Supplementary Table 1. Descriptive characteristics of genome-wide association study cohorts (Stage 1) and follow-up study cohorts (Stage 2)

Study	n	Women	Age (y)*	BMI (kg/m^2)*	Height (cm)*	Weight (kg)*	Normal weight (BMI < 25 kg/m^2)	Overweight - non-obese (25 kg/m^2 ≤ BMI < 30 kg/m^2)	Obese (BMI ≥ 30 kg/m^2)
Stage 1 - Genome-wide association study cohorts									
Population-based cohorts									
CoLaus	5,433	2873 (53%)	53.4 (10.8)	25.9 (4.6)	168.4 (9.3)	73.6 (15.2)	2567 (47%)	1988 (37%)	868 (16%)
SardiNIA	4,301	2416 (56%)	43.6 (17.6)	25.4 (4.7)	160.0 (8.9)	65.0 (13.3)	2185 (51%)	1424 (33%)	692 (16%)
EPIC-Obesity Study	2,415	1284 (53%)	59.2 (9.0)	26.4 (3.9)	167.0 (8.8)	73.8 (13.0)	920 (38%)	1115 (46%)	380 (16%)
NHS (Breast Cancer)	2,265	2265 (100%)	54 (6.0)	25.1 (4.4)	164.1 (6.0)	67.7 (12.6)	1334 (59%)	638 (28%)	293 (13%)
PLCO (Prostate Cancer)	2,235	0 (0%)	64.2 (5.1)	27.5 (3.8)	178.4 (6.6)	87.3 (13.3)	578 (26%)	1155 (52%)	502 (22%)
KORA	1,642	829 (51%)	52.5 (10.0)	27.3 (4.1)	167.3 (8.9)	76.7 (13.5)	467 (28%)	819 (50%)	356 (22%)
WTCCC/UK Blood Services 1	1,437	743 (52%)	43.4 (12.4)	26.2 (4.3)	171.2 (9.5)	77.1 (15.3)	623 (43%)	581 (40%)	233 (16%)
British 1958 Birth Cohort	1,479	738 (50%)	44.5 (0.25)	27.4 (4.8)	169.3 (9.6)	78.7 (16.5)	494 (33%)	617 (42%)	368 (25%)
Control series									
DGI Controls	1,523	791 (52%)	58.6 (10.2)	26.7 (3.8)	168.8 (8.9)	76.2 (12.9)	533 (35%)	736 (48%)	254 (17%)
FUSION Controls	1,291	648 (50%)	60.9 (11.2)	27.0 (3.9)	167.4 (9.3)	75.9 (13.4)	430 (33%)	594 (46%)	267 (21%)
Case series									
WTCCC/HT Cases (Hypertension)	1,895	1138 (60%)	56.7 (11.3)	27.5 (3.8)	166.5 (9.3)	76.5 (13.4)	486 (26%)	927 (49%)	482 (25%)
WTCCC/CAD Cases (Coronory Artery Disease)	1,876	387 (21%)	60.0 (8.1)	27.6 (4.2)	170.9 (8.9)	80.7 (14.2)	492 (26%)	936 (50%)	448 (24%)
WTCCC/T2D Cases (Type 2 diabetes)	1,913	800 (42%)	58.5 (10.1)	31.2 (6.1)	169.6 (9.7)	89.9 (18.9)	264 (14%)	647 (34%)	1002 (52%)
DGI Cases (Type 2 Diabetes)	1,588	793 (50%)	64.4 (10.3)	28.5 (4.4)	167.6 (9.2)	80.2 (14.4)	329 (21%)	750 (47%)	509 (32%)
FUSION Cases (Type 2 Diabetes)	1,094	471 (43%)	62.7 (7.6)	30.2 (4.7)	167.3 (9.0)	84.6 (14.7)	125 (11%)	440 (40%)	529 (48%)
Stage 2 - Follow-up study cohorts									
Populations with direct genotyping data									
EPIC-Norfolk	18,719	9513 (51%)	58.8 (9.3)	26.3 (3.8)	167.4 (9.2)	73.9 (13.1)	7346 (39%)	8676 (46%)	2697 (14%)
FINRISK97	7,670	3831 (50%)	48.7 (13.3)	26.7 (4.5)	168.8 (9.3)	76.2 (14.9)	2981 (39%)	3177 (41%)	1512 (20%)
METSIM	6,225	0 (0%)	58.7 (6.4)	27.3 (4.2)	175.5 (6.3)	84.1 (13.9)	1899 (31%)	3016 (48%)	1310 (21%)
Botnia PPP Study	3,428	1804 (53%)	49.2 (16.0)	26.4 (4.4)	170.1 (9.2)	76.5 (14.8)	1412 (41%)	1423 (42%)	593 (17%)
FUSION stage 2 - Controls	1,266	493 (39%)	58.4 (7.7)	26.9 (3.9)	169.1 (9.3)	76.9 (13.4)	428 (34%)	601 (47%)	237 (19%)
FUSION stage 2 - Cases (Type 2 Diabetes)	1,204	488 (41%)	59.4 (8.7)	30.9 (5.4)	168.9 (9.6)	88.1 (17.6)	136 (11%)	450 (38%)	618 (51%)
Hertfordshire Study**	2,944	1397 (47%)	66.2 (2.9)	27.4 (4.4)	167.9 (9.1)	77.2 (14.2)	903 (31%)	1343 (45%)	698 (24%)
SardiNIA - Stage 2**	1,862	1116 (60%)	44.6 (17.1)	25.4 (4.6)	159.8 (9.1)	64.9 (13.2)	951 (51%)	635 (34%)	276 (15%)
MRC - Ely Study**	1,700	917 (54%)	61.1 (9.1)	27.2 (4.8)	167.4 (9.0)	76.4 (15.1)	588 (35%)	736 (43%)	376 (22%)
Populations with *in silico* data (existing GWA scans)									
Rotterdam Study	5,373	3127 (58%)	68.7 (8.6)	26.3 (3.7)	167.1 (9.4)	73.5 (11.9)	2068 (39%)	2528 (47%)	777 (15%)
Northern Finnish Birth Cohort of 1966	4,478	2243 (50%)	31 (0.0)#	24.7 (4.2)	171.5 (9.3)	72.8 (14.9)	2669 (60%)	1396 (31%)	413 (9%)
TwinsUK	2,218	2,218 (100%)	46.7 (12.2)	25.1 (4.7)	162.4 (6.1)	66.2 (12.8)	1299 (59%)	632 (28%)	287 (13%)
InCHIANTI	1,139	627 (55%)	67.5 (15.2)	27.2 (4.2)	160.1 (9.9)	69.7 (12.9)	373 (33%)	497 (44%)	269 (24%)
Baltimore Longitudinal Study of Aging	856	393 (46%)	67.9 (16.8)	26.6 (4.6)	169.4 (9.6)	76.6 (16.9)	364 (42.5%)	321 (37.5%)	171 (20%)

* Values represent: mean (SD)
** Genotype for *SH2B1* SNPs only
# Birth cohort measures at age 31

Supplementary Table 2. Number of individuals and sample quality control for Stage 1 and 2 genome-wide association study cohorts

Study	Total sample	Sample QC		Samples in analyses	Measured/Self reported height and weight
		Call rate*	other exclusions		
Stage 1 - Genome-wide association study cohorts					
Population-based cohorts					
CoLaus	5,633	≥ 90%	1) gender discrepancy with genetic data from X-linked markers; 2) inconsistent genotypes when compared with control markers; 3) duplicates and first and second degree relatives; 4) Missing body weight and height.	5,433	measured
SardiNIA	6,148	> 95%	1) Missing body weight and height; 2) Not belonging to families where at least one member was genotyped with the 500K Affymetrix chip.	4,301	measured
EPIC-Obesity Study	2,566	≥ 94%	1) heterozygosity <23% or >30%; 2) >5.0% discordance in SNP pairs with r2= 1 in HapMap; 3) ethnic outliers; 4) related individuals and duplicates; 5) Missing body weight and height.	2,415	measured
NHS (Breast Cancer)	2,368	≥ 90%	1) Low genotyping completion (<90%); 2) Unclear identity and admixed origin; 3) Missing body weight and height.	2,265	self reported
PLCO (Prostate Cancer)	2,277	≥ 90%	1) BMI > 4SD from mean; 2) gender discrepancy with X-linked markers; 3) related individuals and duplicates.	2,235	self reported
KORA	1,796	≥ 93%	1) discrepancy for one of the 50 SNPs common to both chips; 2) Missing body weight and height; 3) gender discrepancy with X-linked markers.	1,642	measured
WTCCC/UK Blood Services 1	1,500	≥ 97%	1) heterozygosity <23% or >30%; 2) discrepancy with external identifying information; 3) ethnic outliers; 4) related individuals and duplicates; 5) Missing body weight and height.	1,437	self reported
British 1958 Birth Cohort	1,502	≥ 97%	1) contamination; 2) non-European identity; 3) Missing body weight and height.	1,479	measured
Control series					
DGI Controls	1,595	> 95%	1) Missing body weight and height.	1,523	measured
FUSION Controls	1,295	> 97.5%	1) missing body weight and height; 2) impaired glucose tolerance samples.	1,291	measured
Case series					
WTCCC/HT Cases (Hypertension)	2,000	≥ 97%	1) heterozygosity <23% or >30%; 2) discrepancy with external identifying information; 3) ethnic outliers; 4) related individuals and duplicates 5) Missing body weight and height	1,895	measured
WTCCC/CAD Cases (Coronory Artery Disease)	2,000	≥ 97%	1) heterozygosity <23% or >30%; 2) discrepancy with external identifying information; 3) ethnic outliers; 4) related individuals and duplicates; 5) Missing body weight and height.	1,876	self reported
WTCCC/T2D Cases (Type 2 Diabetes)	1,999	≥ 97%	1) heterozygosity <23% or >30%; 2) discrepancy with external identifying information; 3) ethnic outliers; 4) related individuals and duplicates; 5) Missing body weight and height.	1,913	measured
DGI Cases (Type 2 Diabetes)	1,658	> 95%	1) Missing body weight and height.	1,588	measured
FUSION Cases (Type 2 Diabetes)	1,162	> 97.5%	1) missing body weight and height.	1,094	measured
Stage 2 - *In silico* follow-up study cohorts					
Rotterdam Study	6,449	≥ 97.5%	1) gender discrepancy with genetic data from X-linked markers; 2) excess autosomal heterozygosity > 0.336 ~FDR<0.1%; 3) duplicates and/or 1st or 2nd degree relatives with IBS; probabilities > 97%; 4) ethnic outliers; 5) missing height or weight data.	5,373	measured
Northern Finnish Birth Cohort of 1966	5,408	≥ 95%	1) gender discrepancy with genetic data from X-linked markers; 2) withdrawn consent; 3) duplicates and first and second degree relatives; 4) contaminated samples; 5) missing height and weight data.	4,478	measured
TwinsUK	2,500	≥ 95%	1) heterozygosity <33% or >37%; 2) ethnic outliers (using STRUCTURE); 3) Missing body weight and height.	2,218	measured
InCHIANTI	1,231	≥ 97%	1) gender discrepancy with genetic data from X-linked markers 2) heterozygosity <30%; 3) Missing body weight or height.	1,139	measured
Baltimore Longitudinal Study of Aging	1,230	≥ 98.5%	1) gender discrepancy with genetic data from X-linked markers; 2) ethnic outliers & non-Caucasians (using STRUCTURE); 3) duplicates; 4) missing height or weight data.	856	measured

* Sample genotyping success rate; i.e. percentage of successfully genotyped SNPs per sample

Supplementary Table 3. Information on genotyping methods, quality control of SNPs, imputation, and statistical analysis for Stage 1 and 2 genome-wide association study cohorts

Study	Genotyping						Imputation			Association analyses			
	Platform	Genotype calling algorithm	Inclusion criteria MAF	Call rate	p for HWE	SNPs that met QC criteria	Imputation software	Inclusion criteria MAF	Imputation quality*	SNPs in meta-analysis	λ_{GC}	Analyses software	Analyses model (all additive)
Stage 1 - Genome-wide association study cohorts													
Population-based cohorts													
CoLaus	Affymetrix 500K Array Set	BRLMM	> 0%	≥ 70%	> 10^{-7}	390,631	IMPUTE	≥1%	proper-info ≥ 0.40	2,399,824	1.013	SNPtest	$\log_{10}$-transformed BMI, standardized by gender , age-decades, ancestry (by PCA)
SardiNIA	Affymetrix 500K Array Set & Affymetrix 10K Array Set	BRLMM	> 5%	> 90%	≥ 10^{-6}	356,359	MACH	≥1%	r^2-hat ≥ 0.30	2,248,057	1.000	Merlin	residuals of BMI,regressed against age, age^2, and gender
EPIC-Obesity Study	Affymetrix 500K Array Set	BRLMM	≥ 1%	≥ 90%	> 10^{-6}	397,438	IMPUTE	≥1%	proper-info ≥ 0.40	2,381,011	1.013	SNPtest	$\log_{10}$-transformed BMI, standardized by gender and age-decades
NHS (Breast Cancer)	Illumina HumanHap550	Standard Illumina BeadStudio (GenCall)	≥ 1%	≥ 90%	-	510,073	MACH	≥1%	r^2-hat ≥ 0.30	2,457,986	1.000	MACH2QTL	residuals of BMI regressed against age and ancestery (by PCA)
PLCO (Prostate Cancer)	Illumina HumanHap300 & Illumina HumanHap240	Standard Illumina BeadStudio (GenCall)	≥ 1%	≥ 94%	-	523,231	MACH	≥1%	r^2-hat ≥ 0.30	2,487,997	1.000	MACH2QTL	$\log_{10}$-transformed BMI residuals regressed against age, case-status, and center
KORA	Affymetrix 500K Array Set	BRLMM	>5%	-	-	380,407	MACH	≥1%	r^2-hat ≥ 0.30	2,385,925	1.000	MACH2QTL	residuals of log10-transformed BMI,regressed against age and gender
WTCCC/UK Blood Services 1	Affymetrix 500K Array Set	CHIAMO	> 5%	≥ 95%	> 10^{-6}	387,667	IMPUTE	≥1%	proper-info ≥ 0.40	2,383,762	1.008	SNPtest	$\log_{10}$-transformed BMI, standardized by gender, adjusted for age, and age^2
British 1958 Birth Cohort	Affymetrix 500K Array Set	CHIAMO	> 0%	-	-	490,032	IMPUTE	≥1%	proper-info ≥ 0.40	2,408,712	1.004	SNPtest	$\log_{10}$-transformed BMI, standardized by gender
Control series													
DGI Controls	Affymetrix 500K Array Set	BRLMM	> 5%	≥ 95%	> 10^{-6}	347,010	MACH	≥1%	r^2-hat ≥ 0.30	2,375,087	1.000	Merlin	residuals of population-corrected z-scores standardized by age-decades and gender, regressed against center
FUSION Controls	Illumina HumanHap300	Illumina Beadstudio clustering with FUSION data	>1%	≥ 90%	≥ 10^{-6}	304,581	MACH	≥1%	r^2-hat ≥ 0.30	2,420,814	1.009	Merlin	inverse normally transformed residuals of BMI regressed against age, age^2, gender, study group and birth province
Case series													
WTCCC/HT Cases (Hypertension)	Affymetrix 500K Array Set	CHIAMO	> 5%	≥ 95%	> 10^{-6}	387,667	IMPUTE	≥1%	proper-info ≥ 0.40	2,383,933	1.024	SNPtest	$\log_{10}$-transformed BMI, standardized by gender, adjusted for age
WTCCC/CAD Cases (Coronory Artery Disease)	Affymetrix 500K Array Set	CHIAMO	> 5%	≥ 95%	> 10^{-6}	387,667	IMPUTE	≥1%	proper-info ≥ 0.40	2,383,876	1.007	SNPtest	$\log_{10}$-transformed BMI, standardized by gender, adjusted for age
WTCCC/T2D Cases (Type 2 Diabetes)	Affymetrix 500K Array Set	CHIAMO	> 5%	≥ 95%	> 10^{-6}	387,667	IMPUTE	≥1%	proper-info ≥ 0.40	2,383,902	1.015	SNPtest	$\log_{10}$-transformed BMI, standardized by gender, adjusted for age
DGI Cases (Type 2 Diabetes)	Affymetrix 500K Array Set	BRLMM	> 5%	≥ 95%	> 10^{-6}	347,010	MACH	≥1%	r^2-hat ≥ 0.30	2,375,087	1.047	Merlin	residuals of population-corrected z-scores standardized by age-decades and gender, regressed against center
FUSION Cases (Type 2 Diabetes)	Illumina HumanHap300	Illumina Beadstudio clustering with FUSION data	>1%	≥ 90%	≥ 10^{-6}	304,581	MACH	≥1%	r^2-hat ≥ 0.30	2,420,814	1.023	Merlin	inverse normally transformed residuals of BMI regressed against age, age^2, gender, study group and birth province
Stage 2 - *In silico* follow-up study cohorts													
Rotterdam Study	Illumina HumanHap 550 V.3	Standard Illumina BeadStudio	≥ 5%	≥ 98%	> 10^{-4}	512,349	MACH	≥1%	r^2-hat ≥ 0.30	-	1.048	MACH2QTL	log-transformed BMI standardized residuals adjusted for age, age^2, stratified by gender
Northern Finnish Birth Cohort of 1966	Illumina HumanCNV-370 DUO Analysis BeadChip	Standard Illumina BeadStudio	≥ 5%	≥ 95%	> 10^{-4}	328,007	IMPUTE	≥1%	proper-info ≥ 0.40	-	1.030	SNPtest	$\log_{10}$-transformed BMI, standardized by gender
TwinsUK	Illumina HumanHap300 Duo Infinium	Standard Illumina BeadStudio	≥ 5%	≥ 95%	> 10^{-6}	307,040	IMPUTE	≥1%	proper-info ≥ 0.40	-	1.046	Merlin	untransformed BMI, adjusted for age
InCHIANTI	Illumina HumanHap 550 V.1 & V.3	Standard Illumina BeadStudio	≥1%	≥ 99%	> 10^{-4}	495,343	MACH	≥1%	r^2-hat ≥ 0.30	-	1.000	MerlinOffline	Inverse normal BMI adjusted for age, age^2 and gender
Baltimore Longitudinal Study of Aging	Illmina HumanHap 550 V.3	Standard Illumina BeadStudio	>1%	≥ 99%	> 10^{-4}	514,027	MACH	≥1%	r^2-hat ≥ 0.30	-	1.030	MerlinOffline	Inverse normal BMI adjusted for age, age^2 and gender

* SNPtest calculates the 'proper_info' statistic as a measure of the relative statistical information about the additive genetic effect being estimated. The proper_info statistic has a direct relationship to the power of the test and ranges from 0 to 1, with 1 indicating perfect information. MACH calculates the 'rsq_hat', which is the r^2 between each imputed genotype and its true underlying genotype. Rsq_hat ranges from 0 to 1, with 1 indicating perfect imputation.

Supplementary Table 4. Stage 2 - SNPs taken forward in follow-up studies

							Direct genotyping								*In silico*					
Chr	**Position**	**Gene/Locus**	**Lead SNP for follow-up**	**Proxy SNP**	**Distance (bp) between proxy and lead SNP**	**r^2 between lead and proxy**	**EPIC-Norfolk**	**FINRISK97**	**METSIM**	**PPP-Botnia**	**FUSION - Stage 2**	**Hertfordshire**	**SardiNIA - Stage 2**	**MRC-Ely**	**BLSA**	**InCHIANTI**	**Rotterdam Study**	**TwinsUK**	**NFBC 1966**	**deCODE**
1	72585028	near *NEGR1*	rs2815752	rs3101336/ rs2568958	-121822/ -47324	1 1	proxy 1 + proxy2	lead	lead	lead	lead				imputed	imputed	imputed	imputed	genotyped	genotyped (proxy)
1	230427413	near *SIPA1L2*	rs10910555	rs10797582	4004	0.7	lead	lead	lead	lead	lead				imputed	imputed	imputed	imputed	imputed	genotyped (proxy)
2	624905	near *TMEM18*	rs6548238	rs2867125	-12078	0.85	lead	lead + proxy	lead + proxy	lead + proxy	lead + proxy				imputed	imputed	imputed	imputed	imputed	genotyped (proxy)
2	684759	near *TMEM18*	rs9678953	rs6548251	16862	0.95	proxy	lead		lead					imputed	imputed	imputed	imputed	imputed (proxy)	
2	35257515	near *CRIM1*	rs968059	rs1439845	7414	0.97	lead	lead	lead	lead	lead				genotyped	genotyped	imputed	genotyped	genotyped	genotyped
2	46019476	*PRKCE*	rs10206343	rs2029087	4390	0.63			lead		lead				genotyped	genotyped	imputed	imputed	imputed	genotyped (proxy)
2	142759505	near *LRP1B*	rs17835238												imputed	imputed	imputed	imputed	imputed	
4	44184606	near *KCTD8*	rs752238				lead	lead	lead	lead	lead				genotyped	genotyped	genotyped	genotyped	genotyped	genotyped
4	44877284	near *GNPDA2*	rs10938397				lead	lead	lead	lead	lead				imputed	imputed	imputed	imputed	imputed	
4	183366671	near *ODZ3*	rs2597108	rs2726814	11427	0.74	proxy	lead	lead	lead	lead				imputed	imputed	imputed	imputed	imputed	
5	25486507	near *CDH10*	rs17465346					lead		lead					imputed	imputed	imputed	imputed	imputed	
6	40182179	near *MOCS1*	rs12210863	rs10807218	-124615	0.86		lead		lead					imputed	imputed	imputed	imputed	imputed	genotyped (proxy)
6	46429048	near *DSCR1L1*	rs1867015	rs6911147	-13461	0.92			proxy		proxy				genotyped	genotyped	genotyped	genotyped	genotyped	genotyped
6	55038162	near *FAM83B*	rs7746448				lead	lead	lead	lead	lead				genotyped	genotyped	genotyped	genotyped	genotyped	genotyped
6	71308142	near *C6orf57*	rs9294876					lead	lead	lead	lead				genotyped	genotyped	genotyped	genotyped	genotyped	genotyped
7	25348456	near *NPVF*	rs16873846					lead		lead					imputed	imputed	imputed	imputed	imputed	
7	130084134	near *KLF14*	rs11976955	rs7810507	3882	0.84	lead	lead	lead	lead	lead				imputed	imputed	imputed	imputed	imputed	genotyped (proxy)
8	84787572	near *LOC138046*	rs11773921	rs9650286/ rs4132243	+55129/ +107185	0.77 0.77		lead	proxy	lead	proxy				imputed	imputed	imputed	imputed	imputed	genotyped (proxy)
8	143821485	near *SLURP1*	rs3758082				lead								genotyped	genotyped	genotyped	genotyped	genotyped	
9	97437280	near *PTCH1*	rs4743120				lead		lead		lead				imputed	imputed	imputed	imputed	imputed	
10	7686190	near *ITIH5*	rs11255232	rs4623795/ rs1537626	+5725/ +6940	0.95 0.79	lead + proxy 1	lead	lead	lead	lead				genotyped	genotyped	genotyped	imputed	imputed	genotyped (proxy 2)
11	8440665	near *STK33*	rs10769908	rs10840065/ rs11041990	42554/ 101526	1 0.97	lead	lead	proxy 2	lead	proxy 2				imputed	imputed	imputed	imputed	imputed	genotyped (proxy 1)
11	47619625	near *MTCH2*	rs10838738	rs4752856	-15007	1	lead	lead	lead	lead					genotyped	genotyped	genotyped (proxy)	genotyped	genotyped	genotyped
12	58213837	near *SLC16A7*	rs275982	rs2175950	-7188	0.7									imputed	imputed	imputed	imputed	genotyped (proxy)	genotyped (proxy)
12	121199956	near *MLXIP*	rs925460				lead	lead	lead		lead				genotyped	imputed	genotyped	genotyped	genotyped	genotyped
15	41621610	*MAP1A*	rs2255042	rs2245715	-16266	1		lead	proxy	lead	proxy				genotyped	genotyped	genotyped	imputed	genotyped (proxy)	genotyped (proxy)
15	80139255	near *RKHD3*	rs12324805					lead		lead					genotyped	genotyped	genotyped	genotyped	genotyped	genotyped
15	91375823	near *RGMA*	rs7181095	rs8031888	1912	0.79	lead								imputed	imputed	imputed	imputed	imputed	genotyped (proxy)
15	96988019	near *IGF1R*	rs8024593					lead	lead	lead	lead				imputed	imputed	imputed	imputed	imputed	
16	28790742	*SH2B1*	rs7498665	rs8055982/ rs8061590	-2039/ +11889	1 1	lead + proxy 1	lead + proxy1 + proxy 2	lead + proxy 1		lead + proxy 1	lead + proxy 1	lead + proxy 1	lead + proxy 1	imputed	genotyped	genotyped	genotyped	genotyped	genotyped
16	28806294	*ATP2A1*	rs6565259*	rs9931989	7291	0.95	proxy		proxy		proxy				genotyped	genotyped	genotyped	genotyped	genotyped	genotyped
16	52378028	*FTO*	rs9939609	rs1121980/ rs8050136/ rs12149832	-11280/ -4252/ +22381	0.84 1 0.93	proxy 1	lead	lead + proxy 2	lead	lead + proxy 2				imputed	imputed	genotyped	imputed	imputed	genotyped (proxy 2)
17	8845741	near *NTN1*	rs7219148					lead		lead					genotyped	genotyped	genotyped	genotyped	genotyped	genotyped
18	56002077	near *MC4R*	rs17782313	rs4940927	-118408	0.73	lead	lead	lead		lead				imputed	imputed	imputed	imputed	imputed	genotyped (proxy)
19	8922559	near *MUC16*	rs1423052				lead	lead	lead	lead	lead				imputed	imputed	imputed	imputed	imputed	
19	39013977	near *KCTD15*	rs11084753	rs368794/ rs29941	-1685/ -12605	1 0.65	lead + proxy 1	lead							imputed	imputed	imputed	imputed	imputed	genotyped (proxy 2)
20	6569685	near *BMP2*	rs2145270	rs4813800	-26039	0.77	lead	lead	lead	lead	lead				imputed	imputed	imputed	imputed	imputed	genotyped (proxy)
20	13235505	near *TASP1*	rs1076052				lead		lead		lead				genotyped	genotyped	genotyped	genotyped	genotyped	genotyped
20	50485779	near *ZFP64*	rs6021931				lead								genotyped	genotyped	genotyped	genotyped	genotyped	
22	47209622	near *FAM19A5*	rs4823535				lead	lead		lead					genotyped	genotyped	genotyped	genotyped	genotyped	genotyped

* r^2 between rs6565259 & rs7498665: 0.69; position in Build 35

Supplementary Table 5. All loci followed up

	Best Ranking SNP Results										Lead SNP Results																	
Nearby Gene				Alleles		Imputation quality		Stage 1						Imputation quality		Alleles			Heterogeneity between Stage 1 samples		Stage 1		Stage 2		Stage 1 + Stage 2		With deCODE		deCODE
	SNP	Chr	Position (bp)	Effect	Other	r.sq-hat	Proper-info	n	p-value	SNP	Chr	Position (kb)	r^2 to best	r.sq-hat	Proper-info	Effect	Other	Frequency effect allele (%)	between-study (p-value)	between-gender (p-value)	n	p-value	n	p-value	n	p-value	n	p-value	SNP
FTO	rs1421085	16	52358455	C	T	0.947	1	32530	2.6E-19	rs9939609*	16	52378028	0.931	0.994	GT	A	T	41%	0.02	0.57	32468	6.3E-17	51055	3.29E-36	83523	4.31E-51	113343	4.88E-74	rs8050136
TMEM18	rs6548238	2	624905	C	T	0.931	GT	32394	1.2E-06	rs6548238	2	624905	best	0.931	GT	C	T	84%	0.71	0.67	32394	1.2E-06	52567	1.58E-13	84961	1.35E-18	114781	3.17E-26	rs2867125
MC4R	rs17700144	18	55962962	A	G	0.918	0.956	32528	3.5E-07	rs17782313	18	56002077	0.843	0.962	GT	C	T	21%	0.20	0.49	32519	3.9E-07	48362	1.88E-12	80881	4.56E-18	110701	1.13E-20	rs4940927
GNPDA2	rs10938397	4	45023455	G	A	0.822	0.839	32522	1.0E-05	rs10938397	4	45023455	best	0.822	0.839	G	A	45%	0.76	0.63	32522	1.0E-05	49371	4.45E-12	81893	3.39E-16	ND	ND	ND
SH2B1	rs9931989	16	28813585	G	C	0.771	0.708	32517	1.1E-06	rs7498665*	16	28790742	0.706	0.900	0.992	G	A	41%	0.66	0.65	32502	5.4E-06	54316	1.69E-06	86818	5.06E-11	116638	2.23E-14	rs7498665
MTCH2	rs4752856	11	47604618	A	G	1.004	GT	32450	2.7E-06	rs10838738	11	47619625	1	0.999	0.997	G	A	34%	0.07	0.17	32528	7.1E-06	48530	9.55E-05	81058	4.55E-09	110878	1.90E-11	rs10838738
KCTD15	rs11084753	19	39015054	G	A	0.755	0.722	32466	2.6E-07	rs11084753	19	39013977	best	0.755	0.722	G	A	67%	0.39	0.83	32466	2.6E-07	39371	4.06E-03	71837	2.26E-08	101657	4.50E-12	rs29941
NEGR1	rs2815752	1	72524461	A	G	0.93	0.911	32521	9.3E-06	rs2815752*	1	72524461	best	0.930	0.911	A	G	62%	0.20	0.52	32521	9.3E-06	51112	6.82E-04	83633	5.96E-08	113453	1.02E-12	rs2568958
STK33	rs10769908	11	8440665	C	T	0.983	GT	32428	1.2E-05	rs10769908	11	8440665	best	0.983	GT	C	T	53%	0.99	0.97	32428	1.2E-05	51622	6.96E-03	84050	1.34E-06	113870	2.94E-07	rs10840065
BMP2	rs2145270	20	6569685	T	C	0.936	0.957	32508	3.5E-05	rs2145270	20	6569685	best	0.936	0.957	T	C	65%	0.98	0.68	32508	3.5E-05	50579	1.31E-02	83087	6.02E-06	112907	4.17E-06	rs4813800
RKHD3	rs12324805	15	80139255	C	A	GT	GT	32524	4.4E-06	rs12324805	15	80139255	best	GT	GT	C	A	31%	0.02	0.94	32524	4.4E-06	25820	1.09E-01	58344	6.95E-06	88164	1.74E-03	rs12324805
RGMA	rs7181095	15	91375823	T	C	0.695	0.56	32504	2.5E-07	rs7181095	15	91375823	best	0.695	0.559	T	C	36%	0.36	0.53	32504	2.5E-07	31286	5.43E-01	63789	3.96E-05	93609	8.99E-03	rs8031888
DSCR1L1	rs6911147	6	46415587	T	C	0.78	0.872	32518	4.0E-05	rs1867015*	6	46429048	0.923	1.012	0.999	G	T	61%	0.48	0.08	32523	2.9E-04	22179	6.58E-02	54702	7.26E-05	84522	1.17E-03	rs1867015
SLC16A7	rs275982	12	58213837	A	C	0.968	GT	32504	2.8E-05	rs275982	12	58213837	best	0.968	GT	A	C	23%#	0.22	0.90	32504	2.8E-05	14747	4.76E-01	47251	1.09E-04	77071	1.13E-03	rs2175950
MLXIP	rs925460	12	121158883	C	T	0.658	0.793	32498	2.0E-05	rs925460	12	121158883	best	0.658	0.793	C	T	78%	0.93	0.17	32498	2.0E-05	46481	1.46E-01	78979	1.17E-04	108799	7.63E-03	rs925460
KLF14	rs11976955	7	129890849	C	G	0.877	0.904	32524	3.3E-05	rs11976955	7	129890849	best	0.877	0.904	C	G	69%	0.36	0.45	32524	3.3E-05	50723	1.28E-01	83247	1.53E-04	113067	5.06E-04	rs7810507
CRIM1	rs1439845	2	35323076	G	A	1.015	GT	32511	5.0E-06	rs968059	2	35315662	0.966	0.887	0.982	C	A	63%	0.43	0.84	32506	7.3E-06	50555	2.36E-01	83061	1.92E-04	112881	1.40E-05	rs968059
ODZ3	rs2726814	4	183516253	T	G	0.81	0.96	32529	5.3E-05	rs2597108*	4	183504826	0.737	0.758	0.900	A	T	11%	0.08	0.70	32527	8.3E-05	50167	1.18E-01	82694	2.28E-04	ND	ND	ND
KCTD8	rs752238	4	44330777	G	T	GT	GT	32462	6.7E-06	rs752238	4	44330777	1	GT	GT	G	T	77%	0.85	0.74	32462	6.7E-06	50292	3.06E-01	82754	2.96E-04	112574	4.24E-04	rs752238
LOC138046	rs11773921	8	84787572	A	T	0.894	0.928	32529	6.5E-06	rs11773921*	8	84787572	best	0.894	0.928	A	T	14%	0.25	1.00	32529	6.5E-06	33011	6.07E-01	65540	3.98E-04	95360	1.40E-03	rs4132243
PTCH1	rs4743120	9	95477014	T	C	0.979	0.995	32529	2.1E-05	rs4743120	9	95477014	best	0.979	0.995	T	C	35%	0.46	0.90	32529	2.1E-05	38356	4.20E-01	70885	5.09E-04	ND	ND	ND
MAP1A	rs2245715	15	41605344	G	A	GT	GT	25128	1.5E-06	rs2255042*	15	41621610	1	0.960	0.999	C	T	92%	0.43	0.49	32530	7.1E-05	33296	4.28E-01	65826	7.86E-04	95646	1.00E-03	rs2245715
C6orf143	rs6907460	6	54989006	A	G	1.006	GT	25188	8.8E-06	rs7746448	6	55038162	1	1.011	0.997	C	T	39%	0.06	0.55	32530	2.8E-04	50636	1.86E-01	83166	9.52E-04	112986	1.82E-02	rs7746448
NTN1	rs11078773	17	8845288	C	T	0.832	GT	32398	1.0E-05	rs7219148	17	8845741	0.966	0.864	0.963	G	T	51%	0.97	0.69	32512	2.4E-05	25820	8.26E-01	58332	9.71E-04	88152	1.52E-02	rs7219148
IGF1R	rs8024593	15	96988019	G	A	0.784	GT	32508	2.6E-05	rs8024593	15	96988019	best	0.784	GT	G	A	6%	0.13	0.50	32508	2.6E-05	31032	7.26E-01	63540	1.14E-03	ND	ND	ND
TASP1	rs636887	20	13235035	C	T	0.968	0.996	32530	1.9E-05	rs1076052	20	13235505	1	GT	GT	T	C	45%	0.56	0.84	32478	3.4E-05	39891	5.62E-01	72369	1.34E-03	102189	6.87E-04	rs1076052
MUC16	rs1423052	19	8922559	G	C	0.996	0.974	32530	1.5E-05	rs1423052	19	8922559	best	0.996	0.974	G	C	26%	0.44	0.55	32530	1.5E-05	50679	7.10E-01	83209	2.70E-03	ND	ND	ND
C6orf57	rs9455181	6	71359438	T	C	0.794	0.794	32528	1.5E-05	rs9294876*	6	71308142	0.643	0.988	0.993	C	T	23%	0.92	0.72	32530	1.4E-04	33203	7.20E-01	65733	3.31E-03	95553	2.58E-03	rs9294876
CDH10	rs17465346	5	25486507	T	A	0.932	0.855	32529	3.0E-04	rs17465346	5	25486507	best	0.932	0.855	T	A	95%	0.22	0.65	32529	3.0E-04	25825	9.87E-01	58354	6.77E-03	ND	ND	ND
C7orf9	rs16873846	7	25155171	T	C	0.64	0.672	32528	1.9E-05	rs16873846	7	25155171	best	0.640	0.672	T	C	95%	0.19	0.54	32528	1.9E-05	29586	3.96E-01	62114	1.20E-02	ND	ND	ND
PRKCE	rs10206343	2	46077623	T	C	0.659	0.776	32516	2.1E-05	rs10206343	2	46077623	best	0.659	0.776	T	C	87%	0.10	0.12	32516	2.1E-05	21823	1.98E-01	54339	1.33E-02	84159	4.12E-02	rs2029087
SIPA1L2	rs10910555	1	228667525	T	C	0.836	0.802	32514	1.7E-05	rs10910555	1	228667525	best	0.836	0.802	T	C	65%	0.94	0.82	32514	1.7E-05	50260	7.43E-01	82774	1.46E-02	112594	1.21E-01	rs10797582
ITIH5	rs4623795	10	7691915	G	C	0.855	GT	32512	2.0E-05	rs11255232	10	7686190	0.953	0.842	0.977	A	G	84%	0.20	0.82	32530	2.1E-05	50100	7.03E-01	82629	1.76E-02	112449	1.23E-01	rs1537626
MOCS1	rs12210863	6	40182179	G	A	0.928	0.915	32530	5.3E-05	rs12210863	6	40182179	best	0.928	0.915	G	A	86%	0.73	0.57	32530	5.3E-05	25646	1.73E-01	58175	3.41E-02	87995	1.71E-02	rs10807218
FAM19A5	rs4823535	22	47151477	G	A	0.606	0.846	32526	7.0E-04	rs4823535	22	47151477	best	0.606	0.846	G	A	69%	0.19	0.35	32526	7.0E-04	43246	2.75E-01	75772	1.63E-01	105592	1.64E-01	rs4823535

Best Ranking SNP results: results from SNPs that best associated with BMI
Lead SNP results: results from SNPs that were taken into follow-up analysis
Nearby Gene: gene nearest SNP used in analyses
Chr/Position: the chromosome and position (in Build 35) of SNP noted
Alleles: effect is the BMI increasing allele, other is the BMI decreasing allele
Imputation quality- imputation quality score for MACH (r.sq-hat) or IMPUTE (proper-info) (see Supplementary Information for further details)
R2 to best: R2 of Lead SNP to best SNP
Frequency: Frequency effect allele- weighted frequency of the effect allele (and the proxy SNP for those loci where more than one SNP was typed) in genotyped population based cohorts (EPIC, FINRISK97, BPPP, MetSIM)
Heterogeneity: between samples reports a p-value from an I2 analysis of heterogeneity in the GIANT samples and between genders reports a p-value from a t-test of gender specific effect sizes and their standard errors
Stage 1 (data from Stage 1 GIANT meta-analysis)
Stage 2 (data from meta analysis of genotyped and in silico studies)
With deCODE (data from meta analysis of Stage1, 2 and Decode data)

* SNPs 9939609 [FTO r2=0.96 to rs1121980], rs2568958 [NEGR1 r2=1 to rs2815752], rs6911147[DSSCR1L1 r2=0.93 to rs1867015], rs9931989 [SH2B1 r2=0.71 to rs7498665], rs2245715[MAPA1 r2=1 to rs 2255042], rs2726814 [ODZ3 r2=.74 to rs2597108], rs9650286 [LOC138046 r2=0.89 to rs11773921] and rs9455181 [C6orfrf57 r2=0.64 to rs9294876] were used in some stage 2 samples (see Supplementary Table 4)

deCODE SNP: data from deCODE for listed SNP was used for meta analysis
ND: not determined due to unavailability of good Illumina proxy for SNP used in analysis
frequency from genotyped proxy rs2175950 r2=0.71 in the Northern Finland Birth Cohort

Supplementary Table 6. Association between replicated SNPs and BMI, weight, height and total body fat mass in childhood (age 11) in the ALSPAC study and association with risk of extreme childhood obesity in the SCOOP-UK study.

Nearby gene	SNP	Chr	Position (Build 35)	Alleles		Frequency effect allele (ALSPAC) (%)		BMI*			Weight*			Height*			Total body fat mass (DEXA)*			Risk of extreme childhood obesity**					
				Effect	Other		n	Beta	SE	p-value	Beta	SE	p-value	Beta	SE	p-value	Beta	SE	p-value	N controls	N cases	OR	95%CI		p-value
FTO	rs9939609	16	52378028	A	T	40%	4888	0.125	0.020	***6.3E-10***	0.114	0.020	***1.9E-08***	0.039	0.021	0.06	0.134	0.021	***7.9E-11***	8227	926	1.67	1.51	1.85	***1.4E-23***
TMEM18	rs6548238	2	624905	C	T	84%	4906	0.110	0.027	***3.4E-05***	0.092	0.027	***5.4E-04***	0.018	0.027	0.51	0.094	0.027	***5.1E-04***	8284	1036	1.41	1.23	1.62	***7.9E-07***
MC4R	rs17782313	18	56002077	C	T	24%	5001	0.106	0.023	***6.2E-06***	0.078	0.023	***9.3E-04***	-0.011	0.024	0.65	0.083	0.024	***4.7E-04***	8436	995	1.22	1.10	1.35	***1.8E-04***
GNPDA2	rs10938397	4	45023455	G	A	43%	4896	0.048	0.020	***0.02***	0.043	0.020	***0.04***	0.013	0.021	0.54	0.053	0.021	***0.01***	8293	1016	1.20	1.09	1.31	***1.5E-04***
SH2B1	rs7498665	16	28790742	G	A	41%	4905	0.029	0.020	0.15	0.025	0.020	0.22	0.005	0.021	0.82	0.024	0.021	0.24	8277	1009	1.00	0.91	1.10	0.97
MTCH2	rs10838738	11	47619625	G	A	34%	4917	0.028	0.021	0.19	0.007	0.021	0.74	-0.031	0.021	0.14	0.019	0.022	0.37	8295	1033	1.06	0.97	1.17	0.21
KCTD15	rs368794#	19	39012292	T	A	68%	4838	0.072	0.022	***9.7E-04***	0.069	0.022	***1.7E-03***	0.032	0.022	0.15	0.067	0.022	***2.4E-03***	8156	1000	0.96	0.87	1.06	0.39
NEGR1	rs2568958#	1	72477137	A	G	60%	4948	0.010	0.020	0.63	0.018	0.020	0.38	0.023	0.020	0.27	0.025	0.021	0.23	8369	1038	1.29	1.17	1.42	***2.2E-07***

* gender specific z-score of log10-transformed traits

** Compares children of the UK-SCOOP study (extreme childhood obesity cases) vs all ALSPAC children (controls)

SNP rs2568958 [*NEGR1* r2=1 to rs2815752] and rs368794 [*KCTD15* r2=1 with rs11084753 was used in ALSPAC and SCOOP-UK

The results on rs9939609 were previously reported in Frayling et al (Science 2007) and Wardle et al (JCEM 2008). The results on rs17782313 were previously reported in Loos et al (Nature Genetics 2008)

Supplementary Table 7. Effects of reproducing loci by sample and genotype platform

		rs9939609 (A/T)				rs6548238 (C/T)				rs17782313 (C/T)				rs10938397 (G/A)			
Stage 1 sample	**Genotyping Platform**	**N**	**p-value**	**Effect**	**StdErr**	**N**	**p-value**	**Effect**	**StdErr**	**N**	**p-value**	**Effect**	**StdErr**	**N**	**p-value**	**Effect**	**StdErr**
CoLaus	Affymetrix 500K Array Set	5435	0.0013	0.31	0.09	5435	0.0027	0.37	0.12	5435	0.016	0.19	0.10	5435	0.038	0.21	0.10
SardiNIA	Affymetrix 500K Array Set & Affymetrix 10K Array Set	4301	6.0E-08	0.65	0.12	4301	0.019	0.34	0.16	4301	0.065	0.29	0.18	4301	0.15	0.16	0.13
EPIC-Obesity	Affymetrix 500K Array Set	2357	0.0030	0.36	0.12	2294	0.19	0.20	0.15	2413	0.029	0.28	0.13	2417	0.0034	0.33	0.12
NHS	Illumina HumanHap550	2265	0.0099	0.31	0.13	2265	0.55	0.00	0.18	2265	0.24	0.21	0.15	2265	0.052	0.25	0.13
PLCO	Illumina HumanHap300 & Illumina HumanHap240	2235	1.2E-04	0.44	0.11	2235	0.25	0.17	0.16	2235	0.10	0.25	0.13	2235	0.26	0.14	0.11
KORA	Affymetrix 500K Array Set	1642	0.060	0.24	0.15	1633	0.24	0.24	0.19	1642	0.031	0.40	0.16	1642	0.69	-0.04	0.16
WTCCC UK Blood Services	Affymetrix 500K Array Set	1458	0.14	0.25	0.16	1457	0.47	0.23	0.22	1457	0.072	0.34	0.18	1456	0.48	0.15	0.17
British 1958 Birth Cohort	Affymetrix 500K Array Set	1479	0.28	0.25	0.18	1479	0.16	0.38	0.23	1479	0.030	0.47	0.21	1479	0.64	-0.07	0.19
DGI Controls	Affymetrix 500K Array Set	1523	0.62	0.06	0.15	1523	0.051	0.44	0.20	1523	0.13	0.22	0.18	1523	0.73	0.08	0.17
FUSION Controls	Illumina HumanHap300	1291	0.13	0.22	0.16	1291	0.0052	0.62	0.23	1291	0.91	-0.13	0.20	1291	0.70	-0.07	0.17
WTCCC HT	Affymetrix 500K Array Set	1952	0.93	-0.01	0.13	1951	0.76	0.03	0.16	1950	0.0034	0.41	0.15	1950	0.13	0.22	0.14
WTCCC CAD	Affymetrix 500K Array Set	1925	0.62	0.06	0.14	1925	0.77	0.07	0.19	1922	0.91	0.03	0.16	1924	0.016	0.33	0.15
WTCCC T2D	Affymetrix 500K Array Set	1924	3.6E-05	0.84	0.20	1924	0.36	0.23	0.26	1924	0.70	0.12	0.22	1922	0.25	0.24	0.20
DGI T2D Cases	Affymetrix 500K Array Set	1588	0.053	0.27	0.17	1588	0.58	0.28	0.23	1588	0.69	-0.28	0.21	1588	0.38	0.19	0.19
FUSION T2D Cases	Illumina HumanHap300	1094	0.57	0.07	0.20	1094	0.91	0.09	0.31	1094	0.82	-0.07	0.27	1094	0.13	0.36	0.22
Affymetrix 500K Array Set		25584	3.7E-13	0.29	0.04	25509	1.0E-05	0.26	0.05	25634	4.3E-07	0.23	0.05	25637	8.0E-05	0.19	0.04
Illumina HumanHap		6885	8.1E-06	0.31	0.07	6885	0.027	0.19	0.10	6885	0.15	0.14	0.09	6885	0.031	0.16	0.07
Stage 1		32469	6.3E-17	0.29	0.04	32394	1.2E-06	0.24	0.05	32519	3.9E-07	0.21	0.04	32522	1.0E-05	0.18	0.04

SNPs are listed in the header row with the effect allele first, and the non-effect allele second.
Sample sizes and p-values are from the cohort-specific analyses that were used in the original stage 1 GWAS.
The direction of the effect is relative to the effect allele.
The effect sizes listed are from the uniform inverse normal transformed method, and were scaled into BMI units (kg/m^2) by multiplying by the cohort-specific standard deviation in BMI

Supplementary Table 7 (cont.). Effects of reproducing loci by sample and genotype platform

		rs7498665 (G/A)				rs10838738 (G/A)				rs11084753 (G/A)				rs2815752 (A/G)			
Stage 1 sample	**Genotyping Platform**	**N**	**p-value**	**Effect**	**StdErr**	**N**	**p-value**	**Effect**	**StdErr**	**N**	**p-value**	**Effect**	**StdErr**	**N**	**p-value**	**Effect**	**StdErr**
CoLaus	Affymetrix 500K Array Set	5435	1.10E-04	0.35	0.09	5435	0.096	0.14	0.09	5435	0.33	0.12	0.10	5435	0.94	-0.03	0.09
SardiNIA	Affymetrix 500K Array Set & Affymetrix 10K Array Set	4301	0.16	0.16	0.13	4301	0.075	0.16	0.14	4301	0.061	0.23	0.12	4301	1.6E-04	0.59	0.14
EPIC-Obesity	Affymetrix 500K Array Set	2417	0.076	0.19	0.11	2417	0.97	0.02	0.12	2417	0.00059	0.48	0.14	2417	0.55	0.06	0.12
NHS	Illumina HumanHap550	2265	0.73	-0.04	0.13	2265	0.0066	0.39	0.14	2265	0.078	0.25	0.14	2265	0.11	0.19	0.14
PLCO	Illumina HumanHap300 & Illumina HumanHap240	2207	0.092	0.20	0.11	2233	0.023	0.30	0.12	2181	0.0076	0.28	0.12	2235	0.95	0.00	0.12
KORA	Affymetrix 500K Array Set	1642	0.59	0.08	0.16	1642	0.66	-0.12	0.16	1642	0.22	0.21	0.18	1642	0.80	0.01	0.16
WTCCC UK Blood Services	Affymetrix 500K Array Set	1458	0.53	0.11	0.16	1458	0.63	-0.10	0.17	1456	0.047	0.37	0.19	1456	0.54	0.11	0.17
British 1958 Birth Cohort	Affymetrix 500K Array Set	1479	0.63	0.11	0.17	1479	0.29	-0.18	0.18	1479	0.86	0.01	0.21	1479	0.65	0.09	0.18
DGI Controls	Affymetrix 500K Array Set	1523	0.058	0.14	0.17	1523	0.0095	0.28	0.15	1523	0.31	0.19	0.19	1523	0.42	0.07	0.16
FUSION Controls	Illumina HumanHap300	1291	0.13	0.24	0.15	1291	0.79	-0.02	0.16	1291	0.82	0.01	0.17	1291	0.30	0.11	0.16
WTCCC HT	Affymetrix 500K Array Set	1952	0.46	-0.06	0.13	1952	0.023	0.29	0.13	1949	0.17	0.20	0.15	1950	0.044	0.26	0.13
WTCCC CAD	Affymetrix 500K Array Set	1926	0.11	0.24	0.14	1926	0.32	0.13	0.14	1923	0.66	-0.08	0.17	1924	0.024	0.34	0.15
WTCCC T2D	Affymetrix 500K Array Set	1924	0.46	0.17	0.20	1924	0.0011	0.62	0.20	1922	0.012	0.62	0.24	1922	0.034	0.40	0.20
DGI T2D Cases	Affymetrix 500K Array Set	1588	0.066	0.32	0.18	1588	0.52	0.09	0.18	1588	0.28	0.13	0.21	1588	0.11	0.25	0.18
FUSION T2D Cases	Illumina HumanHap300	1094	0.96	0.04	0.20	1094	0.17	0.21	0.20	1094	0.071	0.46	0.22	1094	0.43	0.18	0.21
Affymetrix 500K Array Set		25645	7.70E-06	0.18	0.04	25645	5.00E-04	0.12	0.04	25635	1.60E-05	0.21	0.05	25636	2.60E-05	0.17	0.04
Illumina HumanHap		6857	0.16	0.12	0.07	6883	0.0013	0.25	0.07	6831	0.002	0.24	0.08	6885	0.094	0.10	0.07
Stage 1		32502	5.40E-06	0.17	0.04	32528	7.1E-06	0.15	0.04	32466	2.6E-07	0.22	0.04	32521	9.3E-06	0.15	0.04

Supplementary Table 8. Association of replicated SNPs with other traits

SNP	Nearby Gene	Alleles		Height			Weight			% Body fat			Risk of overweight n_{cases} = 19,686 vs. $n_{controls}$ = 14,106		Risk of obesity n_{cases} = 5,261 vs. $n_{controls}$ = 14,106		LDL-cholesterol n=8,589		HDL-cholesterol n=8,656		Triglycerides n=8,684		Type 2 diabetes ncases =4,549 vs. ncontrols = 5,579		Coronary Artery Disease ncases = 1,926 vs. ncontrols = 10,377	
		Effect	Other	n	dir	p-value	n	dir	p-value	n	dir	p-value	OR (95%CI)	p-value	OR (95%CI)	p-value	dir	p-value	dir	p-value	dir	p-value	dir	p-value	dir	p-value
rs9939609*	*FTO*	A	T	36325	-	0.84	36325	+	9.31E-23	17984	+	1.49E-06	1.14 [1.10-1.18]	7.0E-15	1.25 [1.19-1.31]	5.5E-20	-	0.43	+	0.12	-	0.45	+	6.8E-07	+	0.47
rs6548238	*TMEM18*	C	T	37961	+	0.81	37962	+	1.82E-11	18279	+	0.0027	1.13 [1.08 - 1.18]	1.2E-08	1.19 [1.11 - 1.26]	1.4E-07	+	0.85	+	0.80	-	0.51	+	7.6E-04	+	0.27
rs17782313	*MC4R*	C	T	33734	+	4.0E-04	33735	+	2.72E-10	15974	+	0.16	1.07 [1.03-1.11]	0.0014	1.15 [1.08-1.21]	4.0E-06	+	0.24	-	0.85	+	0.57	+	0.23	+	0.19
rs10938397	*GNPDA2*	G	A	35523	-	0.11	35523	+	8.66E-08	16515	+	2.33E-05	1.08 [1.05 - 1.12]	4.4E-06	1.12 [1.07 - 1.17]	3.9E-06	+	0.65	+	0.02	+	0.70	+	6.6E-05	+	0.27
rs7498665*	*SH2B1*	G	A	33130	-	0.44	33131	+	7.75E-06	15207	+	0.012	1.07 [1.03 - 1.11]	1.1E-04	1.11 [1.06 - 1.17]	2.2E-05	+	0.056	+	0.07	+	0.80	+	0.35	-	0.22
rs10838738	*MTCH2*	G	A	33786	-	0.04	33786	+	0.27	16632	+	0.069	1.03 [1.00 - 1.07]	0.071	1.03 0.98 - 1.08]	0.24	+	0.23	+	0.14	-	0.18	-	0.88	-	0.03
rs11084753	*KCTD15*	G	A	25141	-	0.11	25140	+	0.37	9696	+	0.76	1.03 [0.99 - 1.07]	0.13	1.04 [0.98 - 1.10]	0.16	+	0.73	+	0.61	-	0.21	-	0.64	-	0.10
rs2815752*	*NEGR1*	A	G	36519	+	0.87	36520	+	9.96E-04	17141	+	0.0072	1.04 [1.01-1.08]	0.020	1.05 [1.01-1.11]	0.026	+	0.24	+	0.08	-	0.22	-	0.50	+	0.25

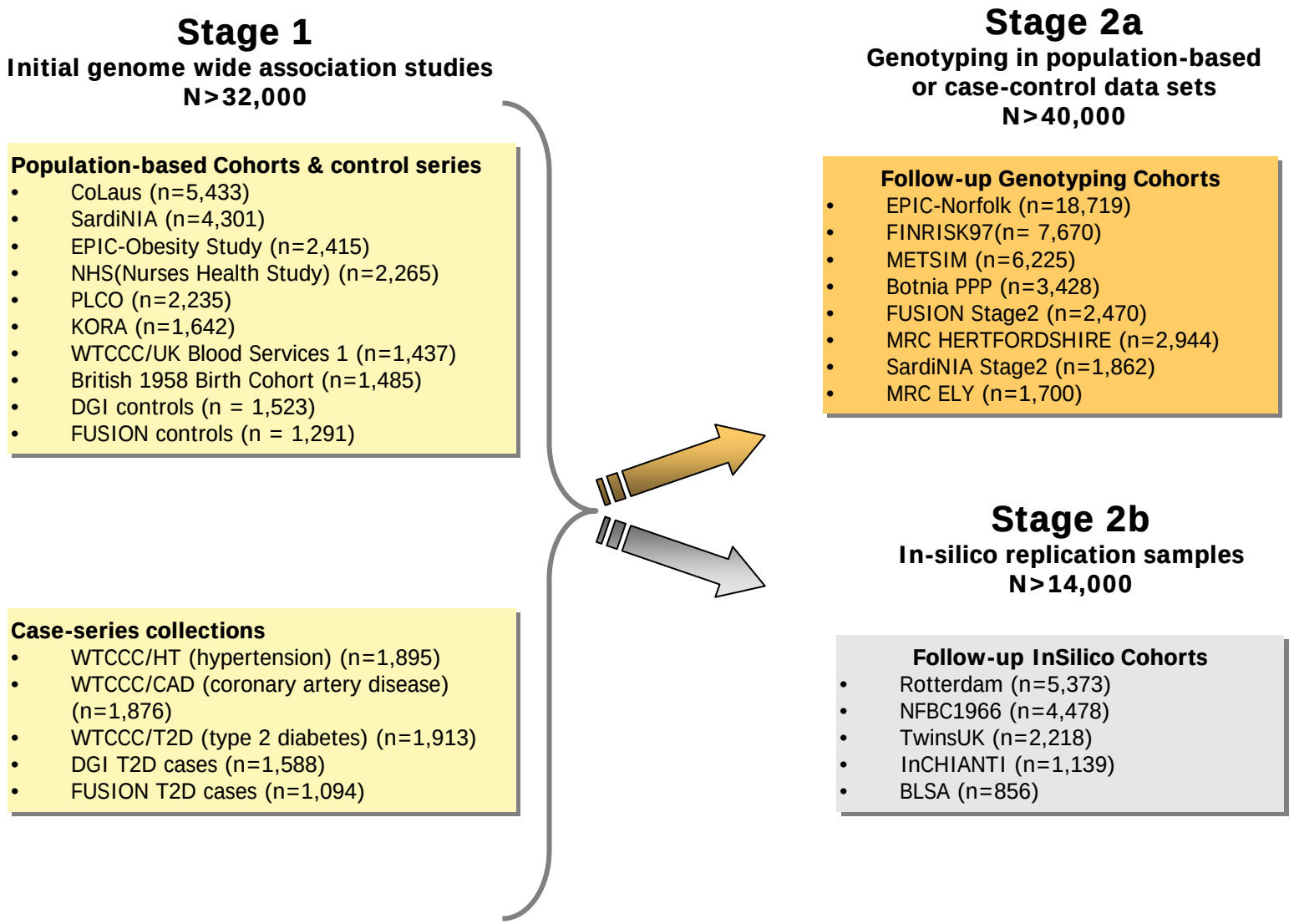
SNP: rs-number of single nuclear polymorphism (SNP) used in analyses
Nearby Gene: gene nearest SNP used in analyses
Alleles: effect is the BMI increasing allele
% fat mass- percent fat mass: number of individuals asesses (n) direction of effect for the effect allele (dir) and p-value of association (p-value) for the traits listed in genotyped Stage 2 samples: shaded boxes have a significant p-value of <0.005
Risk of overweight/obesity compared to normal weight where normal weight= BMI < 25 kg.m-2, overweight= BMI≥ 25 kg.m-2, obese= BMI ≥ 30 kg.m-2 in population based genotyped Stage 2 samples (FINRISK97, BPPP, METSIM, EPIC)
p-value of association and direction (dir) of effect of effect for low density lipoprotein (LDL) cholesterol, High density lipoprotein (HDL) cholesterol, Triglycerdes, Type 2 Diabetes, and Coronary Artery Disease
*rs9939609 [FTO r2=0.96 to rs1121980], rs2568958 [NEGR1 r2=1 to rs2815752] and rs9931989 [SH2B1 r2=0.68 with rs7498665] were used in some stage 2 samples

Supplementary Table 9. Correlation between replicated BMI variants and expression of nearby genes in brain and lymphocytes

Chr	Position (Build 35)	Gene/Locus	Lead SNP for follow-up	Proxy SNP	Distance (bp) between proxy and lead SNP	r^2 between lead and proxy	Symbol	Description	Probe ID (Brain)	p-value for association with proxy SNP (Brain, additive model)	Probes tested (Lymphocytes)	p-value for association with lead SNP (Lymphocytes, additive model)
1	72585028	near *NEGR1*	rs2815752	rs10789340	67266	0.79	NEGR1	Neuronal growth regulator 1, mRNA	GI_27754173	0.81	Average of 1553194_at, 1563316_at, 229461_x_at, 239548_at, 243357_at	0.82
2	624905	near *TMEM18*	rs6548238	rs6548238			TMEM18	Transmembrane protein 18, mRNA	GI_40548390	0.09	Average of 225487_at, 225489_at	0.56
							SH3YL1	SH3 domain containing, Ysc84-like 1 (S. cerevisiae), mRNA	GI_7661669	0.36	Average of 1554198_at, 204019_s_at, 244277_at	0.04
							ACP1	Acid phosphatase 1, soluble, transcript variant 2, mRNA	GI_30090001	0.08	Average of 1554808_at, 201629_s_at, 201630_s_at, 215227_x_at	0.17
							LOC285016	Hypothetical protein LOC285016, mRNA	GI_37547147	0.40	238018_at	0.50
							SNTG2	Syntrophin, gamma 2, mRNA	GI_9507164	0.29	Average of 1570422_at, 220487_at, 234823_at	0.26
4	44877284	near *GNPDA2*	rs10938397	rs348495	+1915	0.562	GNPDA2	Glucosamine-6-phosphate deaminase 2, mRNA	GI_19923880	0.51	Average of 227022_at, 230149_at	0.81
							GUF1	GTPase homolog, mRNA	GI_11345459	0.11		n/a
							KCTD8	Potassium channel tetramerisation domain containing 8, mRNA	GI_38198662	0.98		n/a
							GABRG1	Gamma-aminobutyric acid (GABA) A receptor, gamma 1, mRNA	GI_31742489	0.86	Average of 1552943_at, 241805_at	0.52
							GABRA2	Gamma-aminobutyric acid (GABA) A receptor, alpha 2, mRNA.	GI_4557600	0.74	Average of 1554308_s_at, 207014_at	0.95
							GABRB1	Gamma-aminobutyric acid (GABA) A receptor, beta 1, mRNA	GI_12548775	0.34	Average of 207010_at, 234204_at, 234629_at	0.60
							COX7B2	Cytochrome c oxidase subunit VIIb2, mRNA		n/a	231265_at	0.01
							GABRA4	Gamma-aminobutyric acid A receptor, alpha 4, mRNA		n/a	Average of 208463_at, 233437_at	0.36
11	47619625	near *MTCH2*	rs10838738	rs17788930	+89726	1	SPI1	Spleen focus forming virus proviral integration oncogene spi1, mRNA	GI_4507174	0.33	205312_at	0.32
							SLC39A13	Solute carrier family 39 (zinc transporter), member 13, mRNA	GI_40255100	0.87	Average of 1552295_a_at, 225277_at	0.65
							PSMC3	Proteasome (prosome, macropain) 26S subunit, ATPase, 3, mRNA	GI_24430153	***0.0041***	201267_s_at	**0.02**
							CUGBP1	CUG triplet repeat, RNA binding protein 1, transcript variant 1, mRNA	GI_38570082	0.03	Average of 1555467_a_at, 204113_at, 209489_at, 221742_at, 221743_at, 235297_at, 235865_at	0.18
							KBTBD4	Kelch repeat and BTB (POZ) domain containing 4, transcript variant 1	GI_32189315	0.02	Average of 218569_s_at, 218570_at, 223765_s_at	0.27
							NDUFS3	NADH dehydrogenase (ubiquinone) Fe-S protein 3, 30kDa (NADH-coenzyme Q reductase),	GI_4758787	0.28	201740_at	*0.000081*
							C1QTNF4	C1q and tumor necrosis factor related protein 4, mRNA	GI_13994272	0.31	223708_at	0.12
							MTCH2	mitochondrial carrier homolog 2 (C. elegans), nuclear gene encoding mitochondrial protein, mRNA	GI_40254847	***0.000013***	Average of 217772_s_at, 222403_at, 242645_at	**0.11**
							FNBP4	formin binding protein 4, mRNA	GI_24308032	***0.59***	Average of 212232_at, 229272_at, 235101_at, 239469_at, 242472_x_at	**0.02**
							RAPSN	43 kD receptor-associated protein of the synapse, mRNA		n/a	211570_s_at	0.19
16	28790742	*SH2B1*	rs7498665	rs8062405	-45335	1	CLN3	ceroid-lipofuscinosis, neuronal 3, juvenile (Batten, Spielmeyer-Vogt disease), mRNA	GI_4502888	0.08	Average of 209275_s_at, 210859_x_at	0.07
							SULT1A2	sulfotransferase family, cytosolic, 1A, phenol-preferring, member 2, transcript variant 1, mRNA	GI_29550878	0.36	Average of 207122_x_at, 211385_x_at	0.26
							SULT1A1	sulfotransferase family, cytosolic, 1A, phenol-preferring, member 1, transcript variant 1, mRNA	GI_29540530	0.14	Average of 203615_x_at, 215299_x_at, 238995_at	0.37
							EIF3S8	eukaryotic translation initiation factor 3, subunit 8, 110kDa, mRNA	GI_5579457	***0.0000021***	Average of 200647_x_at, 210949_s_at, 215230_x_at, 236700_at	***0.0000006***
							TUFM	Tu translation elongation factor, mitochondrial (TUFM), mRNA	GI_34147629	0.02	Average of 201113_at, 238190_at	*0.000022*
							SH2B	SH2-B homolog, mRNA	GI_24308080	0.45	Average of 209322_s_at, 40149_at	0.26
							LAT	linker for activation of T cells, mRNA	GI_24475949	0.01	Average of 209881_s_at, 211005_at	0.41
							APOB48R	apolipoprotein B48 receptor, mRNA		n/a	220023_at	0.93
							ATP2A1	ATPase, Ca++ transporting, fast twitch 1 isoform, mRNA		n/a	Average of 205444_at, 230693_at,	0.79
							CD19	CD19 antigen, mRNA		n/a	206398_s_at	0.09
							IL27	interleukin 27, mRNA		n/a	1552995_at	0.37
16	52378028	FTO	rs9939609	rs9939609			FTO	Fat mass and obesity related, mRNA	GI_37541493	0.89	Average of 209702_at, 215385_at	0.98
							KIAA1005	LOC23322, mRNA		n/a	213959_s_at	0.27
18	56002077	near *MC4R*	rs17782313	rs17782313			MC4R	melanocortin 4 receptor, mRNA	GI_5174532	0.28	221467_at	0.86
							PMAIP1	phorbol-12-myristate-13-acetate-induced protein 1, mRNA	GI_10863922	0.08	Average of 204285_s_at, 204286_s_at	0.58
19	39013977	near *KCTD15*	rs11084753	rs11084753			KCTD15	potassium channel tetramerisation domain containing 15, mRNA	GI_13129063	n/a	Average of 218553_s_at, 222664_at, 222668_at, 228683_s_at, 229683_s_at	0.12
							CHST8	carbohydrate (N-acetylgalactosamine 4-0) sulfotransferase 8	GI_21361615	n/a	221065_s_at	0.76

Supplementary Figure 1. Study design. Meta analysis of genome-wide association data was performed in stage 1 across the cohorts shown. SNPs representing the best associating loci were examined in follow up samples (stage 2) where they were genotyped (2a) or tested for replication in-silico (2b).

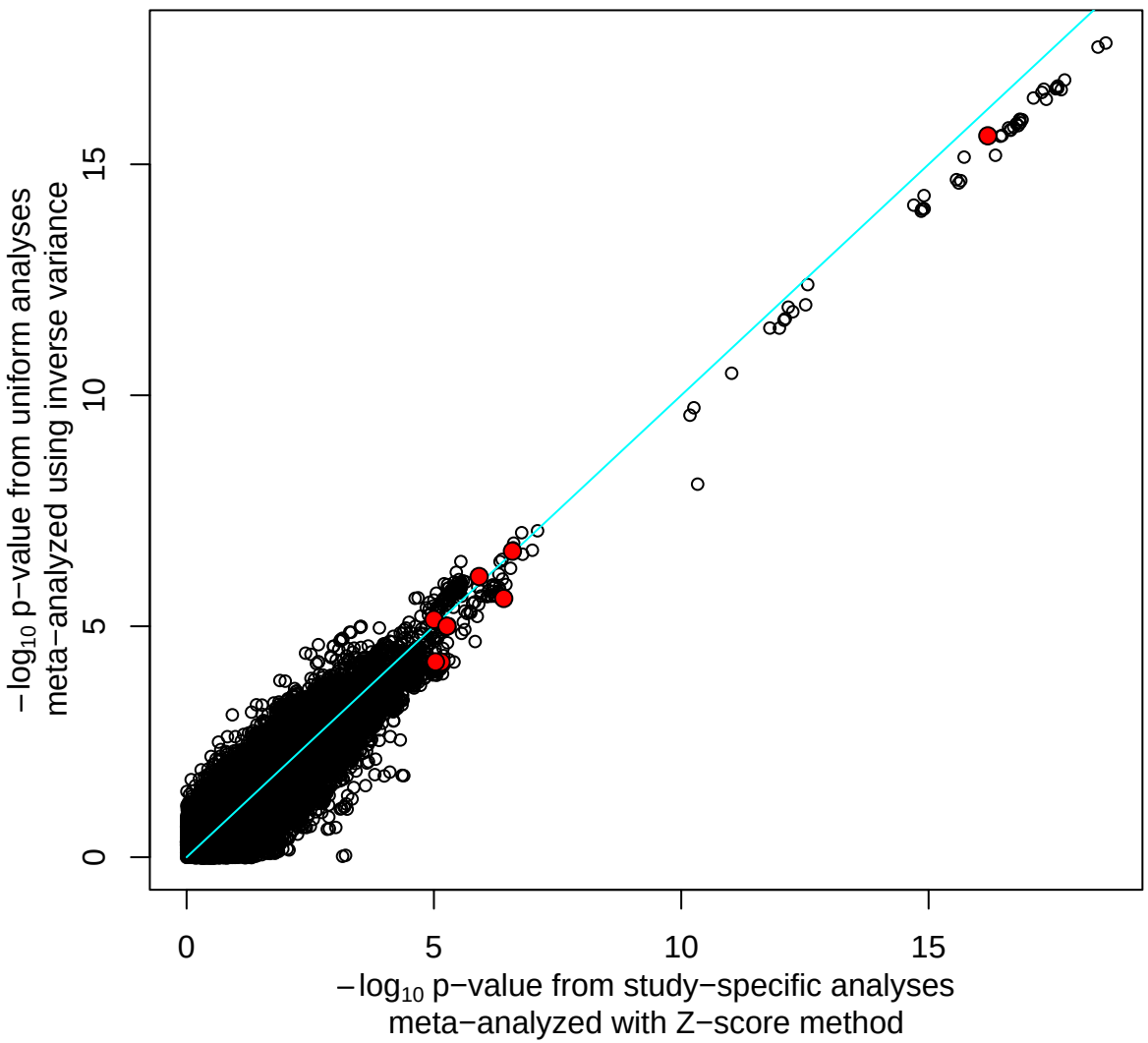
Supplemental Figure 1



Nature Genetics: doi:10.1038/ng.287

Supplementary Figure 2. Comparison of p-values obtained from study-specific and uniform meta-analysis strategies. Plot of $-\log_{10}$ p-value calculated in meta-analysis based on the study-specific z-score and on meta-analysis using inverse variance weighted combination of effect sizes. Blue line has slope of 1. SNPs where association was confirmed in this manuscript, including those near *FTO* and *MC4R*, are shown in red.

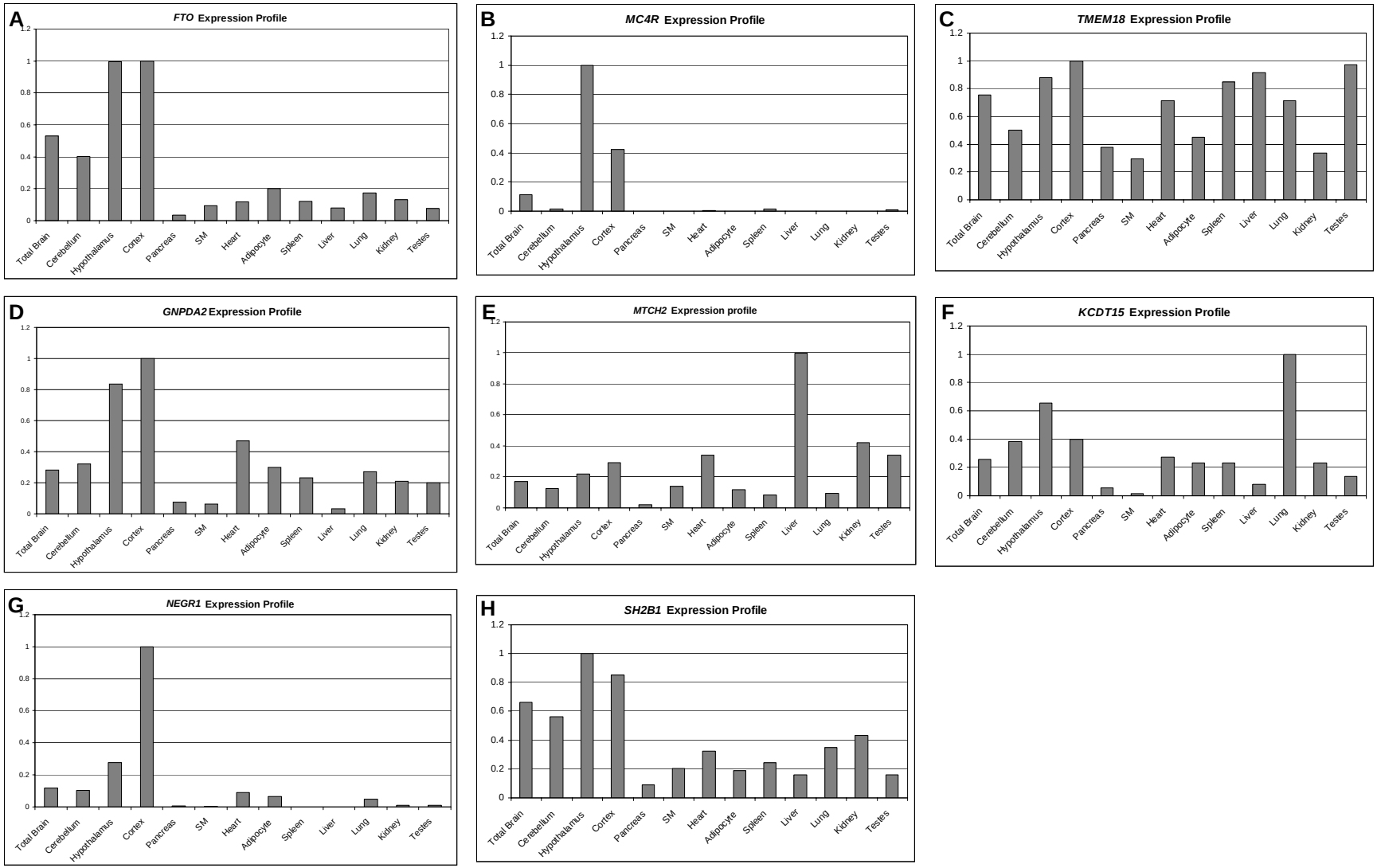
Supplemental Figure 2



Nature Genetics: doi:10.1038/ng.287

Supplementary Figure 3. Expression of genes nearest the most strongly associated validated SNPs in human tissues. Shown are the levels of expression for *FTO, MC4R, TMEM18, GNPDA2, MTCH2, KCTD15, NEGR1,* and *SH2B1* (A-H) in the tissues normalized with respect to *EEF2* levels. Abbreviations: SM- skeletal muscle.

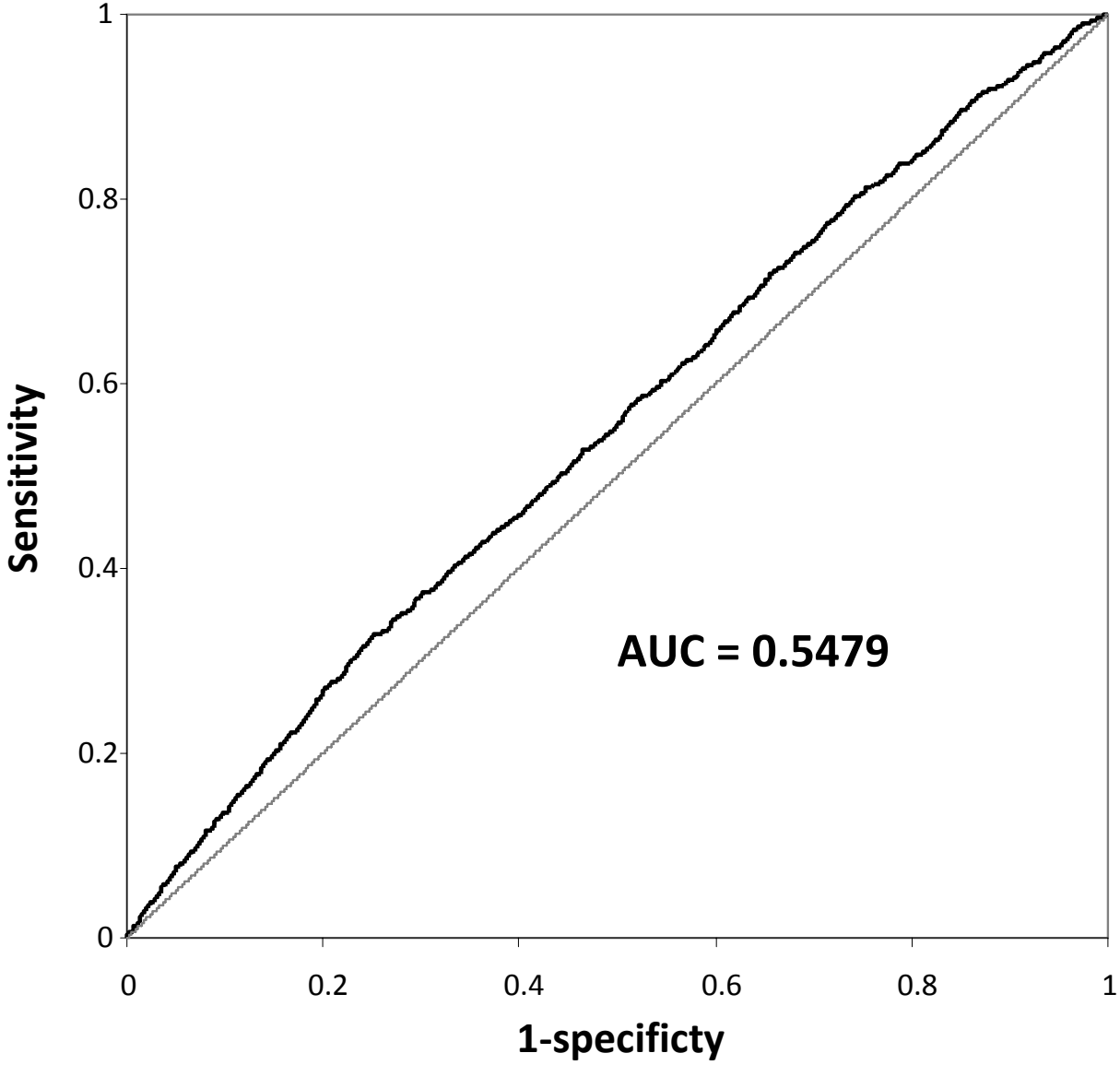
Supplemental Figure 3



Nature Genetics: doi:10.1038/ng.287

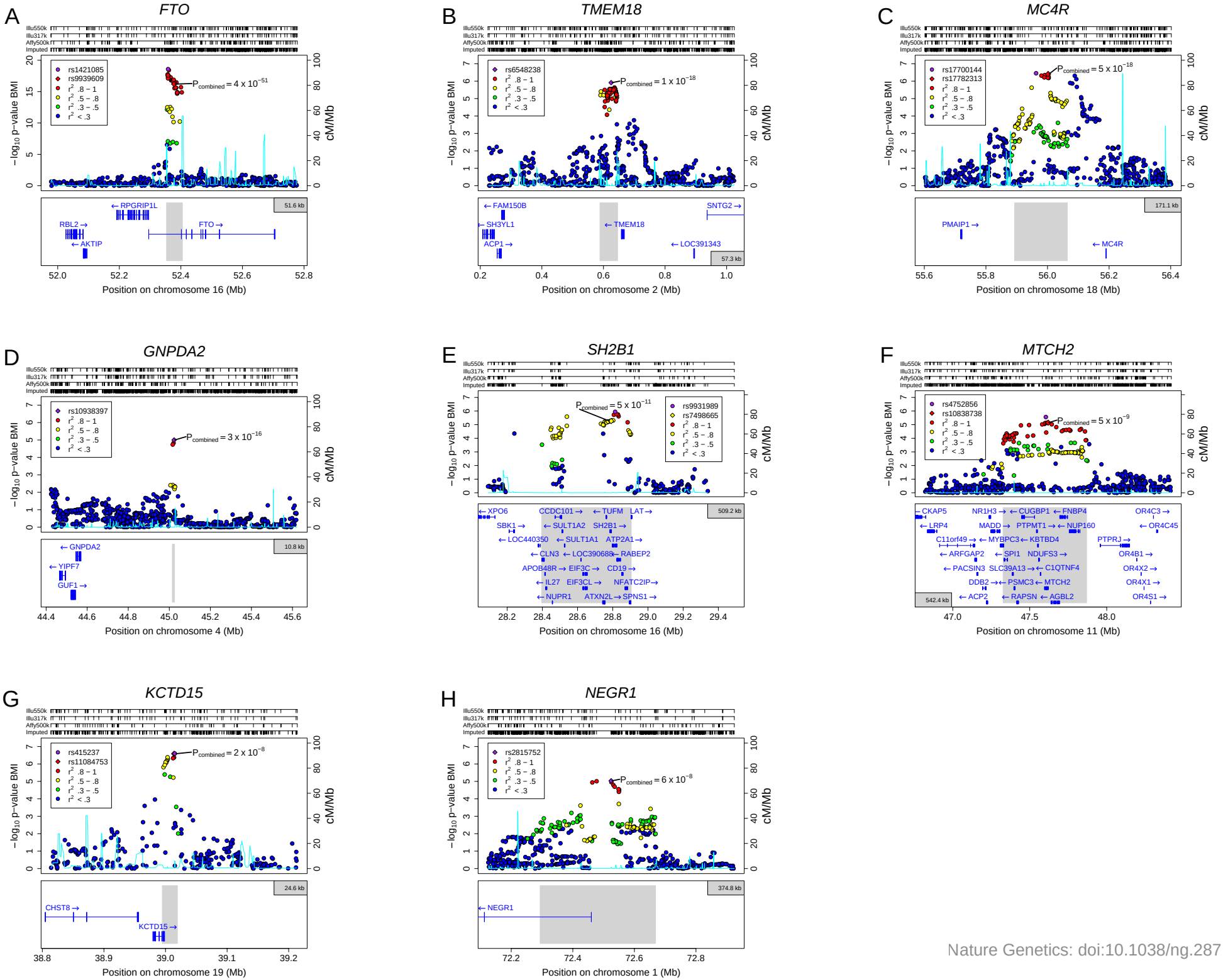
Supplementary Figure 4. Receiver operating characteristics (ROC) curve for the EPIC-Norfolk cohort. The eight confirmed BMI increasing variants (in or near *TMEM18, KCTD15, SH2B1, MTCH2, NEGR1, GNPDA2*, *FTO* and *MC4R*) were genotyped in the EPIC-Norfolk cohort (14,409 individuals with complete genotype data). Although a prediction model incorporating genotypes at these variants is significantly better than the null model ($p=2.32 \times 10^{-12}$) the area under the curve (AUC) of 0.548 indicates that BMI-associated variants as yet have limited predictive value. The null model is represented by the diagonal line and has an AUC of 0.5.

Supplemental Figure 4



Nature Genetics: doi:10.1038/ng.287

Supplementary Figure 5. Regional plots of replicating loci of association in Stage 1 cohorts around *FTO*, *TMEM18, MC4R, GNPDA2, SH2B1, MTCH2, KCTD15, and NEGR1* (A-H). SNPs are plotted by position on chromosome against association with BMI ($-\log_{10}$ p-value). The figures highlight the most significant SNP after Stage 1 meta-analysis (in purple) and the SNP selected for follow-up (diamond) in Stage 2 analyses, labeled with its combined p-value (Stage 1 + Stage 2). In most cases, the SNP followed-up is the most significant SNP in the region (therefore, a purple diamond). Otherwise, the LD between the followed-up SNP and the most significant SNP in the region is indicated by the color of the diamond. Estimated recombination rates (from HapMap) are plotted in cyan to reflect the local LD structure. The SNPs surrounding the most significant SNP (purple diamond) are color-coded to reflect their LD with this SNP as in the inset (taken from pairwise r^2 values from the HapMap CEU database, www.hapmap.org). Genes and the position of exons, as well as the direction of transcription, are noted below the plots (data from UCSC genome browser, genome.ucsc.edu) with a grey area marking the extent of the region that includes any SNP with $r^2 \geq 0.3$ relative to the most significant SNP (shown in purple). Hashmarks represent SNP positions on each genotyping array used by any individual study and also show SNP positions after imputation. In Part G, rs11084753 was selected as the reference SNP for the *KCTD15* region and shows essentially identical results to rs415237. The two SNPs are virtually superimposed on the association plot.



Nature Genetics: doi:10.1038/ng.287

SUPPLEMENTARY NOTE

Nature Genetics: doi:10.1038/ng.287

1. *STAGE 1 – GENOME-WIDE ASSOCIATION META-ANALYSIS*

Stage 1 of our study comprises a genome-wide screen for association between SNPs and BMI (Body Mass Index). We combined the summary statistics of 15 genome-wide association (GWA) studies in a meta-analysis (**Supplementary Figure 1**) amounting to a total of 32,387 individuals. The studies include six population-based studies (CoLaus, SardiNIA, EPIC-Obesity, KORA, WTCCC/UK Blood services and the British 1958 Birth Cohort), cases and controls of two type 2 diabetes (T2D) case-control studies (DGI and FUSION), and five case series (NHS, PLCO, WTCCC/HT, WTCCC/CAD and WTCCC/T2D).

1.1. DESCRIPTION OF POPULATIONS, GENOTYPING & ANALYSES

Basic descriptive characteristics for each study are presented in **Supplementary Table 1**, details on quality control of samples are summarized in **Supplementary Table 2** and on genome-wide genotyping, imputations and analysis models in **Supplementary Table 3.** This section describes the study specific characteristics that were not summarized in **Supplementary Tables 1-3**. All participants provided written informed consent and the studies were approved by the respective Local Research Ethics committees or Institutional Review Boards.

1.1.1. Population-based studies

1.1.1.1. The Cohorte Lausannoise (CoLaus) Study

The CoLaus study has been described in detail previously[1]. Participants were randomly selected from a list of 56,694 individuals aged 35 to 75 years who were permanent residents of the City of Lausanne, Switzerland. Only individuals with four grandparents of European origin were included in the study. The CoLaus study was sponsored in part by GlaxoSmithKline, and all participants were duly informed about this sponsorship.

Principal components were computed to adjust for population stratification using EIGENSOFT (http://genepath.med.harvard.edu/~reich/Software.htm). After using the Akaike Information Criterion (AIC) based stepwise model selection, the 3 principal components significant at $p<0.05$ were included as covariates in the association analyses.

1.1.1.2. The SardiNIA Study of Aging

The SardiNIA study has been described in detail previously[2,3]. The SardiNIA study examined a total of 4,301 related individuals participating in a longitudinal study of aging-related quantitative traits in the Ogliastra region of Sardinia, Italy.

Genotyped individuals had four Sardinian grandparents and were selected for genotyping without regard to their phenotypes. Among the individuals examined, 1,412 were genotyped with the Affymetrix Mapping 500K Array Set. A total of 356,359 SNPs met the quality control criteria and were used as input for the imputation procedure using the software MACH[4,5]. The remaining 2,893 individuals were genotyped with the Affymetrix Mapping 10K Array. These individuals were mostly offspring and siblings of the 1,412 individuals that were genotyped with the Affymetrix Mapping 500K Array Set. We took advantage of the relatedness among individuals to impute missing genotypes in these additional individuals; we identified large stretches of chromosome shared within each family and probabilistically "filled-in" genotypes within each stretch whenever one or more of its carriers was genotyped with the 500K Array Set[5,6].

Nature Genetics: doi:10.1038/ng.287

1.1.1.3. The EPIC-Obesity Study

This EPIC-Obesity study has been described in detail previously[7,8]. The EPIC Obesity cohort includes 2,566 participants randomly selected from the EPIC-Norfolk Study, a population-based cohort study of 25,663 men and women of European descent aged 39-79 years recruited in Norfolk, UK between 1993 and 1997.

1.1.1.4. Nurses Health Study (NHS)

The NHS design has been described in detail elsewhere[9]. The NHS began in 1976 with the recruitment of 121,700 female registered nurses between the ages of 30 and 55 years. The Cancer Genetic Markers of Susceptibility (CGEMS) nested case-control study is derived from 32,826 participants who provided a blood sample between 1989 and 1990 and were free of diagnosed breast cancer at blood collection and followed for incident disease until June 1, 2004. A total of 1,132 cases and 1,133 controls were genotyped and had data on height and weight.

Association between SNPs and BMI was adjusted for age and four statistically significant ($p < 0.05$) principal components that were derived from 10,000 SNP markers distributed across the genome using EIGENSTRAT[10].

1.1.1.5. Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial (PLCO)

The PLCO study design has been described in detail elsewhere[11,12]. PLCO is a population-based, randomized trial to evaluate early detection methods for prostate, lung, colorectal and ovarian cancer[11]. Between 1993 and 2001, over 150,000 men and women ages 55-74 years were enrolled in the trial from ten centers in the United States. A total of 1,172 non-Hispanic white prostate cancer cases and 1,105 matched, non-Hispanic white controls were selected for genotyping[12]. Of those 2,277 participants, information on height and weight was available for 2,235.

1.1.1.6. The KORA Study

The KORA surveys have been described in detail previously[13,14]. The third KORA survey (KORA S3, n=3,996) is a population-based sample from the general population of the South-German city of Augsburg and surrounding counties from 1994/1995. A subsample of 1,642 individuals from this survey with 10-year follow-up (KORA F3) information available was successfully genotyped (the KORA S3/F3 500K Study). All participants had a German passport and were of European origin.

1.1.1.7. The WTCCC/UK Blood Services (UKBS panel 1)

This study was part of the Wellcome Trust Case Control Consortium (WTCCC) and has been described previously[15]. A cohort composed of 1,500 (UK Blood Services [UKBS]) controls were selected from a sample of healthy blood donors ages 18 to 69 (with the majority between ages 40-59), recruited as part of the WTCCC project[15]. The participants were about equally divided into males and females of European ancestry. The current analysis included 1,437 individuals that passed the quality control criteria and that had height and weight data available.

1.1.1.8. British 1958 Birth Cohort (British 1958BC)

This study was part of the Wellcome Trust Case Control Consortium (WTCCC) and has been described previously[15,16]. The British 1958 Birth Cohort is a national population sample followed periodically from birth to age 44-45 years. A total of 1,502 cohort members were used as population controls in the WTCCC Consortium[15]. The current analysis included 1,479 individuals that passed the

Nature Genetics: doi:10.1038/ng.287

quality control criteria and that had height and weight data available.

1.1.2. Case-Control & Case series

The DGI and FUSION studies are case-control studies for type 2 diabetes. Genome-wide analyses were performed for cases and controls separately. The WTCCC-HT, WTCCC-CAD and WTCCC-T2D studies are part of the Wellcome Trust Case Control Consortium (WTCCC) and have been described previously[15].

1.1.2.1. The Diabetes Genetics Initiative (DGI)

The DGI study has been described in detail previously[17,18] and a description of the sample is also available online (http://www.broad.mit.edu/diabetes/). The DGI study is a T2D case-control study that includes 1,588 T2D cases and 1,523 matched controls of European ancestry from Sweden and Finland.

1.1.2.2. The Finland-United States Investigation of NIDDM Genetics (FUSION)

The FUSION study has been described in detail previously[3,19,20]. The FUSION GWA study is a T2D case-control study that includes 1,161 Finnish T2D cases, 1,174 normal glucose tolerant (NGT) controls, and 122 offspring of case/control pairs (1 T2D, 119 NGT, 2 with impaired glucose tolerance). Cases and controls were approximately frequency matched as previously described, taking into account five-year age category, sex, and birth province within Finland[20].

In addition to the QC criteria listed in **Supplementary Table 2**, SNPs were excluded when the total number of duplicate pair discrepancies in 79 duplicate sample pairs and Mendelian inheritance inconsistencies in 122 parent-offspring sets was >3.

1.1.2.3. WTCCC-HT Cases

Participants comprised severely hypertensive probands ascertained from families with multiplex affected sibships or as parent–offspring trios. All cases were of British ancestry and were recruited from the Medical Research Council General Practice Framework and other primary care practices as part of the UK77 study[21]. The current analysis included 1,895 individuals that passed the quality control criteria and that had height and weight data available.

1.1.2.4. WTCCC-CAD Cases

WTCCC-CAD cases were individuals of European descent who had a validated history of either myocardial infarction (MI) or coronary revascularisation before their 66th birthday. Recruitment was carried out on a national basis through responses to a sustained media campaign, responses to posters placed within hospitals and GP surgeries, and as part of a pilot-phase, by contacting patients listed on computer based coronary artery disease databases in the two lead centres. A collection of 2000 unrelated cases affected by premature CAD were selected first for the presence of MI and then on the age of onset of disease for the WTCCC study. The current analysis included 1,876 individuals that passed the quality control criteria and that had height and weight data available.

1.1.2.5. WTCCC/T2D Cases

A total of 1,999 individuals of British/Irish descent were selected from the Diabetes UK Warren 2 repository. Diagnosis of diabetes was made between ages of 25 and 75 and was based on either current prescribed treatment with diabetes-specific medication or, in the case of those treated with diet alone, historical or contemporary laboratory evidence of hyperglycemia. Approximately 30% of cases

Nature Genetics: doi:10.1038/ng.287

were explicitly recruited as part of multiplex sibships, and ~25% represented the T2D offspring within parent-offspring "triads" or "duos". The remainder of the participants were recruited as isolated cases ascertained for early age at diagnosis compared to the population distribution and a high proportion of diabetic relatives. The current analysis included 1,913 individuals that passed the quality control criteria and that had height and weight data available.

1.2. IMPUTATION OF GENOTYPES

Because studies used different genotyping platforms, there was only partial overlap of genotyped SNPs across the different GWA scans. Therefore, we imputed for each GWA scan separately the genotypes of polymorphic autosomal SNPs that were available for the HapMap Phase II - CEU population. This approach facilitates the comparison and meta-analysis of GWA scans, as each scan will have association summary statistics for the same SNPs.

Imputations were based on genotypes of SNPs and samples that passed the study-specific quality control criteria, using either MACH[4] (for 8 GWA scans) or IMPUTE[22] (for 7 GWA scans) as imputation software (**Supplementary Table 3**). Genotypes were imputed for SNPs that were not already genotyped by the genome-wide genotyping chip or for genotyped SNPs that had failed the quality control criteria.

Both imputation software programs rely on the same principles to impute genotypes. Using observed GWAS genotypes as input, a hidden markov model is used to select matching mosaics of haplotypes from a reference panel (the HapMap CEU) and to infer genotypes at SNPs that were studied in the reference panel, but not in the GWAS panel. The certainty of the inference depends on the stretches of haplotypes shared between the reference panel and the study samples. Per individual, the imputation result at each of the imputed SNPs can be expressed as a set of three genotype probabilities, which vary between 0 and 1 per genotype and sum to 1, or as an 'allele dosage', which is defined as the expected number of copies of the minor allele at each SNP and varies between 0 and 2.

Furthermore, each imputation package estimates an overall imputation quality score for each SNP. IMPUTE calculates the 'info' statistic which is a measure of the observed statistical information for the estimate of allele frequency of the SNP. This measure varies between 0 and 1, with 1 indicating perfect information. MACH calculates the 'rsq_hat', which is the estimated r^2 between each imputed genotype and its true underlying genotype. Rsq_hat ranges from 0 to 1, with 1 indicating high quality imputation.

Further details on imputation are available in **Supplementary Table 3**. Both imputation software programs are available online; MACH[4] at http://www.sph.umich.edu/csg/abecasis/MaCH/ and IMPUTE[22] at http://www.stats.ox.ac.uk/~marchini/software/gwas/impute.html.

1.3. GENOME-WIDE ASSOCIATION ANALYSES IN INDIVIDUAL COHORTS

Details on study-specific analysis software and modelling of outcome and covariates are summarised in **Supplementary Table 3**.

Each GWA study performed association analyses between imputed or observed SNPs and BMI, assuming an additive inheritance model. The SardiNIA, and FUSION studies applied a regression-based analysis under a variance

Nature Genetics: doi:10.1038/ng.287

components framework to appropriately account for the relatedness among individuals[5].

The uncertainty of the imputed genotypes was taken into account in the analysis using methods appropriate for the imputation software used. Studies that imputed genotypes using IMPUTE used the SNPtest software[22] for association analyses, which implements frequentist tests to calculate p-values using the imputed genotype probabilities to account for the uncertainty of the genotypes[22]. SNPTEST also calculates the 'proper_info' statistic as a measure of the relative statistical information about the additive genetic effect being estimated. The proper_info statistic has a direct relationship to the power of the test and ranges from 0 to 1, with 1 indicating perfect information for that test. Studies that imputed genotypes with MACH used either MACH2QTL or Merlin, which both use the 'allele dosage' (0.0-2.0) to summarize the imputed genotype probabilities.

For each GWA scan, two series of association analyses were performed; [1] a 'best' analysis that modelled outcome and covariates taking into account study-specific characteristics (**Supplementary Table 3**) and [2] a 'uniform' analysis applying an inverse normal transformation of BMI and adjusted for age, age^2, for men and women separately.

The genomic control inflation factors (λ_{GC}) for each of the GWA scans separately did not exceed 1.10 (**Supplementary Table 3**) and genomic control correction was applied to each study prior to meta-analysis.

1.4. META-ANALYSIS OF GWA STUDIES

A total of 2,399,588 autosomal SNPs with a [1] MAF ≥1%, [2] a proper_info (SNPTEST) of ≥0.40 or a rsq_hat (MACH) of ≥0.30, and [3] data available for at least 16,000 individuals (and up to 32,387 individuals) were included in the meta-analysis (**Supplementary Table 3**).

We used a weighted z-score-based fixed effects meta-analysis to combine the results of the 15 GWA studies[23]. First, a z-statistic for each SNP, which represents the significance of the association (based on p-value) and the direction of effect (sign relative to chosen reference allele), is calculated for each of the 15 GWA scans separately. Next, the study-specific z-statistics for each SNP are combined into one overall z-statistic, computed as the weighted sum of the individual z-statistics (with weights proportional to the square root of the relative sample size for each study). The p-value corresponding to the overall z-statistic was then calculated. Weights are proportional to the square-root of the number of individuals examined in each study and scaled such that the squared weights sum to 1.0.

We performed two meta-analyses using either the p-values of the 'best' analyses or the p-values of the 'uniform' analyses and found that the summary statistics for both analyses were extremely similar (**Supplementary Figure 2**), particularly for strongly associated loci. The correlation between the log of the p-values of the 'best' and 'uniform' analyses was 0.91. SNP selection for follow-up was based on the meta-analysis of the 'best' results. Meta-analyses were performed using the software program METAL (www.sph.umich.edu/csg/abecasis/metal).

The genomic control inflation factor (λ_{GC}) for the 'best' meta-analysis was 1.10 and genomic control correction was conservatively applied to the meta-analysis p-values.

Nature Genetics: doi:10.1038/ng.287

2. *STAGE 2 – FOLLOW UP*

Stage 2 of our study comprises the follow-up of 35 SNPs from the most significant loci of the GWA meta-analysis (Stage 1) (**Supplementary Figure 1 and Supplementary Table 1-4)** in 59,082 individuals from 14 studies. A total of 45,018 individuals from 9 studies (EPIC-Norfolk, FINRISK97, METSIM, PPP-Botnia study, FUSION stage 2, Hertfordshire study, SardiNIA stage2, and MRC Ely study) were directly genotyped, while for 14,064 individuals from 5 studies (The Rotterdam study, the Northern Finnish Birth Cohort 1966, TwinsUK, InChianti and BLSA) the SNPs were extracted from existing GWA scans.

2.1. DIRECT GENOTYPING FOLLOW-UP

Basic descriptive characteristics for each study are presented in **Supplementary Table 1** and details on genotyped SNPs are summarised in **Supplementary Table 4.** SNPs were preferentially chosen for follow-up genotyping if they were the best associating SNP from stage 1. If the best associating SNP from stage 1 did not design, alternative SNPs were chosen to represent the locus. These alternative SNPs were chosen based on the following criteria: 1. If a SNP was directly genotyped on the Illumina platform and had an R2>0.9 to the best SNP that SNP was used 2. Otherwise the SNP with the best r^2 possible to the best SNP at that associating locus was chosen but no SNP with an r^2 below 0.6 was used. All participants provided informed consent and the studies were approved by the respective Local Research Ethics committees or Institutional Review Boards.

2.1.1. Description of populations

2.1.1.1. EPIC-Norfolk

Sample - EPIC-Norfolk is an ongoing prospective cohort study of chronic diseases comprising 25,663 Norfolk residents, an ethnically homogenous European origin population aged 39-79 who were recruited from general practice registers between 1993 and 1997 for a first health examination. Height and weight were measured according to standardized protocols to calculate BMI. A total of 21,583 samples were immediately available for genotyping. Individuals that had been analysed in the context of the GWA-meta-analyses (see 1.1.1.3) were excluded for replication analyses. Eventually, genotypes for 18,719 individuals with height and weight data were available for analyses. Body fat percentage was measured by bioelectrical impedance and was available for 10,096 individuals. More details on the study design of EPIC-Norfolk studies have been reported elsewhere[7,24].

Genotyping & Quality Control – Nine SNPs were genotyped using TaqMan® SNP genotyping assay (Applied Biosystems, Warrington, UK) according to the manufacturer's protocol. A total of 31 SNPs were genotyped using the platform iPLEX™ Sequenom MassARRAY®, of which 6 failed **(Supplementary Table 4)**. Genotype frequencies were in HWE ($p>1 \times 10^{-6}$), call rates were >94%.

2.1.1.2. FINRISK97

Sample - FINRISK 1997 is one of the population-based risk factor surveys carried out by the National Public Health Institute of Finland every five years[25]. The sample was drawn from the National Population register in 1997 for five geographical areas (North Karelia, Kuopio, Oulu province, Helsinki and Vantaa, and Turku/Loimaa area) in Finland, and stratified so that the cell size was 250 per each sex, 10-year age

Nature Genetics: doi:10.1038/ng.287

group (range 25-64 years), and study area. An additional 'senior sample' included 250 men and 500 women aged 65-74 years from the North Karelia and Helsinki/Vantaa area. A total of 11,500 individuals were invited of whom 8,447 (73.5%) participated. Height and weight were measured according to standardized protocols to calculate BMI[25]. DNA sample and BMI data were available for 7,670 participants.

Genotyping & Quality Control – A total of 35 SNPs were genotyped using the platform iPLEX™ Sequenom MassARRAY®, of which 2 failed **(Supplementary Table 4).** Genotype frequencies were in HWE ($p>1 \times 10^{-6}$), call rates were >99%, and concordance of duplicates was >99%.

2.1.1.3. The METabolic Syndrome In Men (METSIM) study

Sample - The METSIM study is an ongoing study that aims to investigate the metabolic syndrome, type 2 diabetes, cardiovascular disease, and cardiovascular risk factors, with the goal of collecting 9,000 men, randomly selected from the population of the town of Kuopio in Eastern Finland (population 95,000). Our sample included the first 6229 participants, aged 45-73 years. Height, weight and body fat percentage (by bioelectrical impedance) were measured during a 1-day visit to the Clinical Research Unit of the University of Kuopio. Height and weight was available in 6225 individuals. The study was approved by the ethics Committee of the University of Kuopio.

Genotyping & Quality Control – A total of 32 SNPs were genotyped using the platform iPLEX™ Sequenom MassARRAY®, of which 3 failed **(Supplementary Table 4).** Genotype frequencies were in HWE ($p>0.01$), call rates were >95%, and concordance of duplicates ($n = 124$) was >99%.

2.1.1.4. Botnia Prevalence, Prediction and Prevention of Diabetes (PPP) study

Sample - Prevalence, Prediction and Prevention of Diabetes (PPP) in the Botnia study – a population-based study started in 2004 aiming 1) to study diabetes prevalence in the Botnia population, 2) to test whether genetic and metabolic risk factors previously identified in the Botnia study can be used to identify high risk individuals and 3) to study whether incidence of T2D can be prevented in these high risk individuals using an intervention program[26]. The study includes 3,428 individuals with data on height, weight and body fat percentage (measured by bioelectrical impedance).

Genotyping & Quality Control – A total of 31 SNPs were genotyped using the platform iPLEX™ Sequenom MassARRAY®, of which 4 failed **(Supplementary Table 4).** Genotype frequencies were in HWE ($p>1 \times 10^{-6}$), call rates were >99%, and concordance of duplicates was >99%.

2.1.1.5. FUSION stage 2

Sample - The FUSION study stage 2 sample includes 1211 T2D cases and 1266 NGT controls approximately frequency matched to take into account five-year age category, sex, and birth province within Finland[19]. FUSION stage 2 samples do not overlap with the individuals used in the initial GWA scan. Height and weight measurements were available for 1204 individuals with type 2 diabetes and 1266 NGT individuals.

Genotyping & Quality Control – A total of 31 SNPs were genotyped using the platform iPLEX™ Sequenom MassARRAY®, of which 5 failed **(Supplementary Table 4).** Genotype frequencies were in HWE ($p>0.008$), call rates were >95%,

Nature Genetics: doi:10.1038/ng.287

and concordance of duplicates (n = 59) was 100%.

2.1.1.6. The Hertfordshire Cohort Study

Sample - The Hertfordshire Cohort Study comprises 2,997 men and women born in the English county of Hertfordshire during the period 1931-1939 and still resident there today. At follow-up (age 60-75 years), all participants attended a clinic for detailed physiological investigations as described previously[27]. Height and weight were measured according to standardized procedures to calculate BMI. Genotypes and phenotype data were available on 2,944 individuals. The study has ethical approval from the Hertfordshire and Bedfordshire Local Research Ethics Committee and all participants have given written informed consent.

Genotyping & Quality Control – Genotyping of rs8055982 and rs7498665 was performed using TaqMan® SNP genotyping assay (Applied Biosystems, Warrington, UK) according to the manufacturer's protocol. Genotype frequencies were in HWE ($p>0.05$), call rates were >98%, and concordance of duplicates (n = 300) was 100%.

2.1.1.7. The SardiNIA study – Stage 2

Sample – The SardiNIA stage 2 sample includes 1,862 individuals selected from the SardiNIA study but genetically unrelated from those individuals analyzed in the GWA scan.

Genotyping & Quality Control – Genotyping of rs8055982 and rs7498665 was performed using TaqMan® SNP genotyping assay (Applied Biosystems, Warrington, UK) according to the manufacturer's protocol. Genotype frequencies were in HWE ($p>0.05$), call rates were >95%, with no Mendelian errors over 431 parent-offspring pairs.

2.1.1.8. MRC-Ely Study

Sample - The MRC Ely Study is a population-based cohort randomly selected from people living in Ely and surrounding villages (East Anglia, UK), an ethnically homogenous European ancestry population. The study design, methods and measurements of the three phases have been described in detail elsewhere[28-31]. The current analyses included 1,700 men and women, aged 35-79 years, from phase 3. Height and weight were measured according to standardized protocols. Ethical permission was granted by the Cambridgeshire Research Ethics Committee, and study participants provided written informed consent.

Genotyping & Quality Control – Genotyping of rs8055982 and rs7498665 was performed using TaqMan® SNP genotyping assay (Applied Biosystems, Warrington, UK) according to the manufacturer's protocol. Genotype frequencies were in HWE ($p>0.05$), call rates >96%, with >98% concordances of duplicates (n = 30).

2.1.2. Association analyses within each of the "stage 2 direct genotyping" cohorts

Association analyses between each of the genotyped SNPs and BMI were performed in each of the "stage 2 – direct genotyping" studies separately, but in an identical manner.

2.1.2.1. BMI association analyses

BMI was first regressed against age, age^2 and sex to obtain residuals. Subsequently, we tested for association between [1] each SNP and the (raw) residuals, as well as between [2] each SNP and the inverse normal transformed residuals after modelling covariates age, age^2 and sex, using linear regression

Nature Genetics: doi:10.1038/ng.287

assuming an additive effect of each additional risk-allele.

2.1.2.2. Additional association analyses

In secondary analyses, we tested for association between each SNP and height, weight and body fat percentage (if available) in a similar way as in the primary analysis; i.e. the residuals (raw or inverse normal transformed), obtained after regressing against age, age^2 and sex, were tested for association with each SNP using a linear regression model that assumed an additive effect.

Logistic regression was used to test for association between each SNP and two discrete variables, overweight status (BMI ≥ 25 kg.m^{-2}) and obese status (BMI ≥ 30 kg.m^{-2}). Age, age^2 and sex were included in the model as covariates. SNPs were coded assuming a log-additive model (multiplicative) model. The non-overweight (BMI < 25 kg.m^{-2}) individuals were considered the reference group, while the overweight or obese individuals were considered the at-risk group.

2.2. IN-SILICO FOLLOW-UP

Basic descriptive characteristics for each study are presented in **Supplementary Table 1** and details on the genotyping and quality control are shown in **Supplementary Tables 2 & 3** and on the extracted SNPs in **Supplementary Table 4**. All participants provided written informed consent and the studies were approved by the respective Local Research Ethics committees or Institutional Review Boards.

2.2.1. Description of populations

2.2.1.1. Rotterdam Study

The Rotterdam Study is an ongoing prospective population-based cohort study, focused on chronic disabling conditions of the elderly. The study comprises an outbred ethnically homogenous population of Dutch Caucasian origin. The rationale of the study has been described in detail elsewhere[32,33]. In summary, 7983 men and women aged 55 years or older, living in Ommoord, a suburb of Rotterdam, the Netherlands, were invited to participate. Body height and weight were measured using standardized procedures. BMI and genome-wide genotype data were available for 5,373 participants. The current analyses include 5,373 individuals.

2.2.1.2. Northern Finland Birth Cohort of 1966 (NFBC1966)

The Northern Finland Birth Cohort of 1966 (NFBC1966) was originally designed to study factors affecting pre-term birth, low birth weight and subsequent morbidity and mortality. Mothers living in the two northern-most provinces of Finland were invited to participate if they had expected delivery dates during 1966. A total of 12,058 live-births were in the study. At age 31 all individuals still living in the Helsinki area or Northern Finland were asked to participate in a detailed biological and medical examination (n=5,923) as well as a questionnaire. Anthropometric measures, including height and weight were taken. Genotype and measured BMI data were available on 4,598 individuals in this study with multiple births being excluded. The University of Oulu ethics committee approved the study.

2.2.1.3. TwinsUK

The TwinsUK cohort (www.twinsuk.ac.uk) is an adult twin registry shown to be representative of the UK singleton population[34]. A total of 2218 women were included in the analysis, including 1018 singletons (754 twin members from monozygotic pairs and 253 from dizygotic pairs and 11 of pairs with unknown

Nature Genetics: doi:10.1038/ng.287

zygosity), and 1200 siblings from 600 dizygotic pairs. Height and weight were measured according to standardised protocols during visit at the clinic. Written informed consent was obtained from every participant to the study, and the study was approved by the Guy's and St. Thomas' Hospital Ethics Committee.

2.2.1.4. InCHIANTI

InCHIANTI is an epidemiological study of risk factors contributing to the decline in physical functioning in late life[35]. Individuals were selected from the population registries of two small towns in Tuscany, Italy. Participants, all of white European origin, were invited to a clinic visit for evaluation of health status as described in detail previously[36]. Height and weight were measured according to standardized procedures. BMI and genotype data were available for 1139 participants. The Italian National Institute of Research and Care of Aging Institutional Review Board ratified the study protocol.

2.2.1.5. The Baltimore Longitudinal Study of Aging (BLSA)

The BLSA is an on-going prospective study that began in 1958 to investigate changes that occur with normal aging. The study consists of volunteers recruited primarily from the Washington DC and Baltimore, MD area. Details of this study have been described elsewhere[37]. Genome-wide data was available for 1230 participants. The analysis was restricted to Caucasian individuals only (N=856). The study was approved by the Medstar Research Institutional Review Board.

2.2.2. Association analyses within each of the "stage 2 *in silico*" cohorts

Each *in silico* study performed its 'best' analyses between each of the follow-up SNPs and BMI adjusting for study-specific covariates (**Supplementary Table 3**). We used p values from these best analyses for meta-analysis.

2.2.2.1. Rotterdam Study

We performed association between imputed or observed SNPs and BMI (log-transformed BMI standardized residuals adjusted for age, age^2 and gender, assuming an additive inheritance model using MACH2QTL.

2.2.2.2. Northern Finnish Birth Cohort of 1966 (NFBC66)

Tests of association between each SNP and BMI (sex- and age- specific z-score of log-transformed BMI) were performed using a Cochran-Armitage test for an additive genetic effect (1 degree of freedom, df) for the presence of each additional effect allele using SNPTEST[22]. Geographical differences in the population structure were taken into account using principal components (PC) analysis based on the genome-wide IBS pairwise distances, as implemented in PLINK.

2.2.2.3. TwinsUK

Genome-wide association analysis sample was carried out using a score test and variance components to account for twin status on untransformed BMI, correcting for age. Analyses were implemented in the Merlin software package[38].

2.2.2.4. InChianti

An inverse normal transformation was applied to BMI and tested for association with imputed allele dosages using Merlin[38]. The analysis was adjusted for age, age^2 and gender.

2.2.2.5. The Baltimore Longitudinal Study of Aging (BLSA)

Associations between imputed allele dosages and BMI were assessed using Merlin[38].

Nature Genetics: doi:10.1038/ng.287

An inverse normal transformation was applied to BMI and the analysis was adjusted for age, age^2 and gender.

3. *META-ANALYSES OF STAGE 1 AND STAGE 2 RESULTS*

3.1. BMI META-ANALYSIS

We used a weighted z-score-based fixed effects meta-analysis to combine the results of the 15 GWA studies of stage 1, with the results of 6 studies that performed direct genotyping and the 5 studies for which *in silico* results were available from stage 2. The details of the method were described above (see 1.4). Meta-analyses were performed using the software program METAL (www.sph.umich.edu/csg/abecasis/metal).

3.2. ADDITIONAL ANALYSES IN STAGE 2 SAMPLES

3.2.1. Estimation of effect size

To estimate an additive effect for SNPs that showed significant evidence of association after follow-up studies, we performed secondary meta-analyses including only the six directly genotyped stage 2 samples. We used the inverse variance-weighted method, which combines the effect sizes (beta or log odds-ratios) and standard errors estimated within each study using linear or logistic regression. Effect sizes were pooled using the fixed effect model that assumed no heterogeneity. Heterogeneity was tested using the Q statistic and was found to be negligible.

We performed a meta-analysis for the associations with BMI residuals (raw) to estimate an overall effect size for each SNP. Similar analyses were performed for the residuals of height, weight and body fat percentage as well as for the risk of being overweight and obese. All analyses were initially performed within each study and adjusted for age, age^2 and sex. When the lead SNP was not genotyped in a particular sample due to assay limitations, we use a proxy SNP in strong linkage disequilibrium (if available, $r^2 > 0.80$).

Meta-analyses were performed using the software program METAL (www.sph.umich.edu/csg/abecasis/metal).

3.2.2. Heterogeneity

We tested for heterogeneity among studies and men and women in stage 1 samples for the 35 SNPs that were taken forward to stage 2. The I^2-statistic[39] and the Cochran's heterogeneity statistic[40] were used to estimate between-study heterogeneity. To test for gender differences in effect sizes we first meta-analysed the results of the 'uniform' analyses for men and women separately. The gender-specific effect sizes and their standard errors were compared by using a t-test.

4. *ANALYSES IN CHILDHOOD COHORTS*

4.1. DESCRIPTION OF POPULATIONS

4.1.1. Population-based sample – The ALSPAC Population Children's cohort

Sample **-** The Avon Longitudinal Study of Parents and Children (ALSPAC) is a

Nature Genetics: doi:10.1038/ng.287

prospective study, which recruited pregnant women with expected delivery dates between April 1991 and December 1992 from Bristol, UK[41]. Individuals of known non-white ethnic origin were excluded from all analyses. DNA was collected from mothers and children as described previously[42]. After exclusion for ethnicity, genotypes were available for 8,369 children. Where the dataset included multiple singleton siblings born to the same mother, only the first-born was included in the analyses of children. Children were invited annually to attend a specially-designed clinic, at which anthropometric measures, including height and weight, were taken[43]. A total of 4,951 children had BMI and genotype data available at age 11. Of these, 4,876 children agreed to undergo a whole-body dual energy X-ray absorptiometry (DEXA) scan to assess fat and lean mass, described in detail previously[44]. All aspects of the study are reviewed and approved by the ALSPAC Law and Ethics Committee, which is registered as an Institutional Review Board. Approval was also obtained from the Local Research Ethics Committees, which are governed by the Department of Health.

Genotyping & Quality Control – Genotyping of the 6 replicated SNPs were performed by KBiosciences (Hoddesdon, UK) using their own system of fluorescence-based competitive allele-specific PCR (KASPar). Details of assay design are available from the KBiosciences website (http://www.kbioscience.co.uk). Genotype frequencies were in HWE ($p>0.01$), call rates were 97%, with 100% concordances of duplicates.

4.1.2. Severe Childhood Onset Obesity Project UK (SCOOP-UK)

Sample - The Severe Childhood Onset Obesity Project UK (SCOOP-UK) comprises 1,038 UK participants of European descent with severe early onset obesity of unknown aetiology. This cohort has emerged out of the Genetics of Obesity Study (GOOS) (n=2,800). The entry criteria for the GOOS cohort comprise a BMI > 3 SDs from the age-specific mean and an onset of obesity before the age of 10 years[45]. SCOOP-UK represents a subgroup of GOOS patients of UK European ancestry in whom all the known monogenic obesity syndromes have been excluded by direct nucleotide sequencing. The ALSPAC children (as described above) were used as controls (n = 8,369) in this analysis. To exclude potential population stratification, as the ALSPAC cohort is geographically-based, we also used the EPIC-Norfolk population as controls (as described above).

Genotyping & Quality Control – Genotyping of the 6 SNPs was performed using TaqMan® SNP genotyping assay (Applied Biosystems, Warrington, UK) according to the manufacturer's protocol. Genotype frequencies were in HWE ($p>0.01$) and call rates >96%.

4.2. ASSOCIATION ANALYSES

4.2.1. Population-based sample – The ALSPAC Population Children's cohort

BMI, weight, height and DEXA fat mass at age 11 were log-transformed before calculating sex-specific Z-scores. Association between these anthropometric measures and SNPs was tested using linear regression assuming an additive effect for the presence of each additional minor allele. All analyses were performed with Stata/SE 9.2 for Windows (StataCorp LP, College Station, Texas, USA).

4.2.2. Case-control studies - SCOOP-UK

Nature Genetics: doi:10.1038/ng.287

For the case-control analysis, all SCOOP-UK children were considered cases, while ALSPAC children, irrespective of their BMI were used as controls (n = 8,369). Association between SNP and case-control status was performed using Fisher's exact test. Similar results were obtained when the EPIC-Norfolk cohort was used as controls for SCOOP-UK.

5. *ACKNOWLEDGEMENTS*

CoLaus: The CoLaus study was supported by research grants from GlaxoSmithKline and from the Faculty of Biology and Medicine of Lausanne, Switzerland. The authors would like to express their gratitude to the participants in the Lausanne CoLaus study, to the investigators who have contributed to the recruitment, in particular Yolande Barreau, Anne-Lise Bastian, Binasa Ramic, Martine Moranville, Martine Baumer, Marcy Sagette, Jeanne Ecoffey and Sylvie Mermoud for data collection and to Allen Roses, and Lefkos T. Middleton for their support. Some computation was carried out on the Vital-IT system at the Swiss Institute of Bioinformatics. We thank Paul Matthews, Lefkos Middleton, Dan Burns, Eric Lai and Allen Roses for their support. We thank Jacques Beckmann and Sven Bergmannfor assistance with initial analyses. Computation for genotype imputation in CoLaus was done on the Vital-IT system at the Swiss Institute of Bioinformatics.

SardiNIA: We are indebted to the many volunteers who generously participated in these studies. This work was supported in part by the Intramural Research Program of the National Institute on Aging (NIA), by extramural grants from National Human Genome Research Institute (NHGRI HG02651) and the National Heart Lung and Blood Institute (NHLBI HL084729). The SardiNIA team was supported by Contract NO1-AG-1-2109 from the NIA. Additional support for the SardiNIA study was provided by the mayors, administration and residents of Lanusei, Ilbono, Arzana and Elini and the head of Public Health Unit ASL4 in Sardinia. G.R.A. is a Pew Scholars for the Biomedical Sciences.

DGI, PPP-Botnia: The studies were supported by a Linné grant from Swedish Research Council, Wallenberg Foundation, the Juselius Foundation, and the Folkhälsan Research Foundation. DGI was funded by a grant from Novartis. E.K.S. was supported by an NIH T32 DK07191 grant to Daniel K. Podolsky in the Department of Gastroenterology at Massachusetts General Hospital and NIH grants F32 DK079466 and K23 DK080145. H.N.L. was supported by NIH grant K23 DK067288. Genotyping and analysis in PPP-Botnia was supported by NIH grant DK075787 and an ADA Smith Family Foundation Pinnacle Program Project Award, to J.N.H.

The EPIC-Obesity Study & MRC Ely study: The EPIC Norfolk Study is funded by Cancer Research United Kingdom and the Medical Research Council. The MRC Ely Study was funded by the Medical Research Council and the Wellcome Trust. I.B. and S.O. acknowledge support from EU FP6 funding (contract no LSHM-CT-2003-503041). S.L. is supported by a grant from Unilever Corporate Research. We thank Sarah Dawson, Farzana Shah, Sofie Ashford, Larissa Richardson, and Steven Knighton for their rapid and accurate large-scale sample preparation and genotyping, and Stephen Sharp for his statistical support.

Nature Genetics: doi:10.1038/ng.287

FUSION and METSIM: The FUSION and METSIM studies thank the Finnish citizens who generously participated in these studies. Support for FUSION was provided by NIH grants DK062370 (M.B.) and DK072193 (K.L.M.), intramural project number 1Z01 HG000024 (F.S.C.), and a postdoctoral fellowship award from the American Diabetes Association (C.J.W.). Genome-wide genotyping was performed by the Johns Hopkins University Genetic Resources Core Facility (GRCF) SNP Center at the Center for Inherited Disease Research (CIDR) with support from CIDR NIH Contract Number N01-HG-65403. Support for METSIM was provided by grant 124243 from the Academy of Finland (M.L.).

NHS: We thank Constance Chen for data management and analysis. We thank the participants in the Nurses' Health Studies. The Nurses' Health Studies are supported by US NIH grants CA65725, CA87969, CA49449, CA67262, CA50385 and 5UO1CA098233.

PLCO: This research was supported by the Intramural Research Program of the Division of Cancer Epidemiology and Genetics and by contracts from the Division of Cancer Prevention, National Cancer Institute, National Institutes of Health (NIH), Department of Health and Human Services (DHHS). The authors thank Drs. Christine Berg and Philip Prorok at the Division of Cancer Prevention, National Cancer Institute, the investigators and staff at the screening centers for the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial, Mr. Tom Riley and staff at Information Management Services, Inc., and Ms. Barbara O'Brien and staff at Westat, Inc. Most importantly, we acknowledge the study participants for their generous contribution in making this study possible.

WTCCC-CAD: Collection of the CAD used in the WTCCC was funded by the British Heart Foundation and the Medical Research Council. N.J.S. hold a BHF Chair.

WTCCC-T2D: For the WTCCC samples, genome wide genotypes for cases and controls were generated through funding from the Wellcome Trust (Strategic Award 076113/B/04/Z). Additional support for collection, genotyping and analysis was provided by the British Heart Foundation, Diabetes UK, the Medical Research Council (G0601261, G0000649), the European Commission (ENGAGE:HEALTH-F4-2007-201413), and the U.K. National Institute for Health Research (via the Oxford Biomedical Research Centre). N.J.S. holds a BHF Chair, E.Z. a Wellcome Trust Research Career Development Fellowship, C.M.L. a Nuffield Department of Medicine Scientific Leader Fellowship, M.W. is Vandervell Foundation Research Fellow and C.W. holds a intermediate BHF research fellow (Grant no. FS/05/061/19501).

KORA: This work was funded and supported by the National Genome Research Net (HE Wichmann, J Hebebrand, T Meitinger), Germany, EU FP6 funding (contract No. LSHM-CT-2003-503041, J Hebebrand), the Munich Center of Health Sciences (MC Health) as part of LMUinnovativ (HE Wichmann), Germany and the NIH-subcontract from the Children's Hospital, Boston, USA, under the prime grant 1 R01 DK075787-01A1, CFDA 93.848 (HE Wichmann, IM Heid). The KORA research platform was financed by the Helmholtz Center Munich.

Nature Genetics: doi:10.1038/ng.287

The British 1958 BC: We acknowledge use of genotype data from the British 1958 Birth Cohort DNA collection, funded by the Medical Research Council grant G0000934 and the Wellcome Trust grant 068545/Z/02.

FINRISK97: The Study was supported by European Community's Seventh Framework Programme grant agreement HEALTH-F4-2007- 201413, The Center Of Excellence of the Academy of Finland and the Finnish Heart Association. Genotyping in FINRISK97 was supported by NHLBI grant 5R01HL087679-02 through the STAMPEED program and NIDDK grant 5R01DK075787. V.S. was supported by a grant from the Sigrid Juselius Foundation. P.J. was supported by the Academy of Finland, grant number 118065. M.K. was supported by The Finnish Cultural Foundation.

ALSPAC - We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. The UK Medical Research Council, the Wellcome Trust (including grant number 076467/Z/05/Z) and the University of Bristol provide core support for ALSPAC.

SCOOP-UK: S.O. and I.S.F. are funded by the Wellcome Trust, MRC and the NIHR Biomedical Research Centre.

The Hertfordshire Study: This study was supported by the Medical Research Council UK and the University of Southampton UK.

Rotterdam Study: The GWAS database of the Rotterdam Study was funded through the Netherlands Organisation of Scientific Research NWO (nr. 175.010.2005.011). We thank Dr Michael Moorhouse, Pascal Arp and Mila Jhamai for their help in creating the database. The Rotterdam Study is supported by the Erasmus Medical Center and Erasmus University, Rotterdam; the Netherlands organization for scientific research (NWO), the Netherlands Organization for the Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commision (DG XII), and the Municipality of Rotterdam.

NFBC66: We acknowledge the support of NHLBI grant 5R01HL087679-02 through the STAMPEED program (1RL1MH083268-01), the Medical Research Council, UK, Academy of Finland, EURO-BLCS, QLG1-CT-2000-01643, University Hospital Oulu, Finland, European Community's Seventh Framework Programme (FP7/2007-2013), ENGAGE project, grant agreement HEALTH-F4-2007-201413**.** Genotyping of NFBC was supported on NHLBI grant 5R01HL087679-02 through the STAMPEED program.

TwinsUK: TwinsUK thanks NIHR Biomedical Research Centre (BRC) at Kings College London and Guys and St Thomas' Foundation Hospitals' Trust for support, and acknowledges support of a BBSRC project grant.

Nature Genetics: doi:10.1038/ng.287

InChianti and BLSA: This work was supported in part by the Intramural Research Program of the National Institute on Aging (NIA). We thank Drs. Andrew Singleton, Jack Guralnik, and Stefania Bandinelli for their kind help with the study.

CNV analysis: S.A.M. was supported by a Life Sciences Research Fellowship. A.E. was supported by a Sarnoff Cardiovascular Research Foundation Fellowship.

Membership of the Wellcome Trust Case Control Consortium

Management Committee: Paul R Burton[1], David G Clayton[2], Lon R Cardon[3], Nick Craddock[4], Panos Deloukas[5], Audrey Duncanson[6], Dominic P Kwiatkowski[3,5], Mark I McCarthy[3,7], Willem H Ouwehand[8,9], Nilesh J Samani[10], John A Todd[2], Peter Donnelly (Chair)[11]
Data and Analysis Committee: Jeffrey C Barrett[3], Paul R Burton[1], Dan Davison[11], Peter Donnelly[11], Doug Easton[12], David M. Evans[3], Hin-Tak Leung[2], Jonathan L Marchini[11], Andrew P Morris[3], Chris CA Spencer[11], Martin D Tobin[1], Lon R Cardon (Co-chair)[3], David G Clayton (Co-chair)[2]
UK Blood Services & University of Cambridge Controls: Antony P Attwood[5,8], James P Boorman[8,9], Barbara Cant[8], Ursula Everson[13], Judith M Hussey[14], Jennifer D Jolley[8], Alexandra S Knight[8], Kerstin Koch[8], Elizabeth Meech[15], Sarah Nutland[2], Christopher V Prowse[16], Helen E Stevens[2], Niall C Taylor[8], Graham R Walters[17], Neil M Walker[2], Nicholas A Watkins[8,9], Thilo Winzer[8], John A Todd[2], Willem H Ouwehand[8,9]
1958 Birth Cohort Controls: Richard W Jones[18], Wendy L McArdle[18], Susan M Ring[18], David P Strachan[19], Marcus Pembrey[18,20]
Bipolar Disorder (Aberdeen): Gerome Breen[21], David St Clair[21]; (**Birmingham):** Sian Caesar[22], Katherine Gordon-Smith[22,23], Lisa Jones[22]; **(Cardiff):** Christine Fraser[23], Elaine K Green[23], Detelina Grozeva[23], Marian L Hamshere[23], Peter A Holmans[23], Ian R Jones[23], George Kirov[23], Valentina Moskvina[23], Ivan Nikolov[23], Michael C O'Donovan[23], Michael J Owen[23], Nick Craddock[23]; **(London):** David A Collier[24], Amanda Elkin[24], Anne Farmer[24], Richard Williamson[24], Peter McGuffin[24]; **(Newcastle):** Allan H Young[25], I Nicol Ferrier[25]
Coronary Artery Disease (Leeds): Stephen G Ball[26], Anthony J Balmforth[26], Jennifer H Barrett[26], D Timothy Bishop[26], Mark M Iles[26], Azhar Maqbool[26], Nadira Yuldasheva[26], Alistair S Hall[26]; **(Leicester):** Peter S Braund[10], Paul R Burton[1], Richard J Dixon[10], Massimo Mangino[10], Suzanne Stevens[10], Martin D Tobin[1], John R Thompson[1], Nilesh J Samani[10]
Crohn's Disease (Cambridge): Francesca Bredin[27], Mark Tremelling[27], Miles Parkes[27]; **(Edinburgh):** Hazel Drummond[28], Charles W Lees[28], Elaine R Nimmo[28], Jack Satsangi[28]; **(London):** Sheila A Fisher[29], Alastair Forbes[30], Cathryn M Lewis[29], Clive M Onnie[29], Natalie J Prescott[29], Jeremy Sanderson[31], Christopher G Mathew[29]; **(Newcastle):** Jamie Barbour[32], M Khalid Mohiuddin[32], Catherine E Todhunter[32], John C Mansfield[32]; **(Oxford):** Tariq Ahmad[33], Fraser R Cummings[33], Derek P Jewell[33]
Hypertension (Aberdeen): John Webster[34]; **(Cambridge):** Morris J Brown[35], David G Clayton[2]; **(Evry, France): G** Mark Lathrop[36]; **(Glasgow):** John Connell[37], Anna Dominiczak[37]; **(Leicester):** Nilesh J Samani[10]; **(London):** Carolina A Braga

Nature Genetics: doi:10.1038/ng.287

Marcano[38], Beverley Burke[38], Richard Dobson[38], Johannie Gungadoo[38], Kate L Lee[38], Patricia B Munroe[38], Stephen J Newhouse[38], Abiodun Onipinla[38], Chris Wallace[38], Mingzhan Xue[38], Mark Caulfield[38]; **(Oxford):** Martin Farrall[39]
Rheumatoid Arthritis: Anne Barton[40], Ian N Bruce[40], Hannah Donovan[40], Steve Eyre[40], Paul D Gilbert[40], Samantha L Hider[40], Anne M Hinks[40], Sally L John[40], Catherine Potter[40], Alan J Silman[40], Deborah PM Symmons[40], Wendy Thomson[40], Jane Worthington[40]
Type 1 Diabetes: David G Clayton[2], David B Dunger[2,41], Sarah Nutland[2], Helen E Stevens[2], Neil M Walker[2], Barry Widmer[2,41], John A Todd[2]
Type 2 Diabetes (Exeter): Timothy M Frayling[42,43], Rachel M Freathy[42,43], Hana Lango[42,43], John R B Perry[42,43], Beverley M Shields[43], Michael N Weedon[42,43], Andrew T Hattersley[42,43]; **(London):** Graham A Hitman[44]; **(Newcastle):** Mark Walker[45]; **(Oxford):** Kate S Elliott[3,7], Christopher J Groves[7], Cecilia M Lindgren[3,7], Nigel W Rayner[3,7], Nicholas J Timpson[3,46], Eleftheria Zeggini[3,7], Mark I McCarthy[3,7]
Tuberculosis (Gambia): Melanie Newport[47], Giorgio Sirugo[47]; **(Oxford):** Emily Lyons[3], Fredrik Vannberg[3], Adrian VS Hill[3]
Ankylosing Spondylitis: Linda A Bradbury[48], Claire Farrar[49], Jennifer J Pointon[48], Paul Wordsworth[49], Matthew A Brown[48,49]
AutoImmune Thyroid Disease: Jayne A Franklyn[50], Joanne M Heward[50], Matthew J Simmonds[50], Stephen CL Gough[50]
Breast Cancer: Sheila Seal[51], Michael R Stratton[51,52], Nazneen Rahman[51]
Multiple Sclerosis: Maria Ban[53], An Goris[53], Stephen J Sawcer[53], Alastair Compston[53]
Gambian Controls (Gambia): David Conway[47], Muminatou Jallow[47], Melanie Newport[47], Giorgio Sirugo[47]; **(Oxford):** Kirk A Rockett[3], Dominic P Kwiatkowski[3,5]
DNA, Genotyping, Data QC and Informatics (Wellcome Trust Sanger Institute, Hinxton): Suzannah J Bumpstead[5], Amy Chaney[5], Kate Downes[2,5], Mohammed JR Ghori[5], Rhian Gwilliam[5], Sarah E Hunt[5], Michael Inouye[5], Andrew Keniry[5], Emma King[5], Ralph McGinnis[5], Simon Potter[5], Rathi Ravindrarajah[5], Pamela Whittaker[5], Claire Widden[5], David Withers[5], Panos Deloukas[5]; **(Cambridge):** Hin-Tak Leung[2], Sarah Nutland[2], Helen E Stevens[2], Neil M Walker[2], John A Todd[2]
Statistics (Cambridge): Doug Easton[12], David G Clayton[2]; **(Leicester):** Paul R Burton[1], Martin D Tobin[1]; **(Oxford):** Jeffrey C Barrett[3], David M Evans[3], Andrew P Morris[3], Lon R Cardon[3]; **(Oxford):** Niall J Cardin[11], Dan Davison[11], Teresa Ferreira[11], Joanne Pereira-Gale[11], Ingeleif B Hallgrimsdóttir[11], Bryan N Howie[11], Jonathan L Marchini[11], Chris CA Spencer[11], Zhan Su[11], Yik Ying Teo[3,11], Damjan Vukcevic[11], Peter Donnelly[11]
PIs: David Bentley[5,54], Matthew A Brown[48,49], Lon R Cardon[3], Mark Caulfield[38], David G Clayton[2], Alistair Compston[53], Nick Craddock[23], Panos Deloukas[5], Peter Donnelly[11], Martin Farrall[39], Stephen CL Gough[50], Alistair S Hall[26], Andrew T Hattersley[42,43], Adrian VS Hill[3], Dominic P Kwiatkowski[3,5], Christopher G Mathew[29], Mark I McCarthy[3,7], Willem H Ouwehand[8,9], Miles Parkes[27], Marcus Pembrey[18,20], Nazneen Rahman[51], Nilesh J Samani[10], Michael R Stratton[51,52], John A Todd[2], Jane Worthington[40]

[1] Genetic Epidemiology Group, Department of Health Sciences, University of Leicester, Adrian Building, University Road, Leicester, LE1 7RH, UK; [2] Juvenile Diabetes Research Foundation/Wellcome Trust Diabetes and Inflammation Laboratory, Department of Medical Genetics, Cambridge Institute for Medical Research, University of Cambridge, Wellcome Trust/MRC Building, Cambridge, CB2 0XY, UK; [3]

Nature Genetics: doi:10.1038/ng.287

Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, UK; [4] Department of Psychological Medicine, Henry Wellcome Building, School of Medicine, Cardiff University, Heath Park, Cardiff CF14 4XN, UK; [5] The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK; [6] The Wellcome Trust, Gibbs Building, 215 Euston Road, London NW1 2BE, UK; [7] Oxford Centre for Diabetes, Endocrinology and Medicine, University of Oxford, Churchill Hospital, Oxford, OX3 7LJ, UK; [8] Department of Haematology, University of Cambridge, Long Road, Cambridge, CB2 2PT, UK; [9] National Health Service Blood and Transplant, Cambridge Centre, Long Road, Cambridge, CB2 2PT, UK; [10] Department of Cardiovascular Sciences, University of Leicester, Glenfield Hospital, Groby Road, Leicester, LE3 9QP, UK; [11] Department of Statistics, University of Oxford, 1 South Parks Road, Oxford OX1 3TG, UK; [12] Cancer Research UK Genetic Epidemiology Unit, Strangeways Research Laboratory, Worts Causeway, Cambridge CB1 8RN, UK; [13] National Health Service Blood and Transplant, Sheffield Centre, Longley Lane, Sheffield S5 7JN, UK; [14] National Health Service Blood and Transplant, Brentwood Centre, Crescent Drive, Brentwood, CM15 8DP, UK; [15] The Welsh Blood Service, Ely Valley Road, Talbot Green, Pontyclun, CF72 9WB, UK; [16] The Scottish National Blood Transfusion Service, Ellen's Glen Road, Edinburgh, EH17 7QT, UK; [17] National Health Service Blood and Transplant, Southampton Centre, Coxford Road, Southampton, SO16 5AF, UK; [18] Avon Longitudinal Study of Parents and Children, University of Bristol, 24 Tyndall Avenue, Bristol, BS8 1TQ, UK; [19] Division of Community Health Services, St George's University of London, Cranmer Terrace, London SW17 0RE, UK; [20] Institute of Child Health, University College London, 30 Guilford St, London WC1N 1EH, UK; [21] University of Aberdeen, Institute of Medical Sciences, Foresterhill, Aberdeen, AB25 2ZD, UK; [22] Department of Psychiatry, Division of Neuroscience, Birmingham University, Birmingham, B15 2QZ, UK; [23] Department of Psychological Medicine, Henry Wellcome Building, School of Medicine, Cardiff University, Heath Park, Cardiff CF14 4XN, UK; [24] SGDP, The Institute of Psychiatry, King's College London, De Crespigny Park Denmark Hill London SE5 8AF, UK; [25] School of Neurology, Neurobiology and Psychiatry, Royal Victoria Infirmary, Queen Victoria Road, Newcastle upon Tyne, NE1 4LP, UK; [26] LIGHT and LIMM Research Institutes, Faculty of Medicine and Health, University of Leeds, Leeds, LS1 3EX, UK; [27] IBD Research Group, Addenbrooke's Hospital, University of Cambridge, Cambridge, CB2 2QQ, UK; [28] Gastrointestinal Unit, School of Molecular and Clinical Medicine, University of Edinburgh, Western General Hospital, Edinburgh EH4 2XU UK; [29] Department of Medical & Molecular Genetics, King's College London School of Medicine, 8th Floor Guy's Tower, Guy's Hospital, London, SE1 9RT, UK; [30] Institute for Digestive Diseases, University College London Hospitals Trust, London, NW1 2BU, UK; [31] Department of Gastroenterology, Guy's and St Thomas' NHS Foundation Trust, London, SE1 7EH, UK; [32] Department of Gastroenterology & Hepatology, University of Newcastle upon Tyne, Royal Victoria Infirmary, Newcastle upon Tyne, NE1 4LP, UK; [33] Gastroenterology Unit, Radcliffe Infirmary, University of Oxford, Oxford, OX2 6HE, UK; [34] Medicine and Therapeutics, Aberdeen Royal Infirmary, Foresterhill, Aberdeen, Grampian AB9 2ZB, UK; [35] Clinical Pharmacology Unit and the Diabetes and Inflammation Laboratory, University of Cambridge, Addenbrookes Hospital, Hills Road, Cambridge CB2 2QQ, UK; [36] Centre National de Genotypage, 2, Rue Gaston Cremieux, Evry, Paris 91057.; [37] BHF Glasgow Cardiovascular Research Centre, University of Glasgow, 126 University Place, Glasgow, G12 8TA, UK; [38] Clinical Pharmacology and Barts and The London Genome Centre, William Harvey Research Institute, Barts and The London, Queen Mary's School of Medicine, Charterhouse Square, London EC1M 6BQ, UK; [39] Cardiovascular Medicine, University of Oxford, Wellcome Trust Centre for Human Genetics, Roosevelt Drive, Oxford OX3 7BN, UK; [40]arc Epidemiology Research Unit, University of Manchester, Stopford Building, Oxford Rd, Manchester, M13 9PT, UK; [41] Department of Paediatrics, University of Cambridge, Addenbrooke's Hospital, Cambridge, CB2 2QQ, UK; [42] Genetics of Complex Traits, Institute of Biomedical and Clinical Science, Peninsula Medical School, Magdalen Road, Exeter EX1 2LU UK; [43] Diabetes Genetics, Institute of Biomedical and Clinical Science, Peninsula Medical School, Barrack Road, Exeter EX2 5DU UK; [44] Centre for Diabetes and Metabolic Medicine, Barts and The London, Royal London Hospital, Whitechapel, London, E1 1BB UK; [45] Diabetes Research Group, School of Clinical Medical Sciences, Newcastle University, Framlington Place, Newcastle upon Tyne NE2 4HH, UK; [46] The MRC Centre for Causal Analyses in Translational Epidemiology, Bristol University, Canynge Hall, Whiteladies Rd, Bristol BS2 8PR, UK; [47] MRC Laboratories, Fajara, The Gambia; [48] Diamantina Institute for Cancer, Immunology and Metabolic Medicine, Princess Alexandra Hospital, University of Queensland, Woolloongabba, Qld 4102, Australia; [49] Botnar Research Centre, University of Oxford, Headington, Oxford OX3 7BN, UK; [50] Department of Medicine, Division of Medical Sciences, Institute of Biomedical Research, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK; [51] Section of Cancer Genetics, Institute of Cancer Research, 15 Cotswold Road, Sutton, SM2 5NG, UK; [52] Cancer Genome Project, The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus,

Nature Genetics: doi:10.1038/ng.287

Hinxton, Cambridge CB10 1SA, UK; [53] Department of Clinical Neurosciences, University of Cambridge, Addenbrooke's Hospital, Hills Road, Cambridge CB2 2QQ, UK; [54] PRESENT ADDRESS: Illumina Cambridge, Chesterford Research Park, Little Chesterford, Nr Saffron Walden, Essex, CB10 1XL, UK.

6. *COMPETING FINANCIAL INTERESTS:*

Competing financial interests:

Peter Vollenweider and Gérard Waeber received financial support from GlaxoSmithKline to build the CoLaus study; Dawn Waterworth, Kijoung Song, Noha Lim, Vincent Mooser are full-time employees of GlaxoSmithKline (GSK); Inês Barroso and her spouse own stock in the companies GlaxoSmithKline (GSK) and Incyte (INCY).

7. *AUTHOR CONTRIBUTIONS*

Writing Group (drafted and edited manuscript)

Gonçalo R. Abecasis, Inês Barroso, Michael Boehnke, Iris M. Heid, Joel N. Hirschhorn, Shengxu Li, Cecilia M. Lindgren, Ruth J.F. Loos (chair), Mark I. McCarthy, Elizabeth K. Speliotes (lead), Cristen J. Willer

Steering Committee (oversaw the project)

Gonçalo R. Abecasis, Inês Barroso, Michael Boehnke, Panagiotis Deloukas, Timothy M. Frayling, Leif C. Groop, Richard B. Hayes, Joel N. Hirschhorn (chair), David J. Hunter, Ruth J.F. Loos, Mark I. McCarthy, Karen L. Mohlke, David Schlessinger, David P. Strachan, H-Erich Wichmann

Analysis Working Group (decided on and implemented analytic plans for meta-analysis)

Gonçalo R. Abecasis (chair), Sonja I. Berndt, Timothy M. Frayling, Christian Gieger, Iris M. Heid, Joel N. Hirschhorn, Guillaume Lettre, Shengxu Li, Cecilia M. Lindgren, Ruth J.F. Loos, Helen N. Lyon, Lu Qi, Paul Scheet, Elizabeth K. Speliotes, Michael N. Weedon, Eleanor Wheeler, Cristen J. Willer (lead), Jing Hua Zhao

Obesity Working Group (decided on phenotype modeling of BMI for meta-analysis)

Gonçalo R. Abecasis (chair), Iris M. Heid, Joel N. Hirschhorn, Frank B. Hu, Ruth J.F. Loos, Helen N. Lyon (lead), Karen L. Mohlke, Lu Qi, Paul Scheet, Elizabeth K. Speliotes, Michael N. Weedon, Cristen J. Willer, Jing Hua Zhao

Copy Number Polymorphism Analysis

David Altshuler, Amanda L. Elliott (lead), Steve A. McCarroll (chair), Elizabeth K. Speliotes

Expression Analysis

Inês Barroso (chair), Zorica Jovanovic, Rosa Maria Roccasecca (lead), Y.C. Loraine Tung

Oversight of contributing cohorts

CoLaus: Vincent Mooser, Peter Vollenweider

Nature Genetics: doi:10.1038/ng.287

SardiNIA: David Schlessinger, Manuela Uda
DGI: David Altshuler, Leif C. Groop
EPIC-Obesity: Nicholas J. Wareham
FUSION: Richard N. Bergman, Michael Boehnke, Francis S. Collins, Karen L. Mohlke, Jaakko Tuomilehto
NHS: David J. Hunter
PLCO: Stephen J. Chanock, Richard B. Hayes
WTCCC-CAD: Nilesh J. Samani
WTCCC-HT: Mark J. Caulfield
WTCCC-T2D: Timothy M. Frayling, Andrew T. Hattersley, Mark I. McCarthy, Eleftheria Zeggini
KORA: H-Erich Wichmann
WTCCC-NBD: Willem H. Ouwehand, Jonathan Stephens
BC58: David P. Strachan
EPIC-Norfolk: Sheila A. Bingham, Kay-Tee Khaw
FINRISK97: Pekka Jousilahti, Leena Peltonen, Veikko Salomaa
METSIM: Johanna Kuusisto, Markku Laakso
ALSPAC: George Davey Smith, Ken K. Ong
PPP-Botnia: Leif C. Groop, Bo Isomaa, Tiinamaija Tuomi
FUSION (stage 2): Jaakko Tuomilheto
SCOOP: I. Sadaf Farooqi, Stephen O'Rahilly
Hertfordshire: Cyrus Cooper
Rotterdam Study: Albert Hofman, André G. Uitterlinden, Jacqueline C.M. Witteman
NFBC 1966: Marjo-Riitta Jarvelin, Jaana Laitinen, Mark I. McCarthy, Aimo Ruokonen
TwinsUK: Panagiotis Deloukas, Timothy D. Spector
InChianti: Luigi Ferrucci
BLSA: Jack M. Guralnik

Phenotype modeling and genotype-phenotype association analysis in contributing cohorts
CoLaus: Toby Johnson, Noha Lim, Kijoung Song, Gerard Waeber, Dawn M. Waterworth
SardiNIA: Edward G. Lakatta, Serena Sanna, Angelo Scuteri, Paul Scheet
DGI: Peter Almgren, Helen N. Lyon, Martin Ridderstråle, Elizabeth K. Speliotes
EPIC-Obesity: Shengxu Li, Ruth J.F. Loos, Manjinder S. Sandhu, Eleanor Wheeler, Jing Hua Zhao
FUSION: Peter Chines, Laura J. Scott, Timo T. Valle, Richard M. Watanabe, Cristen L. Willer
NHS: Frank B. Hu, Peter Kraft, Lu Qi
PLCO: Sonja I. Berndt, Kevin B. Jacobs
WTCCC-CAD: Alistair S. Hall, Massimo Mangino, Suzanne Stevens
WTCCC-HT: Patricia B. Munroe, Chris Wallace
WTCCC-T2D: Inga Prokopenko, Joshua C. Randall, Michael N. Weedon
KORA: Johannes Hebebrand, Iris M. Heid, Claudia Lamina
WTCCC-NBD: David M. Evans, Nicholas Watkins
BC58: David Hadley, Wendy L. McArdle, Konstantinos Papadakis, Susan M. Ring

Nature Genetics: doi:10.1038/ng.287

EPIC-Norfolk: Shengxu Li, Ruth J.F. Loos, Robert N. Luben, Karani S. Vimaleswaran
FINRISK97: Aki S. Havulinna, Mikko Kuokkannen, Kaisa Silander, Elizabeth K. Speliotes
METSIM: Anne U. Jackson
ALSPAC: Andrew R. Ness, Kate Northstone, Nicholas J. Timpson
PPP-Botnia: Peter Almgren, Bo Isomaa, Elizabeth K. Speliotes
FUSION (stage 2): Narisu Narisu, Jouko Saramies, Heather M. Stringham, Timo T. Valle
SCOOP: Carolin Purmann
Hertfordshire: Elaine M. Dennison, Ruth J.F. Loos
MRC-Ely: Jian'an Luan, Ruth J.F. Loos
Rotterdam Study: Karol Estrada, Leonie C. Jacobs, Fernando Rivadeneira, Cornelia M. Van Duijn, M. Carola Zillikens
NFBC 1966: Amandat Bennett, Lachlan Coin, Paul Elliott, Leena Peltonen
TwinsUK: Nicole Soranzo, Guangju Zhai
InChianti: Toshiko Tanaka
BLSA: Toshiko Tanaka

Expression data: Zorica Jovanovic, Rosa Maria Roccasecca, Y.C. Loraine Tung
CNV Analysis: David Altshuler, Amanda L. Elliott, Steven A. McCarroll

Generating genotype data not published prior to this study
KORA: Thomas Meitinger
EPIC-Norfolk: Christopher J. Gillson, Matthew A. Sims
FINRISK97: Candace Guiducci, Rachel Hackett
METSIM: Lori L. Bonnycastle, Matthew G. Rees
ALSPAC: Susan M. Ring
PPP-Botnia: Lauren Gianniny, Noël Burtt
FUSION (stage 2): Michael R. Erdos, Parimal Deodhar
Hertfordshire: Christopher J. Gillson, Matthew A. Sims
SardiNIA (stage 2): Antonella Mulas
MRC-Ely: Christopher J. Gillson, Matthew A. Sims
Rotterdam Study: Karol Estrada

8. *REFERENCES*

1. Firmann,M. *et al.* The CoLaus study: a population-based study to investigate the epidemiology and genetic determinants of cardiovascular risk factors and metabolic syndrome. *BMC Cardiovascular Disorders* 8, online (2008).
2. Pilia,G. *et al.* Heritability of Cardiovascular and Personality Traits in 6,148 Sardinians. *PLos Genetics* 2, e132 (2006).
3. Willer,C.J. *et al.* Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nat Genet* 40, 161-169 (2008).
4. Li,Y., Willer,C.J., Ding,J., Sheet,P., & Abecasis,G.R. Rapid Markov Chain Haplotyping and Genotype Inference. *Submitted*(2008).
5. Chen,W.M. & Abecasis,G. Family-Based Association Tests for

Nature Genetics: doi:10.1038/ng.287

Genomewide Association Scans. *Am J Hum Genet* 81, 913-926 (2007).
6. Burdick,J.T., Chen,W.M., Abecasis,G.R., & Cheung,V.G. In silico method for inferring genotypes in pedigrees. *Nat Genet* 38, 1002-1004 (2006).
7. Day,N.E. *et al.* EPIC-Norfolk: study design and characteristics of the cohort. European Prospective Investigation of Cancer. *British Journal of Cancer* 80, 95-103 (1999).
8. Loos,R.J. *et al.* Common variants near MC4R are associated with fat mass, weight and risk of obesity. *Nat Genet* 40, 768-775 (2008).
9. Hunter,D.J. *et al.* A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. *Nat Genet* 39, 870-874 (2007).
10. Price,A.L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 38, 904-909 (2006).
11. Prorok,P.C. *et al.* Design of the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial. *Control Clin Trials* 21, 273S-309S (2000).
12. Yeager,M. *et al.* Genome-wide association study of prostate cancer identifies a second risk locus at 8q24. *Nat Genet* 39, 645-649 (2007).
13. Wichmann,H.E., Gieger,C., & Illig,T. KORA-gen - Resource for Population Genetics, Controls and a Broad Spectrum of Disease Phenotypes. *Gesundheitswesen*26-30 (2005).
14. Heid,I.M. *et al.* Association of the 103I MC4R allele with decreased body mass in 7937 participants of two population based surveys. *J. Med. Genet.* 42, e21 (2005).
15. The Wellcome Trust Case Control Consortium Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 447, 661-678 (2007).
16. Power,C. & Elliott,J. Cohort profile: 1958 British birth cohort (National Child Development Study). *Int. J. Epidemiol.* 35, 34-41 (2006).
17. Diabetes Genetics Initiative of Broad Institute of Harvard and MIT Lund University and Novartis Institutes of BioMedical Research *et al.* Genome-Wide Association Analysis Identifies Loci for Type 2 Diabetes and Triglyceride Levels. *Science* 316, 1331-1336 (2007).
18. Groop,L. *et al.* Metabolic consequences of a family history of NIDDM (the Botnia study): evidence for sex-specific parental effects. *Diabetes* 45, 1585-1593 (1996).
19. Scott,L.J. *et al.* A Genome-Wide Association Study of Type 2 Diabetes in Finns Detects Multiple Susceptibility Variants. *Science* 316, 1341-1345 (2007).
20. Valle,T. *et al.* Mapping genes for NIDDM. Design of the Finland-United States Investigation of NIDDM Genetics (FUSION) Study. *Diabetes Care* 21, 949-958 (1998).
21. Caulfield,M. *et al.* Genome-wide mapping of human loci for essential hypertension. *Lancet* 361, 2118-2123 (2003).
22. Marchini,J., Howie,B., Myers,S., McVean,G., & Donnelly,P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet* 39, 906-913 (2007).
23. Mosteller,F. & Bush,R. Selected quantitative techniques in *Handbook of*

Nature Genetics: doi:10.1038/ng.287

Social Psychology - Volume I (ed. Lindzey,G.) 289-334 (Addison-Wesley, Cambridge, MA, 1954).
24. Harding,A.H. *et al.* Dietary Fat and the Risk of Clinical Type 2 Diabetes: The European Prospective Investigation of Cancer-Norfolk Study. *Am. J. Epidemiol.* 159, 73-82 (2004).
25. Vartiainen,E. *et al.* Cardiovascular risk factor changes in Finland, 1972-1997. *Int. J. Epidemiol.* 29, 49-56 (2000).
26. Lyssenko,V. *et al.* Predictors of and Longitudinal Changes in Insulin Sensitivity and Secretion Preceding Onset of Type 2 Diabetes. *Diabetes* 54, 166-174 (2005).
27. Syddall,H.E. *et al.* Cohort Profile: The Hertfordshire Cohort Study. *Int. J. Epidemiol.* 34, 1234-1242 (2005).
28. Williams,D.R. *et al.* Undiagnosed glucose intolerance in the community: the Isle of Ely Diabetes Project. *Diabet. Med.* 12, 30-35 (1995).
29. Wareham,N.J., Byrne,C.D., Williams,R., Day,N.E., & Hales,C.N. Fasting proinsulin concentrations predict the development of type 2 diabetes. *Diabetes Care* 22, 262-270 (1999).
30. Mesa,J.L. *et al.* Lamin A/C polymorphisms, type 2 diabetes and the metabolic syndrome: case-control and quantitative trait studies. *Diabetes* 56, 884-889 (2007).
31. Forouhi,N.G., Luan,J., Hennings,S., & Wareham,N.J. Incidence of type 2 diabetes in England and its association with baseline impaired fasting glucose: The Ely study 1990-2000. *Diabet. Med.*in press (2007).
32. Hofman,A., Grobbee,D.E., de Jong,P.T., & van den Ouweland,F.A. Determinants of disease and disability in the elderly: the Rotterdam Elderly Study. *Eur. J. Epidemiol.* 7, 403-422 (1991).
33. Hofman,A. *et al.* The Rotterdam Study: objectives and design update. *Eur. J. Epidemiol.* 22, 819-829 (2007).
34. Andrew,T. *et al.* Are twins and singletons comparable? A study of disease-related and lifestyle characteristics in adult women. *Twin Res.* 4, 464-477 (2001).
35. Ferucci,L. *et al.* Subsystems contributing to the decline in ability to walk: bridging the gap between epidemiology and geriatric practice in the InCHIANTI study. *J. Am. Geriatr. Soc.* 48, 1618-1625 (2000).
36. Bartalli,B. *et al.* Changes in anthropometric measures in men and women across the life-span: findings from the InCHIANTI study. *Soz. Praventivmed.* 47, 336-348 (2002).
37. Shock,R.C., Greulich,R.C., & Andres,R.A. Normal Human Aging: The Baltimore Longitudinal Study of Aging. NIH publication no. 84-2450, 45. 1984. Washington D.C., U.S. Government Printing Office.
38. Abecasis,G.R., Cherny,S.S., Cookson,W.O., & Cardon,L.R. Merlin—rapid analysis of dense genetic maps using sparse gene flow trees. *Nature Genet.* 30, 97-101 (2002).
39. Higgins,J.P.T., Thompson,S.G., Deeks,J.J., & Altman,D.G. Measuring inconsistency in meta-analyses. *BMJ* 327, 557-560 (2003).
40. Sutton,A., Abrams,K.R., Jones,D.R., Sheldon,T.A., & Song,F. *Methods for meta-analysis in medical research*(John Wiley & Sons, Chichester,

Nature Genetics: doi:10.1038/ng.287

2000).
41. Golding, Pembrey, Jones, & The Alspac Study Team ALSPAC-The Avon Longitudinal Study of Parents and Children. *Paediatr. Perinat. Epidemiol.* 15, 74-87 (2001).
42. Jones,R.W. *et al.* A new human genetic resource: a DNA bank established as part of the Avon longitudinal study of pregnancy and childhood (ALSPAC). *Eur J Hum Genet* 8, 653-660 (2000).
43. Leary,S.D. *et al.* Smoking during Pregnancy and Offspring Fat and Lean Mass in Childhood. *Obesity Res* 14, 2284-2293 (2006).
44. Rogers,I.S. *et al.* Associations of size at birth and dual-energy X-ray absorptiometry measures of lean and fat mass at 9 to 10 y of age. *Am. J. Clin. Nutr.* 84, 739-747 (2006).
45. Cole,T.J., Freeman,J.V., & Preece,M.A. Body mass index reference curves for the UK, 1990. *Arch. Dis. Child.* 73, 25-29 (1995).

Nature Genetics: doi:10.1038/ng.287