SUPPLEMENTARY DATA

Supplementary Table 1. Traits associated with LP SNP rs4988235 or proxies (R^2>0.8) according to GWAS studies included in phenoscanner and GWAS catalog.

Trait	SNP	Proxy SNP	r^2	Consortium	N	Effect allele	Other allele	EAF	Beta	SE	P value	Unit
Hip circumference	rs4988235			GIANT	207,713	T	C	0.76	0.022	0.0039	$2*10^{-8}$	IVNT
LDL cholesterol	rs4988235			GLGC	169,531	T	C	0.52	-0.028	0.0042	$3*10^{-11}$	IVNT
Total cholesterol	rs4988235			GLGC	183,761	T	C	0.52	-0.031	0.004	$4*10^{-14}$	IVNT
Height		rs6754311	0.965	GIANT	250,456	A	G	0.76	0.018	0.0034	$5*10^{-8}$	Z-score

SNP = single nucleotide polymorphism, EAF = effect allele frequency, SE = standard error, GIANT = Genetic Investigation of Antrhropometric Traits, GLGC = Global Lipids Genetics Consortium, IVNT = inverse normally rank transformed phenotype

SUPPLEMENTARY DATA

Supplementary Table 2. Dietary intake* of participants from France and Italy in the EPIC-InterAct study subcohort.

	France, n= 332		Italy, n=1467	
	Dietary intake	No intake	Dietary intake	No intake
Energy (kcal/day)	2152 (1789-2569)		2241 (1801-2729)	
Milk (g/day)	15 (0-150)	47.6	120 (0-189)	26.7
Non-milk dairy (g/day)	209 (144-269)	0	82 (47-138)	0
Milk beverages (g/day)	0 (0-0)	88.6		
Milk for coffee (g/day)				
Dairy creams (g/day)	1 (1-5)	10.2	0 (0-1)	39.9
Cream desserts (g/day)	10 (0-29)	32.2		
Curd (g/day)	19 (0-64)	33.4		
Yoghurt (g/day)	81 (32-113)	11.7	8 (0-54)	40.5
Ice cream (g/day)	2 (0-7)	32.2	11 (4-21)	8.5
Cheese (g/day)	53 (30-86)	0.9	58 (35-85)	0
Potatoes (g/day)	53 (32-81)	4.2	22 (12-38)	0.4
Pasta/rice (g/day)	68 (42-106)	2.7	138 (83-210)	0.1
Bread (g/day)	111 (73-166)	0.9	131 (95-203)	0.1
Cereal (g/day)	202 (148-279)	0.3	305 (216-418)	0
Vegetables (g/day)	269 (193-375)	0	155 (105-216)	0
Legumes (g/day)	13 (3-29)	19.6	5 (2-12)	12.1
Fruits (g/day)	245 (158-340)	1.8	293 (199-411)	0.1
Nuts (g/day)	3 (0-7)	29.5	0 (0-1)	11
Red meat (g/day)	45 (22-73)	6.6	45 (25-66)	0.7
Poultry (g/day)	17 (3-30)	22.3	22 (12-34)	1.5
Processed meat (g/day)	25 (15-39)	1.5	19 (11-32)	0.4
Vegetable oils (g/day)	6 (3-8)	1.8	29 (21-38)	0
Margarine (g/day)	0 (0-4)	53.9	0 (0-0)	16.5
Butter (g/day)	4 (0-11)	16.3	1 (0-2)	9.5
Sugar (g/day)	35 (17-59)	5.7	34 (20-54)	0.8
Cake/cookies (g/day)	28 (14-56)	3.3	35 (15-66)	5.0
Beverages (g/day) †	1356 (958-1829)	0	172 (110-272)	1.0
Softdrinks (g/day)	0 (0-0)	80.7	0 (0-29)	54.3
Juice (g/day)	23 (0-107)	33.1	8 (0-36)	42.3
Coffee (g/day)	250 (75-413)	14.5	90 (60-130)	6.5
Tea (g/day)	41 (0-301)	40.1	5 (0-43)	42.1
Alcohol (g/day)	7 (2-17)	3.3	7 (0-24)	6.5
Wine (g/day)	38 (4-119)	23.5	54 (2-250)	24.3

*Dietary intake is expressed as median (p25-p75), no intake is expressed as percentage of non-consumers with available data. † Sum of all non-alcoholic beverages, excluding water and milk.

SUPPLEMENTARY DATA

Supplementary Table 3. Dietary intake* of participants from Spain and the United Kingdom in the EPIC-InterAct study subcohort.

	Spain, n=2544		United Kingdom n=1079	
	Dietary intake	No intake	Dietary intake	No intake
Energy (kcal/day)	2057 (1670-2556)		1945 (1597-2387)	
Milk (g/day)	200 (107-307)	7.9	295 (149-440)	2.4
Non-milk dairy (g/day)	50 (15-106)	10.4	67 (34-113)	0.6
Milk beverages (g/day)	0 (0-0)	99.7		
Milk for coffee (g/day)				
Dairy creams (g/day)	0 (0-0)	97.2	1 (0-3)	33.7
Cream desserts (g/day)	0 (0-0)	78.2	11 (1-22)	24.9
Curd (g/day)	0 (0-0)	95	1 (0-4)	48.6
Yoghurt (g/day)	0 (0-49)	58.6	18 (0-55)	29.6
Ice cream (g/day)	0 (0-0)	83.5	6 (1-11)	22.3
Cheese (g/day)	16 (1-43)	20.9	15 (5-19)	3.5
Potatoes (g/day)	71 (44-106)	1.5	107 (70-133)	0.4
Pasta/rice (g/day)	57 (38-81)	4.2	32 (14-60)	13.1
Bread (g/day)	130 (84-191)	1.0	79 (39-112)	0
Cereal (g/day)	195 (139-265)	0.2	167 (114-241)	0
Vegetables (g/day)	234 (147-340)	0.1	237 (172-319)	0
Legumes (g/day)	45 (25-69)	3.8	11 (6-21)	12.3
Fruits (g/day)	274 (161-438)	5.2	198 (118-305)	0.6
Nuts (g/day)	0 (0-4)	57.3	2 (0-4)	40.6
Red meat (g/day)	34 (16-58)	4.1	32 (16-58)	4.3
Poultry (g/day)	30 (17-49)	4.0	16 (8-49)	13.1
Processed meat (g/day)	30 (14-53)	3.2	18 (9-31)	3.2
Vegetable oils (g/day)	26 (18-36)	0	3 (2-5)	3.6
Margarine (g/day)	0 (0-1)	73.2	10 (4-25)	10.7
Butter (g/day)	0 (0-0)	93.3	0 (0-4)	28.0
Sugar (g/day)	19 (8-32)	11.6	41 (20-69)	0.7
Cake/cookies (g/day)	21 (0-55)	30.2	46 (24-86)	2.1
Beverages (g/day) †	169 (74-289)	1.9	1168 (951-1414)	0.1
Softdrinks (g/day)	0 (0-0)	79.1	32 (0-158)	25.9
Juice (g/day)	1 (0-7)	38.6	17 (1-95)	23.0
Coffee (g/day)	91 (12-175)	11.1	475 (95-488)	10.0
Tea (g/day)	0 (0-0)	96.3	475 (475-855)	5.8
Alcohol (g/day)	3 (0-19)	16.8	5 (1-11)	1.1
Wine (g/day)	2 (0-150)	47.8	18 (1-54)	23.1

*Dietary intake is expressed as median (p25-p75), no intake is expressed as percentage of non-consumers with available data. † Sum of all non-alcoholic beverages, excluding water and milk.

SUPPLEMENTARY DATA

Supplementary Table 4. Dietary intake* of participants from the Netherlands and Germany in the EPIC-InterAct study subcohort.

	Netherlands, n=1176		Germany, n=1783	
	Dietary intake	No intake	Dietary intake	No intake
Energy (kcal/day)	1903 (1581-2275)	0	1949 (1590-2430)	0
Milk (g/day)	203 (81-382)	4.7	33 (2-100)	1.1
Non-milk dairy (g/day)	169 (106-229)	0	137 (82-222)	0
Milk beverages (g/day)	15 (1-31)	22.8	1 (0-16)	40.0
Milk for coffee (g/day)	5 (0-18)	39.1	0 (0-10)	64.8
Dairy creams (g/day)	2 (1-4)	2.6	1 (1-3)	0.3
Cream desserts (g/day)	18 (6-41)	4.8	6 (2-12)	7.6
Curd (g/day)	6 (2-13)	4.8	13 (6-29)	2
Yoghurt (g/day)	49 (15-90)	3.1	53 (13-108)	7.6
Ice cream (g/day)	7 (2-12)	4.8	3 (1-6)	4.4
Cheese (g/day)	31 (20-47)	0.9	32 (17-40)	0.4
Potatoes (g/day)	82 (51-130)	0.2	83 (57-116)	0
Pasta/rice (g/day)	32 (14-54)	2.3	20 (11-35)	0.8
Bread (g/day)	120 (91-160)	0.1	150 (100-199)	0.1
Cereal (g/day)	166 (128-220)	0	187 (139-239)	0
Vegetables (g/day)	125 (99-160)	0	116 (88-154)	0
Legumes (g/day)	7 (2-13)	9.1	2 (1-5)	5.7
Fruits (g/day)	189 (118-273)	0.2	106 (81-182)	0.1
Nuts (g/day)	5 (1-10)	4.8	1 (0-4)	9.9
Red meat (g/day)	55 (33-80)	0.4	25 (15-39)	0.1
Poultry (g/day)	8 (4-15)	5.2	9 (5-17)	0.4
Processed meat (g/day)	21 (10-36)	0.9	54 (31-81)	0.1
Vegetable oils (g/day)	3 (1-5)	6.0	5 (3-7)	0
Margarine (g/day)	13 (6-21)	1.1	6 (1-20)	0.1
Butter (g/day)	3 (1-5)	0	5 (1-12)	1.2
Sugar (g/day)	36 (21-56)	0.2	27 (15-47)	0.5
Cake/cookies (g/day)	27 (14-45)	1.1	47 (23-78)	0.7
Beverages (g/day) †	1403 (1120-1715)	0	1314 (983-1754)	0.1
Softdrinks (g/day)	58 (11-138)	17.1	2 (0-29)	45.2
Juice (g/day)	54 (10-134)	16.3	115 (37-238)	1.9
Coffee (g/day)	500 (375-750)	4	392 (261-580)	4.3
Tea (g/day)	238 (68-430)	8.4	24 (2-150)	22.5
Alcohol (g/day)	5 (1-16)	0.3	9 (3-22)	0
Wine (g/day)	13 (0-57)	28.2	37 (7-94)	7

*Dietary intake is expressed as median (p25-p75), no intake is expressed as percentage of non-consumers with available data. † Sum of all non-alcoholic beverages, excluding water and milk.

SUPPLEMENTARY DATA

Supplementary Table 5. Dietary intake of participants from Sweden and Denmark in the EPIC-InterAct study subcohort.

	Sweden, n=2451		Denmark, n=1890	
	Dietary intake	No intake	Dietary intake	No intake
Energy (kcal/day)	2035 (1652-2491)	0	2175 (1812-2595)	0
Milk (g/day)	197 (80-403)	2.4	182 (30-509)	0
Non-milk dairy (g/day)	124 (61-228)	0.2	97 (50-210)	0.2
Milk beverages (g/day)			3 (3-7)	19.1
Milk for coffee (g/day)				
Dairy creams (g/day)	4 (1-9)	14.5	2 (1-4)	2.2
Cream desserts (g/day)	0 (0-0)	81.6	3 (0-7)	35.5
Curd (g/day)	0 (0-0)	86.2	0 (0-1)	50.2
Yoghurt (g/day)	71 (3-175)	23.8	28 (5-156)	5.2
Ice cream (g/day)	4 (1-11)	10.0	1 (1-3)	3.2
Cheese (g/day)	29 (17-48)	1.3	25 (18-53)	1.3
Potatoes (g/day)	126 (78-180)	0.5	134 (89-194)	0
Pasta/rice (g/day)	30 (9-54)	21.3	36 (21-61)	1.1
Bread (g/day)	98 (69-142)	0	140 (103-187)	0
Cereal (g/day)	182 (127-261)	0	207 (157-265)	0
Vegetables (g/day)	109 (60-178)	0	163 (104-232)	0
Legumes (g/day)	0 (0-1)	58.6	0 (0-2)	26.9
Fruits (g/day)	147 (87-241)	0.7	143 (71-242)	0
Nuts (g/day)	0 (0-1)	44.1	1 (1-2)	24.8
Red meat (g/day)	23 (11-43)	4.1	73 (52-99)	0.1
Poultry (g/day)	7 (0-16)	27.9	18 (11-27)	1.2
Processed meat (g/day)	36 (22-55)	2.1	25 (14-41)	0.2
Vegetable oils (g/day)	0 (0-2)	45.1	1 (0-5)	14.8
Margarine (g/day)	31 (19-46)	0.4	17 (6-29)	0.2
Butter (g/day)	0 (0-0)	65.0	0 (0-0)	79.3
Sugar (g/day)	36 (19-61)	0.5	54 (31-95)	0.1
Cake/cookies (g/day)	41 (22-65)	1.8	13 (6-26)	2.1
Beverages (g/day) †	1032 (700-1488)	0	1850 (1454-2324)	0
Softdrinks (g/day)	39 (0-127)	31.5	16 (0-89)	28.6
Juice (g/day)	13 (0-79)	30.4	8 (2-43)	4.9
Coffee (g/day)	400 (300-601)	3.8	900 (500-1300)	3.7
Tea (g/day)	0 (0-16)	65.7	86 (3-500)	17.2
Alcohol (g/day)	4 (1-11)	13.1	14 (6-32)	1.9
Wine (g/day)	1 (0-34)	38.3	54 (10-98)	4.0

*Dietary intake is expressed as median (p25-p75), no intake is expressed as percentage of non-consumers with available data. † Sum of all non-alcoholic beverages, excluding water and milk.

SUPPLEMENTARY DATA

Supplementary Table 6. Observed allele frequencies and Hardy Weinberg equilibrium for rs4988235 in the EPIC-InterAct study subcohort.

	n	CC	CT	TT	p value*
France	332	71	174	87	0.40
Italy	1,467	636	723	108	$5*10^{-7}$
Spain	2,544	562	1,405	577	$2*10^{-7}$
UK	1,079	85	502	492	0.006
Netherlands	1,176	110	543	523	0.08
Germany	1,783	253	944	586	$5*10^{-5}$
Sweden	2,451	155	974	1,322	0.18
Denmark	1,890	92	847	951	$2*10^{-8}$
Total	12,722	1,964	6,122	4,646	0.54

C = lactase non-persistence allele. T = lactase persistence allele. SNP dosage for rs4899235 were rounded to 0,1 or 2 to examine HWE for the total subcohort, and stratified to country. * P value is derived from chi squared test comparing observed versus expected LP genotypes based on the HWE equation.

SUPPLEMENTARY DATA

Supplementary Table 7. Baseline characteristics of the EPIC-InterAct study subcohort by rs4988235 genotype.

Baseline characteristics	*C/C**	*C/T**	*T/T**
n (persons)	1,964	6,112	4,646
age	51 ± 9	52 ± 9	53 ± 9
male sex	37.6%	38.4%	38.6%
current smoker	24.9%	26.1%	26.8%
former smoker	25.2%	27.1%	27.7%
hypertension	19.1%	18.9%	17.4%
systolic blood pressure (mmHg)	129 ± 18	133 ± 20	134 ± 20
diastolic blood pressure (mmHg)	81 ± 10	82 ± 11	82 ± 11
total cholesterol (mmol/L)	5.9 ± 1.1	5.9 ± 1.1	6.0 ± 1.1
HDL cholesterol (mmol/L)	1.5 ± 0.4	1.5 ± 0.4	1.5 ± 0.4
LDL cholesterol (mmol/L)	3.8 ± 1.0	3.8 ± 1.0	3.8 ± 1.0
triglycerides (mmol/L)	1.0 (0.7-1.5)	1.1 (0.8-1.6)	1.2 (0.8-1.7)
lipoprotein a (mg/L)	385 (201-658)	394 (200-702)	367 (202-668)
physically inactive	26.4%	23.1%	20.3%
BMI (kg/m2)	26.4 ± 4.3	26.0 ± 4.1	25.8 ± 4.2
waist-hip ratio	0.85 ± 0.09	0.85 ± 0.09	0.85 ± 0.09
HbA1c > 6.5% (>48 mmol/mol)	1.5%	1.3%	1.5%
low level of education †	46.7%	41.3%	37.0%
premenopausal status ‡	38.5%	35.1%	26.1%
history of myocardial infarction	1.2%	1.3%	1.7%
history of stroke	1.1%	0.6%	0.2%

C/C = lactase non-persistent, C/T = intermediate, T/T = lactase persistent. Most probable LP genotype based on rounded SNP dosage. * Data are expressed as mean ± standard deviation, median (quartile 1 – quartile 4) or percentage of participants with available data for variable.† no education or only primary school education. ‡ Percentage among women.

SUPPLEMENTARY DATA

Supplementary Table 8. Dietary intakes of the EPIC-InterAct study subcohort by rs4988235 genotype.

Dietary intakes	*C/C**	*C/T**	*T/T**
Energy (kcal/day)	2084 (1690–2580)	2059 (1686-2521)	2038 (1663-2482)
Milk (g/day)	146 (19-225)	160 (36-299)	188 (48-373)
Non-milk dairy (g/day)	82 (40-151)	100 (47-183)	109 (55-195)
Milk beverages (g/day)	0 (0-3)	0 (0-7)	3 (0-15)
Milk for coffee (g/day)	0 (0-14)	0 (0-13)	0 (0-14)
Dairy creams (g/day)	0 (0-2)	1 (0-3)	2 (0-4)
Cream desserts (g/day)	0 (0-11)	3 (0-16)	3 (0-16)
Curd (g/day)	0 (0-7)	0 (0-8)	0 (0-6.4)
Yoghurt (g/day)	15 (0-63)	27 (0-97)	36 (3-108)
Ice cream (g/day)	3 (0-11)	3 (0-9)	3 (1-8)
Cheese (g/day)	34 (16-61)	28 (14-51)	26 (14-45)
Potatoes (g/day)	59 (26-102)	84 (46-132)	100 (63-150)
Pasta/rice (g/day)	59 (29-111)	42 (21-77)	34 (17-60)
Bread (g/day)	129 (90-192)	123 (82-178)	115 (78-166)
Cereal (g/day)	220 (153-310)	198 (141-273)	186 (134-251)
Vegetables (g/day)	160 (105-251)	158 (102-240)	145 (93-222)
Legumes (g/day)	8 (1-31)	5 (1-23)	2 (0-11)
Fruits (g/day)	241 (136-370)	188 (102-304)	163 (90-272)
Nuts (g/day)	0 (0-3)	1 (0-3)	1 (0-3)
Red meat (g/day)	37 (19-64)	38 (19-66)	39 (18-67)
Poultry (g/day)	20 (9-35)	16 (7-31)	12 (5-25)
Processed meat (g/day)	25 (13-48)	29 (15-51)	29 (16-49)
Vegetable oils (g/day)	18 (5-29)	6 (1-20)	3 (0-7)
Margarine (g/day)	0 (0-8)	5 (0-21)	14 (2-30)
Butter (g/day)	0 (0-2)	0 (0-3)	0 (0-3)
Sugar (g/day)	29 (15-50)	32 (16-54)	35 (18-60)
Cake/cookies (g/day)	31 (12-63)	31 (12-60)	30 (12-57)
Beverages (g/day) †	311 (138-1157)	957 (256-1520)	1179 (697-1686)
Softdrinks (g/day)	0 (0-43)	3 (0-67)	16 (0-90)
Juice (g/day)	8 (0-63)	14 (0-90)	17 (1-95)
Coffee (g/day)	130 (58-300)	271 (90-541)	450 (172-675)
Tea (g/day)	0 (0-86)	5 (0-190)	12 (0-250)
Alcohol (g/day)	5.7 (1-21)	7 (1-19)	7 (1-17)
Wine (g/day)	21 (0-125)	18 (0-98)	18 (1-75)

C/C = lactase non-persistent, C/T = intermediate, T/T = lactase persistent. Most probable LP genotype based on rounded SNP loading *Data are expressed as median (quartile 1 – quartile 4).
† All non-alcoholic beverages, excluding water and milk.

SUPPLEMENTARY DATA

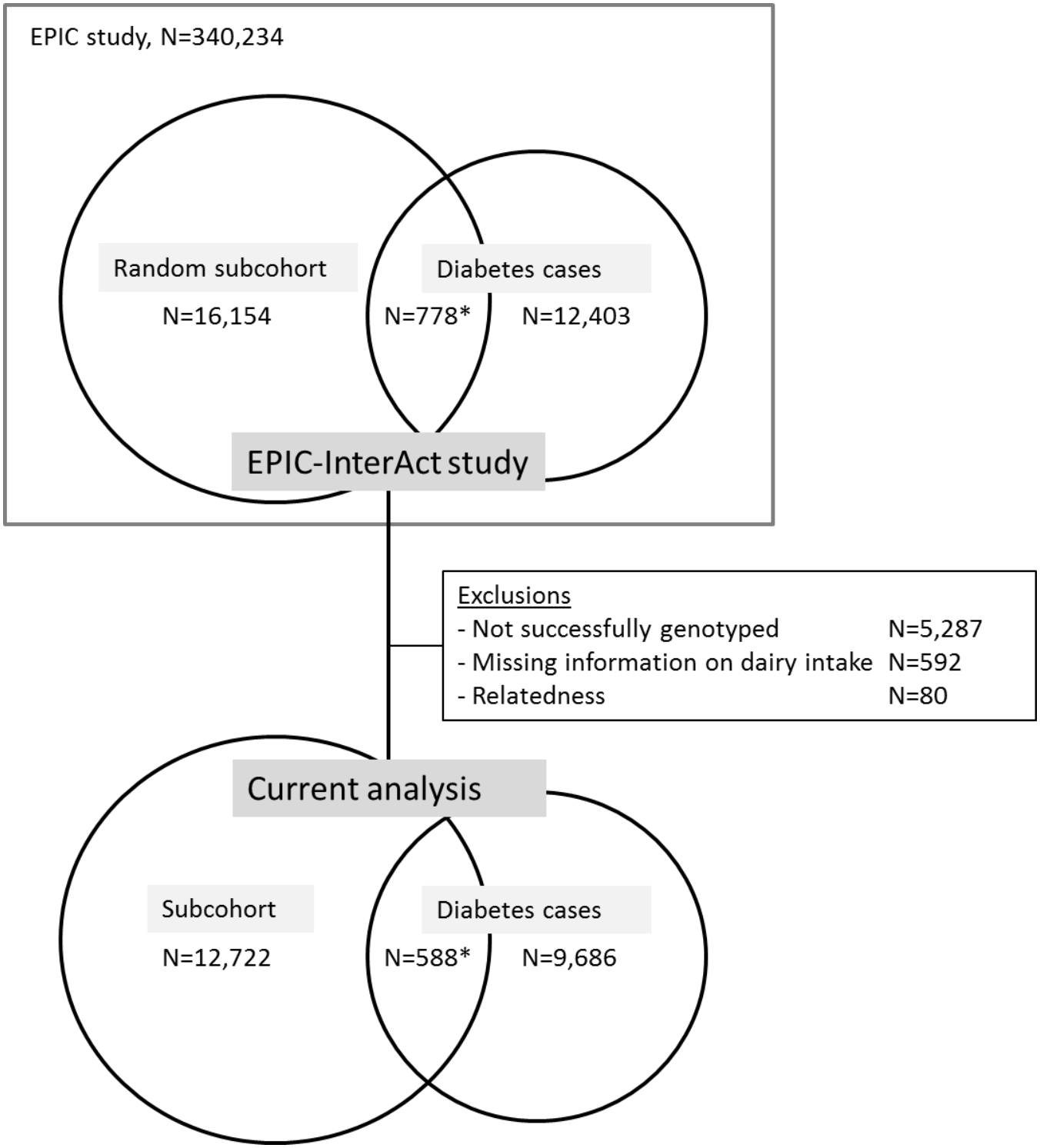
Supplementary Table 9. Association between genetic lactase persistence and diabetes (gene-outcome) and association between genetically predicted milk intake and diabetes in sensitivity analyses.

	HR*	95% confidence interval		N cases	N total
Percentile bootstrap (n samples = 10,000)	0.99	0.94	1.04	9,686	21,820
Excluding HbA1c> 6.5% (48 mmol/mol)	0.98	0.92	1.05	9,686	21,747
Only Illumina 660W quad chip	0.97	0.90	1.05	4,543	8,955
Only countries in HWE	1.04	0.81	1.33	3,789	6,679
LP dominant model	0.99	0.92	1.07	9,686	21,820

* HR for diabetes in gene-outcome is expressed per additional lactase persistence allele. HR for diabetes in IV analyses is expressed per 15 gram of genetically predicted milk intake. Analyses are performed per country, using age as underlying time scale and are adjusted for sex, genetic variability (first three principal components), study centre and genotyping platform. The reported overall estimates were obtained by pooling country-specific results in random-effects meta-analysis.

SUPPLEMENTARY DATA

Supplementary Figure S1. EPIC-InterAct case-cohort study before and after exclusions.
*N for overlap between diabetes cases and subcohort



SUPPLEMENTARY DATA

Supplementary Figure S2. Association between genetically predicted milk intake (per 15g) and diabetes per country and pooled.*

* Hazard ratio for diabetes is expressed per 15 gram of genetically predicted milk intake. Analyses are performed per country, using age as underlying time scale and are adjusted for sex, genetic variability (first three principal components), study centre and genotyping platform. The reported overall estimates were obtained by pooling country-specific results in random-effects meta-analysis using restricted maximum-likelihood estimation, adjusting for study size with inverse variance weights.