment of insulin resistance; HOMA-B, homeostatic model assessment of beta cell function; HbA1c, hemoglobin A1c.
†indicates that the p-value passed the correction for multiple testing based on Bonferroni correction (p-value of threshold: 0.05/7=0.007).

Online Supplemental Files

Supplemental Table 7 Lead SNPs identified in GWAS and their proxy SNPs used in the Mendelian randomization analyses*

Lead SNPs	Proxy SNPs	r^2	Lead SNPs			Proxy SNPs				
			Beta	SE	p	Beta	SE	p	Vitamin C-raising allele/ other allele	Vitamin C-raising allele frequency
rs6693447	rs1123571	0.996	0.039	0.006	6.25×10^{-10}	0.039	0.006	6.26×10^{-10}	G/A	0.552
rs13028225	rs17655123	0.885	0.102	0.009	2.38×10^{-30}	0.097	0.009	3.03×10^{-27}	G/A	0.861
rs33972313	rs17131975	0.889	0.360	0.018	4.61×10^{-90}	0.349	0.018	8.63×10^{-87}	G/A	0.967
rs10051765	rs12654812	0.836	0.039	0.007	3.64×10^{-9}	0.032	0.007	8.41×10^{-07}	A/G	0.356
rs7740812	rs10948728	1	0.038	0.006	1.88×10^{-9}	0.036	0.006	2.39×10^{-08}	G/A	0.607
rs117885456	NA		0.078	0.012	1.70×10^{-11}					
rs2559850	rs3809260	0.971	0.058	0.006	6.30×10^{-20}	0.057	0.006	1.24×10^{-19}	T/G	0.595
rs10136000	rs1130214	0.98	0.040	0.007	1.33×10^{-8}	0.038	0.007	4.72×10^{-08}	A/C	0.293
rs56738967	rs17689024	0.996	0.041	0.007	7.62×10^{-10}	0.041	0.007	8.36×10^{-10}	C/G	0.321
rs9895661	no		0.063	0.008	1.05×10^{-14}					

*LD (r^2) of the proxy SNPs and the lead SNPs is derived from the 1000 Genomes project phase 3 (version 5). "NA" means that proxy SNP (r^2>0.8) was not found and applied in the MR on glycemic traits with public available consortium datasets, where the results from Fenland was used for the MR estimate; "no" means that the summary statistics of the original lead SNPs are available in all the publicly available large consortium datasets used in the present study so no proxy SNP is needed. SNP, single nucleotide polymorphism.

Online Supplemental Files

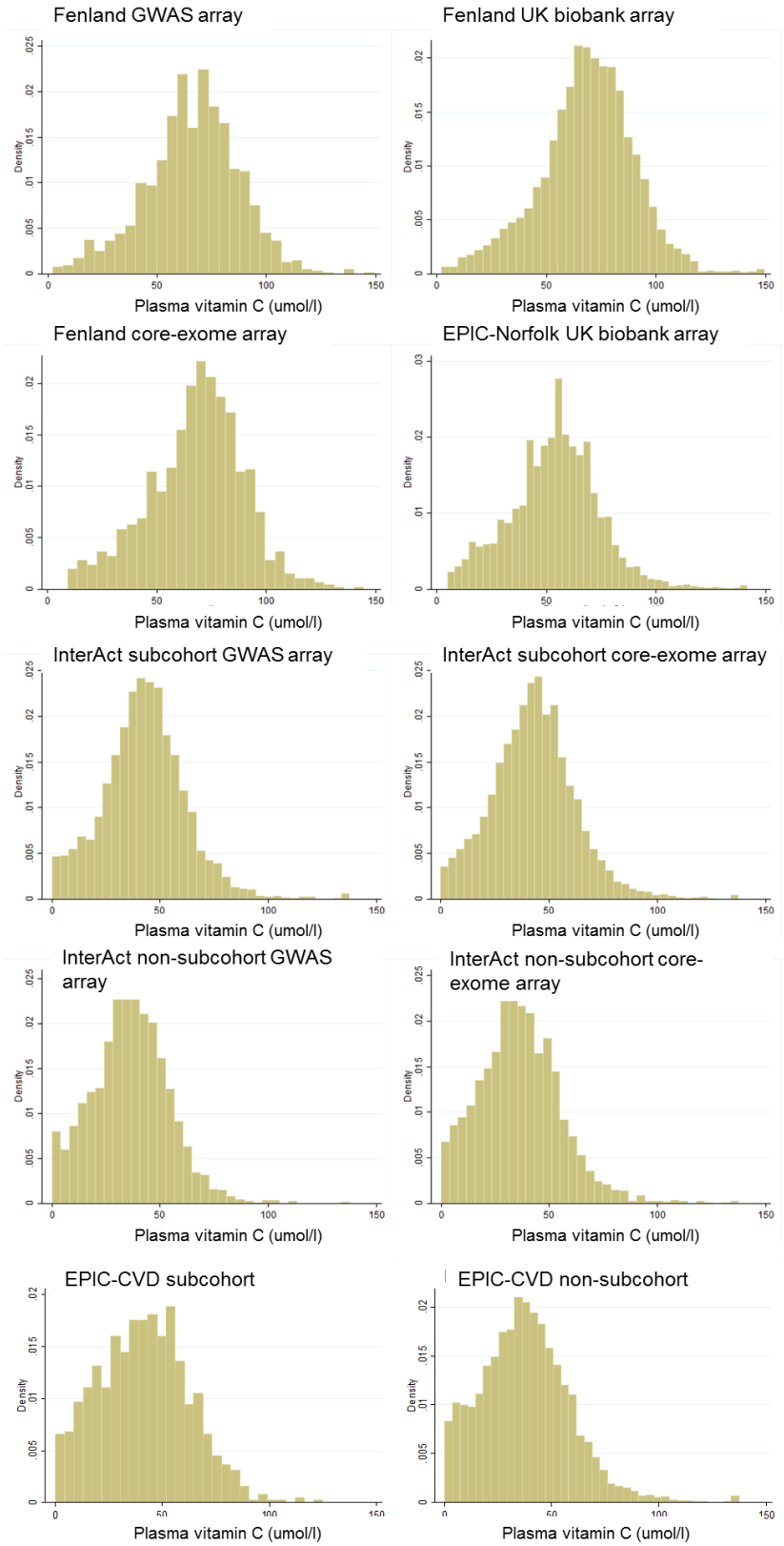
Supplemental Table 8 Sensitivity and heterogeneity analyses for Mendelian randomization estimate between plasma vitamin C and type 2 diabetes and related glycemic traits*

	Inverse-variance weighted		Cochran's Q test	MR-Egger regression		MR-Egger intercept		Weighted mean	
	Beta (se)	p	p	Beta (se)	p	Beta (se)	p	Beta (se)	p
Type 2 diabetes	0.027 (0.033)	0.43	0.34	-0.012 (0.056)	0.84	0.004 (0.004)	0.41	0.014 (0.041)	0.73
Fasting insulin	-0.011 (0.021)	0.60	1.0	-0.029 (0.033)	0.39	0.002 (0.002)	0.49	-0.018 (0.025)	0.47
Fasting glucose	0.004 (0.019)	0.85	1.0	-0.034 (0.028)	0.23	0.004 (0.002)	0.05	-0.013 (0.023)	0.56
2-h glucose	0.103 (0.124)	0.40	0.69	0.105 (0.207)	0.61	0 (0.015)	0.99	0.076 (0.133)	0.57
HOMA-IR	0.019 (0.029)	0.50	1.0	-0.036 (0.041)	0.38	0.005 (0.003)	0.08	-0.006 (0.03)	0.83
HOMA-B	0.008 (0.022)	0.72	1.0	-0.023 (0.034)	0.49	0.003 (0.002)	0.23	-0.012 (0.024)	0.63
HbA1c	-0.014 (0.01)	0.17	1.0	-0.015 (0.015)	0.31	0 (0.001)	0.86	-0.015 (0.012)	0.22

*Abbreviation: HOMA-IR, homeostatic model assessment of insulin resistance; HOMA-B, homeostatic model assessment of beta cell function; HbA1c, hemoglobin A1c.

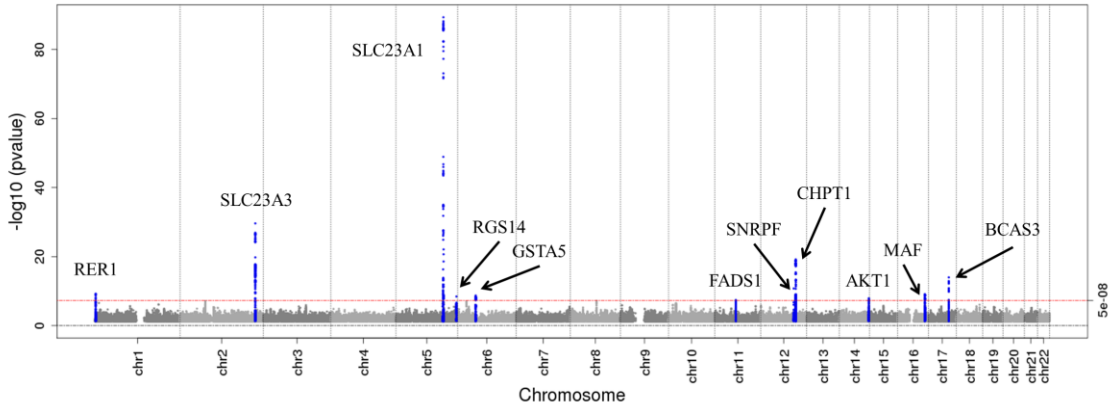
Online Supplemental Files

Supplemental Figure 1 Distribution of plasma vitamin C in each GWAS



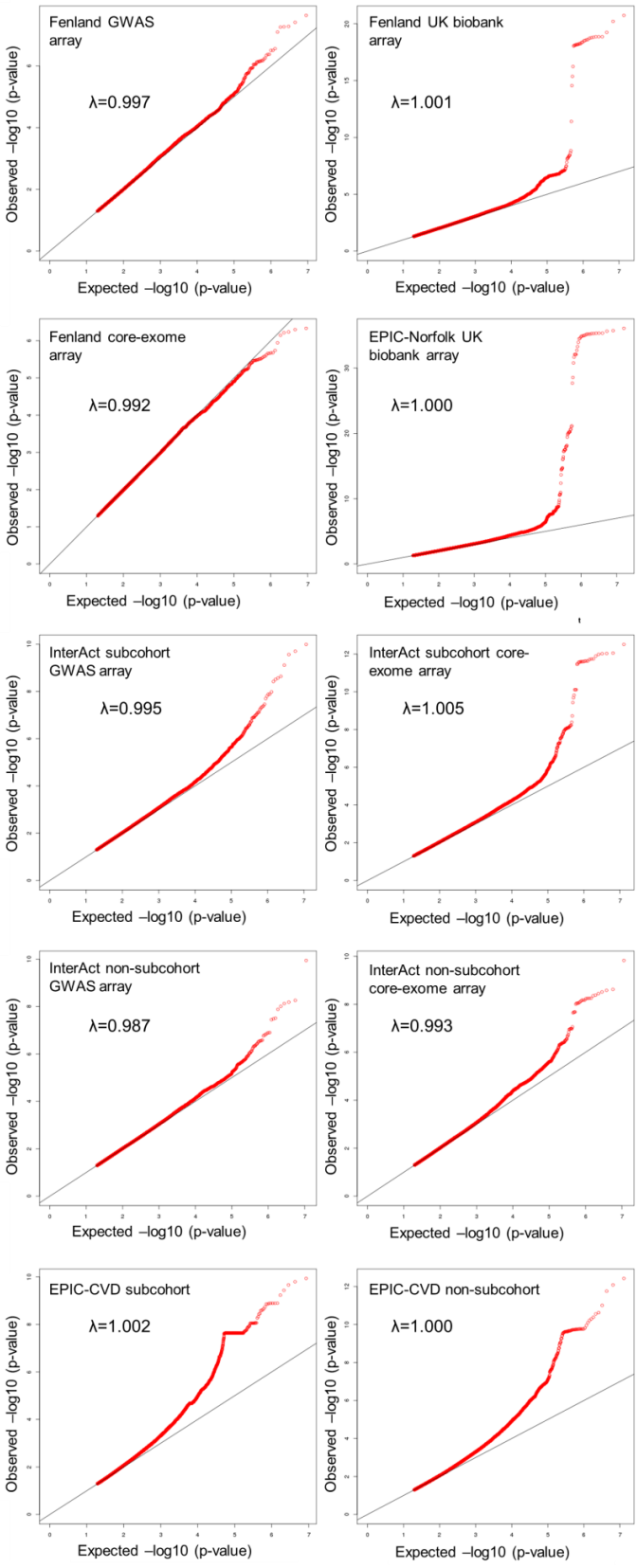
Online Supplemental Files

Supplemental Figure 2 Manhattan plot of all SNPs from the genome-wide meta-analysis of plasma vitamin C concentration. SNPs are plotted on the x-axis according to their positions on each chromosome. The red line indicates the threshold for genome-wide significance ($p<5\times10^{-8}$). Blue points represent SNPs in a 100-kb region centered on the genome-wide significant hits. Loci are annotated with names of the genes closest to the significant SNPs.



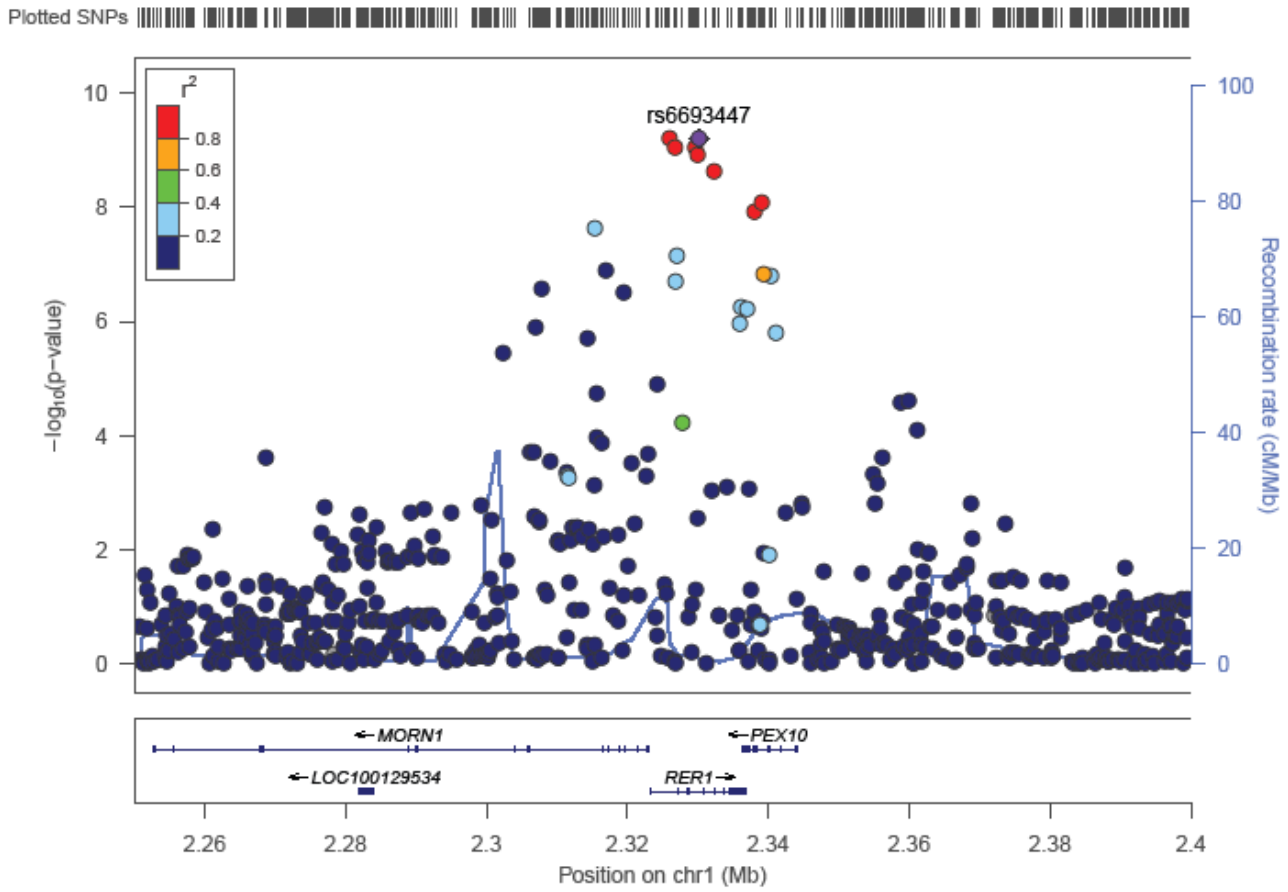
Online Supplemental Files

Supplemental Figure 3 Quantile-quantile plot for the genome-wide meta-analysis results of plasma vitamin C levels in each participating cohort



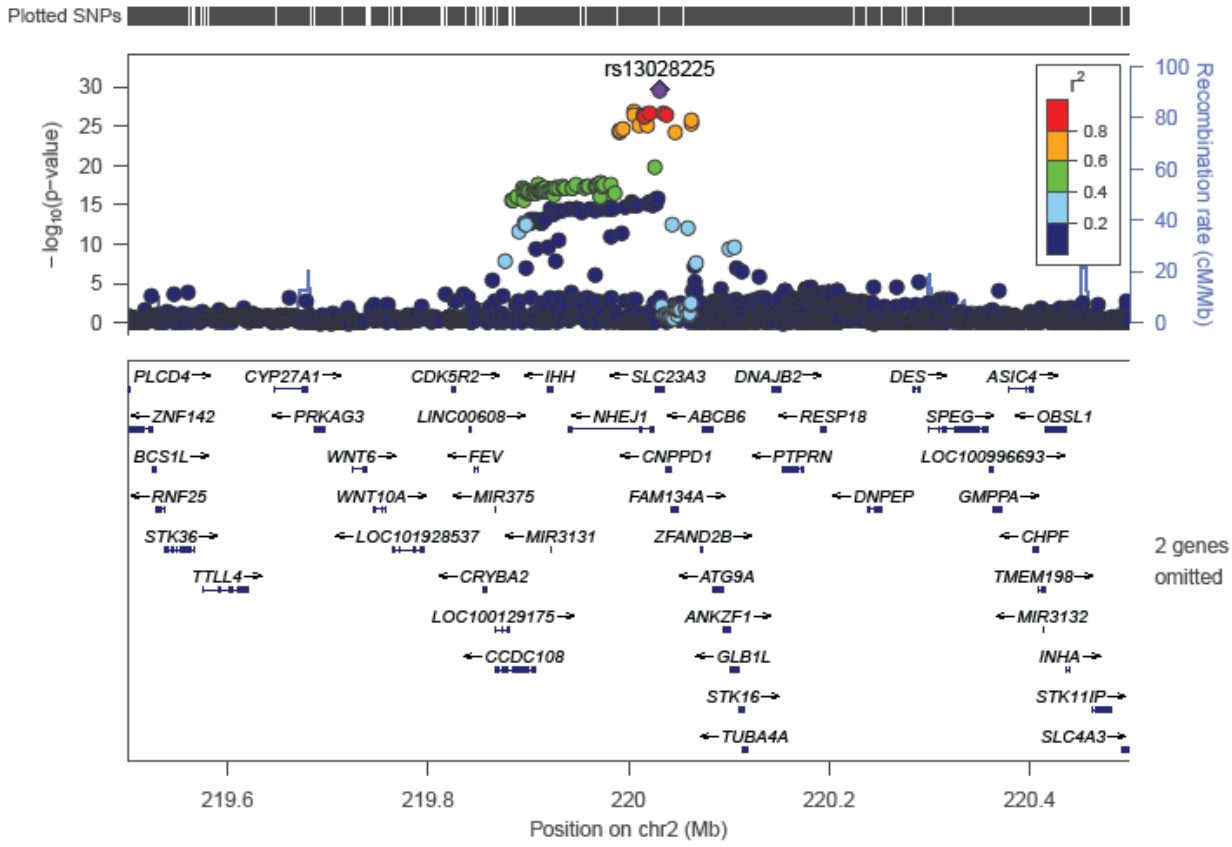
Online Supplemental Files

Supplemental Figure 4 Regional plots for SNPs on chromosome 1 identified in the genome-wide meta-analysis



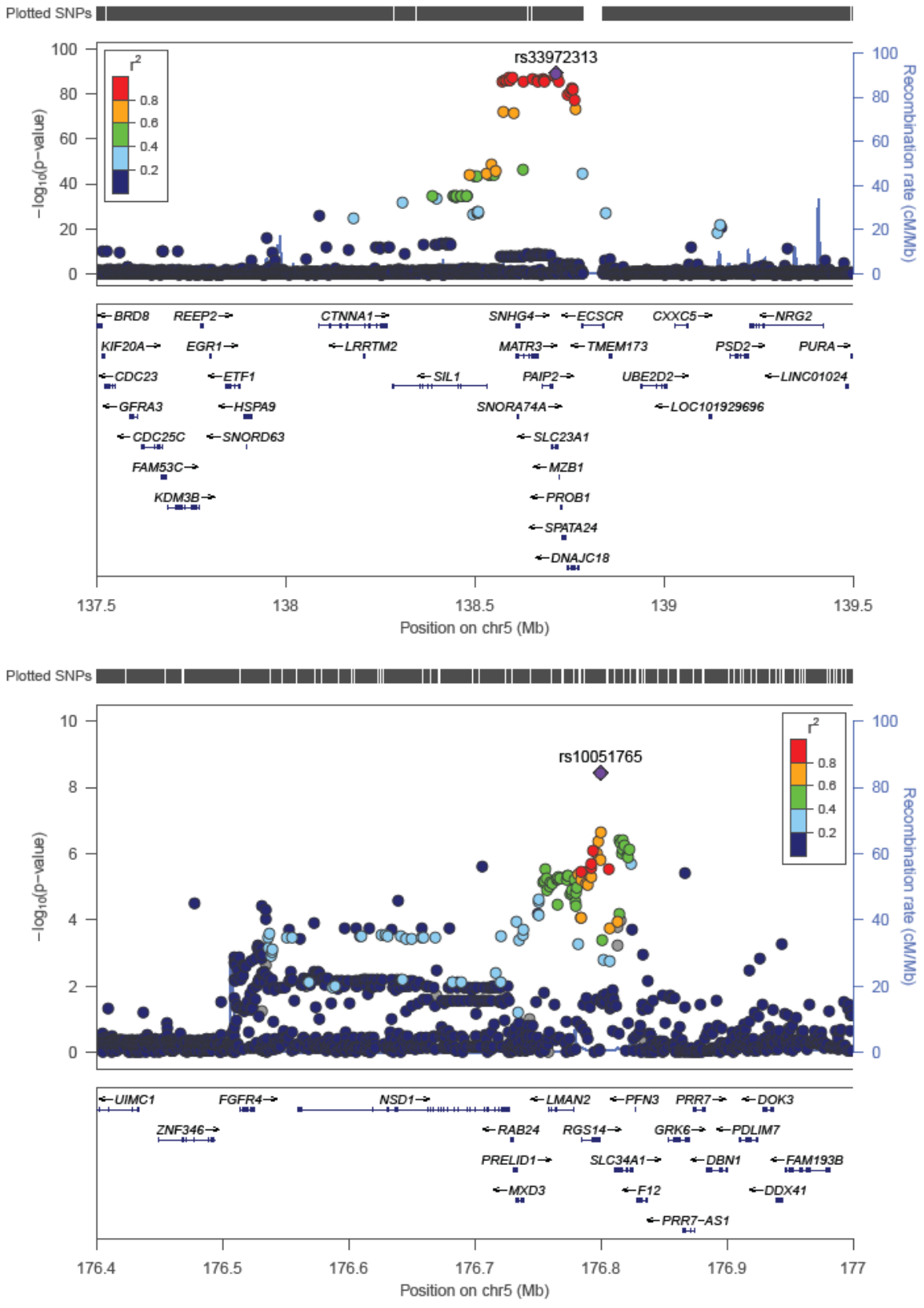
Online Supplemental Files

Supplemental Figure 5 Regional plots for SNPs on chromosome 2 identified in the genome-wide meta-analysis



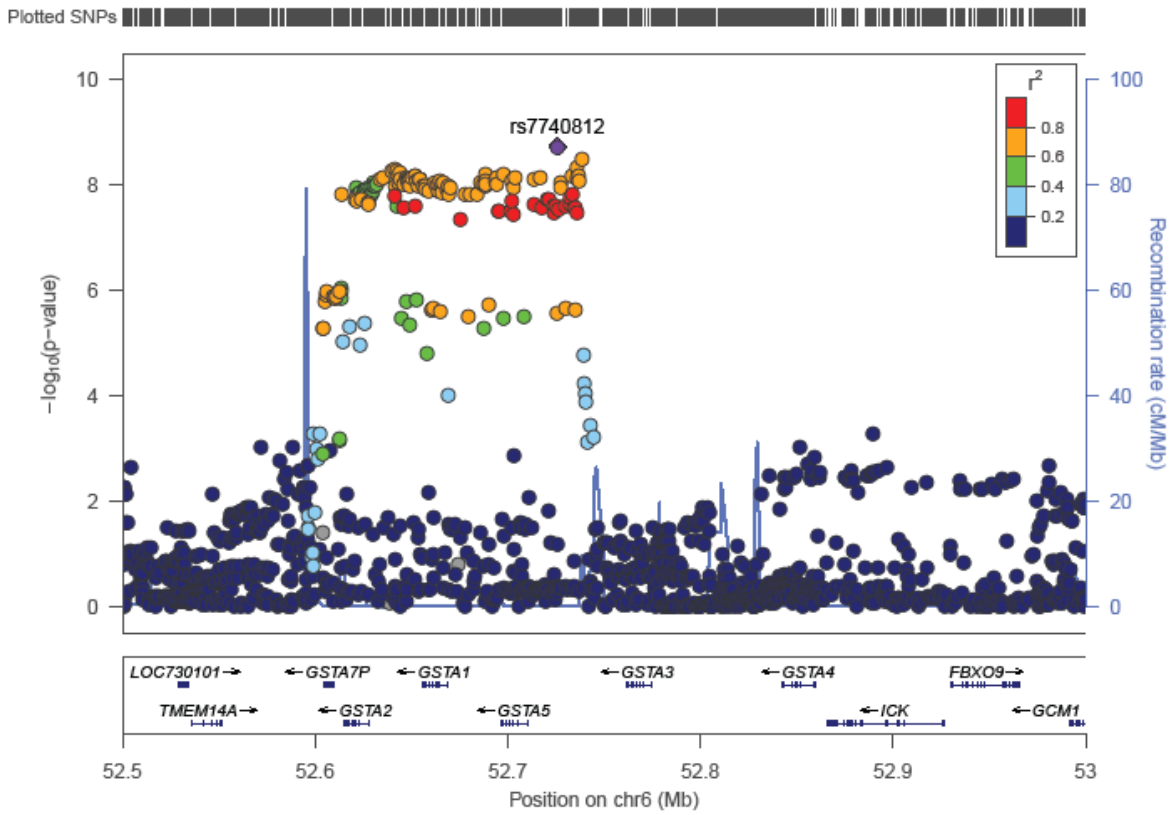
Online Supplemental Files

Supplemental Figure 6 Regional plots for SNPs on chromosome 5 identified in the genome-wide meta-analysis



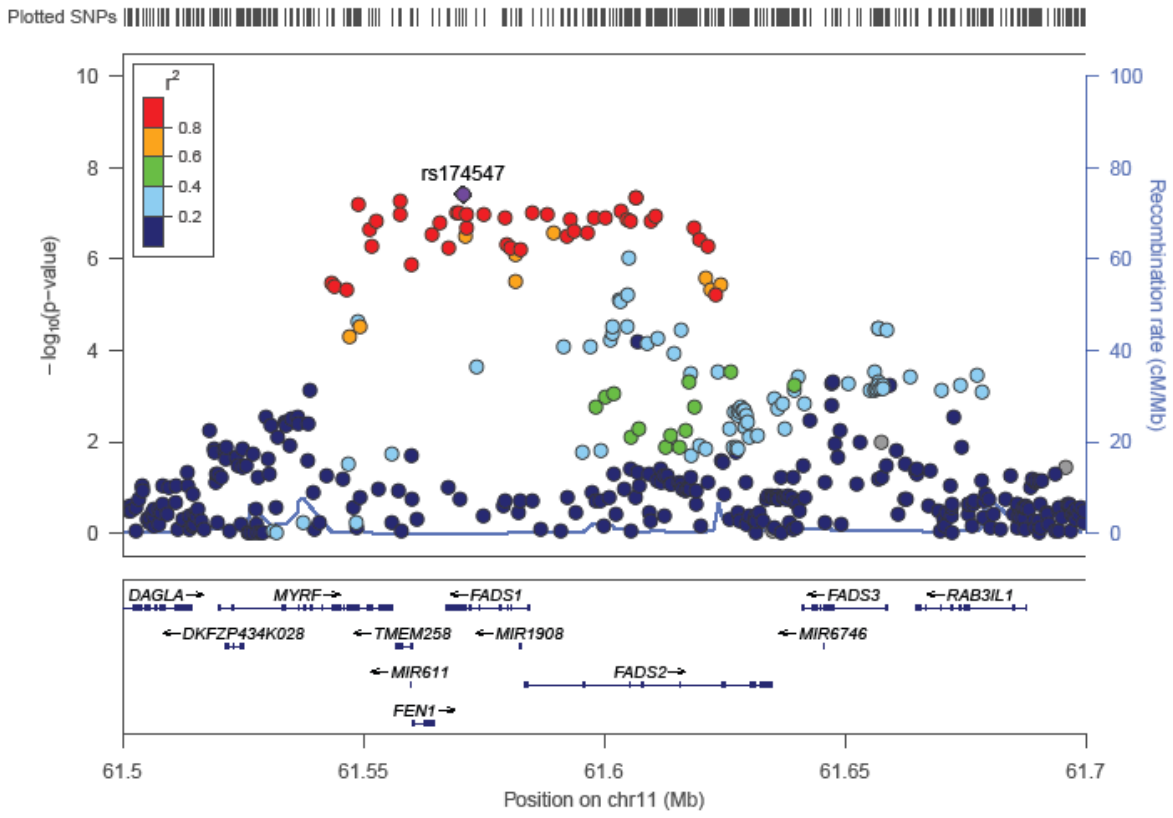
Online Supplemental Files

Supplemental Figure 7 Regional plots for SNPs on chromosome 6 identified in the genome-wide meta-analysis



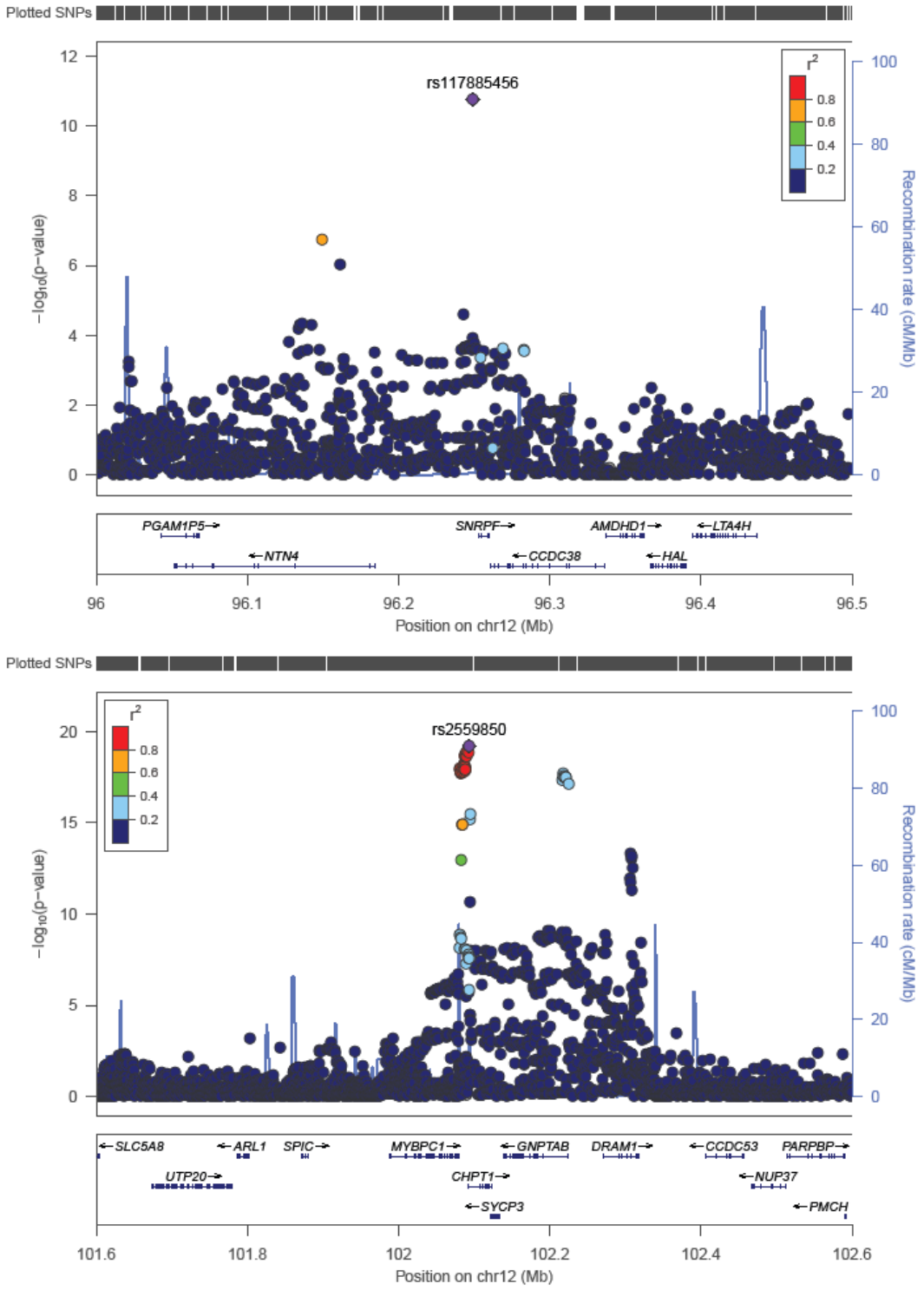
Online Supplemental Files

Supplemental Figure 8 Regional plots for SNPs on chromosome 11 identified in the genome-wide meta-analysis



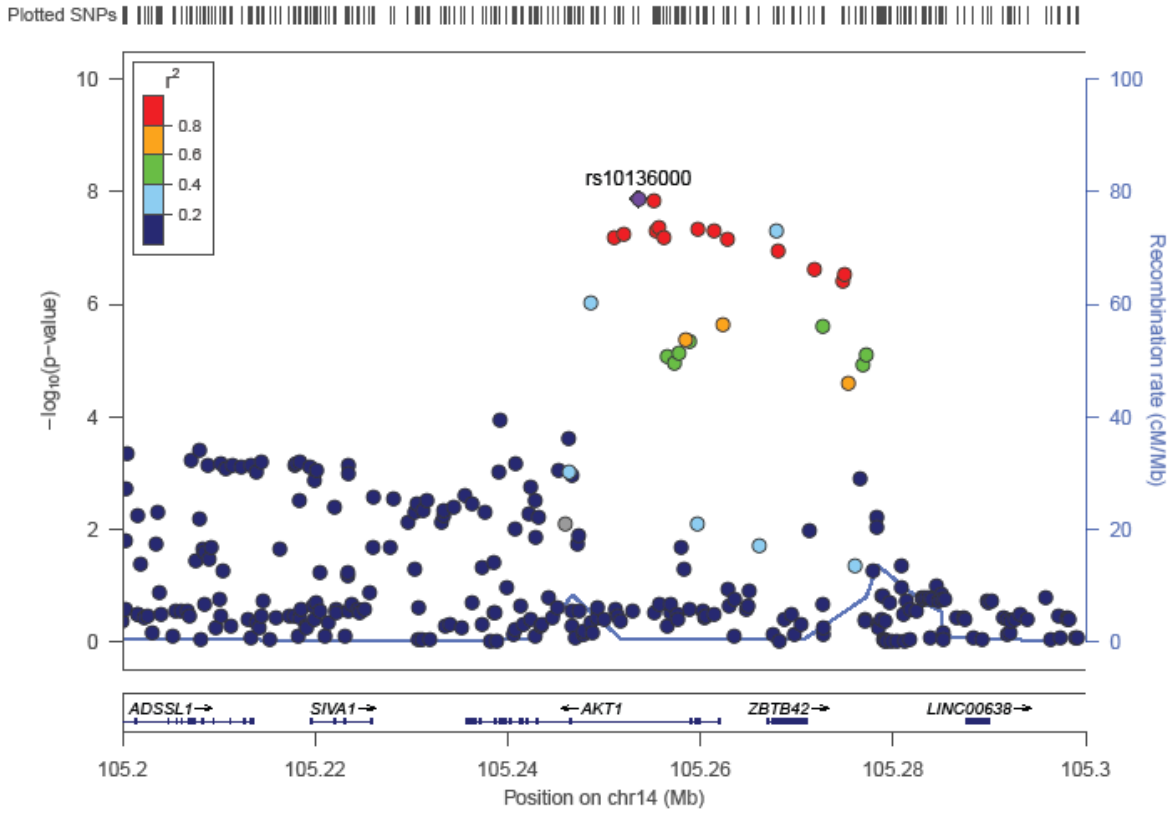
Online Supplemental Files

Supplemental Figure 9 Regional plots for SNPs on chromosome 12 identified in the genome-wide meta-analysis



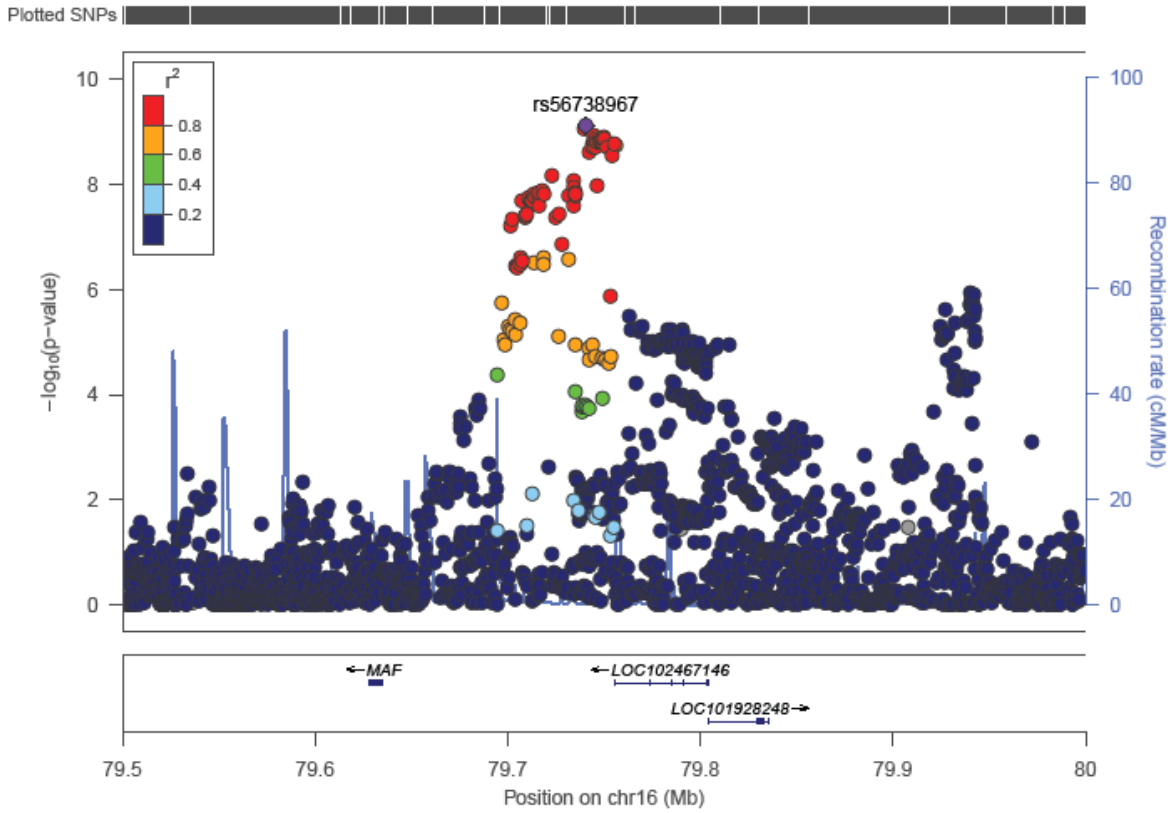
Online Supplemental Files

Supplemental Figure 10 Regional plots for SNPs on chromosome 14 identified in the genome-wide meta-analysis



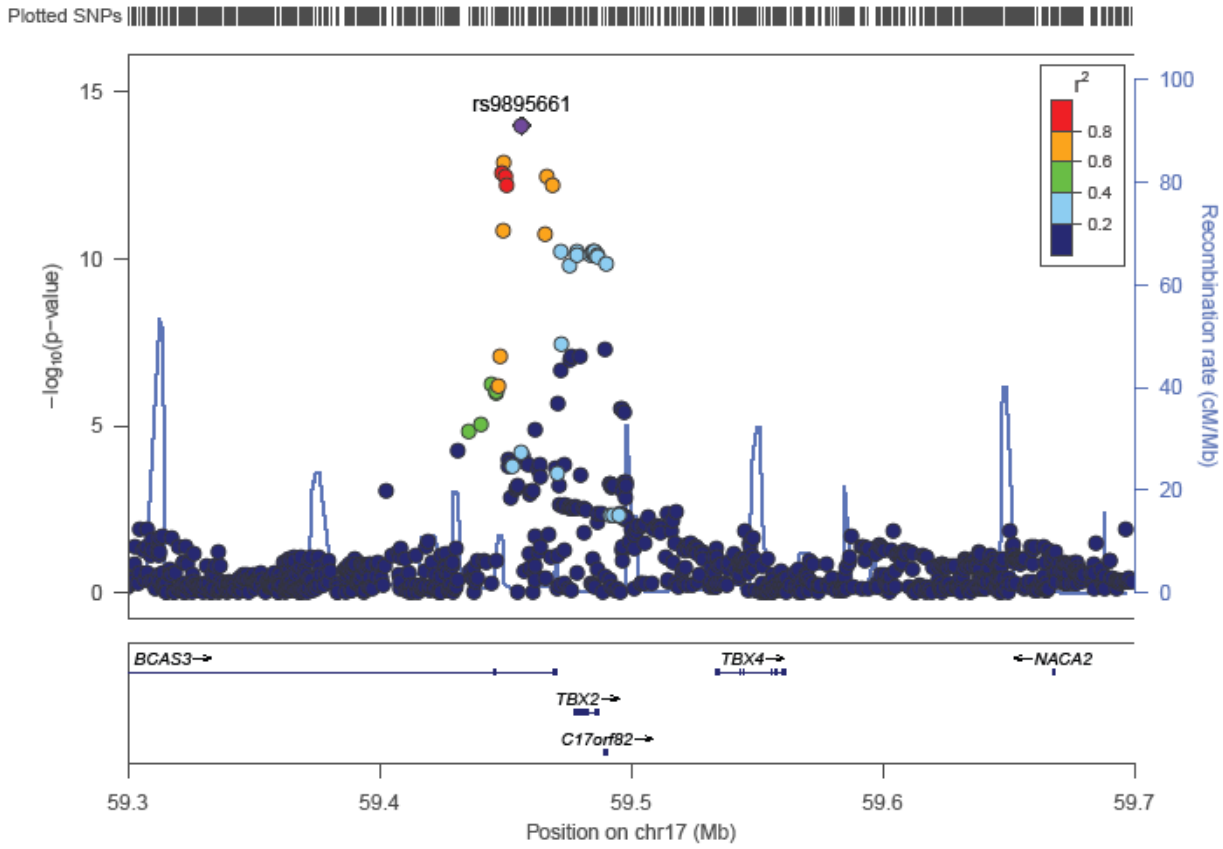
Online Supplemental Files

Supplemental Figure 11 Regional plots for SNPs on chromosome 16 identified in the genome-wide meta-analysis



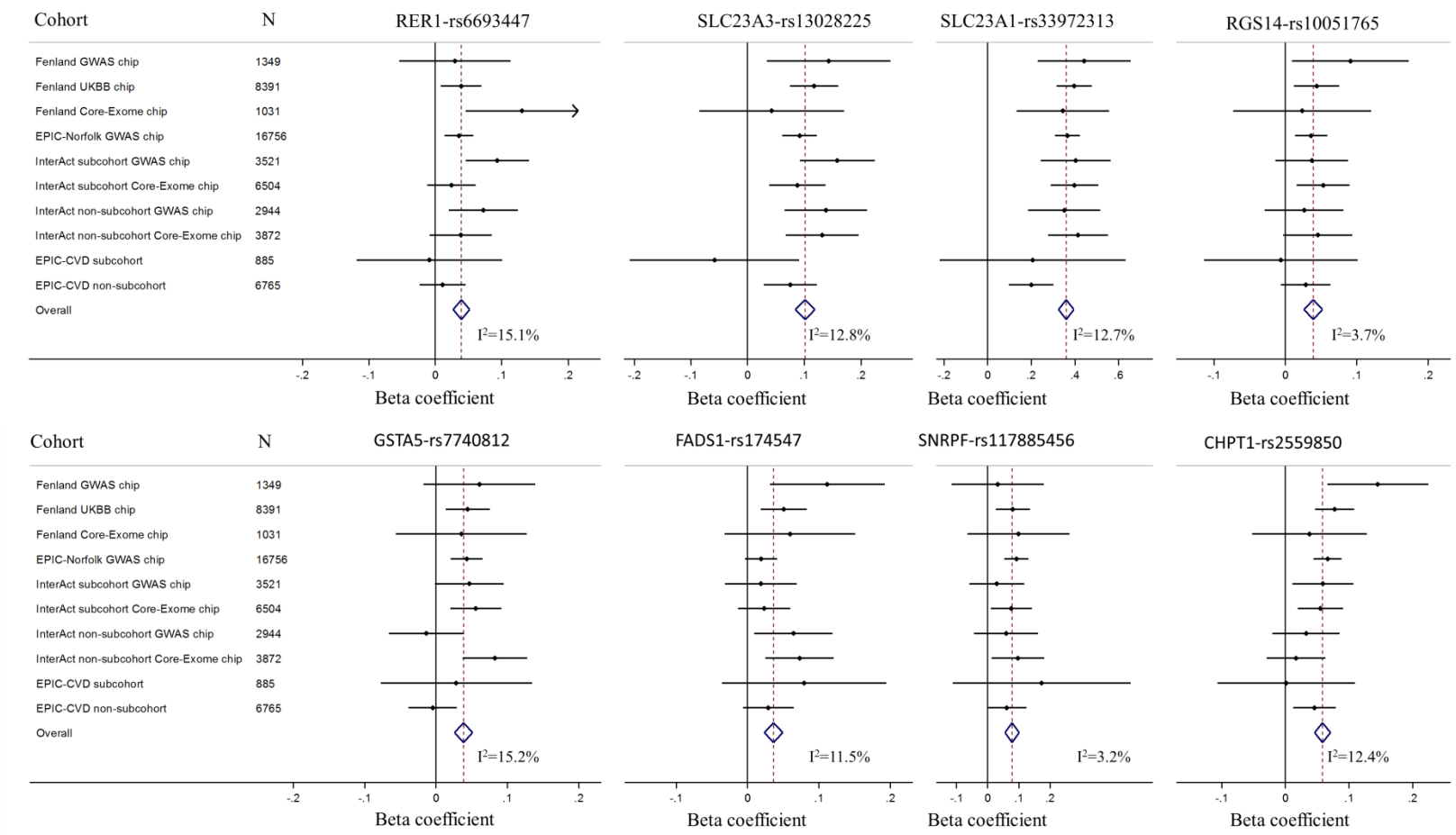
Online Supplemental Files

Supplemental Figure 12 Regional plots for SNPs on chromosome 17 identified in the genome-wide meta-analysis

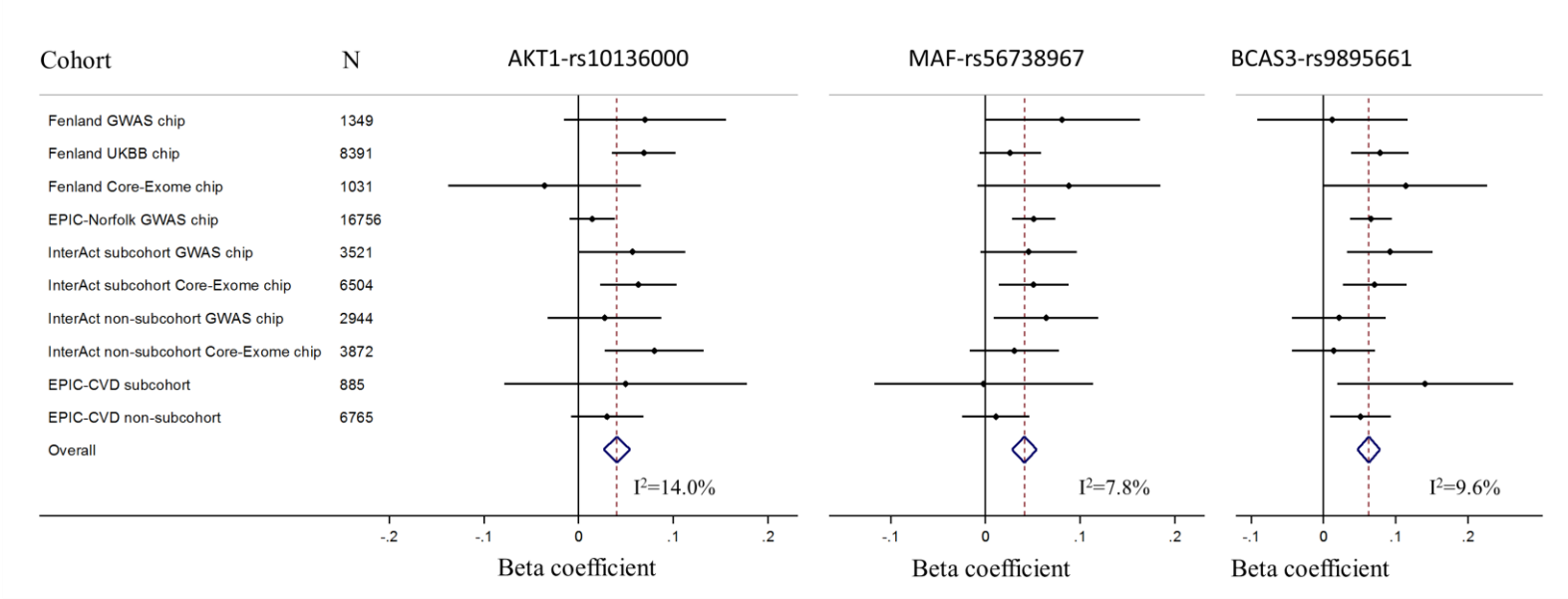


Online Supplemental Files

Supplemental Figure 13 Association of GWAS-identified lead SNPs with plasma vitamin C by individual cohorts

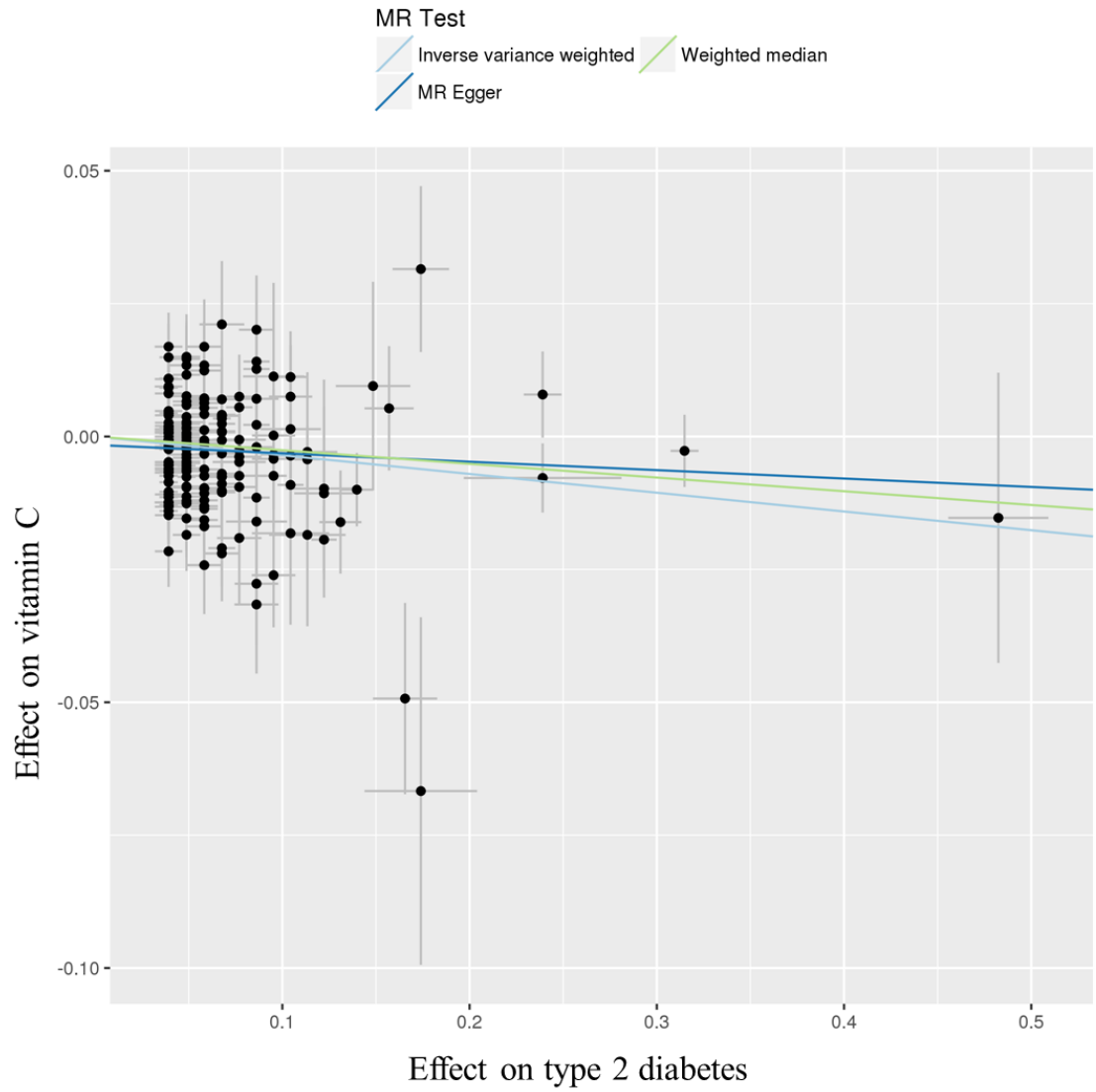


Online Supplemental Files



Online Supplemental Files

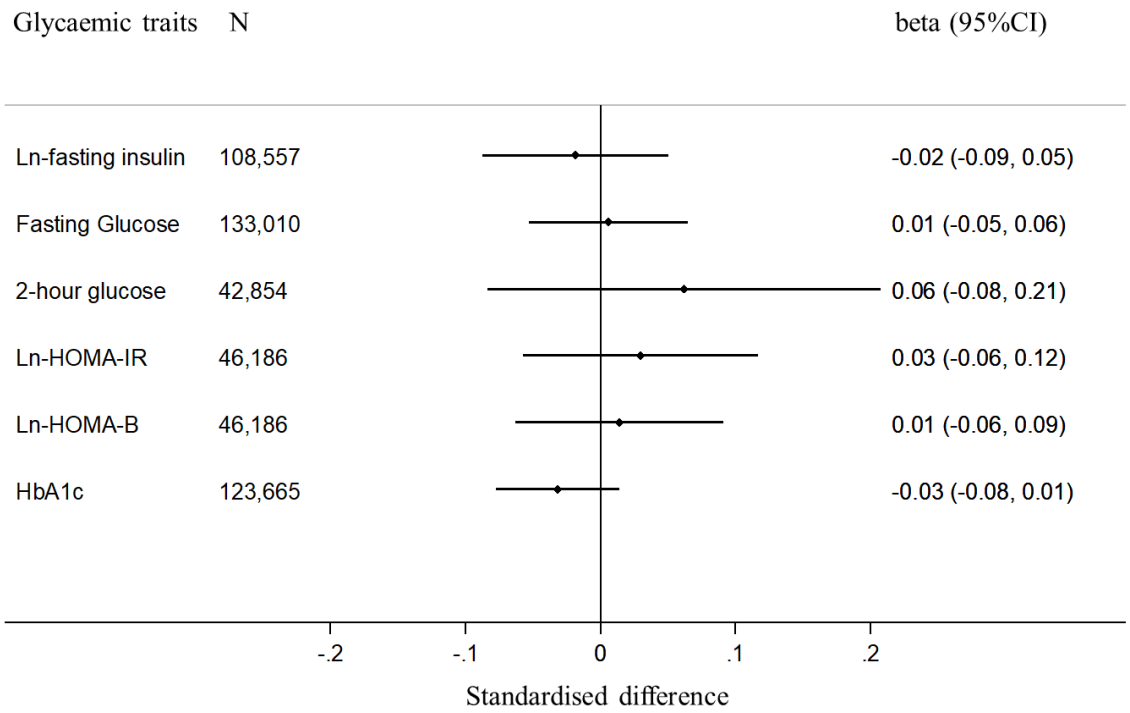
Supplemental Figure 14 Mendelian randomization estimate for the causal effect of type 2 diabetes on vitamin C. In the Mendelian randomization (MR) analysis of the association of genetically predicted type 2 diabetes risk with vitamin C, we used three methods: inverse variance weighted, MR-Egger regression and weighted median, and presented the results below. Significant heterogeneity (P<0.001) was observed in the MR analysis and results from the weighted median method were highlighted in the manuscript. A total of 231 independent SNPs of type 2 diabetes were used as genetic instrument (Nat Genet. 2018 Nov;50(11):1505-1513. doi: 10.1038/s41588-018-0241-6).



Inverse-variance weighted		Cochran's Q test	MR-Egger regression		MR-Egger intercept		Weighted mean	
Beta (se)	p	p	Beta (se)	p	Beta (se)	p	Beta (se)	p
-0.035 (0.009)	0.0001	0.0001	-0.016 (0.018)	0.37	-0.002 (0.001)	0.20	-0.026 (0.013)	0.056

Online Supplemental Files

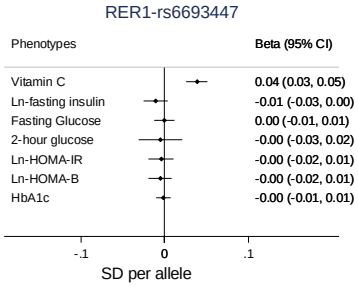
Supplemental Figure 15 Mendelian randomization for effect of a 1-SD higher plasma vitamin C concentration on glycemic traits. Estimates and 95% CIs are shown for beta, i.e. the difference in SD units of the outcome per 1-SD genetic predicted increase of plasma vitamin C concentration. Fasting glucose, 2-hour glucose, ln-transformed fasting insulin, ln-transformed HOMA-IR, ln-transformed HOMA-B and HbA1c are from the Meta-Analysis of Glucose and Insulin-related traits Consortium [MAGIC]; SD, standard deviation; HOMA-IR, homeostatic model assessment of insulin resistance; HOMA-B, homeostatic model assessment of beta-cell function; HbA1c, hemoglobin A1c.

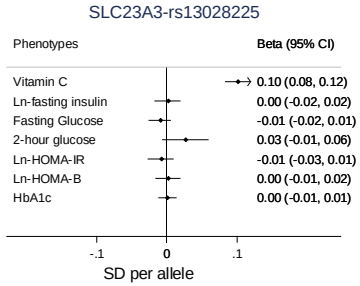


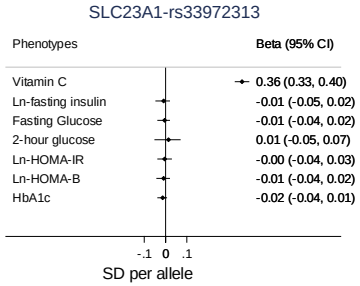
Online Supplemental Files

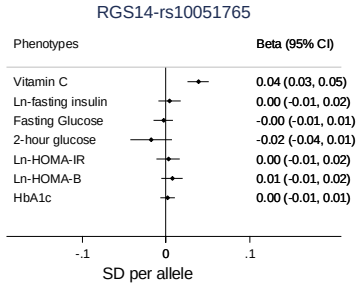
Supplemental Figure 16 Association of vitamin C-raising alleles with continuous glycemic traits. Associations are in standardized units (in SD unit) per vitamin C-raising allele. Fasting glucose (n=133,010), 2-hour glucose (n=42,854), ln-transformed fasting insulin (n=108,557), ln-transformed HOMA-IR (n=46,186), ln-transformed HOMA-B (n=46,186) and HbA1c (n=123,665) are from the Meta-Analysis of Glucose and Insulin-related traits Consortium (MAGIC). 10 SNPs used as genetic instrument in the Mendelian randomization were assessed for their associations with glycemic traits. For SNRPF-rs117885456, we used the estimates of this SNP derived from Fenland, as the information of this SNP was not available in the MAGIC and no proxy SNP was found.

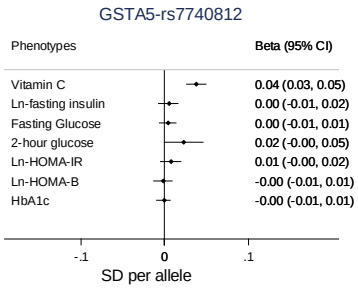
Online Supplemental Files

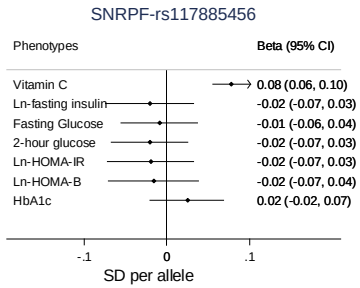


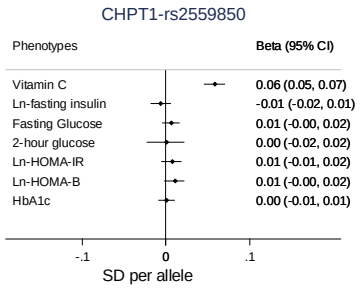


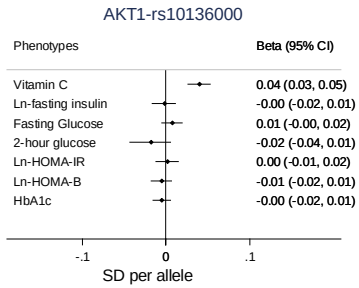


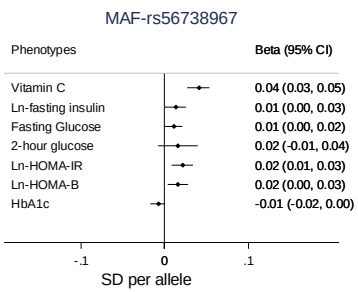


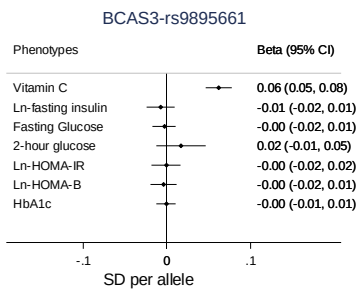












Online Supplemental Files

Supplemental Text - Supplemental materials

1. Meta-analysis of genome-wide association studies

Fenland Study

The Fenland study is an ongoing, population-based cohort study including 12,435 non-diabetic participants aged 29-62 years (born between 1950 and 1975) at baseline in Cambridgeshire, UK (http://www.mrc-epid.cam.ac.uk/research/studies/fenland/). The recruitment of Fenland participants began in 2005 from general practice lists in Ely, Wisbech, Cambridge and surrounding villages, and finished in 2015. In Fenland study, genome-wide genotyping was performed in three waves, and three arrays were used: Affymetrix genome-Wide Human SNP Array 5.0 (n=1,404), Affymetrix UK Biobank Axiom Array (n=8,994), and Illumina Human CoreExome array (n=1,060). Genotype imputation was performed to the Haplotype Reference Consortium (HRC) reference panel using IMPUTE4 or the Sanger imputation server. There were 10,771 participants with both genotyping and plasma vitamin C data (n=1,349 for Affymetrix genome-Wide Human SNP Array 5.0; n=8,391 for Affymetrix UK Biobank Axiom Array; n=1,031 for Illumina Human CoreExome array) included in the present analysis and a GWAS for each genotyping array was conducted.

Measurement of 175 metabolites in the Fenland study was conducted by the AbsoluteIDQ® Biocrates p180 Kit (Biocrates Life Sciences AG, Innsbruck, Austria) as reported previously (1, 2). Briefly, we used a Waters Acquity ultra-performance liquid chromatography (UPLC; Waters ltd, Manchester, UK) system coupled to an ABSciex 5500 Qtrap mass spectrometer (Sciex ltd, Warrington, UK). Samples were derivatized and extracted using a Hamilton STAR liquid handling station (Hamilton Robotics Ltd, Birmingham, UK). Flow injection analysis coupled with tandem mass spectrometry (FIA-MS/MS) using multiple reaction monitoring (MRM) in positive mode ionisation was performed to measure the relative levels of acylcarnitines, phosphatidylcholines, lysophosphatidylcholines and sphingolipids. The level of hexose was measured in negative ionisation mode. Ultra-performance liquid chromatography coupled with tandem mass spectrometry (UPLC-MS/MS) using MRM was performed to measure the concentration of amino acids and biogenic amines. We applied extensive quality control procedures for the raw metabolite readings. Firstly, we excluded from any further analysis metabolites for which the number of measurements below the limit of quantification (LOQ) exceeded 5% of measured samples. Excluded metabolites were carnosine, dopamine, putrescine, asymmetric dimethyl arginine, dihydroxyphenylalanine, nitrotyrosine, spermine, sphingomyelines SM(22:3), SM(26:0), SM(26:1), SM(24:1-OH), phosphatidylcholine acyl-alky 44:4, and phosphatidylcholine diacyl C30:2. Secondly, in samples with detectable but not quantifiable peaks, we assigned random values between 0 and the run-specific LOQ of a given metabolite. Finally, we corrected for batch-effects with a "location-scale" approach, i.e. with normalization for mean and standard deviation of batches (3). There were 9,237 individuals with both genotype and metabolomics data.

EPIC-InterAct study

EPIC-InterAct is a case-cohort study nested within the European Prospective Investigation into Cancer and Nutrition study, including 12,403 incident type 2 diabetes cases verified from among 340,234 participants across eight European countries (France, Italy, Spain, UK, Netherlands, Germany, Sweden and Denmark), and 16,154 subcohort (with 778 verified

Online Supplemental Files

incident type 2 diabetes cases as feature of case-cohort design) members randomly selected from these participants (4).

Ascertainment of incident type 2 diabetes cases up until Dec 31, 2007, was conducted through a review of multiple sources of evidence, including self-report, linkage to primary care registers, secondary care registers, medication use (drug registers), hospital admissions and mortality data. There were no type 2 diabetes cases ascertained solely by self-report and we sought further evidence for all cases with information on incident type 2 diabetes from fewer than two independent sources at minimum. type 2 diabetes cases in Denmark and Sweden were identified through local and national diabetes and pharmaceutical registers, and were considered to be verified (4).

Genome-wide genotyping in the EPIC-InterAct study was performed among 23,019 participants with two arrays: Illumina HumanCoreExome array (n=13,725) and Illumina 660W-Quad BeadChip (n=9,294). A total of 16,841 participants (n= 10,376 for Illumina HumanCoreExome chip; n=6,465 for Illumina 660W-Quad BeadChip) had available genome-wide genotyping and plasma vitamin C data and contributed to the present GWAS. Genotype imputation was performed to the Haplotype Reference Consortium reference panel using IMPUTE v2 software. The GWAS in EPIC-InterAct was conducted stratified by the subcohort status (subcohort and non-subcohort) and GWAS arrays used respectively. Therefore, four GWAS was performed in InterAct (Supplemental Table 1)

EPIC-Norfolk study

EPIC-Norfolk is one of the two UK constituents of the European Prospective Investigation into Cancer and Nutrition study. Between 1993 and 1997, a total of 25,639 men and women aged 40-79 were recruited. Genome-wide genotyping in EPIC-Norfolk was conducted among 21,044 participants using the Affymetrix UK Biobank Axiom Array. Genotype imputation was performed to the Haplotype Reference Consortium (HRC) reference panel using IMPUTE v2 software. The EPIC-Norfolk case-cohort study (n=1503) was part of the InterAct study, and therefore these participants were excluded to avoid duplication with InterAct for the GWAS analysis. Finally, 16,756 participants with both genome-wide genotyping and plasma vitamin C data contributed to the vitamin C GWAS in the EPIC-Norfolk study.

EPIC-CVD study

EPIC-CVD study is a large, prospective, case-cohort study nested within the EPIC study (5), with a random subcohort of 18,249 participants and 24,557 participants who later developed CVD during the follow-up, stratified by center and selected from the EPIC participants with a store blood sample available. EPIC-CVD involved participants from ten European countries (France, Italy, Spain, UK, Netherlands, Germany, Sweden, Denmark, Norway and Greece). Genome-wide genotyping in EPIC-CVD was conducted using Illumina Human Exome v1.1 SNP array. Genotype imputation was performed to the Haplotype Reference Consortium reference panel using IMPUTE v2 software. As EPIC-CVD shared the random subcohort with the EPIC-InterAct for eight participating countries, we excluded any overlapping participants with InterAct for the GWAS. A GWAS was separately performed for subcohort (n=885) and non-subcohort (n=6,765) participants in EPIC-CVD, with 7,650 participants included in the analyses.

Online Supplemental Files

Meta-analysis of genome-wide association studies
As part of the quality control procedures, we dropped variants for which the standard error exceeded 5, imputation quality ('info score')<0.4, the p-value for violations of Hardy-Weinberg equilibrium was below 10^{-6}, the absolute value of the beta coefficient exceeded 5 or the minor allele frequency was below 1%.

Genome-wide association results from the 10 GWAS analyses from EPIC-InterAct, EPIC-Norfolk, Fenland and EPIC-CVD were meta-analyzed using METAL software (6). We performed a meta-analysis of beta coefficients and standard errors with genomic control correction. A Manhattan plot and QQ plot for the association of genetic variants with plasma vitamin C levels was generated using R package EasyStrata version 8.5 (7). A conventional threshold for genome-wide significance $p < 5\times10^{-8}$ was used to define loci associated with plasma vitamin C levels, and a lead single nucleotide polymorphism (SNP) at a given genomic locus was identified as the SNP with the lowest p-value within a 1 million base-pair window. Regional association plots were drawn using LocusZoom software (8).

Functional annotation and pathway analyses
We used MAGENTA (9) (https://software.broadinstitute.org/mpg/magenta/) to examine the genome-wide genetic associations with biological pathways defined by Gene Ontology, PANTHER, KEGG and Ingenuity. We used DEPICT (10) to prioritise genes and pathways responsible for the genetic associations of SNPs with a P-value less than 1×10^{-5}, and to highlight the relevant tissues/ cell types where associated genes are highly expressed.

We used Haploreg v4.1 (11) (http://archive.broadinstitute.org/mammals/haploreg/haploreg.php) to explore annotations of the identified lead SNPs, including eQTL, GRASP QTL, and previous GWAS hits; and we used BIOS QTL database to explore the influence of genetic variants on the DNA methylation status (http://genenetwork.nl/biosqtlbrowser/) (12, 13).

2. Plasma vitamin C-raising alleles and type 2 diabetes
UK Biobank
The UK Biobank study is a population-based cohort of around 0.5 million UK individuals aged 40-69 years recruited between 2006 and 2010 across UK (14). Both genotype data and prevalent/incident type 2 diabetes information was available among a total 449,333 individuals in the initial UK Biobank dataset (24,758 cases and 424,575 non-cases). Type 2 diabetes was defined on the basis of self-reported physician diagnosis at nurse interview or digital questionnaire, age at diagnosis>36 years, and use of oral anti-diabetic medications. We meta-analyzed DIAMANTE (European) plus EPIC-Norfolk GWAS association results with UK Biobank GWAS results with fixed effect models by using the STATA software.

DIAbetes Meta-ANalysis of Trans-Ethnic association studies (DIAMANTE) (European)
DIAMANTE is a consortium which published meta-analysis of genome-wide association studies of type 2 diabetes in individuals of different ethnicities, and the meta-analysis results from DIAMANTE (European ethnicity) including a total of 74,124 type 2 diabetes cases and 824,006 controls were published and publicly available(15). In the present study, we used the

Online Supplemental Files

summary statistics from the DIAMANTE European excluding UK biobank study, thus including 55,005 type 2 diabetes cases and 400,308 controls.

EPIC-Norfolk
We used the type 2 diabetes GWAS results from EPIC-Norfolk excluding type 2 diabetes cases already included in the DIAMANTE study, and finally included 1,220 type 2 diabetes cases and 18,026 controls.

The estimates of lead vitamin C-related SNPs with type 2 diabetes risk were extracted from the meta-analysis results of the above UK Biobank, DIAMANTE and EPIC-Norfolk study, including up to 80,983 type 2 diabetes cases and 842,909 non-cases.

3. Prospective association of plasma vitamin C with incident type 2 diabetes
In the EPIC-InterAct study, we used Prentice-weighted Cox regression to estimate the country-specific hazard ratio (HRs) and 95% CIs for associations per 1-SD (calculated from the subcohort, 19.2 μmol/L) of plasma vitamin C with incident type 2 diabetes, which allows for over-representation of cases in a case-cohort design, and then pooled the results via random-effects meta-analysis. The adjusted covariates included sex, center, physical activity (inactive, moderately inactive, moderately active, active), smoking status (never, former, current), employment (no, yes), marital status (single, married, separated/divorced, widowed), education (low, middle, high), alcohol drinking (never, 0 to < 6, 6 to < 12, 12 to < 24, and ≥ 24 g/day), total energy intake (continuous), individual plasma carotenoids, BMI (continuous) and waist circumference (continuous). A total of 8,133 type 2 diabetes cases and 11,073 non-cases were included. We only used the EPIC-InterAct study in our estimate for the prospective association of vitamin C with type 2 diabetes, because EPIC-InterAct is so far the largest study on incident type 2 diabetes and to the best of our knowledge, there is only one previous study (EPIC-Norfolk) reporting the prospective association (16). The EPIC-InterAct study already included all the type 2 diabetes cases from that EPIC-Norfolk study (16).

Supplemental text reference

1. Illig T, Gieger C, Zhai G, Romisch-Margl W, Wang-Sattler R, Prehn C, et al. A genome-wide perspective of genetic variation in human metabolism. Nature genetics. 2010;42(2):137-41. Epub 2009/12/29.
2. Lotta LA, Scott RA, Sharp SJ, Burgess S, Luan J, Tillin T, et al. Genetic Predisposition to an Impaired Metabolism of the Branched-Chain Amino Acids and Risk of Type 2 Diabetes: A Mendelian Randomisation Analysis. PLoS medicine. 2016;13(11):e1002179. Epub 2016/11/30.
3. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics. 2007;8(1):118-27.
4. Langenberg C, Sharp S, Forouhi NG, Franks PW, Schulze MB, et al. 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(9):2272-82. Epub 2011/07/01.
5. Danesh J, Saracci R, Berglund G, Feskens E, Overvad K, Panico S, et al. 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. European journal of epidemiology. 2007;22(2):129-41. Epub 2007/02/14.

Online Supplemental Files

6. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics. 2010;26(17):2190-1. Epub 2010/07/10.
7. Winkler TW, Kutalik Z, Gorski M, Lottaz C, Kronenberg F, Heid IM. EasyStrata: evaluation and visualization of stratified genome-wide association meta-analysis data. Bioinformatics. 2015;31(2):259-61. Epub 2014/09/28.
8. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, et al. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics. 2010;26(18):2336-7. Epub 2010/07/17.
9. Segre AV, Groop L, Mootha VK, Daly MJ, Altshuler D, Consortium D, et al. Common Inherited Variation in Mitochondrial Genes Is Not Enriched for Associations with Type 2 Diabetes or Related Glycemic Traits. Plos Genet. 2010;6(8).
10. Pers TH, Karjalainen JM, Chan Y, Westra HJ, Wood AR, Yang J, et al. Biological interpretation of genome-wide association studies using predicted gene functions. Nat Commun. 2015;6.
11. Ward LD, Kellis M. HaploReg v4: systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease. Nucleic Acids Res. 2016;44(D1):D877-D81.
12. Bonder MJ, Luijk R, Zhernakova DV, Moed M, Deelen P, Vermaat M, et al. Disease variants alter transcription factor levels and methylation of their binding sites. Nature genetics. 2017;49(1):131-8.
13. Zhernakova DV, Deelen P, Vermaat M, van Iterson M, van Galen M, Arindrarto W, et al. Identification of context-dependent expression quantitative trait loci in whole blood. Nature genetics. 2017;49(1):139-45.
14. Collins R. What makes UK Biobank special? Lancet. 2012;379(9822):1173-4.
15. Mahajan A, Taliun D, Thurner M, Robertson NR, Torres JM, Rayner NW, et al. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. Nature genetics. 2018;50(11):1505-+.
16. Harding AH, Wareham NJ, Bingham SA, Khaw K, Luben R, Welch A, 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. Archives of internal medicine. 2008;168(14):1493-9. Epub 2008/07/30.