nature genetics

Genome-wide meta-analysis identifies 11 new loci for anthropometric traits and provides insights into genetic architecture

Approaches exploiting trait distribution extremes may be used to identify loci associated with common traits, but it is unknown whether these loci are generalizable to the broader population. In a genome-wide search for loci associated with the upper versus the lower 5th percentiles of body mass index, height and waist-to-hip ratio, as well as clinical classes of obesity, including up to 263,407 individuals of European ancestry, we identified 4 new loci (*IGFBP4*, *H6PD*, *RSRC1* and *PPP2R2A*) influencing height detected in the distribution tails and 7 new loci (*HNF4G*, *RPTOR*, *GNAT2*, *MRPS33P4*, *ADCY9*, *HS6ST3* and *ZZZ3*) for clinical classes of obesity. Further, we find a large overlap in genetic structure and the distribution of variants between traits based on extremes and the general population and little etiological heterogeneity between obesity subgroups.

Twin studies have established a strong heritable component to body mass index (BMI; h^2 of ~40–70%)[1,2] and height (h^2 of ~70–90%)[3]. Previous meta-analyses of genome-wide association studies (GWAS) have identified 36 genetic loci associated with BMI[4–6], 14 loci associated with waist-to-hip ratio adjusted for BMI (WHR), reflecting fat distribution[7,8], and 180 loci associated with height[9] and contributed to our understanding of the genetic architecture of complex traits. However, established loci for complex traits only account for a small proportion of trait heritability, as discussed recently[10,11]. Some postulated explanations for this include undiscovered low-frequency variants with larger effects, imperfect tagging of causal variants, epistasis, gene-environment interaction and phenotype heterogeneity. This has led to increasing interest in approaches exploiting extremes of the trait distribution where there may be less locus heterogeneity, greater genetic contribution and enrichment for highly penetrant variants. The use of population extremes has also been proposed to improve cost-efficiency, as effect sizes may be larger, fewer subjects may be needed for genotyping and a smaller proportion of the variance may be attributable to environmental factors. Indeed, several previous studies have used population extreme designs to discover loci for various complex traits, such as obesity and lipid fractions, using microarray genotyping[12–16] or sequencing methods[17–20]. However, the few previous studies that have systematically addressed differences between the genetic architecture of the overall distribution with that of distribution extremes for complex traits have been small[21–23], and, hence, it remains largely unknown whether genetic loci identified as associated with the extremes of a trait can be extended to the general population.

Studies of extremely obese individuals have reported 13 loci at or near genome-wide significance ($P < 5 \times 10^{-7}$)[14–16,22–26], but not all have shown evidence of association with BMI in the general population[4,27]. For example, variants in *PCSK1* (rs6232) and *PTER* have been convincingly associated with severe obesity[14,25] but have at best shown nominal evidence of association with BMI in large-scale meta-analyses[4,28]. Although it is possible that other genetic or environmental factors modify the manifestations of these variants, producing an extreme phenotype only in selected individuals, it is also conceivable that the trait extremes are, at least in part, etiologically distinct and different from those acting in the general population. Within the extremes of the trait distribution, there may be etiologically discrete subgroups or enrichment for less common causal variants[19]. Although analyzing the full distribution is generally more powerful, in cases where there is heterogeneity, analyzing extremes by case-control design may offer superior power[29].

The extremes for anthropometric traits, particularly BMI, have been defined in numerous ways, including using the tails of the full population distribution (for example, >95th or >97th percentile) and absolute cutoffs (for example, ≥40 kg/m^2) based on clinical or standard references, and some studies have used a combination of definitions for their discovery and replication analyses. The common denominator for studies addressing trait extremes (herein used as a more generic term) is that they dichotomize the trait distribution and analyzed data using a case-control design. Studies suggest that the percentile cutoff choice and ascertainment strategy used may affect the observed risk and subsequent power[30,31]; however, the consequences of the definitions of trait extremes on the discovery and characterization of loci for complex traits have not been systematically evaluated. In the present study, we have used the terms 'distribution tails' to describe analyses comparing the upper and lower 5th percentiles of the trait distributions; 'clinical classes of obesity' to describe analyses where controls were subjects with BMI < 25 kg/m^2 and cases were defined as having BMI ≥ 25 kg/m^2 for the overweight class, BMI ≥ 30 kg/m^2 for obesity class I, BMI ≥ 35 kg/m^2 for obesity class II and BMI ≥ 40 kg/m^2 for obesity class III (ref. 32); and 'extremely obese' to describe studies using different sampling designs for selecting their extremely obese cases and controls.

The overall aim of the present study was to use and compare different distribution cutoffs for the identification of genetic loci for

A full list of authors and affiliations appears at the end of the paper.

Received 16 October 2012; accepted 14 March 2013; published online 7 April 2013; doi:10.1038/ng.2606

anthropometric traits. The two specific aims were (i) to systematically compare findings using these cutoffs with those from the full population distribution, as well as with those from studies using a different ascertainment strategy, and (ii) to draw inferences about the value of these different approaches for sampling within a population-based study. Our focus was primarily on BMI, which is a major risk factor for multiple chronic diseases and of important public health relevance[33], but we also examined height and WHR adjusted for BMI (as a measure of body fat distribution) to determine whether our findings could be generalized to other traits. To address these aims, we performed a genome-wide search for genetic determinants of the distribution tails (defined as the upper versus lower 5th percentiles of the trait distribution) of BMI, height and WHR and, for comparison, of clinical classes of obesity drawn from populations within the Genetic Investigation of ANthropometric Traits (GIANT) Consortium. Association analyses were conducted in a study base (or sampling frame) of up to 168,267 individuals with follow-up of the 273 most significantly associated loci in a study base of up to 109,703 additional individuals. Further, systematic comparisons were conducted to assess differences in genetic inheritance and the distribution of risk variants between the population extremes and the general population for these anthropometric traits.

RESULTS

To first evaluate the contribution of common SNPs to the anthropometric trait distribution tails and clinical classes of obesity and discover new loci, we conducted meta-analyses of GWAS of six obesity-related traits (distribution tails of BMI and WHR, overweight class and obesity classes I–III), as well as the distribution tails of height, using results for ~2.8 million genotyped or imputed SNPs. Stage 1 analyses included 51 studies with study bases of 158,864 (BMI), 168,267 (height) and 100,605 (WHR) individuals of European ancestry (see **Supplementary Table 1** for the numbers of cases and controls per phenotype and **Supplementary Tables 2–5** for study characteristics). We observed an enrichment of SNPs with small association *P* values compared to the null distribution for all seven traits (quantile-quantile plots; **Supplementary Figs. 1** and **2**). The excess was diminished after the exclusion of loci previously established for the overall distributions or population extremes of these traits, but some enrichment remained, especially for the distribution tails of height and, to a lesser extent, for the overweight class and obesity classes I and II. In total, 69 loci (defined as separated by at least 1 Mb) were associated at $P < 5 \times 10^{-8}$ with at least 1 trait (**Supplementary Figs. 3** and **4**).

To identify and validate loci for these traits, SNPs for which associations reached $P < 5 \times 10^{-6}$ in the stage 1 analyses were taken forward for follow-up (stage 2) in 12 studies with *in silico* GWAS data and 24 studies with Metabochip data with study bases of 109,703 (BMI), 107,740 (height) and 75,220 (WHR) individuals (**Supplementary Tables 1–5**).

BMI-related traits

Seventeen SNPs were taken forward to stage 2 in up to 4,900 and 4,891 individuals from the upper and lower distribution tails of BMI, respectively. Ten SNPs reached genome-wide significance ($P < 5 \times 10^{-8}$) in the joint meta-analysis of stage 1 and stage 2, but all had been previously identified as loci associated with BMI in the general population[4]. A total of 118 SNPs were included in stage 2 for clinical classes of obesity, which included up to 1,162 cases and 22,307 controls for obesity class III and 65,332 cases and 39,294 controls for the overweight class. Of the 62 SNPs that showed association $P < 5 \times 10^{-8}$ in the joint meta-analyses for at least 1 obesity class (**Supplementary Table 6**), 7 were new, explaining an additional 0.09% of the variability in BMI (**Supplementary Table 7**). These included one locus for the overweight class (*RPTOR*), three loci for obesity class I (*GNAT2*, *MRPS33P4* and *ADCY9*), two loci for obesity class II (*HS6ST3* and *ZZZ3*) and one locus associated with both the overweight class and obesity class I (*HNF4G*) (**Table 1** and **Supplementary Figs. 5–7**). Although these loci were identified for specific clinical classes of obesity, all newly associated loci showed consistent effect direction across the distribution tails of BMI and the other classes of obesity, and most *P* values were significant ($P < 0.007$, Bonferroni corrected for seven SNPs), except for those for obesity class III and the distribution tails of BMI (presumably owing to lower statistical power for these traits; **Table 2**).

Of the new obesity loci, at least four are located near genes of high biological relevance. In particular, rs7503807 for the overweight class is located within the *RPTOR* gene (encoding regulatory-associated protein of the MTOR, complex 1), which regulates cell growth in response to nutrient and insulin levels[34], and within 500 kb of the *BAIAP2* gene (encoding BAI1-associated protein 2), a brain-specific angiogenesis inhibitor (BAI1)-binding protein that regulates insulin uptake in the central nervous system. The rs4735692 SNP associated with the overweight class and obesity class I is located downstream of the *HNF4G* gene (encoding hepatocyte nuclear factor 4γ). Mutations in *HNF4A*, a closely related gene encoding a factor that forms a heterodimer with HNF4G to activate gene transcription[35], cause maturity-onset diabetes of the young type 1 (ref. 36), and a common variant near *HNF4A* was found to be associated with type 2 diabetes (T2D) in east Asians[37]. The rs2531995 SNP associated with obesity class I is located within *ADCY9* (encoding adenylate cyclase 9), which catalyzes the formation of cyclic AMP from ATP. This SNP was found to be associated with *ADCY9* expression in several tissue types (**Supplementary Table 8**). Loci near other adenylate cyclase genes have been associated with several T2D-related traits, such as glucose homeostasis and susceptibility to T2D (*ADCY5*)[38,39]. The rs17024258 SNP associated with obesity class II is located 207 kb away from the lipid-related gene *SORT1* (encoding sortilin), which is expressed in multiple cell types and has been reported to be involved in insulin responsiveness in adipose cells[40]. Lower amounts of sortilin have been observed in the adipose tissues of morbidly obese humans and mice and in the skeletal muscle of obese mice[41]. A more comprehensive summary of the biological relevance of the genes nearest to all newly associated loci is given in the **Supplementary Note**.

Distribution tails of height

A total of 134 SNPs from stage 1 were taken forward to stage 2 in up to 4,872 and 4,831 individuals from the upper and lower distribution tails of height, respectively. Of the 95 SNPs that reached association $P < 5 \times 10^{-8}$ in the joint meta-analysis of stage 1 and stage 2 (**Supplementary Table 6**), 4 new loci (*IGFBP4*, *H6PD*, *RSRC1* and *PPP2R2A*) were identified for the distribution tails of height (**Table 1** and **Supplementary Fig. 8**). The contribution of the four new loci to overall height variability was ≤0.02% (**Supplementary Table 7**).

Two of the new loci are located near genes that seem particularly relevant to height. rs584438 is located approximately 500 bp upstream of *IGFBP4*, which codes for insulin-like growth factor–binding protein 4 (IGFBP4), and is in linkage disequilibrium (LD, $r^2 = 0.87$) with another SNP (rs598892) that results in a synonymous amino-acid change in IGFBP4. IGFBP4 binds to IGF1 and IGF2 (ref. 42), which have an important role in childhood growth. In blood, this same SNP showed a significant association with the expression of *TNS4* (**Supplementary Table 8**), which encodes a factor that interacts

Table 1 New loci reaching genome-wide significance ($P < 5 \times 10^{-8}$) for the tails of anthropometric traits and clinical classes of obesity

SNP	Chr.	Position	Nearby gene	Effect allele	Other allele	Effect allele freq.	Stage 1				Stage 2[a]				Stage 1 + stage 2			
							Cases (*n*)	Controls (*n*)	OR	*P*	Cases (*n*)	Controls (*n*)	OR	*P*	Cases (*n*)	Controls (*n*)	OR	*P*
Height tails																		
rs584438	17	35852698	*IGFBP4*	C	A	0.617	7,830	7,850	1.18	1.11×10^{-9}	1,814	1,814	1.19	0.001	9,644	9,664	1.18	5.22×10^{-12}
rs6662509	1	9240191	*H6PD*	T	C	0.1463	5,462	5,461	1.23	2.21×10^{-6}	3,615	3,566	1.23	3.37×10^{-5}	9,077	9,027	1.23	3.19×10^{-10}
rs2362965	3	159592073	*RSRC1-SHOX2*	T	A	0.50	7,989	7,993	1.14	1.45×10^{-7}	4,819	4,775	1.10	0.002	12,808	12,768	1.12	2.14×10^{-9}
rs1594829	8	26261994	*PPP2R2A*	C	T	0.7688	6,693	6,697	1.18	5.51×10^{-7}	4,166	4,115	1.11	0.01	10,859	10,812	1.15	3.88×10^{-8}
Obesity class II																		
rs7989336	13	95815549	*HS6ST3*	A	G	0.4704	9,825	62,114	1.12	5.88×10^{-9}	1,664	17,113	1.04	0.25	11,489	79,226	1.10	1.06×10^{-8}
rs17381664	1	77820919	*ZZZ3*	C	T	0.3923	9,833	62,114	1.11	7.61×10^{-8}	5,351	33,841	1.05	0.04	15,184	95,955	1.09	2.85×10^{-8}
Obesity class I																		
rs17024258	1	109948844	*GNAT2*	T	C	0.0364	18,662	38,427	1.23	1.41×10^{-6}	8,956	15,471	1.28	1.12×10^{-6}	27,618	53,898	1.25	8.66×10^{-12}
rs4735692	8	76778218	*HNF4G*	A	G	0.5834	32,675	65,697	1.07	5.03×10^{-8}	22,086	38,352	1.04	0.005	54,761	104,049	1.06	2.48×10^{-9}
rs13041126	20	50526403	*MRPS33P4*	T	C	0.7179	32,020	64,015	1.07	3.05×10^{-7}	22,088	37,595	1.04	0.007	54,108	101,610	1.06	2.16×10^{-8}
rs2531995	16	3953468	*ADCY9*	T	C	0.6146	32,433	65,542	1.06	3.17×10^{-6}	6,680	16,602	1.07	0.004	39,113	82,144	1.07	4.04×10^{-8}
Overweight class																		
rs4735692	8	76778218	*HNF4G*	A	G	0.5839	92,703	65,698	1.05	6.13×10^{-9}	65,323	39,290	1.03	0.003	158,026	104,988	1.04	3.51×10^{-10}
rs7503807	17	76205706	*RPTOR*	A	C	0.5654	92,855	65,723	1.04	4.20×10^{-6}	64,535	38,813	1.03	0.0009	157,390	104,536	1.04	1.98×10^{-8}

Chr., chromosome; freq., frequency; OR, odds ratio.
[a]Stage 2 consists of studies with either GWAS or Metabochip data. Not all SNPs were present on the Metabochip.

with β-catenin[43], a critical component of the canonical Wnt pathway related to bone formation[44]. The height-associated SNP rs2362965 lies 285 kb away from *SHOX2*, a homolog to the X-linked, pseudoautosomal *SHOX* (short stature homeobox) gene family, which has a major role in skeletal limb development.

Distribution tails of WHR

Ten SNPs were taken forward to stage 2 in 3,351 and 3,352 individuals from the upper and lower distribution tails of WHR, respectively. The four SNPs that reached genome-wide significance ($P < 5 \times 10^{-8}$; **Supplementary Table 6**) have previously been identified as WHR-associated loci in the general population[7].

Effects of new loci in the full distribution and previously identified loci in distribution tails and obesity classes

We assessed the impact of our newly associated loci on the full distribution of these anthropometric traits using data from studies included in stage 1 and stage 2. In the full distribution, evidence of association ($P < 0.005$, Bonferroni corrected for 11 SNPs) with consistent effect direction was observed with BMI for all new obesity-related trait loci and with height for all new loci identified for the distribution tails of height (**Table 2**). None of the loci were associated with WHR, suggesting that these obesity loci are primarily associated with overall adiposity rather than with fat distribution.

In the GIANT Consortium, we previously identified 32 loci associated with BMI[4]. There is considerable overlap of samples with the current study, so it is not unexpected that we observed that the effects of all established BMI loci were directionally consistent between the previous study of overall BMI and the present study of obesity-related traits (**Supplementary Table 9**). Twenty-seven of 32 SNPs were significantly associated with the distribution tails of BMI ($P < 0.0016$, Bonferroni corrected). Although only half of the SNPs were significantly associated with obesity class III, presumably owing to smaller sample size and reduced power, the majority of SNPs were significantly associated with obesity class II, and all were associated with obesity class I and the overweight class.

Effects of our new loci in other studies of extreme obesity

Both empirical[16] and theoretical[29] work has shown that genetic architecture may differ the more extreme the selection (the further out in the distribution), suggesting that the ascertainment strategy may affect the observed results[31]. To evaluate the impact of the ascertainment strategy, we also performed analyses of all SNPs we found to be associated with BMI-related traits in five studies that applied other ascertainment strategies to define the extremely obese class (**Supplementary Tables 2–5**, bottom; $n_{cases} = 6{,}848$, $n_{controls} = 7{,}023$). Four studies recruited participants from specialized clinics or hospitals on the basis of absolute or percentile-derived cutoffs, and one study used liability-based (women) and standard-based (men) percentile cutoffs. We performed a meta-analysis of these five studies and observed directionally consistent associations for all BMI-associated SNPs between studies (**Supplementary Table 10**). The effect sizes in these extreme obesity studies were similar to those observed for the distribution tails of BMI in our analysis (heterogeneity *P* value (P_{het}) > 0.007 for all SNPs, Bonferroni corrected). Four out of seven new obesity-related loci showed significance at $P < 0.007$ (Bonferroni corrected) in these studies of extremely obese individuals.

Effects of known extreme obesity loci in our study

Previous studies of extreme childhood and/or adult obesity using different ascertainment strategies have reported genome-wide significant or near genome-wide significant associations ($P < 5 \times 10^{-7}$)

Table 2 Association results for new SNPs associated with height- and obesity-related traits at genome-wide significance ($P < 5 \times 10^{-8}$)

SNP	Gene	Effect allele	Other allele	BMI tails		Obesity class III		Obesity class II		Obesity class I		Overweight class		BMI (continuous)[a]		Height tails		Height (continuous)[a]	
				OR	*P*	OR	*P*	OR	*P*	OR	*P*	OR	*P*	Effect	*P*	OR	*P*	Effect	*P*
Height tails																			
rs584438	*IGFBP4*	C	A	0.98	0.52	1.02	0.64	1.01	0.47	1.00	0.75	1.00	0.59	0.005	0.22	1.18	5.22×10^{-12}	0.025	9.43×10^{-11}
rs6662509	*H6PD*	T	C	1.00	0.95	1.11	0.07	1.01	0.83	0.99	0.34	0.99	0.35	−0.006	0.27	1.23	3.19×10^{-10}	0.031	7.76×10^{-12}
rs2362965	*RSRC1-SHOX2*	T	A	0.95	0.02	0.97	0.25	0.98	0.20	0.99	0.37	0.99	0.21	−0.007	0.05	1.12	2.14×10^{-9}	0.017	7.07×10^{-8}
rs1594829	*PPP2R2A*	C	T	1.03	0.33	1.06	0.11	1.01	0.58	1.00	0.82	1.01	0.32	0.004	0.33	1.15	3.88×10^{-8}	0.016	4.29×10^{-5}
Obesity class II																			
rs7989336	*HS6ST3*	A	G	1.09	0.0001	1.11	0.0006	1.10	1.06×10^{-8}	1.04	9.38×10^{-5}	1.04	2.33×10^{-6}	0.016	8.80×10^{-6}	1.00	0.89	−0.001	0.71
rs17381664	*ZZZ3*	C	T	1.08	0.0001	1.12	5.41×10^{-5}	1.09	2.85×10^{-8}	1.05	6.80×10^{-8}	1.04	2.23×10^{-7}	0.022	2.50×10^{-11}	1.06	0.005	0.010	0.004
Obesity class I																			
rs17024258	*GNAT2*	T	C	1.27	0.02	1.45	0.002	1.26	7.73×10^{-5}	1.25	8.66×10^{-12}	1.13	1.41×10^{-8}	0.067	4.34×10^{-14}	1.21	0.08	0.010	0.36
rs4735692	*HNF4G*	A	G	1.09	1.97×10^{-5}	1.08	0.006	1.05	0.0005	1.06	2.48×10^{-9}	1.04	3.51×10^{-10}	0.019	9.94×10^{-10}	1.02	0.50	0.006	0.10
rs13041126	*MRPS33P4*	T	C	1.08	0.001	1.05	0.16	1.06	0.0002	1.06	2.16×10^{-8}	1.04	1.43×10^{-6}	0.017	8.52×10^{-7}	1.02	0.56	0.002	0.65
rs2531995	*ADCY9*	T	C	1.06	0.01	1.09	0.006	1.06	0.001	1.07	4.04×10^{-8}	1.03	5.57×10^{-5}	0.021	6.58×10^{-8}	1.07	0.005	0.018	1.87×10^{-6}
Overweight class																			
rs4735692	*HNF4G*	A	G	1.09	1.97×10^{-5}	1.08	0.006	1.05	0.0005	1.06	2.48×10^{-9}	1.04	3.51×10^{-10}	0.019	9.94×10^{-10}	1.02	0.50	0.006	0.10
rs7503807	*RPTOR*	A	C	1.08	7.07×10^{-5}	1.13	9.44×10^{-6}	1.07	2.46×10^{-6}	1.05	1.12×10^{-7}	1.04	1.98×10^{-8}	0.020	3.00×10^{-10}	0.99	0.55	−0.001	0.85

[a]The β value represents the difference in standardized effects.

with *FTO*, *MC4R*, *TMEM18*, *FAIM2*, *TNKS*, *HOXB5*, *OLFM4*, *NPC1*, *MAF*, *PTER*, *SDCCAG8*, *PCSK1* (rs6235 and rs6232) and *KCNMA1* (refs. 14–16,22–26). With the exception of *PCSK1* (rs6232) for the distribution tails of BMI and *MAF* for the distribution tails of BMI and obesity class II, all associations showed consistent directions of effect across the BMI-related outcomes (**Supplementary Table 11**). Of the 13 loci, replication at a significance level of $P < 0.004$ (Bonferroni corrected) was observed for 4 SNPs (*FTO*, *MC4R*, *TMEM18* and *FAIM2*) for the distribution tails of BMI and all clinical classes of obesity. Two loci, *MAF* and *KCNMA1*, which have thus far only been reported for extreme obesity, were not significantly associated with any of our traits at either a Bonferroni-corrected or nominal significance threshold ($P < 0.05$).

Empirical power comparison of the population extremes and the full distribution

If the trait extremes have different genetic inheritance or are etiologically more homogenous than the full distribution, analyzing extremes or tails of the distribution by case-control design may offer superior power. To test this empirically, we conducted meta-analyses of the full distributions of BMI and height with all studies included in stage 1 and stage 2. Only two loci (*IGFBP4* and *H6PD*) out of the four new loci for the distribution tails of height reached genome-wide significance ($P < 5 \times 10^{-8}$) using the full height distribution (**Table 2**). Four loci (*GNAT2*, *ZZZ3*, *HNF4G* and *RPTOR*) out of the seven new loci identified for the clinical classes of obesity achieved genome-wide significance for the full BMI distribution. The remaining loci had *P* values of $<5 \times 10^{-5}$ in the full distribution and, thus, would likely have been detected with a larger sample size.

Genetic architecture in the distribution tails and full distribution

To investigate differences in genetic architecture between the distribution tails and the full distributions, we estimated whether the observed genetic effects in the distribution tails of BMI, height and WHR were different from what would be expected based on the full distributions of the corresponding traits. To do this, we first estimated the expected effect for each SNP in the distribution tails on the basis of the full distribution in each study and then carried out meta-analysis of the expected associations across studies. The quantile-quantile plots of *P* values testing differences between the observed and expected effects (**Fig. 1** and **Supplementary Fig. 9**) did not show any enrichment,

Figure 1 Quantile-quantile plot of the $-\log_{10}$ *P* values for the difference between the observed association for the distribution tails of BMI and the expected association based on the overall BMI distribution.

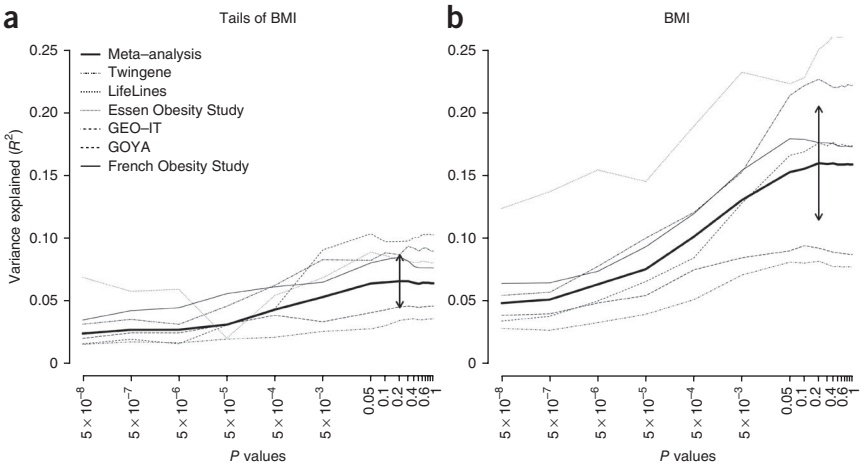

Figure 2 Variance in extreme obesity explained by common genetic variants. The phenotypic variance explained is higher when SNPs with lower degrees of significance are included in the polygenetic prediction model. The *y* axis represents the proportion of variance explained (Nagelkerke R^2) for extreme obesity in six studies not included in the discovery meta-analysis. The thicker lines represent the weighted average; 95% confidence intervals are reported as double-headed arrows. (**a**) The prediction model was based on the results from the stage 1 meta-analysis of the distribution tails of BMI. (**b**) The prediction model was based on BMI from the full distribution (modified version of the previous GIANT meta-analysis by Speliotes *et al.*[4]). The Essen Obesity Study was not adjusted by age.

indicating that the effect sizes observed in distribution tails and those expected based on the overall distribution were similar. Further, comparable results were observed for the 32 SNPs previously associated with BMI in Speliotes *et al.*[4], as well as for previously published and new loci for extreme obesity (**Supplementary Table 12**).

To further compare genetic inheritance in the distribution tails with that in the full distribution, we used a polygenic approach[45]. The meta-analysis results of the distribution tails and full distribution were used to create two polygenetic scores (by summing the number of risk-associated alleles at each SNP) in six studies (**Supplementary Table 13**). We found that the polygenic score based on the full BMI distribution consistently explained more of the variance than the score based on the distribution tails (for example, 15.3% versus 6.4% at $P < 0.05$) (**Fig. 2** and **Supplementary Table 14**). Similar results were observed for height and WHR (**Supplementary Fig. 10**). On liability scale, the variance explained by the two polygenic scores was similar for different BMI-related outcomes (**Supplementary Fig. 11**) and different percentile cutoffs used to define the distribution tails (data not shown), suggesting that the fraction of the overall variance explained by SNPs is not influenced by outcome categorization but by the ability to accurately rank and estimate the β coefficients of the association, which is better achieved by using the entire study population instead of the distribution tails. Our results also indicate that genetic determinants for the distribution tails are similar to those for the full distribution and that common variant loci contribute to extreme phenotypes. However, it should be noted that our analyses of the upper and lower 5th percentiles of the distribution (tails) does not necessarily extend to more extreme cutoffs, such as the top and bottom 1st percentiles.

Allelic heterogeneity at new and previously identified loci

To explore enrichment for allelic heterogeneity in the distribution tails and clinical classes of obesity, we performed conditional analyses using a recently described method[46]. In these analyses, we found secondary signals that reached genome-wide significance ($P < 5 \times 10^{-8}$) at 17 loci, including 1 locus for the distribution tails of BMI (*FTO*), 13 loci for the distribution tails of height (*PTCH1* (2 signals), *GHSR*, *EDEM2*, *C6orf106*, *CRADD*, *EFEMP1*, *HHIP*, *FBXW11*, *NPR3*, *LINC00471* (also known as *C2orf52*), *BCKDHB* and *EFR3B*), 1 locus for the distribution tails of WHR (*RSPO3*), 2 loci for the overweight class (*MC4R* and *FANCL*) and 1 locus for obesity class I (*FANCL*) (**Supplementary Table 15**). Whereas the secondary signals for the distribution tails of BMI (*FTO*) and WHR (*RSPO3*) and the overweight class and obesity class I (*FANCL*) have not been established previously, all 13 height-related loci identified here, as well as the *MC4R* locus, have previously been shown to have allelic heterogeneity in the general population[7,9], suggesting that there is no enrichment in the distribution tails for secondary signals (**Supplementary Figs. 12–14**).

We also looked for evidence of enrichment of unobserved low-frequency variants by conducting haplotype analyses within known and new loci, as haplotypes constructed from common SNPs may tag low-frequency variants that are enriched in the tails of the trait distributions but are rarer in the general population. Using genotype data from the largest studies, three signals of association were observed for the distribution tails of height that exceeded conservative prior odds of association of 1 in 30,000: *ID4* (Bayes factor of 118,839), *LIN28B* (Bayes factor of 105,478) and *DLEU7* (Bayes factor of 66,599) (**Supplementary Table 16**). However, for all three loci, association signals were characterized by two clusters of haplotypes (both common and rare) and were not consistent with enrichment of unobserved low-frequency causal variants in the distribution tails.

DISCUSSION

In our meta-analysis of GWAS of up to 263,407 individuals of European ancestry, we identified 165 loci associated with distribution tails (the upper versus lower 5th percentiles) of BMI, height and WHR and/or clinical classes of obesity. Eleven of these loci have not previously been associated with anthropometric traits. Several of the new loci were located near strong biological candidate genes, such as *IGFBP4* and *SHOX2* for the distribution tails of height and *HNF4G* and *ADCY9* for the overweight class and/or obesity class I, suggesting future areas of research. Although by using different distribution cutoffs we discovered additional loci that would not have been identified as genome-wide significant using the full distribution of the same study samples, there is no evidence to suggest that the clinical classes of obesity are etiologically distinct, and the majority of evidence indicates that the population extremes share many of the same loci with the general population.

To assess the impact of different distribution cutoffs on genetic variants associated with the population extremes, we chose to evaluate the 5% tails of trait distribution and clinical classes of obesity, specifically obesity classes II and III. Although others have ascertained population extremes differently, all variants associated with obesity-related traits in our meta-analysis were found to have directionally consistent results in five independent studies of extremely

obese samples. Of the 13 loci previously identified as associated with extreme obesity[14–16,22–26], nearly all (except *PCSK1* (rs6232) and *MAF*) showed a consistent direction of effect in the distribution tails of BMI. Only two loci (*MAF* and *KCNMA1*), originally identified for early-onset and morbid adult obesity[14,26], did not replicate for any of our BMI-related outcomes. Although it is possible that we had insufficient power if there was a substantial winner's curse present in the initial publications, it is also conceivable that these susceptibility loci are population specific, only contribute to risk at younger ages[47], represent false positive findings or tag rare causal variants that are difficult to detect in population-based samples.

Because our study was based on GWAS data, we were not well suited to address the role of rare variants in extreme traits. Although haplotype-based analyses identified strong associations of haplotypes in three genes with the distribution tails of height, which could suggest that they are tagged by rare variants, such putative variants could not be established using our approach. The suggestion that rare variants could be more important in the distribution extremes of complex traits needs to be addressed using other study designs, such as resequencing projects or using the new Exome Chip microarrays that are currently being analyzed in many large study samples.

Our systematic comparisons between distribution extremes and the full distribution yielded several important insights that also may be informative for other complex traits. When comparing observed genetic effects in distribution tails with the expected effects extrapolated from the overall distributions of the corresponding traits, we did not observe any systematic differences. Further, we showed that the polygenic score based on the full distribution explained a larger proportion of variance than the score based on the distribution tails. Taken together with the finding that half of our new loci were associated at a genome-wide significant level in the overall distribution, this implies that there is limited etiological heterogeneity in these anthropometric traits. Our analysis shows that, whereas some common variants can have larger effects in the distribution extremes, these effects as a whole are not larger than expected based on their effects in the overall distribution. Further, whereas rare variants specific to the distribution extremes may still exist, the extremes share most of the common loci with the overall distribution.

Conclusions that can be drawn from these observations are that, when access is available to data for the full distribution, case-control analyses using population extremes can be useful to find additional loci. Although analyzing the full distribution is generally more powerful, small amounts of heterogeneity in the distribution may allow for the identification of additional loci by analyzing the data using different cutoffs, such as the distribution tails. Further, as in most cases when resources are limited, our results indicate that a strategy with the selection of individuals from the population extremes for genetic analyses could be a cost-effective approach and will likely yield loci that are relevant and can largely be extended to the general population. Compatible with the findings from recent, smaller studies[21–23], our results show that this theoretically appealing approach also holds empirically.

In conclusion, in our large GWAS meta-analysis including up to 263,407 individuals, we identified 4 new loci influencing height detected at the distribution tails, as well as 7 new loci for clinical classes of obesity. Consistent with theoretical predictions and previous smaller studies, our results show that there is a large overlap in terms of genetic structure and the distribution of variants between traits based on different distribution cutoffs with those from population-level studies, but additional insight may still be gained from evaluating the population extremes. Our results are informative for designing future genetic studies of obesity as well as other complex traits.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Supplementary information is available in the online version of the paper.

ACKNOWLEDGMENTS

A full list of acknowledgments appears in the **Supplementary Note**. Funding was provided by the Aarno Koskelo Foundation; the Academy of Finland; the Agency for Science, Technology and Research of Singapore; the Australian National Health and Medical Research Council; the Australian Research Council; BDA Research; the BioSHaRE Consortium; the British Heart Foundation; the Cedars-Sinai Board of Governors' Chair in Medical Genetics; the Centre for Clinical Research at the University of Leipzig; the Centre of Excellence in Genomics and the University of Tartu; the Chief Scientist Office of the Scottish government; the City of Kuopio and the Social Insurance Institution of Finland; the Department of Educational Assistance, the University and Research of the Autonomous Province of Bolzano; the Donald W. Reynolds Foundation; the Dutch Ministry for Health, Welfare and Sports; the Dutch Ministry of Education, Culture and Science; Dutch BBRMI-NL; the Dutch Brain Foundation; the Dutch Centre for Medical Systems Biology; the Dutch Diabetes Research Foundation; the Dutch Government Economic Structure–Enhancing Fund; the Dutch Inter-University Cardiology Institute; the Dutch Kidney Foundation; the Dutch Ministry of Economic Affairs; the Dutch Ministry of Justice; the Dutch Research Institute for Diseases in the Elderly; Eleanor Nichols endowments; the Emil Aaltonen Foundation; Erasmus Medical Center and Erasmus University; the Estonian government; the European Commission; the European Regional Development Fund; the European Research Council; the European Science Foundation; the Faculty of Biology and Medicine of Lausanne; Finland's Slot Machine Association; the Finnish Cultural Foundation; the Finnish Diabetes Research Foundation; the Finnish Foundation for Cardiovascular Research; the Finnish Funding Agency for Technology and Innovation; the Finnish Heart Association; the Finnish Medical Society; the Finnish Ministry of Education and Culture; the Finnish Ministry of Health and Social Affairs; the Finnish National Institute for Health and Welfare; the Finnish Social Insurance Institution; Finska Läkaresällskapet; the Folkhälsan Research Foundation; the Foundation for Life and Health in Finland; the French Ministry of Research; the French National Research Agency; the Genetic Association Information Network; the German Diabetes Association; the German Federal Ministry of Education and Research; the German Ministry of Cultural Affairs; the German National Genome Research Network; the German Research Foundation; GlaxoSmithKline; the Göteborg Medical Society; the Greek General Secretary of Research and Technology; the Gyllenberg Foundation; Health Care Centers in Vasa, Närpes and Korsholm; the Heinz Nixdorf Foundation; Helmholtz Zentrum München–German Research Center for Environmental Health; the Icelandic Heart Association; the Icelandic Parliament; the Intramural Research Program of the Division of Cancer Epidemiology and Genetics, the National Cancer Institute, NIH; Italian Ministry of Education, Universities and Research; Italian Ministry of Health; Juho Vainio Foundation; Juvenile Diabetes Research Foundation International; the Knut and Alice Wallenberg Foundation; Kuopio, Tampere and Turku University Hospital Medical Funds; the Leducq Foundation; the Lundberg Foundation; the March of Dimes; the Munich Center of Health Sciences as part of LMUinnovativ; the Municipal Health Care Center and Hospital in Jakobstad; the Municipality of Rotterdam; the Närpes Health Care Foundation; National Alliance for Research on Schizophrenia and Depression Young Investigator Awards; the Netherlands Genomics Initiative; the Netherlands Organization for Health Research and Development; UK NHSBT; US National Institutes of Health; the Nordic Center of Cardiovascular Research; the Nordic Center of Excellence in Disease Genetics; the Nordic Centre of Excellence on Systems Biology in Controlled Dietary Interventions and Cohort Studies; the Northern Netherlands Collaboration of Provinces; the Novo Nordisk Foundation; the Ollqvist Foundation; the Orion-Farmos Research Foundation; the Paavo Nurmi Foundation; the Päivikki and Sakari Sohlberg Foundation; the Perklen Foundation; the Petrus and Augusta Hedlunds Foundation; the Province of Groningen; the Republic of Croatia Ministry of Science, Education and Sport; the Reynold's Foundation; the Royal Society; Samfundet Folkhälsan; the Signe and Ane Gyllenberg Foundation; the Sigrid Juselius Foundation; the Social Ministry of the Federal State of Mecklenburg–West Pomerania; the Sophia Foundation for Medical Research; the South Tyrolean Sparkasse Foundation; the Southern California Diabetes Endocrinology Research Center; the Stockholm County Council; the Strategic Cardiovascular Program of Karolinska Institutet; Strategic Support for Epidemiological Research at Karolinska Institutet; the Susan G. Komen Breast Cancer Foundation; the Swedish Ministry for Higher Education; the Swedish Cancer Society; the Swedish Cultural Foundation in Finland; the Swedish Diabetes Association; the Swedish Foundation for Strategic Research; the Swedish Heart-Lung Foundation; the Swedish Medical Research Council; the Swedish Ministry of

Education; the Swedish Research Council; the Swedish Royal Academy of Science; the Swedish Society for Medical Research; the Swedish Society of Medicine; the Swiss National Science Foundation; the Tampere Tuberculosis Foundation; The Great Wine Estates of the Margaret River Region of Western Australia; The Paul Michael Donovan Charitable Foundation; the Torsten and Ragnar Söderberg Foundation; Cancer Research UK; the UK Diabetes Association; the UK Heart Foundation; the UK MRC; the UK NIHR, Biomedical Research Centre; UK West Anglia Primary and Community Care; the University Medical Center Groningen and the University of Groningen; the Västra Götaland Foundation; VU University: the Institute for Health and Care Research and the Neuroscience Campus Amsterdam; the Wellcome Trust; and the Yrjö Jahnsson Foundation.

AUTHOR CONTRIBUTIONS

Steering committee (oversaw the consortium): G.R.A., T.L.A., I.B., S.I.B., M. Boehnke, I.B.B., P.D., C.S.F., T. Frayling, L.C.G., T.H., I.M.H., D.H., E.I., R.C.K., R.J.F.L., M.I.M., K.L. Mohlke, K.E.N., J.R.O., D. Schlessinger, D.P.S., U.T. and C.M.v.D. Writing group (drafted and edited the manuscript): S.I.B., M.F.F., A. Ganna, S.G., E.I., A.E.J., C.M.L., R.J.F.L., R.M., M.I.M., D. Meyre, K.L. Monda, A.P.M., K.E.N., A. Scherag, E.K.S., E. Wheeler and C.J.W. Data cleaning and preparation: S.I.B., D.C.C.-C., F.R.D., T.E., T. Fall, T. Ferreira, S.G., I.M.H., E.I., A.U.J., C.M.L., J. Luan, R.M., J.C.R., A. Scherag, E.K.S., G.T., S.V., T.W.W. and A.R.W. Statistical advisors: S.H.L., B.M.N., Y.P., P.M.V., J.Y., D.-Y.L. and Y.-J.H. Gene expression (eQTL) analyses: L. Liang, W.O.C., M.F.M., G.R.A., V. Steinthorsdottir, G.T., J.L.M., G. Nicholson, F. Karpe., M.I.M. and E.E.S.

Project design, management and coordination of contributing studies

Stage 1—GWAS: (ADVANCE) T.L.A. and C.I.; (AGES) V.G., T.B.H. and L.J.L.; (ARIC) E.B. and K.E.N.; (B58C) D.P.S.; (BRIGHT Study) M.J.C. and P.B.M.; (CAPS) E.I.; (CHS) B.M. and B.M.P.; (CoLaus) V.M., P.V. and G. Waeber; (COROGENE) M.S.N. and J. Sinisalo; (deCODE) K. Stefansson and U.T.; (DGI) L.C.G. and J.N.H.; (EGCUT) A.M.; (EPIC) K.-T.K., R.J.F.L. and N.J.W.; (ERF) B.A.O. and C.M.v.D.; (FamHS) I.B.B. and M.A.P.; (Fenland) R.J.F.L. and N.J.W.; (FRAM) L.A.C. and C.S.F.; (FUSION GWAS) M. Boehnke and K.L. Mohlke; (Genmets) A.J., S.R. and V. Salomaa; (GerMIFS1) J.E. and H. Schunkert; (GerMIFS2) C. Hengstenberg and K. Stark; (GOOD) C.O.; (HBCS) J.G.E.; (KORA S3) H.-E.W.; (KORA S4) C.G., T.I., W.K. and A. Peters; (MGS) P.V.G. and D.F.L.; (MICROS (SOUTH TYROL)) P.P.P.; (MIGEN) J.N.H. and S. Kathiresan; (NESDA) B.P.; (NFBC 1966) M.-R.J.; (NHS) L.Q.; (Nijmegen Biomedical Study) L.A.K.; (NSPHS) U.G.; (NTR) D.I.B.; (ORCADES) J.F.W. and A.F.W.; (PLCO) S.J.C. and S.I.B.; (PROCARDIS) M.F. and H. Watkins; (RS-I) F.R. and A.G.U.; (RUNMC) L.A.K.; (SardiNIA) D. Schlessinger; (SASBAC) E.I.; (SHIP) H. Wallaschofski; (Sorbs) M.S. and A. Tönjes; (TwinsUK) T.D.S.; (VIS) I.R.; (WGHS) P.M.R.; (WTCC-T2D) M.I.M.; (WTCCC-CAD) A.J.B., A.S.H. and N.J.S.; and (YFS) M. Kähönen, T.L., O.R. and J. Viikari. Stage 2—Metabochip and *in silico* replication: (AMC-PAS) K.G.H.; (B58C) C. Power; (BHS) L.J.P.; (DILGOM) K. Kuulasmaa and V. Salomaa; (DPS) M.U.; (DR's EXTRA) T.A.L. and R.R.; (EPIC, Fenland and Ely) C. Langenberg, R.J.F.L. and N.J.W.; (FIN-D2D 2007) S.M.K.-K. and T.E.S.; (GLACIER) P.W.F.; (Go-DARTS (Dundee)) A.D.M. and C.N.A.P.; (HNR) K.-H.J.; (HUNT 2) K.H.; (Hypergenes) D.C.; (IMPROVE) U.d.F., A. Hamsten and E.T.; (KORA S3) I.M.H.; (LifeLines Cohort Study) H. Snieder, M.M.V.d.K. and B.H.R.W.; (LURIC) B.O.B., W.M. and B.R.W.; (METSIM) J.K. and M. Laakso; (MORGAM) P.A., P.B., M.M.F., J.F., F. Kee, D.-A.T. and J. Virtamo; (NSHD) D.K.; (PIVUS) E.I.; (PLCO2) S.I.B. and S.J.C.; (PREVEND) P.v.d.H.; (QIMR) N.G.M., G.W.M., A.C.H. and P.M.; (RS-II) A. Hofman and J.B.J.v.M.; (RS-III) C.M.v.D. and J.C.M.W.; (Swedish Twin Registry) E.I.; (THISEAS/AMCPAS/CARDIOGENICS) P.D.; (THISEAS) G.V.D.; (TRAILS) A.J.O.; (Tromsø 4) I.N.; (TWINGENE) E.I.; (UKBS2) W.H.O.; (ULSAM) E.I.; (Whitehall II) A. Hingorani and M. Kivimäki; and (WTCC-T2D) M.I.M. and C.M.L. Other contributing studies, clinical extremes: (French Extreme Obesity Study) D. Meyre and P.F.; (GEO-IT) A.M.D.B.; (Essen Obesity Study, Essen Case-Control & Essen Obesity Trio GWAS) J.H. and A. Hinney; and (GOYA) T.I.A.S. and E.A.N.

Genotyping of contributing studies

Stage 1—GWAS: (ADVANCE) D.A.; (ARIC) E.B.; (B58C) W.L.M.; (CAPS) H. Grönberg; (CHS) T.H.; (CoLaus) V.M.; (COROGENE) M. Perola; (EGCUT) T.E. and L.M.; (EPIC) I.B.; (ERF) B.A.O. and C.M.v.D.; (FamHS) I.B.B., M.A.P. and A.T.K.; (Fenland) J. Luan; (Genmets) S.R.; (GOOD) C.O., J.-O.J. and M. Lorentzon; (HBCS) A. Palotie and E. Widén; (MGS) P.V.G. and A.R.S.; (MICROS (SOUTH TYROL)) A.A.H.; (NHS) F.B.H. and D.H.; (NSPHS) Å.J.; (NTR and NESDA) J.-J.H.; (ORCADES) J.F.W.; (PLCO) S.J.C. and K.B.J.; (RS-I) F.R., A.G.U., K.E. and C.M.-G.; (SardiNIA) M. Dei; (SASBAC) P.H. and J. Liu; (SHIP) G.H.; (TwinsUK) M.M., S.-Y.S. and N.S.; (VIS) C. Hayward and V.V.; (WGHS) D.I.C.; (WTCC-T2D) M.I.M.; (WTCCC-CAD) A.J.B., A.S.H. and N.J.S.; and (YFS) T.L. Stage 2—Metabochip and *in silico* replication: (BHS) L.J.P. and J.B.; (CARDIOGENICS) S.E. and S.E.H.; (DPS) A.J.S.; (DR's EXTRA) M.A.; (EPIC, Fenland and Ely) J. Luan and K.K.O.; (FIN-D2D 2007) P.S.C.; (FUSION) F.S.C., J. Saramies and J. Tuomilehto; (GLACIER) I.B. and S.E.; (Go-DARTS (Dundee)) C.N.A.P.; (HNR) T.W.M.; (HUNT 2) N.N.; (Hypergenes) F.F.; (KORA S3) H. Grallert; (KORA S4) T.I.; (LifeLines Cohort Study) H. Snieder, B.H.R.W., M. Bruinenberg and L.F.; (NSHD) D.K., K.K.O. and A.W.; (PIVUS) E.I. and L. Lind; (PLCO2) S.J.C., K.B.J. and Z.W.; (PREVEND) P.v.d.H. and F.W.A.; (QIMR) N.G.M., G.W.M., A.C.H. and P.A.M.; (RS-II) M.J.P. and M. Dei; (Swedish Twin Registry) E.I., P.K.M. and N.P.; (THISEAS/AMCPAS/CARDIOGENICS) K. Stirrups; (TRAILS) A.J.O., I.M.N. and J.V.V.V.-O.; (Tromsø 4) L.L.B.; (TWINGENE) E.I., A. Hamsten and N.P.; (ULSAM) E.I.; (Whitehall II) C. Langenberg; and (WTCC-T2D) M.I.M. Other contributing studies, clinical extremes: (French Extreme Obesity Study) D. Meyre and P.F.; (GEO-IT) D.G.; and (GOYA) L.P. and D.M.E.

Phenotyping of contributing studies

Stage 1—GWAS: (ARIC) E.B.; (B58C) D.P.S.; (BRIGHT Study) J.M.C.; (CAPS) H. Grönberg; (CHS) B.M.P.; (CoLaus) P.V. and G. Waeber; (COROGENE) J. Sinisalo, M.-L.L.; (EGCUT) A.M. and K.F.; (EPIC) R.J.F.L.; (ERF) B.A.O. and C.M.v.D.; (FamHS) I.B.B., M.A.P. and M.F.F.; (Fenland) R.J.F.L.; (FRAM) C.S.F.; (Genmets) A.J. and V. Salomaa; (GerMIFS1) S. Schreiber; (GerMIFS2) A. Peters; (GOOD) C.O., J.-O.J., M. Lorentzon and L.V.; (MGS) P.V.G., A.R.S. and D.F.L.; (NFBC 1966) A.-L.H., J.H.L. and A. Pouta; (NHS) L.Q.; (Nijmegen Biomedical Study) F.d.V., M.d.H. and S.H.V.; (NSPHS) Å.J. and U.G.; (NTR and NESDA) G. Willemsen; (ORCADES) H.C. and S.H.W.; (PLCO) S.I.B.; (RS-I) F.R. and A.G.U.; (SASBAC) P.H.; (SHIP) S. Schipf; (Sorbs) A. Tönjes; (TwinsUK) M.M. and T.D.S.; (VIS) O.P.; (WTCC-T2D) M.I.M.; (WTCCC-CAD) A.J.B., A.S.H. and N.J.S.; and (YFS) M. Kähönen, O.R. and J. Viikari. Stage 2—Metabochip and *in silico* replication: (AMC-PAS) H.B. and M.D.T.; (B58C) C. Power and E.H.; (BHS) L.J.P., J.B. and A.W.M.; (DPS) J. Lindström; (EPIC, Fenland and Ely) R.J.F.L.; (GLACIER) P.W.F. and D. Shungin; (Go-DARTS (Dundee)) C.N.A.P. and A.D.M.; (Hypergenes) D.C. and P.M.; (KORA S3) B.T.; (KORA S4) A. Peters; (LifeLines Cohort Study) B.H.R.W. and M.M.V.d.K.; (METSIM) A. Stančáková and P.V.G.; (NSHD) D.K.; (PIVUS) E.I. and L. Lind; (PLCO2) S.I.B.; (PREVEND) G. Navis; (QIMR) N.G.M., A.C.H. and P.M.; (RS-II) M.C.Z.; (RS-III) J.C.M.W.; (Swedish Twin Registry) E.I., P.K.M. and N.P.; (THISEAS) M. Dimitriou and E.V.T.; (TRAILS) R.P.S.; (TWINGENE) E.I. and N.P.; (UKBS2) A. Rendon; (ULSAM) E.I.; (Whitehall II) M. Kumari; and (WTCC-T2D) M.I.M. Other contributing studies, clinical extremes: (Essen Obesity Study, Essen Case-Control GWAS & Essen Obesity Trio GWAS) J.H. and A. Hinney; (GEO-IT) A.L. and S. Signorini; and (GOYA) T.I.A.S. and E.A.N.

Analyses of contributing studies

Stage 1—GWAS: (ADVANCE) L.L.W.; (AGES) A.V.S.; (ARIC) K.E.N., A.E.J. and K.L. Monda; (B58C) D.P.S.; (BRIGHT Study) T.J.; (CAPS) E.I. and R.M.; (CHS) B.M. and G.L.; (CoLaus) D. Marek; (COROGENE) M. Perola; (deCODE) V. Steinthorsdottir and G.T.; (DGI) E.K.S. and S.V.; (EGCUT) K.F., T.E. and E.M.; (EPIC) J.H.Z.; (ERF) N.A.; (FamHS) M.F.F.; (Fenland) J. Luan; (FRAM) L.A.C., N.L.H.-C. and J.S.N.; (Genmets) I.S.; (GerMIFS1) M. Preuss; (GerMIFS2) I.R.K.; (GOOD) C.O., J.-O.J., M. Lorentzon and L.V.; (HBCS) N.E.; (KORA S3) C. Lamina; (KORA S4) E.A.; (MGS) D.F.L. and J. Shi; (MICROS (SOUTH TYROL)) J.E.H. and Å.J.; (MIGEN) E.K.S. and S.V.; (NFBC 1966) A. Pouta, R.M. and J.C.R.; (NHS) L.Q. and T.W.; (NSPHS) Å.J.; (NTR and NESDA) J.- J.H.; (ORCADES) Å.J.; (PLCO) K.B.J. and S.I.B.; (PROCARDIS) M.F., A. Goel and J.F.P.; (RS-I) F.R., K.E. and C.M.-G.; (SardiNIA) J.L.B.-G. and S. Sanna; (SASBAC) E.I. and R.M.; (SEARCH) J. Tyrer; (SHIP) A. Teumer; (Sorbs) R.M. and I.P.; (TwinsUK) M.M. and N.S.; (VIS) Å.J.; (WGHS) D.I.C.; (WTCC-T2D) C.M.L., R.M. and J.C.R.; (WTCCC-CAD) R.M. and J.C.R.; (WTCCC-NBS (UKBS-CC)) A.P.A., R.M., J.C.R., J.G.S. and J.C.S.; and (YFS) O.R. and T.L. Stage 2—Metabochip and *in silico* replication: (B58C) E.H. and T. Ferreira; (BHS) G.C.; (DILGOM) K. Kristiansson and K. Kuulasmaa; (DPS) A.U.J.; (DR's EXTRA) A.U.J.; (EPIC, Fenland and Ely) J. Luan and K.K.O.; (FIN-D2D 2007) A.U.J.; (FUSION) A.U.J.; (GLACIER) P.W.F. and D. Shungin; (HNR) S.P. and C. Pütter; (HUNT 2) A.U.J.; (Hypergenes) F.F. and Z.K.; (IMPROVE) R.J.S.; (KORA S3) I.M.H. and T.W.W.; (KORA S4) M.M.-N.; (LifeLines Cohort Study) M. Bruinenberg and L.F.; (LURIC) M.E.K.; (METSIM) A.U.J.; (NSHD) A.W. and J. Luan; (PIVUS) E.I., S.G.; (PLCO2) S.I.B., Z.W.; (PREVEND) P.v.d.H. and I.M.L.; (QIMR) S.E.M., J.Y.; (RS-II) M.J.P.; (Swedish Twin Registry) E.I. and S.G.; (THISEAS/AMCPAS/CARDIOGENICS) S. Kanoni; (TRAILS) I.M.N. and J.V.V.V.-O.; (Tromsø 4) A.U.J.; (TWINGENE) E.I. and S.G.; (UKBS2) A. Radhakrishnan; (ULSAM) E.I., A. Ganna and S.G.; (WGHS) L.M.R.; and (WTCC-T2D) R.M. and T. Ferreira. Other contributing studies, clinical extremes: (French Extreme Obesity Study) D. Meyre, C. Lecoeur and B.S.; (GEO-IT) A.M.D.B. and D.G.; (Essen Obesity Study, Essen Case-Control GWAS and Essen Obesity Trio GWAS) A. Scherag and I.J.; and (GOYA) L.P. and D.M.E.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available in the online version of the paper.

1. Maes, H.H., Neale, M.C. & Eaves, L.J. Genetic and environmental factors in relative body weight and human adiposity. *Behav. Genet.* **27**, 325–351 (1997).
2. Stunkard, A.J., Foch, T.T. & Hrubec, Z. A twin study of human obesity. *J. Am. Med. Assoc.* **256**, 51–54 (1986).
3. Silventoinen, K. *et al.* Heritability of adult body height: a comparative study of twin cohorts in eight countries. *Twin Res.* **6**, 399–408 (2003).
4. Speliotes, E.K. *et al.* Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat. Genet.* **42**, 937–948 (2010).
5. Okada, Y. *et al.* Common variants at *CDKAL1* and *KLF9* are associated with body mass index in east Asian populations. *Nat. Genet.* **44**, 302–306 (2012).
6. Wen, W. *et al.* Meta-analysis identifies common variants associated with body mass index in east Asians. *Nat. Genet.* **44**, 307–311 (2012).
7. Heid, I.M. *et al.* Meta-analysis identifies 13 new loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. *Nat. Genet.* **42**, 949–960 (2010).
8. Lindgren, C.M. *et al.* Genome-wide association scan meta-analysis identifies three loci influencing adiposity and fat distribution. *PLoS Genet.* **5**, e1000508 (2009).
9. Lango Allen, H. *et al.* Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* **467**, 832–838 (2010).
10. Lee, S.H., Wray, N.R., Goddard, M.E. & Visscher, P.M. Estimating missing heritability for disease from genome-wide association studies. *Am. J. Hum. Genet.* **88**, 294–305 (2011).
11. Zuk, O., Hechter, E., Sunyaev, S.R. & Lander, E.S. The mystery of missing heritability: genetic interactions create phantom heritability. *Proc. Natl. Acad. Sci. USA* **109**, 1193–1198 (2012).
12. Duncan, E.L. *et al.* Genome-wide association study using extreme truncate selection identifies novel genes affecting bone mineral density and fracture risk. *PLoS Genet.* **7**, e1001372 (2011).
13. Edmondson, A.C. *et al.* Dense genotyping of candidate gene loci identifies variants associated with high-density lipoprotein cholesterol. *Circ. Cardiovasc. Genet.* **4**, 145–155 (2011).
14. Meyre, D. *et al.* Genome-wide association study for early-onset and morbid adult obesity identifies three new risk loci in European populations. *Nat. Genet.* **41**, 157–159 (2009).
15. Scherag, A. *et al.* Two new loci for body-weight regulation identified in a joint analysis of genome-wide association studies for early-onset extreme obesity in French and German study groups. *PLoS Genet.* **6**, e1000916 (2010).
16. Bradfield, J.P. *et al.* A genome-wide association meta-analysis identifies new childhood obesity loci. *Nat. Genet.* **44**, 526–531 (2012).
17. Cohen, J.C. *et al.* Multiple rare alleles contribute to low plasma levels of HDL cholesterol. *Science* **305**, 869–872 (2004).
18. Emond, M.J. *et al.* Exome sequencing of extreme phenotypes identifies *DCTN4* as a modifier of chronic *Pseudomonas aeruginosa* infection in cystic fibrosis. *Nat. Genet.* **44**, 886–889 (2012).
19. Harismendy, O. *et al.* Population sequencing of two endocannabinoid metabolic genes identifies rare and common regulatory variants associated with extreme obesity and metabolite level. *Genome Biol.* **11**, R118 (2010).
20. Romeo, S. *et al.* Population-based resequencing of *ANGPTL4* uncovers variations that reduce triglycerides and increase HDL. *Nat. Genet.* **39**, 513–516 (2007).
21. Chan, Y. *et al.* Common variants show predicted polygenic effects on height in the tails of the distribution, except in extremely short individuals. *PLoS Genet.* **7**, e1002439 (2011).
22. Cotsapas, C. *et al.* Common body mass index–associated variants confer risk of extreme obesity. *Hum. Mol. Genet.* **18**, 3502–3507 (2009).
23. Paternoster, L. *et al.* Genome-wide population-based association study of extremely overweight young adults—the GOYA study. *PLoS ONE* **6**, e24303 (2011).
24. Hinney, A. *et al.* Genome wide association (GWA) study for early onset extreme obesity supports the role of fat mass and obesity associated gene (*FTO*) variants. *PLoS ONE* **2**, e1361 (2007).
25. Benzinou, M. *et al.* Common nonsynonymous variants in *PCSK1* confer risk of obesity. *Nat. Genet.* **40**, 943–945 (2008).
26. Jiao, H. *et al.* Genome wide association study identifies *KCNMA1* contributing to human obesity. *BMC Med. Genomics* **4**, 51 (2011).
27. den Hoed, M. *et al.* Evaluation of common genetic variants identified by GWAS for early onset and morbid obesity in population-based samples. *Int. J. Obes. (Lond).* **37**, 191–196 (2013).
28. Willer, C.J. *et al.* Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. *Nat. Genet.* **41**, 25–34 (2009).
29. Pütter, C. *et al.* Missing heritability in the tails of quantitative traits? A simulation study on the impact of slightly altered true genetic models. *Hum. Hered.* **72**, 173–181 (2011).
30. Williams, P.T. Quantile-specific penetrance of genes affecting lipoproteins, adiposity and height. *PLoS ONE* **7**, e28764 (2012).
31. Guey, L.T. *et al.* Power in the phenotypic extremes: a simulation study of power in discovery and replication of rare variants. *Genet. Epidemiol.* Published online doi:10.1002/gepi.20572 (9 February 2011).
32. World Health Organization. Obesity: preventing and managing the global epidemic. Report of a WHO Consultation. in *WHO Technical Report Series 8949* (World Health Organization, Geneva, 2000).
33. Kumanyika, S.K. *et al.* Population-based prevention of obesity: the need for comprehensive promotion of healthful eating, physical activity, and energy balance: a scientific statement from American Heart Association Council on Epidemiology and Prevention, Interdisciplinary Committee for Prevention (formerly the expert panel on population and prevention science). *Circulation* **118**, 428–464 (2008).
34. Sarbassov, D.D. & Sabatini, D.M. Redox regulation of the nutrient-sensitive raptor-mTOR pathway and complex. *J. Biol. Chem.* **280**, 39505–39509 (2005).
35. Daigo, K. *et al.* Proteomic analysis of native hepatocyte nuclear factor-4α (HNF4α) isoforms, phosphorylation status, and interactive cofactors. *J. Biol. Chem.* **286**, 674–686 (2011).
36. Nakajima, H. *et al.* Hepatocyte nuclear factor-4α gene mutations in Japanese non-insulin dependent diabetes mellitus (NIDDM) patients. *Res. Commun. Mol. Pathol. Pharmacol.* **94**, 327–330 (1996).
37. Cho, Y.S. *et al.* Meta-analysis of genome-wide association studies identifies eight new loci for type 2 diabetes in east Asians. *Nat. Genet.* **44**, 67–72 (2012).
38. Dupuis, J. *et al.* New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat. Genet.* **42**, 105–116 (2010).
39. Saxena, R. *et al.* Genetic variation in *GIPR* influences the glucose and insulin responses to an oral glucose challenge. *Nat. Genet.* **42**, 142–148 (2010).
40. Shi, J. & Kandror, K.V. Sortilin is essential and sufficient for the formation of Glut4 storage vesicles in 3T3-L1 adipocytes. *Dev. Cell* **9**, 99–108 (2005).
41. Kaddai, V. *et al.* Involvement of TNF-α in abnormal adipocyte and muscle sortilin expression in obese mice and humans. *Diabetologia* **52**, 932–940 (2009).
42. Zhang, M. *et al.* Paracrine overexpression of IGFBP-4 in osteoblasts of transgenic mice decreases bone turnover and causes global growth retardation. *J. Bone Miner. Res.* **18**, 836–843 (2003).
43. Liao, Y.C., Chen, N.T., Shih, Y.P., Dong, Y. & Lo, S.H. Up-regulation of C-terminal tensin-like molecule promotes the tumorigenicity of colon cancer through β-catenin. *Cancer Res.* **69**, 4563–4566 (2009).
44. Milat, F. & Ng, K.W. Is Wnt signalling the final common pathway leading to bone formation? *Mol. Cell Endocrinol.* **310**, 52–62 (2009).
45. Purcell, S.M. *et al.* Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* **460**, 748–752 (2009).
46. Yang, J. *et al.* Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat. Genet.* **44**, 369–375 (2012).
47. Kilpeläinen, T.O., Bingham, S.A., Khaw, K.T., Wareham, N.J. & Loos, R.J. Association of variants in the *PCSK1* gene with obesity in the EPIC-Norfolk study. *Hum. Mol. Genet.* **18**, 3496–3501 (2009).

Sonja I Berndt[1,244], Stefan Gustafsson[2,3,244], Reedik Mägi[4,5,244], Andrea Ganna[3,244], Eleanor Wheeler[6], Mary F Feitosa[7], Anne E Justice[8], Keri L Monda[8,9], Damien C Croteau-Chonka[10], Felix R Day[11], Tõnu Esko[5,12], Tove Fall[3], Teresa Ferreira[4], Davide Gentilini[13], Anne U Jackson[14], Jian'an Luan[11], Joshua C Randall[4,6], Sailaja Vedantam[15–17], Cristen J Willer[18–20], Thomas W Winkler[21], Andrew R Wood[22], Tsegaselassie Workalemahu[23,24], Yi-Juan Hu[25], Sang Hong Lee[26], Liming Liang[27,28], Dan-Yu Lin[29], Josine L Min[4], Benjamin M Neale[30], Gudmar Thorleifsson[31], Jian Yang[32,33], Eva Albrecht[34], Najaf Amin[35], Jennifer L Bragg-Gresham[14], Gemma Cadby[36–38], Martin den Heijer[39], Niina Eklund[40], Krista Fischer[5], Anuj Goel[41], Jouke-Jan Hottenga[42], Jennifer E Huffman[43], Ivonne Jarick[44], Åsa Johansson[45,46], Toby Johnson[47,48], Stavroula Kanoni[6], Marcus E Kleber[49,50], Inke R König[51], Kati Kristiansson[40], Zoltán Kutalik[52,53], Claudia Lamina[54], Cecile Lecoeur[55,56], Guo Li[57], Massimo Mangino[58], Wendy L McArdle[59],

Carolina Medina-Gomez[35,60,61], Martina Müller-Nurasyid[34,62–64], Julius S Ngwa[65], Ilja M Nolte[66], Lavinia Paternoster[67], Sonali Pechlivanis[68], Markus Perola[5,40,69], Marjolein J Peters[35,60,61], Michael Preuss[51,70], Lynda M Rose[71], Jianxin Shi[1], Dmitry Shungin[72–74], Albert Vernon Smith[75,76], Rona J Strawbridge[77], Ida Surakka[40,69], Alexander Teumer[78], Mieke D Trip[79,80], Jonathan Tyrer[81], Jana V Van Vliet-Ostaptchouk[82,83], Liesbeth Vandenput[84], Lindsay L Waite[85], Jing Hua Zhao[11], Devin Absher[85], Folkert W Asselbergs[86], Mustafa Atalay[87], Antony P Attwood[88], Anthony J Balmforth[89], Hanneke Basart[79], John Beilby[90,91], Lori L Bonnycastle[92], Paolo Brambilla[93], Marcel Bruinenberg[83], Harry Campbell[94], Daniel I Chasman[71,95], Peter S Chines[92], Francis S Collins[92], John M Connell[96,97], William O Cookson[98], Ulf de Faire[99], Femmie de Vegt[100], Mariano Dei[101], Maria Dimitriou[102], Sarah Edkins[6], Karol Estrada[35,60,61], David M Evans[67], Martin Farrall[41], Marco M Ferrario[103], Jean Ferrières[104], Lude Franke[83,105], Francesca Frau[106], Pablo V Gejman[107,108], Harald Grallert[109], Henrik Grönberg[3], Vilmundur Gudnason[75,76], Alistair S Hall[110], Per Hall[3], Anna-Liisa Hartikainen[111], Caroline Hayward[43], Nancy L Heard-Costa[112], Andrew C Heath[113], Johannes Hebebrand[114], Georg Homuth[78], Frank B Hu[23], Sarah E Hunt[6], Elina Hyppönen[115], Carlos Iribarren[116], Kevin B Jacobs[1,117], John-Olov Jansson[118], Antti Jula[119], Mika Kähönen[120], Sekar Kathiresan[16,121–123], Frank Kee[124], Kay-Tee Khaw[125], Mika Kivimäki[126], Wolfgang Koenig[127], Aldi T Kraja[7], Meena Kumari[126], Kari Kuulasmaa[128], Johanna Kuusisto[129], Jaana H Laitinen[130], Timo A Lakka[87,131], Claudia Langenberg[11,126], Lenore J Launer[132], Lars Lind[133], Jaana Lindström[134], Jianjun Liu[135], Antonio Liuzzi[136], Marja-Liisa Lokki[137], Mattias Lorentzon[84], Pamela A Madden[113], Patrik K Magnusson[3], Paolo Manunta[138], Diana Marek[52,53], Winfried März[50,139], Irene Mateo Leach[140], Barbara McKnight[141], Sarah E Medland[33], Evelin Mihailov[5,12], Lili Milani[5], Grant W Montgomery[33], Vincent Mooser[142], Thomas W Mühleisen[143,144], Patricia B Munroe[47,48], Arthur W Musk[145–147], Narisu Narisu[92], Gerjan Navis[148], George Nicholson[149,150], Ellen A Nohr[151], Ken K Ong[11,152], Ben A Oostra[61,153,154], Colin N A Palmer[155], Aarno Palotie[6,69], John F Peden[156], Nancy Pedersen[3], Annette Peters[109,157,158], Ozren Polasek[159], Anneli Pouta[111,160], Peter P Pramstaller[161–163], Inga Prokopenko[4,164], Carolin Pütter[68], Aparna Radhakrishnan[6,165,166], Olli Raitakari[167,168], Augusto Rendon[88,165,166,169], Fernando Rivadeneira[35,60,61], Igor Rudan[94], Timo E Saaristo[170,171], Jennifer G Sambrook[165,166], Alan R Sanders[107,108], Serena Sanna[101], Jouko Saramies[172], Sabine Schipf[173], Stefan Schreiber[174], Heribert Schunkert[175], So-Youn Shin[6], Stefano Signorini[176], Juha Sinisalo[177], Boris Skrobek[55,56], Nicole Soranzo[6,58], Alena Stančáková[178], Klaus Stark[179], Jonathan C Stephens[165,166], Kathleen Stirrups[6], Ronald P Stolk[66,83], Michael Stumvoll[180,181], Amy J Swift[92], Eirini V Theodoraki[102], Barbara Thorand[157], David-Alexandre Tregouet[182], Elena Tremoli[183], Melanie M Van der Klauw[82,83], Joyce B J van Meurs[35,60,61], Sita H Vermeulen[100,184], Jorma Viikari[185], Jarmo Virtamo[128], Veronique Vitart[43], Gérard Waeber[186], Zhaoming Wang[1,117], Elisabeth Widén[69], Sarah H Wild[94], Gonneke Willemsen[42], Bernhard R Winkelmann[187], Jacqueline C M Witteman[35,61], Bruce H R Wolffenbuttel[82,83], Andrew Wong[152], Alan F Wright[43], M Carola Zillikens[60,61], Philippe Amouyel[188], Bernhard O Boehm[189], Eric Boerwinkle[190,191], Dorret I Boomsma[42], Mark J Caulfield[47,48], Stephen J Chanock[1], L Adrienne Cupples[65], Daniele Cusi[106,192], George V Dedoussis[102], Jeanette Erdmann[70,175], Johan G Eriksson[193–195], Paul W Franks[23,72,73], Philippe Froguel[55,56,196], Christian Gieger[34], Ulf Gyllensten[45], Anders Hamsten[77], Tamara B Harris[132], Christian Hengstenberg[179], Andrew A Hicks[161], Aroon Hingorani[126], Anke Hinney[114], Albert Hofman[35,61], Kees G Hovingh[79], Kristian Hveem[197], Thomas Illig[109,198], Marjo-Riitta Jarvelin[160,199–201], Karl-Heinz Jöckel[68], Sirkka M Keinanen-Kiukaanniemi[201,202], Lambertus A Kiemeney[100,203,204], Diana Kuh[152], Markku Laakso[129], Terho Lehtimäki[205], Douglas F Levinson[206], Nicholas G Martin[33], Andres Metspalu[5,12], Andrew D Morris[155], Markku S Nieminen[177], Inger Njølstad[207,208], Claes Ohlsson[85], Albertine J Oldehinkel[209], Willem H Ouwehand[6,88,165,166], Lyle J Palmer[36,37], Brenda Penninx[210], Chris Power[115], Michael A Province[7], Bruce M Psaty[57,211–214], Lu Qi[23,24], Rainer Rauramaa[131,215], Paul M Ridker[71,95], Samuli Ripatti[6,41,69], Veikko Salomaa[128], Nilesh J Samani[216,217], Harold Snieder[66,83], Thorkild I A Sørensen[218,219], Timothy D Spector[58], Kari Stefansson[31,76], Anke Tönjes[180,181], Jaakko Tuomilehto[134,220,221], André G Uitterlinden[35,60,61], Matti Uusitupa[222,223], Pim van der Harst[105,140], Peter Vollenweider[186], Henri Wallaschofski[224], Nicholas J Wareham[11], Hugh Watkins[41], H-Erich Wichmann[63,225,226], James F Wilson[94], Goncalo R Abecasis[14], Themistocles L Assimes[227], Inês Barroso[6,228], Michael Boehnke[14], Ingrid B Borecki[7], Panos Deloukas[6], Caroline S Fox[229], Timothy Frayling[22], Leif C Groop[230], Talin Haritunian[231], Iris M Heid[21,225], David Hunter[23,24], Robert C Kaplan[232], Fredrik Karpe[164,233], Miriam F Moffatt[98], Karen L Mohlke[10], Jeffrey R O'Connell[234],

Yudi Pawitan[3], Eric E Schadt[235,236], David Schlessinger[237], Valgerdur Steinthorsdottir[31], David P Strachan[238], Unnur Thorsteinsdottir[31,76], Cornelia M van Duijn[35,61,154], Peter M Visscher[26,32], Anna Maria Di Blasio[13], Joel N Hirschhorn[15–17], Cecilia M Lindgren[4], Andrew P Morris[4], David Meyre[55,56,239], André Scherag[68], Mark I McCarthy[4,164,233,245], Elizabeth K Speliotes[20,240,245], Kari E North[8,245], Ruth J F Loos[11,241–243,245] & Erik Ingelsson[2–4,22]

[1]US Department of Health and Human Services, Division of Cancer Epidemiology and Genetics, National Cancer Institute, US National Institutes of Health, Bethesda, Maryland, USA. [2]Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden. [3]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden. [4]Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK. [5]Estonian Genome Center, University of Tartu, Tartu, Estonia. [6]Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK. [7]Department of Genetics, Washington University School of Medicine, St. Louis, Missouri, USA. [8]Department of Epidemiology, School of Public Health, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA. [9]Center for Observational Research, Amgen, Thousands Oaks, California, USA. [10]Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA. [11]Medical Research Council (MRC) Epidemiology Unit, Institute of Metabolic Science, Addenbrooke's Hospital, Cambridge, UK. [12]Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia. [13]Molecular Biology Department, Istituto Auxologico Italiano, Milan, Italy. [14]Department of Biostatistics, Center for Statistical Genetics, University of Michigan, Ann Arbor, Michigan, USA. [15]Divisions of Genetics and Endocrinology and Center for Basic and Translational Obesity Research, Children's Hospital, Boston, Massachusetts, USA. [16]Metabolism Initiative and Program in Medical and Population Genetics, Broad Institute, Cambridge, Massachusetts, USA. [17]Department of Genetics, Harvard Medical School, Boston, Massachusetts, USA. [18]Department of Internal Medicine (Cardiovascular), University of Michigan, Ann Arbor, Michigan, USA. [19]Department of Human Genetics, University of Michigan, Ann Arbor, Michigan, USA. [20]Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, Michigan, USA. [21]Public Health and Gender Studies, Institute of Epidemiology and Preventive Medicine, Regensburg University Medical Center, Regensburg, Germany. [22]Genetics of Complex Traits, Peninsula College of Medicine and Dentistry, University of Exeter, Exeter, UK. [23]Department of Nutrition, Harvard School of Public Health, Boston, Massachusetts, USA. [24]Channing Laboratory, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA. [25]Department of Biostatistics and Bioinformatics, Emory University, Atlanta, Georgia, USA. [26]The Queensland Brain Institute, The University of Queensland, Brisbane, Queensland, Australia. [27]Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts, USA. [28]Department of Biostatistics, Harvard School of Public Health, Boston, Massachusetts, USA. [29]Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA. [30]Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, Massachusetts, USA. [31]deCODE Genetics, Reykjavik, Iceland. [32]University of Queensland Diamantina Institute, The University of Queensland, Princess Alexandra Hospital, Brisbane, Queensland, Australia. [33]Queensland Institute of Medical Research, Brisbane, Queensland, Australia. [34]Institute of Genetic Epidemiology, Helmholtz Zentrum München–German Research Center for Environmental Health, Neuherberg, Germany. [35]Department of Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands. [36]Genetic Epidemiology and Biostatistics Platform, Ontario Institute for Cancer Research, Toronto, Ontario, Canada. [37]Prosserman Centre for Health Research, Samuel Lunenfeld Research Institute, Toronto, Ontario, Canada. [38]Centre for Genetic Epidemiology and Biostatistics, The University of Western Australia, Crawley, Western Australia, Australia. [39]Department of Internal Medicine, VU University Medical Centre, Amsterdam, The Netherlands. [40]Unit of Public Health Genomics, Department of Chronic Disease Prevention, National Institute for Health and Welfare, Helsinki, Finland. [41]Cardiovascular Medicine, University of Oxford, Wellcome Trust Centre for Human Genetics, Oxford, UK. [42]Department of Biological Psychology, VU University Amsterdam, Amsterdam, The Netherlands. [43]MRC Human Genetics Unit, MRC Institute for Genetics and Molecular Medicine, Western General Hospital, Edinburgh, UK. [44]Institute of Medical Biometry and Epidemiology, University of Marburg, Marburg, Germany. [45]Department of Immunology, Genetics and Pathology, Uppsala University, Uppsala, Sweden. [46]Uppsala Clinical Research Center, Uppsala University Hospital, Uppsala, Sweden. [47]Genome Centre, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK. [48]Clinical Pharmacology, William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK. [49]Ludwigshafen Risk and Cardiovascular Health (LURIC) Study, Freiburg, Germany. [50]Mannheim Institute of Public Health, Social and Preventive Medicine, Medical Faculty of Mannheim, University of Heidelberg, Mannheim, Germany. [51]Institut für Medizinische Biometrie und Statistik, Universität zu Lübeck, Universitätsklinikum Schleswig-Holstein, Campus Lübeck, Lübeck, Germany. [52]Department of Medical Genetics, University of Lausanne, Lausanne, Switzerland. [53]Swiss Institute of Bioinformatics, Lausanne, Switzerland. [54]Department of Medical Genetics, Division of Genetic Epidemiology, Molecular and Clinical Pharmacology, Innsbruck Medical University, Innsbruck, Austria. [55]University Lille Nord de France, Lille, France. [56]Centre National de la Recherche Scientifique (CNRS) Unité Mixte de Recherche (UMR) 8199, Institut de Biologie de Lille (IBL)–Institut Pasteur de Lille, Lille, France. [57]Cardiovascular Health Research Unit, University of Washington, Seattle, Washington, USA. [58]Department of Twin Research and Genetic Epidemiology, King's College London, London, UK. [59]School of Social and Community Medicine, University of Bristol, Bristol, UK. [60]Department of Internal Medicine, Erasmus Medical Center, Rotterdam, The Netherlands. [61]Netherlands Genomics Initiative (NGI)-sponsored Netherlands Consortium for Healthy Aging (NCHA), The Netherlands. [62]Department of Medicine I, University Hospital Grosshadern, Ludwig-Maximilians-Universität, Munich, Germany. [63]Chair of Epidemiology, Institute of Medical Informatics, Biometry and Epidemiology, Ludwig-Maximilians-Universität, Munich, Germany. [64]Chair of Genetic Epidemiology, Institute of Medical Informatics, Biometry and Epidemiology, Ludwig-Maximilians-Universität, Munich, Germany. [65]Department of Biostatistics, Boston University School of Public Health, Boston, Massachusetts, USA. [66]Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands. [67]MRC Centre for Causal Analyses in Translational Epidemiology, School of Social and Community Medicine, University of Bristol, Bristol, UK. [68]Institute for Medical Informatics, Biometry and Epidemiology (IMIBE), University Hospital of Essen, University of Duisburg-Essen, Essen, Germany. [69]Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, Finland. [70]Medizinische Klinik II, Universität zu Lübeck, Lübeck, Germany. [71]Division of Preventive Medicine, Brigham and Women's Hospital, Boston, Massachusetts, USA. [72]Genetic and Molecular Epidemiology Unit, Department of Clinical Sciences, Skåne University Hospital Malmö, Lund University, Malmö, Sweden. [73]Department of Public Health & Clinical Medicine, Umeå University, Umeå, Sweden. [74]Department of Odontology, Umeå University, Umeå, Sweden. [75]Icelandic Heart Association, Kopavogur, Iceland. [76]Faculty of Medicine, University of Iceland, Reykjavik, Iceland. [77]Atherosclerosis Research Unit, Department of Medicine, Solna, Karolinska Institutet, Karolinska University Hospital, Stockholm, Sweden. [78]Interfaculty Institute for Genetics and Functional Genomics, Ernst-Moritz-Arndt-University Greifswald, Greifswald, Germany. [79]Department of Vascular Medicine, Academic Medical Center, Amsterdam, The Netherlands. [80]Heart Failure Research Centre, Department of Clinical and Experimental Cardiology, Academic Medical Center, Amsterdam, The Netherlands. [81]Department of Oncology, University of Cambridge, Cambridge, UK. [82]Department of Endocrinology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands. [83]LifeLines Cohort Study, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands. [84]Department of Internal Medicine, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden. [85]Hudson Alpha Institute for Biotechnology, Huntsville, Alabama, USA. [86]Department of Cardiology, Division of Heart & Lungs, University Medical Center Utrecht, Utrecht, The Netherlands. [87]Institute of Biomedicine/Physiology, University of Eastern Finland, Kuopio Campus, Kuopio, Finland. [88]National Institute for Health Research (NIHR) Cambridge Biomedical Research Centre, Cambridge, UK. [89]Division of Epidemiology, Multidisciplinary Cardiovascular Research Centre (MCRC), Leeds Institute of Genetics, Health and Therapeutics (LIGHT), University of Leeds, Leeds, UK. [90]Department of Molecular Genetics, PathWest Laboratory of Western Australia, Queen Elizabeth II Medical Centre, Nedlands, Western Australia, Australia. [91]Department of Surgery and Pathology, University of Western Australia, Nedlands, Western Australia, Australia. [92]Genome Technology Branch, National Human Genome Research Institute, US National Institutes of Health, Bethesda, Maryland, USA. [93]Dipartimento di Medicina Sperimentale. Università degli Studi Milano-Bicocca, Monza, Italy. [94]Centre for Population Health Sciences, University of Edinburgh, Edinburgh, UK. [95]Harvard Medical School, Boston, Massachusetts, USA. [96]British Heart Foundation Glasgow Cardiovascular Research Centre, University of Glasgow, Glasgow, UK. [97]University of Dundee, Ninewells Hospital & Medical School, Dundee, UK. [98]National Heart and Lung Institute, Imperial College London, London, UK. [99]Division of Cardiovascular Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden. [100]Department of

Epidemiology, Biostatistics and HTA, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands. [101]Istituto di Ricerca Genetica e Biomedica del Consiglio Nazionale delle Ricerche (CNR), Monserrato, Italy. [102]Department of Dietetics-Nutrition, Harokopio University, Athens, Greece. [103]Epidemiology and Preventive Medicine Research Center, Department of Clinical and Experimental Medicine, University of Insubria, Varese, Italy. [104]Department of Cardiology, Toulouse University School of Medicine, Rangueil Hospital, Toulouse, France. [105]Department of Genetics, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands. [106]Department of Health Sciences, University of Milan, Ospedale San Paolo, Milan, Italy. [107]University of Chicago, Chicago, Illinois, USA. [108]Northshore University Health System, Evanston, Illinois, USA. [109]Research Unit for Molecular Epidemiology, Helmholtz Zentrum München–German Research Center for Environmental Health, Neuherberg, Germany. [110]Division of Cardiovascular and Neuronal Remodelling, Multidisciplinary Cardiovascular Research Centre, LIGHT, University of Leeds, Leeds, UK. [111]Department of Clinical Sciences/Obstetrics and Gynecology, University of Oulu, Oulu, Finland. [112]Department of Neurology, Boston University School of Medicine, Boston, Massachusetts, USA. [113]Department of Psychiatry, Washington University School of Medicine, St. Louis, Missouri, USA. [114]Department of Child and Adolescent Psychiatry, University of Duisburg-Essen, Essen, Germany. [115]Centre For Paediatric Epidemiology and Biostatistics, MRC Centre of Epidemiology for Child Health, University College London Institute of Child Health, London, UK. [116]Division of Research, Kaiser Permanente Northern California, Oakland, California, USA. [117]Core Genotyping Facility, SAIC-Frederick, Inc., National Cancer Institute (NCI)-Frederick, Frederick, Maryland, USA. [118]Department of Physiology, Institute of Neuroscience and Physiology, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden. [119]Population Studies Unit, Department of Chronic Disease Prevention, National Institute for Health and Welfare, Turku, Finland. [120]Department of Clinical Physiology, University of Tampere and Tampere University Hospital, Tampere, Finland. [121]Cardiovascular Research Center, Massachusetts General Hospital, Boston, Massachusetts, USA. [122]Cardiology Division, Massachusetts General Hospital, Boston, Massachusetts, USA. [123]Center for Human Genetic Research, Massachusetts General Hospital, Boston, Massachusetts, USA. [124]UK Clinical Research Collaboration (UKCRC) Centre of Excellence for Public Health (Northern Ireland), Queens University, Belfast, UK. [125]Department of Public Health and Primary Care, Institute of Public Health, University of Cambridge, Cambridge, UK. [126]Department of Epidemiology and Public Health, University College London, London, UK. [127]Department of Internal Medicine II–Cardiology, University of Ulm Medical Center, Ulm, Germany. [128]Chronic Disease Epidemiology and Prevention Unit, Department of Chronic Disease Prevention, National Institute for Health and Welfare, Helsinki, Finland. [129]Department of Medicine, University of Eastern Finland, Kuopio Campus and Kuopio University Hospital, Kuopio, Finland. [130]Finnish Institute of Occupational Health, Oulu, Finland. [131]Kuopio Research Institute of Exercise Medicine, Kuopio, Finland. [132]Laboratory of Epidemiology, Demography, Biometry, National Institute on Aging, US National Institutes of Health, Bethesda, Maryland, USA. [133]Department of Medical Sciences, Uppsala University, Akademiska Sjukhuset, Uppsala, Sweden. [134]Diabetes Prevention Unit, National Institute for Health and Welfare, Helsinki, Finland. [135]Human Genetics, Genome Institute of Singapore, Singapore. [136]Department of Internal Medicine, Istituto Auxologico Italiano, Verbania, Italy. [137]Transplantation Laboratory, Haartman Institute, University of Helsinki, Helsinki, Finland. [138]Chair of Nephrology San Raffaele Scientific Institute, Università Vita-Salute San Raffaele, Unità Ospedaliera Nephrology and Dialysis, Milan, Italy. [139]Synlab Academy, Mannheim, Germany. [140]Department of Cardiology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands. [141]Department of Biostatistics, University of Washington, Seattle, Washington, USA. [142]Genetics Division, GlaxoSmithKline, King of Prussia, Pennsylvania, USA. [143]Institute of Human Genetics, University of Bonn, Bonn, Germany. [144]Department of Genomics, Life & Brain Center, University of Bonn, Bonn, Germany. [145]School of Population Health, The University of Western Australia, Nedlands, Western Australia, Australia. [146]Department of Respiratory Medicine, Sir Charles Gairdner Hospital, Nedlands, Western Australia, Australia. [147]Busselton Population Medical Research Foundation, Sir Charles Gairdner Hospital, Nedlands, Western Australia, Australia. [148]Department of Internal Medicine, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands. [149]MRC Harwell, Harwell, UK. [150]Department of Statistics, University of Oxford, Oxford, UK. [151]Section of Epidemiology, Department of Public Health, Aarhus University, Aarhus, Denmark. [152]MRC Unit for Lifelong Health & Ageing, London, UK. [153]Department of Clinical Genetics, Erasmus Medical Center, Rotterdam, The Netherlands. [154]Center of Medical Systems Biology, Leiden University Medical Center, Leiden, The Netherlands. [155]Medical Research Institute, University of Dundee, Ninewells Hospital and Medical School, Dundee, UK. [156]Illumina, Inc., Cambridge, UK. [157]Institute of Epidemiology II, Helmholtz Zentrum München–German Research Center for Environmental Health, Neuherberg, Germany. [158]Munich Heart Alliance, Munich, Germany. [159]Faculty of Medicine, University of Split, Split, Croatia. [160]National Institute for Health and Welfare, Oulu, Finland. [161]Center for Biomedicine, European Academy Bozen/Bolzano (EURAC), Bolzano/Bozen, Italy (affiliated institute of the University of Lübeck). [162]Department of Neurology, General Central Hospital, Bolzano, Italy. [163]Department of Neurology, University of Lübeck, Lübeck, Germany. [164]Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Oxford, UK. [165]Department of Haematology, University of Cambridge, Cambridge, UK. [166]National Health Service (NHS) Blood and Transplant, Cambridge Centre, Cambridge, UK. [167]Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, Finland. [168]The Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku, Finland. [169]MRC Biostatistics Unit, Institute of Public Health, Cambridge, UK. [170]Finnish Diabetes Association, Tampere, Finland. [171]Pirkanmaa Hospital District, Tampere, Finland. [172]South Karelia Central Hospital, Lappeenranta, Finland. [173]Institute for Community Medicine, University Medicine Greifswald, Greifswald, Germany. [174]Institute for Clinical Molecular Biology, Christian-Albrechts University, Kiel, Germany. [175]Deutsches Herzzentrum München and DZHK (German Centre for Cardiovascular Research), partner site Munich Heart Alliance, Munich, Germany. [176]Azienda Ospedaliera di Desio e Vimercate, Milan, Italy. [177]Division of Cardiology, Cardiovascular Laboratory, Helsinki University Central Hospital, Helsinki, Finland. [178]University of Eastern Finland and Kuopio University Hospital, Kuopio, Finland. [179]Klinik und Poliklinik für Innere Medizin II, Universitätklinikum Regensburg, Regensburg, Germany. [180]Department of Medicine, University of Leipzig, Leipzig, Germany. [181]Integrierte Forschungs- und Behandlungszentren (IFB) Adiposity Diseases, University of Leipzig, Leipzig, Germany. [182]Institut National de la Santé et de la Recherche Médicale (INSERM) Unité Mixte de Recherche Scientifique (UMRS) 937, Institute of Cardiometabolism And Nutrition (ICAN), Pierre et Marie Curie Medical School, Paris, France. [183]Dipartimento di Scienze Farmacologiche e Biomolecolari, Università di Milano, Centro Cardiologico Monzino, Istituto di Ricovero e Cura a Carattere Scientifico (IRCCS), Milan, Italy. [184]Department of Human Genetics, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands. [185]Department of Medicine, University of Turku and Turku University Hospital, Turku, Finland. [186]Department of Internal Medicine, Centre Hospitalier Universitaire Vaudois (CHUV) University Hospital, Lausanne, Switzerland. [187]Cardiology Group, Frankfurt-Sachsenhausen, Germany. [188]Institut Pasteur de Lille, INSERM U744, Université Lille Nord de France, Lille, France. [189]Department of Medicine, Division of Endocrinology and Diabetes, University Hospital, Ulm, Germany. [190]Human Genetics Center, University of Texas Health Science Center, Houston, Texas, USA. [191]Institute of Molecular Medicine, University of Texas Health Science Center, Houston, Texas, USA. [192]Fondazione Filarete, Milan, Italy. [193]Department of General Practice and Primary Health Care, University of Helsinki, Helsinki, Finland. [194]National Institute for Health and Welfare, Helsinki, Finland. [195]Unit of General Practice, Helsinki University Central Hospital, Helsinki, Finland. [196]Department of Genomics of Common Disease, School of Public Health, Imperial College London, London, UK. [197]Helseundersøkelsen i Nord-Trøndelag (HUNT) Research Centre, Department of Public Health and General Practice, Norwegian University of Science and Technology, Levanger, Norway. [198]Hannover Unified Biobank, Hannover Medical School, Hannover, Germany. [199]Department of Epidemiology and Biostatistics, School of Public Health, Faculty of Medicine, Imperial College London, London, UK. [200]Institute of Health Sciences, University of Oulu, Oulu, Finland. [201]Biocenter Oulu, University of Oulu, Oulu, Finland. [202]Unit of General Practice, Oulu University Hospital, Oulu, Finland. [203]Department of Urology, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands. [204]Comprehensive Cancer Center East, Nijmegen, The Netherlands. [205]Department of Clinical Chemistry, Fimlab Laboratories, University of Tampere and Tampere University Hospital, Tampere, Finland. [206]Stanford University School of Medicine, Stanford, California, USA. [207]Department of Clinical Medicine, Faculty of Health Sciences, University of Tromsø, Tromsø, Norway. [208]Department of Community Medicine, Faculty of Health Sciences, University of Tromsø, Tromsø, Norway. [209]Interdisciplinary Center Psychopathology and Emotion Regulation, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands. [210]Department of Psychiatry, University Medical Centre Groningen, Groningen, The Netherlands. [211]Department of Epidemiology, University of Washington, Seattle, Washington, USA. [212]Department of Medicine, University of Washington, Seattle, Washington, USA. [213]Department of Health Services, University of Washington, Seattle, Washington, USA. [214]Group Health Research Institute, Group Health, Seattle, Washington, USA. [215]Department of Clinical Physiology and Nuclear Medicine, Kuopio University Hospital, Kuopio, Finland. [216]Department of Cardiovascular Sciences, University of Leicester, Glenfield Hospital, Leicester, UK. [217]Leicester NIHR Biomedical Research Unit in Cardiovascular Disease, Glenfield Hospital, Leicester, UK. [218]Institute of Preventive Medicine, Bispebjerg University Hospital, Copenhagen, Denmark. [219]Novo Nordisk Foundation Center for Basic Metabolic Research, University of Copenhagen, Copenhagen, Denmark. [220]Centre for Vascular Prevention, Danube-University Krems, Krems, Austria. [221]King Abdulaziz University, Jeddah, Saudi Arabia. [222]Institute of Public Health and Clinical Nutrition, University of Eastern Finland, Kuopio, Finland. [223]Research Unit, Kuopio University Hospital, Kuopio, Finland. [224]Institute of Clinical Chemistry and Laboratory Medicine, University Medicine Greifswald, Greifswald, Germany. [225]Institute of Epidemiology I, Helmholtz Zentrum München–German Research Center for Environmental Health, Neuherberg, Germany. [226]Klinikum Grosshadern, Munich, Germany. [227]Department of Medicine, Stanford University School of Medicine, Stanford, California, USA. [228]University of Cambridge Metabolic Research Laboratories, Institute of Metabolic Science Addenbrooke's Hospital, Cambridge, UK. [229]Division of Intramural Research, National Heart, Lung, and Blood Institute,

Framingham Heart Study, Framingham, Massachusetts, USA. [230]Lund University Diabetes Centre, Department of Clinical Sciences, Lund University, Malmö, Sweden. [231]Medical Genetics Institute, Cedars-Sinai Medical Center, Los Angeles, California, USA. [232]Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, New York, USA. [233]Oxford NIHR Biomedical Research Centre, Churchill Hospital, Oxford, UK. [234]Department of Medicine, University of Maryland School of Medicine, Baltimore, Maryland, USA. [235]Department of Genetics and Genomic Sciences, Mount Sinai School of Medicine, New York, New York, USA. [236]Institute of Genomics and Multiscale Biology, Mount Sinai School of Medicine, New York, New York, USA. [237]Laboratory of Genetics, National Institute on Aging, Baltimore, Maryland, USA. [238]Division of Population Health Sciences and Education, St George's, University of London, London, UK. [239]Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario, Canada. [240]Department of Internal Medicine, Division of Gastroenterology, University of Michigan, Ann Arbor, Michigan, USA. [241]The Charles Bronfman Institute of Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, New York, USA. [242]The Mindich Child Health and Development Institute, Icahn School of Medicine at Mount Sinai, New York, New York, USA. [243]Department of Preventive Medicine, Icahn School of Medicine at Mount Sinai, New York, New York, USA. [244]These authors contributed equally to this work. [245]These authors jointly directed this work. Correspondence should be addressed to E.I. (erik.ingelsson@medsci.uu.se).

ONLINE METHODS

Study design. We conducted a two-stage study for the distribution tails of three anthropometric traits (BMI, WHR adjusted for BMI and height) and four clinical classes of obesity (overweight class and obesity classes I, II and III), followed by a combined analysis of the two stages. Stage 1 consisted of a meta-analysis of GWAS using data from a study base (or sampling frame) of up to 168,267 adult individuals of European ancestry from 51 studies participating in the GIANT Consortium (**Supplementary Tables 1–5**). In stage 2, 273 SNPs with P value $< 5 \times 10^{-6}$ were followed up in up to 109,703 additional individuals of European descent, who included 67,243 individuals from 24 studies with data from the Metabochip (a custom-designed array of ~200,000 SNPs with previous evidence of suggestive association with metabolic traits) and 42,460 individuals from 12 studies with *in silico* replication GWAS data (**Supplementary Tables 1–5**). This gave us a study base of up to 276,007 individuals of European descent for the joint meta-analysis of stage 1 and stage 2. Full details of the discovery and replication stages, analysis of data and meta-analyses are given in the **Supplementary Note**. All studies were approved by local research ethics committees, and all participants gave informed consent.

Phenotype definitions. The distribution tails of the three anthropometric traits (BMI, WHR adjusted for BMI and height) were defined as the upper 5th percentile (cases) and lower 5th percentile (controls) of the distribution stratified by sex and disease status after controlling for the following covariates: age, age^2 and principal components for BMI; age and principal components for height; and age, age^2, BMI and principal components for WHR. For the clinical obesity classes, cases were defined by BMI ≥ 25 kg/m^2 for the overweight class, BMI ≥ 30 kg/m^2 for obesity class I, BMI ≥ 35 kg/m^2 for obesity class II and BMI ≥ 40 kg/m^2 for obesity class III. Controls were subjects with BMI < 25 kg/m^2. A minimum of 30 cases and 30 controls for each study-specific stratum was required.

Association analyses and meta-analyses. Each study conducted single-marker association analyses assuming an additive genetic model taking genotype imputation uncertainty into account. Analyses were stratified by sex (except for studies with related individuals, where analyses accounted for family structure) and disease status for studies that ascertained participants on the basis of a relevant disease (for example, diabetes). Before carrying out meta-analysis of the data, the results from each study were extensively reviewed using standardized quality control procedures to identify potential problems, such as strand issues, discrepancies between the reported standard errors and P values and allele frequency differences. SNPs with poor imputation quality scores or estimated minor allele count ≤ 20 ($2 \times N \times$ minor allele frequency) in each stratum (men/women or pooled for family-based studies) of each study were removed from analysis. For discovery stage 1, each stratum- and study-specific GWAS was corrected for genomic control. Meta-analyses were performed for each phenotype in METAL[48] using the fixed-effects inverse variance method based on β estimates and standard errors from each study. The results of the discovery meta-analysis were corrected by an additional genomic control step. Similar methods were employed for the replication and joint discovery and replication analyses.

Association testing in studies of extremely obese individuals. We tested the association of all SNPs reaching $P < 5 \times 10^{-8}$ in the joint analysis of stage 1 and stage 2 results for the BMI-related traits in five studies of extremely obese individuals (**Supplementary Tables 2–5**). For the four case-control studies (French Extreme Obesity Study, Essen Case-Control GWAS, GEO and GOYA), a fixed-effects inverse variance method was used to perform meta-analysis of the results. With the four case-control studies, meta-analysis was performed including a fifth study (Essen Obesity Trio GWAS) that has a nuclear structure, using a weighted z-score method that takes into account the direction but not the magnitude of the association.

Systematic comparison of the genetic structure between distribution tails and the overall distribution. For these analyses, we included all GWAS that provided genome-wide results for both the full distribution and the distribution tails of BMI, height and WHR. First, we used the results for the full distribution to calculate, for each genotype, the expected number of individuals in the upper and lower 5% tails. We used these values to perform logistic regression, comparing the upper and lower tails, and obtained the expected β values and expected standard error.

Second, we tested the differences between the expected β values and the observed β values obtained from the meta-analyses of the tails of the distributions. The standard error of the each difference was estimated as sqrt((expected standard error)^2 + (observed standard error)^2 − 2 × 0.65 × (expected standard error × observed standard error)), where 0.65 is the correlation between the expected β values and the observed β values obtained from TWINGENE by bootstrapping. Finally, meta-analysis was performed on the differences between expected β values and observed β values using the inverse variance method in METAL.

Polygenic comparison of genetic determinants of the BMI distribution tails and overall distribution. Within each trait (BMI, height and WHR), we aimed to compare variance explained in the distribution tails of the trait by two genetic (polygenic) scores obtained from (i) the meta-analyses of the distribution tails of the trait and (ii) the meta-analyses of the full distribution. To make the scores comparable, we limited polygenic score construction to the studies that provided genome-wide meta-analysis results for both distribution tails and the overall distribution. After LD filtering (using $r^2 \geq 0.05$ and 1 Mb of distance) and excluding SNPs present in <50% of samples, we created polygenic scores as the weighted sum of risk alleles using the method proposed by the International Schizophrenia Consortium[45]. For the BMI analysis, the association between the polygenic scores and the distribution tails of BMI was investigated in four samples of extremely obese individuals and two independent cohort studies using the same definition of distribution tails (**Supplementary Table 13**). Only the two independent cohorts were used for the height and WHR analyses. To estimate the phenotypic variance explained, we fit logistic or linear regression models including age, sex, study-specific covariates and the polygenic score as predictors and the distribution tails of the trait or the overall trait as outcomes in separate models. The phenotypic variance explained by the polygenic scores was defined as the difference in R^2 (linear regression) or Nagelkerke R^2 (logistic regression) between these models and a basic model including only age, sex and study-specific covariates as predictors.

Secondary signals analysis. To identify potential secondary signals, we used the approximate conditional and joint analysis previously proposed[46], which uses summary-level statistics and the LD structure from a reference sample to approximate conditional P values. The meta-analysis results for each trait were analyzed separately with LD correction between SNPs estimated from 6,654 unrelated individuals from the ARIC cohort.

Haplotype-based analyses. Using data from ten of the largest studies, we tested the association between the distribution tails of height, BMI and WHR and haplotypes across each established and new locus separately for males and females within each study using GENEBPM[49]. Haplotypes were estimated from GWAS SNP data by means of an expectation-maximization algorithm and then clustered according to their allelic similarity. Within a logistic regression modeling framework, haplotypes within the same cluster were assigned the same allelic effect, reducing the required number of parameters. Markov-chain Monte Carlo techniques were employed to sample over the space of haplotype clusters and regression model parameters. Evidence in favor of a haplotype association with the trait was assessed by summing $\log_{10}$ (Bayes factors) across studies.

Expression quantitative trait locus (eQTL) analyses. We examined the *cis* associations between SNPs that reached genome-wide significance ($P < 5 \times 10^{-8}$) and the expression of nearby genes in multiple tissues from 5 studies described previously: (i) subcutaneous adipose tissue ($n = 603$) and whole blood ($n = 747$) from deCODE[50]; (ii) lymphoblastoid cell lines ($n = 830$) from a childhood asthma study[51]; (iii) liver ($n = 707$), subcutaneous fat ($n = 870$) and omental fat ($n = 916$) tissue from a bariatric surgery study[52]; (iv) subcutaneous abdominal ($n = 52$) and gluteal ($n = 62$) adipose tissue and whole blood ($n = 65$) from MolOBB[53]; and (v) cortical brain tissue ($n = 193$) from a survey study[54]. SNPs were tested for *cis* associations with transcripts

within 500 kb or 1 Mb, assuming an additive effect of the BMI-related allele or using an ANOVA test with study-specific *P*-value thresholds used to account for multiple testing. Conditional analyses were performed for all expression data, except for cortical tissue, by conditioning the trait-associated SNP on the most significant *cis*-associated SNP for the particular gene transcript and vice versa.

A list of SNPs identified in other GWAS near new loci is given in **Supplementary Table 17**.

48. Willer, C.J., Li, Y. & Abecasis, G.R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
49. Morris, A.P. Direct analysis of unphased SNP genotype data in population-based association studies via Bayesian partition modelling of haplotypes. *Genet. Epidemiol.* **29**, 91–107 (2005).
50. Emilsson, V. *et al.* Genetics of gene expression and its effect on disease. *Nature* **452**, 423–428 (2008).
51. Dixon, A.L. *et al.* A genome-wide association study of global gene expression. *Nat. Genet.* **39**, 1202–1207 (2007).
52. Zhong, H., Yang, X., Kaplan, L.M., Molony, C. & Schadt, E.E. Integrating pathway analysis and genetics of gene expression for genome-wide association studies. *Am. J. Hum. Genet.* **86**, 581–591 (2010).
53. Min, J.L. *et al.* Coexpression network analysis in abdominal and gluteal adipose tissue reveals regulatory genetic loci for metabolic syndrome and related phenotypes. *PLoS Genet.* **8**, e1002505 (2012).
54. Myers, A.J. *et al.* A survey of genetic human cortical gene expression. *Nat. Genet.* **39**, 1494–1499 (2007).

doi:10.1038/ng.2606