Appendix

Genetic architecture of human thinness compared to severe obesity

Fernando Riveros-McKay[1*], Vanisha Mistry[2*], Rebecca Bounds[2], Audrey Hendricks[1,3], Julia M. Keogh[2], Hannah Thomas[2], Elana Henning[2], Laura J. Corbin[4,5], Understanding Society Scientific Group, Stephen O'Rahilly[2], Eleftheria Zeggini[1], Eleanor Wheeler[1], Inês Barroso[1,2], I. Sadaf Farooqi[2].

Affiliations

[1]Wellcome Sanger Institute, Cambridge, UK; [2]University of Cambridge Metabolic Research Laboratories and NIHR Cambridge Biomedical Research Centre, Wellcome Trust-MRC Institute of Metabolic Science, Addenbrooke's Hospital, Cambridge, UK; [3]Department of Mathematical and Statistical Sciences, University of Colorado-Denver, Denver, CO 80204, USA; [4]MRC Integrative Epidemiology Unit at University of Bristol, Bristol, BS8 2BN, UK; [5]Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, BS8 2BN, UK. Correspondence should be addressed to: I. Sadaf Farooqi (isf20@cam.ac.uk) and Inês Barroso (ib1@sanger.ac.uk). *These authors contributed equally.

Abstract

The variation in weight within a shared environment is largely attributable to genetic factors. Whilst many genes/loci confer susceptibility to obesity, little is known about the genetic architecture of healthy thinness. Here, we characterise the heritability of thinness which we found was comparable to that of severe obesity (h^2=28.07 vs 32.33% respectively), although with incomplete genetic overlap (r=-0.49, 95% CI [-0.17, -0.82], p=0.003). In a genome-wide association analysis of thinness (n=1,471) vs severe obesity (n=1,456), we identified 10 loci previously associated with obesity, and demonstrate enrichment for established BMI-associated loci ($p_{binomial}$=3.05x10^{-5}). Simulation analyses showed that different association results between the extremes were likely in agreement with additive effects across the BMI distribution, suggesting different effects on thinness and obesity could be due to their different degrees of extremeness. In further analyses, we detected a novel obesity and BMI-associated locus at *PKHD1* (rs2784243, obese vs. thin p=5.99x10^{-6}, obese vs. controls p=2.13x10^{-6} p_{BMI}=2.3x10^{-13}), associations at loci recently discovered with much larger sample sizes (e.g. *FAM150B* and *PRDM6-CEP120*), and novel variants driving associations at previously established signals (e.g. rs205262 at the *SNRPC*/*C6orf106* locus and rs112446794 at the *PRDM6-CEP120* locus). Our ability to replicate loci found with much larger sample sizes demonstrates the value of clinical extremes and suggest that characterisation of the genetics of thinness may provide a more nuanced understanding of the genetic architecture of body weight regulation and may inform the identification of potential anti-obesity targets.

38 **Author Summary**

39 Obesity-associated disorders are amongst the leading causes of morbidity and mortality
40 worldwide. Most genome-wide association studies (GWAS) have focused on body mass index (BMI=
41 weight in Kg divided by height squared (m^2)) and obesity, but to date no genetic association study
42 testing thin and healthy individuals has been performed. In this study, we recruited a first of its kind
43 cohort of 1,471 clinically ascertained thin and healthy individuals and contrasted the genetic
44 architecture of the trait with that of severe early onset obesity. We show that thinness, like obesity,
45 is a heritable trait with a polygenic component. In a GWAS of persistent healthy thinness vs. severe
46 obesity with a total sample size of 2,927, we are able to find evidence of association in loci that
47 have only been recently discovered using large cohorts with >40,000 individuals. We also find a
48 novel BMI-associated locus at *PKHD1* in UK Biobank highlighted by our association study. This work
49 illustrates the value and increased power brought upon by using clinically ascertained extremes to
50 study complex traits and provides a valuable resource on which to study resistance to obesity in an
51 increasingly obesogenic environment.

Introduction

The rising prevalence of obesity is driven by changes in the environment including the consumption of high calorie foods and reduced levels of physical activity [1]. However, within a given environment, there is considerable variation in body weight; some people are particularly susceptible to severe obesity, whilst others remain thin [2,3]. Family, twin and adoption studies have consistently demonstrated that 40-70% of the variation in body weight can be attributed to heritable factors [4]. As a result, many studies have focused on the genetic basis of body mass index (BMI) and/or obesity. To date >250 common and low-frequency obesity-susceptibility loci have been identified [5-10]. Additionally, studies of people at one extreme of the distribution (severe obesity) have led to the identification of rare, penetrant genetic variants that affect key molecular and neural pathways involved in human energy homeostasis [11-14]. These findings have provided a rationale for targeting these pathways for therapeutic benefit. In contrast, little is known about the specific genetic characteristics of persistently thin individuals (thinness defined using WHO criteria BMI$\leq$18kg/m^2). Understanding the mechanisms underlying thinness/resistance to obesity may highlight novel anti-obesity targets for future drug development.

A small number of previous studies have found that thinness appears to be a trait that is at least as stable and heritable as obesity [15-18]. A large study of 7,078 UK children and adolescents, found that the strongest predictor of child/adolescent thinness was parental weight status. The prevalence of thinness was highest (16.2%) when both parents were thin and progressively lower when both parents were normal weight, overweight or obese [19].

One approach to studying thinness is to study individuals from a population-based cohort for a quantitative or continuous trait. For example, it is possible to generate a "case-control" study by taking the extremes of the population distribution for a continuous trait such as BMI, an approach used effectively by Berndt *et al.* 2013 [20] who analysed the top and bottom 5% in cohorts participating in the GIANT Consortium. However, by their very definition, such population-based cohorts often contain a limited number of people at the "extremes" (i.e. severe obesity and thinness) [20]. To date, other GWAS approaches that included thin individuals have either used them exclusively as controls to contrast with extreme obesity [21], or have not ascertained for healthy thinness [22]. Here, we use a different study design, and one that has been used to increase power to detect genetic association, in particular for disorders where there is a large environmental component (e.g. asthma, type 2 diabetes and obesity), enriching our case series with affected individuals that may be more genetically loaded. This selection is usually done by selecting individuals who may have a more extreme form of disease, are younger (less time for environment to impact their disease) and perhaps have family members also affected with the same condition. To complement this approach to the selection of cases, controls are also selected to increase the chances that they do not have the disease or are unlikely to develop the disease later in life [21]. This is normally done by selecting contrasting controls, or "super-controls". However, the low prevalence of thinness in countries such as the UK and the fact that people who are well but constitutionally thin do not routinely come to medical attention, poses challenges to recruitment of a cohort of healthy thin individuals. We were able to take advantage of the UK National Health Service (NHS) research infrastructure to recruit from primary care (**Methods**) using body mass index

(BMI: weight in kg/height in metres^2) criteria and personal review of individual case files to identify a cohort of approximately 2000 UK European descent thin adults (STudy Into Lean and Thin Subjects, STILTS cohort; mean BMI = 17.6 kg/m^2) who are well, without medical conditions or eating disorders (**Methods**). 74% of the STILTS cohort have a family history of persistent thinness throughout life, suggesting we have enriched for genetically driven thinness.

Here, we present a new, and the largest-to-date, GWAS focused on persistent healthy thinness and contrast the genetic architecture of this trait with that of severe early onset obesity ascertained in the clinic. We explored whether the genetic loci influencing thinness are the same as those influencing obesity, i.e., are these two clinically ascertained traits reverse sides of the same "coin", or whether there are important genetic differences between them. We show that persistent thinness and severe early onset obesity are both heritable traits (h^2=28.07% and h^2=32.33%, respectively) that share a number of associated loci, and both are enriched for established BMI associated loci (binomial p=3.05x10^{-5} and 9.09x10^{-13}, respectively). Nonetheless, we also detected important differences, with some loci more strongly associated at the upper clinical end of the BMI distribution (e.g. *FTO*), some at the lower end (e.g. *CADM2*), whilst other loci are equivalently associated with both clinical ends of the BMI spectrum (e.g. *MC4R*). Simulation tests showed that these results did not significantly deviate from additive effects and most likely reflect the different degrees of extremeness present in our clinically ascertained cohorts, where severely obese individuals represent a more significant deviation from the mean than healthy thin individuals do (the same degree of thinness may not be compatible with healthy human life). These data support expansion of genetic studies of persistent thinness as an approach to gain further insights into the

biology underlying human energy homeostasis, and as an alternative approach to uncovering potential anti-obesity targets for drug development.

Results

Heritability of persistent thinness and severe early onset obesity

To investigate the heritability of healthy thinness and contrast it with that of severe early onset childhood obesity we obtained genotype data for 1,622 persistently thin healthy individuals (STILTS), 1,985 severe childhood onset obesity cases (SCOOP; European ancestry individuals from the GOOS cohort) and 10,433 population-based individuals (UKHLS) used as a common set of controls (**Methods, S1 Table**). All participants were genotyped on the Illumina Core Exome array, including 551,839 markers. After sample and variant quality control, we retained 1,471 thin individuals, 1,456 obese individuals, 6,460 control individuals in the BMI range 19-30 kg/m^2 (non-extremes). 477,288 directly genotyped variants were included in the analysis (**Methods**); 54% common variants (minor allele frequency (MAF) ≥1% amongst controls) and 46% rare variants (MAF<1% amongst controls), of which most were protein-coding (96.8%). We then imputed genotypes to a combined UK10K+1000G reference panel and, using LD score regression, we estimated that a subset of 1,197,969 HapMap3 markers accounted for 32.33% (95% CI 23.75%-40.91%) of the phenotypic variance on the liability scale in severe early onset obesity, and 28.07% (95% CI 13.80%-42.34%) in persistent thinness, suggesting both traits are similarly heritable (**Methods**). The heritability estimates reported here were used mainly to establish the fact that

thinness is a heritable trait; we expect our liability scale estimates to be mostly unbiased given the study design [23]. However, given the low prevalence of the traits presented here, these estimates may represent upper bounds.

Contribution of known BMI associated loci to thinness and severe early onset obesity

To investigate the role of established common variant European BMI associated loci, we studied the 97 loci from GIANT [24] in persistent thinness vs severe early onset obesity and performed three-way association analyses: obese vs. thin, obese vs controls, controls vs. thin (**Methods, S1 Table**). After quality control, 41,266,535 variants remained for association analyses in the three cohorts: SCOOP vs STILTS, SCOOP vs UKHLS and UKHLS vs STILTS. Of the 97 established BMI associated loci from GIANT [24], we found that 40 were nominally significant (p<0.05) in SCOOP vs UKHLS and 15 in UKHLS vs STILTS (**S2 Table**). Direction of effect was consistent for all of these loci, which was more than expected by chance (binomial p=9.09x10^{-13} and binomial p=3.05x10^{-5}, respectively). Overall, the proportion of phenotypic variance explained by the 97 established BMI associated loci was 10.67% in SCOOP vs UKHLS, and 4.33% in STILTS vs UKHLS (**Methods**). Evaluation of association results in thin (STILTS) and obese (SCOOP) individuals, compared to the same controls (UKHLS), suggested that the results are not a mirror image of each other (**Figs 1-2**), **however we found little evidence of non-additive effects at the loci explaining this discrepancy (see below).** We observed a striking difference in association results in the *FTO* locus where the lead intronic obesity risk variant, rs1558902, showed a moderate effect size and modest evidence of association in controls

compared to thin individuals from STILTS (p=0.00027, OR=1.17, 95% CI [1.08,1.28], EAF=0.39), despite having a large effect and being associated at genome-wide significance levels in SCOOP (p=1.25x10^{-17}, OR=1.43, 95% CI [1.32,1.55], EAF=0.41), and *GNAT2* also showed a larger effect and significance in the analysis of obese compared to control individuals (p=1.26x10^{-4}, OR=1.57, 95% CI [1.25, 1.97], EAF=0.03), than in the thin analysis (p=0.52, OR=1.10, 95% CI [0.82, 1.47], EAF=0.02, **Fig 1**, **S2 Table**). This discrepancy in association strength and effect size was also seen at the opposite end of the BMI spectrum in *CADM2* where the lead SNP, rs13078960, showed evidence of association in STILTS (p= 9.48x10^{-4}, OR=1.2, 95% CI [1.08, 1.33], EAF=0.20) but no association in SCOOP (p>0.05). In contrast to results at the *FTO* and *CADM2* loci, for *MC4R* the results are more comparable, with genome-wide significant association in obese individuals (rs6567160, p=7.91x10^{-9}, OR=1.31, 95% CI [1.19, 1.43], EAF=0.25) and highly significant association results in thin individuals (p=1.38x10^{-5}, OR=1.26, 95% CI [1.13, 1.39], EAF=0.23, **S2 Table**)**.** To formally test if these results were significantly different from those expected under a model where loci act additively across the BMI distribution, we simulated 10,000 different populations of 1 million individuals with genotypes for the 97 established BMI loci using allele frequencies in the European population, and then simulated a phenotype using the effect sizes in GIANT (**Methods**). These simulations detected fourteen loci with nominally significant deviation from an additive model, however none remained significant after correction for the number of tests (p=0.05/97*2 = ~0.0002, **S3 Table**), though *CADM2* was nominally significant in both SCOOP and STILTS analyses, with slightly lower OR detected in SCOOP compared to simulated data, and slightly higher OR detected in STILTS compared to simulated data (**S3 Table**). **Recent work in mouse knockouts has shown *CADM2* plays**

175 **an important role in systemic energy homeostasis [25] and variants near the gene have also been**
176 **recently linked to habitual physical activity in humans [26].** Since SCOOP participants are
177 significantly younger than UKHLS, we used summary statistics from a subset of the ALSPAC cohort
178 [27] which consists of 4,964 children aged 13-16 to test if the observed OR differences in SCOOP vs
179 UKHLS, compared to STILTS vs UKHLS, were due to age effects in SCOOP (**Methods**). For the 97
180 GIANT loci overall there were no significant differences in the ORs when comparing SCOOP to
181 UKHLS or SCOOP to ALSPAC (z-test, p>0.05) except for rs2245368 (*PMS2L11* locus, z-test
182 p=3.81x10^{5}, **S4 Table**). In combination, these results suggest that the observed differences in ORs
183 and p-values could have arisen because our severe obese cases are much more extreme (i.e.
184 deviate more from the mean) than the healthy thin individuals, and that our obese and thin sample
185 sizes gave us limited power to detect significant differences compared to the additive model.

186 **Fig 1. Odds ratio comparison for established BMI associated loci.** Odds ratios for SCOOP vs UKHLS
187 (x-axis) and UKHLS vs STILTS (y-axis) comparisons are shown for the 97 known BMI loci from GIANT
188 [24]. Colours of data points represent nominal significance in both analyses (red), only SCOOP vs.
189 UKHLS (green), only STILTS vs UKHLS (blue) or in neither analysis (purple). Error bars represent 95%
190 confidence intervals for the odds ratios for SCOOP vs UKHLS (x-axis) and for UKHLS vs STILTS (y-
191 axis). A subset of data points with larger separation from the red diagonal line (x=y) are labelled.

192

193 Next we investigated the association of a genetic risk score, generated from the 97 BMI associated
194 loci from GIANT [24] on BMI category (i.e. thin, normal, obese) using an ordinal logistic regression

(**Methods**). As expected, the standardised BMI genetic risk score was strongly associated with BMI category (weighted score p=8.59x10^{-133}). We found that the effect of a one standard deviation increase in the standardised BMI genetic risk score was significantly larger for obese vs. (thin & normal) than for (obese & normal) vs. thin (p=7.48x10^{-11}, **S1 Appendix**) with odds ratio and 95% confidence intervals of 1.94 (1.83, 2.07) and 1.50 (1.42, 1.59) respectively. However, using the simulations described above (**Methods**), we confirm that the larger OR for obese vs. (thin & normal) is not significantly different (p=0.41) than what we would expect given an additive genetic model, and the different degrees of extremeness in our thin and obese cases. Mean GRS in each BMI category was also not significantly different from that predicted via simulations (**S1 Fig**, **Methods**).

Genetic Correlation between persistent thinness, severe early onset childhood obesity and BMI

Given the observed differences in association results from thin and obese individuals, compared to the same set of control individuals, we next explored the genetic correlation of severe early onset obesity, persistent thinness and BMI using LD score regression (**Methods**). For this, we used summary statistics from the SCOOP vs UKHLS, STILTS vs UKHLS and BMI data from participants in UK Biobank (UKBB, **Methods**). As expected from the association results, the genetic correlation of severe early onset obesity and BMI was high (r=0.79, 95% CI [0.69, 0.89], p=1.14x10^{-52}). We also observed weaker negative correlation between persistent thinness and BMI (r=-0.69, 95% CI [-0.86, -0.51], p= 1.17x10^{-14}), and between persistent thinness and severe obesity (r=-0.49, 95% CI [-0.17, -0.82], p=0.003). As an inverse genetic correlation between BMI, obesity and anorexia nervosa (a

215 disorder that is characterised by thinness and complex behavioural manifestations) has recently
216 been reported [28], we also tested for genetic correlation with anorexia nervosa, and found that
217 neither severe early onset obesity, nor persistent thinness, were significantly correlated with
218 anorexia nervosa (r=-0.05, 95% CI [-0.15,0.05], p=0.33 and r=0.13, 95% CI [-0.02,0.28], p=0.09,
219 respectively; **Methods**).

220

221 **Association signals for persistent thinness and severe early onset obesity replicate established**
222 **BMI associated loci**

223 Given available genome-wide directly genotyped and imputed data we sought evidence for novel
224 signals associated with either end of the BMI distribution (persistent thinness or severe early onset
225 obesity; **Methods**) **but found no novel replicating loci (details below).** In all three **discovery**
226 analyses, in addition to loci mapping to established BMI and obesity loci, we identified *PIGZ* and
227 *C3orf38*, two **putative** novel loci in the thin vs control analysis, that reached conventional genome-
228 wide significance (GWS) ($p \leq 5x10^{-8}$) (**Tables S5-S7, Fig 2**). However, an additional 125 SNPs, in 118
229 distinct loci, reached the arbitrary threshold of $p \leq 10^{-5}$ in at least one analysis, for which we sought
230 replication (**Tables S5-S7**)**.**

231 **Fig 2. Miami plot of SCOOP vs. UKHLS and STILTS vs. UKHLS.** Miami plot produced in EasyStrata
232 [29], Red=SCOOP vs. UKHLS; Blue=STILTS vs. UKHLS. Red lines indicate genome-wide significance
233 threshold at $p=5x10^{-08}$. Orange lines indicate discovery significance threshold at $p=1x10^{-05}$. Black
234 labels highlight known BMI/obesity loci that were taken forward for replication and yellow peaks

indicate those that met genome-wide significance after replication. Grey labels highlight novel loci with $p<5x10^{-08}$ that did not replicate.

As our obese and thin cases (SCOOP and STILTS) lie at the very extreme tails of the BMI distribution, there are few comparable replication datasets. We therefore used the UKBB dataset and selected individuals at the top (BMI>=40, N= 7,526) and bottom end of the distribution (BMI≤19, N= 3,532) to more closely match the BMI criteria of our clinically ascertained thin and obese individuals. We used 20,720 samples from the rest of the UKBB cohort as a control set (**Methods**, **S2 Fig**). In cases where lead variants or proxies ($r^2>0.8$) were not currently available in the full UKBB genetic release we used results from the interim release using 2,799 individuals with BMI>=40, 1,212 with BMI<=19 and 8,193 controls (**Methods**). We noted a significant negative genetic correlation for our obese replication cohort with anorexia nervosa (r= -0.24, 95% CI [-0.37,-0.11], p=0.01) and a positive genetic correlation for our thin cohort (r=0.49, 95% CI [0.22-0.76] p=0.0003). We also observed significant genetic correlation between obesity in the discovery and replication cohorts (r=0.84, 95% CI [0.65-1] $p=5.05x10^{-17}$) and between thinness in the discovery and replication cohorts (r= 0.62, 95% CI [0.20-1] p=0.004).

To further increase power, we took advantage of publicly available summary statistics from the GIANT Extremes obesity meta-analysis [20], the EGG childhood obesity study [30], and our own previous study on non-overlapping SCOOP participants (SCOOP 2013) [31], as additional replication datasets. For SCOOP vs. STILTS we used the GIANT BMI tails meta-analysis results [20] (up to 7,962

255 cases/8,106 controls from the upper/lower 5th percentiles of the BMI trait distribution). For SCOOP
256 vs. UKHLS we used the GIANT obesity class III summary statistics [20] (up to 2,896 cases with BMI
257 ≥40kg/m^2 vs 47,468 controls with BMI <25 kg/m^2), the EGG childhood obesity study [30] (children
258 with BMI ≥95th percentile of BMI vs 8,318 children with BMI <50th percentile of BMI) and SCOOP
259 2013 [31]. Fixed effect meta-analyses yielded genome-wide significant signals at well-known BMI
260 associated loci in both the obese vs. thin, and obese vs. control analyses, and both the *PIGZ* and
261 *C3orf38* loci identified at the discovery stage failed to replicate when combined with additional data
262 **(Table 1, S7 Table).** However, the *SNRPC* locus described here (rs75398113), though not
263 independent from the previously described *SNRPC*/*C6orf106* locus (rs205262, r^2= 0.29) [24],
264 appears to be driving the previously reported association at this locus (rs205262 conditioned on
265 rs75398113, $p_{conditioned}$=0.7, **S8 Table**). Both SNPs are eQTLs for *C6or106* and *UHRF1BP1* in multiple
266 tissues including brain and colon tissues on GTEx however neither of these are obvious biological
267 candidates linked to energy homeostasis.

268

269 **Table 1 - GWAS results for SNPs meeting *p*<5x10^{-8} in all three analyses.** EA= Effect allele (BMI
270 increasing allele); NEA= Non-effect allele; OR = Odds ratio; 95% CI = 95% confidence interval for the
271 odds ratio; EAF = effect allele frequency. Positions mapped to hg19, Build 37. [a]rs12995480 used as
272 proxy in GIANT. [b]rs2384054 used as proxy in GIANT. [c]rs12641981 used as proxy in GIANT. [d]rs663129
273 used as proxy in GIANT, EGG and SCOOP 2013. [e]rs13007080 used as proxy in GIANT, EGG and
274 SCOOP 2013. [f]rs7138803 used as proxy in SCOOP 2013. [g]rs6722587 used as proxy in GIANT, EGG
275 and SCOOP 2013. [h]rs4132288 used as proxy in GIANT, EGG and SCOOP 2013. [i]rs1460940 used as

proxy in GIANT, EGG and SCOOP 2013. [j]rs1366333 used as proxy in GIANT, EGG and SCOOP 2013. [k]GIANT BMI tails [20]. [l]GIANT obesity class III [20].

Finally, we used the independent BMI dataset from UKBB (**Methods**) to investigate whether any of the loci meeting our arbitrary $p \leq 10^{-5}$ in discovery efforts, were independently associated with BMI as a continuous trait. This identified a novel BMI-associated locus near *PKHD1* (SCOOP vs. STILTS $p=5.99\times10^{-6}$, SCOOP vs. UKHLS $p=2.13\times10^{-6}$, BMI $p=2.3\times10^{-13}$, **S9 Table**). Furthermore, we note that when comparing the signals we took for replication (based on case control analyses) with association results with BMI as a continuous trait derived from an independent set of samples from UKBB, there are more directionally consistent and nominally significant associations with BMI than expected by chance suggesting that amongst these loci, there may be additional real associations (binomial $p=4.88\times10^{-4}$, and binomial $p=9.77\times10^{-3}$, respectively, Methods, S9 Table)."

Despite the smaller sample size, the obese vs thin comparison had increased power to detect some loci (**S3 Fig**), including a recently discovered variant near *FAM150B* [32] (rs62107261, MAF= ~5%), which did not meet our $p<10^{-5}$ threshold to be taken forward for replication in obese vs controls analysis ($p=2.36\times10^{-4}$).

Discussion

Here we present results from the largest to-date GWAS performed on healthy individuals with persistent thinness and provide the first insights into the genetic architecture of this trait. To our

296 knowledge, there are only two other studies using thin individuals with comparable mean BMIs
297 **[21,22]**. The study by Hinney *et al*. **[21]** (N=442), was only able to detect *FTO* at genome-wide
298 significance level with rs1121980 having a similar effect to that which we report (OR=1.66 vs OR=
299 1.69 in our data). In the Scannell Bryan *et al***. [22]** study, Bangladeshi individuals were reportedly
300 thin and malnourished, and a single suggestive association was found with an intronic variant in
301 *NRXN3* (rs12882679, $p=9.57x10^{-7}$) which is not significant in our study (p=0.77).

302 Using genome-wide genotype data we show that persistent healthy thinness, similar to severe
303 obesity (h^2=32.33%), is a heritable trait (h^2=28.07%). Persistent healthy thinness and severe
304 childhood obesity are negatively correlated (r=-0.49, 95% CI [-0.17, -0.82], *p*=0.003), and share a
305 number of genetic risk loci. Nonetheless, the genetic overlap between the two clinically ascertained
306 traits appears to be incomplete, as highlighted by some loci which were more strongly associated
307 at one end of the BMI distribution (e.g. *CADM2*), while others, appeared to exert effects across the
308 entire BMI spectrum (e.g. *MC4R* [9,33,34]). Further exploration by simulation demonstrated that
309 these differences are likely to be due to the different degrees of extremeness of the two clinical
310 cohorts (i.e. a similar degree of thinness to that of the obese cohort may not be compatible with
311 healthy human life) and not due to a deviation from additive effects of the tested loci on BMI, with
312 the possible exception of *CADM2* which deviated from expectation with nominal significance in
313 both the obese and the thin analysis (**S3 Table**). This is in contrast with earlier studies which
314 suggested larger effects at the higher end of the BMI distribution [35,36] but in agreement with
315 more recent observations contrasting the bottom 5% and top 5% of the BMI tails where associated
316 loci were also consistent with additive effects [20]. This is also in contrast with a previous study on

317 height, where a deviation from additivity was found, but only for short individuals in the bottom
318 1.5% of the distribution [37], which suggests that analysis focused just on the most extreme
319 individuals may be warranted.

320

321 Focusing on the 97 previously established BMI associated loci [24], we show that the percentage of
322 phenotypic variance explained by these loci is lower in persistently thin (4.33%) compared to obese
323 individuals (10.67%), and that the effect of an increase/decrease in the BMI genetic risk score was
324 much larger, on average, for obese individuals than for thin individuals (one standard deviation
325 increase in the standardised BMI genetic risk score of 1.94, 95% CI (1.83, 2.07) and 1.50, 95% CI
326 (1.42, 1.59), respectively) which is consistent with the difference in BMI units amongst categories.
327 And, although our analysis using age-matched controls from ALSPAC suggested that the observed
328 differences in ORs, comparing obese vs control individuals to controls vs thin individuals, was
329 unlikely to be due to age effects, we cannot completely exclude the possibility that different effects
330 of age and sex in our discovery cohorts (**S1 Table**), and gene-by-environment interactions, could be
331 influencing some of the results we observe. For example, gene-by-environment interactions and
332 age effects have been previously reported at the *FTO* locus [38-41] where a larger effect is detected
333 in younger adults. **It is worth noting though that non-additive effects have also been observed in**
334 **the *FTO* locus [42].**

335

In studying thin individuals there are often concerns regarding the prevalence of eating disorders, notably anorexia nervosa amongst participants. We sought to carefully exclude eating disorders at two phases of recruitment (by medical history and by questionnaire). Additionally, we demonstrate that in our cohort of healthy thin individuals, anorexia nervosa is unlikely to be a confounder as the two traits are genetically only weakly correlated (r=0.13, 95% CI [-0.02,0.28], p=0.09). This was not the case for the UKBB replication cohort where a positive genetic correlation was observed (r= 0.49 95% CI [0.22-0.76] p=0.0003). The positive genetic correlation with anorexia was still observed after removing individuals with medical conditions that could explain their low BMI (r=0.62, 95% CI [0.30,0.92], p=0.0001, **Methods)**. These results highlight the importance of the careful phenotyping performed in the recruitment phase and the utility of the STILTS cohort as a resource to study healthy and persistent thinness.

In the genome-wide association analyses amongst the signals we took forward for replication, in addition to detecting established BMI-associated loci, we find a novel BMI-association at *PKHD1* in the UKBB BMI dataset (rs10456655, β=0.10, p=2.3x10^{-13}, **S9 Table**), where a proxy for this variant (rs2579994, r^2=1 in 1000G Phase 3 CEU) has been previously nominally associated with waist and hip circumference (p=5.60x10^{-5} and p=4.40x10^{-4} respectively) [43]. In addition, we found associations at loci that have only recently been established using very large sample sizes. *FAM150B*, was only suggestively associated at discovery stage in Tachmazidou *et al.* (2017) [32] (n=47,476, p=2.57×10^{-5}) whereas it reached genome-wide significance when contrasting SCOOP vs STILTS (n=2,927, p=2.07x10^{-8}, **S5 Table**). Also, *PRDM6-CEP120* [5] was recently discovered in a Japanese study with a sample size of 173,430 and has not been previously reported in a European

357 population. In our study, a signal near the locus (rs112446794, r^2=0.36) showed suggestive evidence
358 of association in SCOOP vs UKHLS (p=2.08x10^{-6}, **S6 Table**) with a significantly smaller sample size.
359 Conditional analysis reveals the lead SNP in this study drives the association of the previously
360 established signal (**S8 Table**). *CEP120* codes for centrosomal protein 120. Variants near this locus
361 have been previously associated with height [44] and waist circumference in East Asians [45].
362 Missense variants in the gene itself have been associated with rare ciliopathies [46,47]. Lastly,
363 amongst the signals we took for replication, and after removing known and newly established loci,
364 we still observe an enrichment of directionally consistent and nominal associations in the analysis
365 of BMI as a continuous trait, suggesting that some of these results may warrant additional
366 investigation, in particular in similarly ascertained thin and obese cohorts. One such example is
367 rs4447506, near *PIK3C3,* which was not only nominally significant and consistent in the
368 independent UKBB BMI analysis (p=1.5x10^{-6}, **S9 Table**), but also in the Locke *et al*. (2015) [24] BMI
369 results (p= 0.01), and in the GIANT BMI tails analysis we used as replication (**S5 Table**). We also
370 note, that despite not reaching genome-wide significance in our discovery cohorts, we observe
371 directionally consistent suggestive associations at a number of loci previously associated with BMI
372 tails and with different obesity classes [20] (**S10 Table**)**.** Altogether, these results highlight some
373 power advantages of using clinically ascertained extremes of the phenotype distribution to detect
374 associations and suggest that healthy thinness falls at the lower end of the polygenic BMI spectrum.
375 It is worth noting though that these clinically ascertained extremes display evidence of incomplete
376 genetic correlation with BMI, in contrast to previously described obesity classes (S4 Fig), so it is
377 plausible that additional loci might be uncovered by focusing on clinical extremes.

As our results were based on clinically ascertained participants which met very specific criteria, it is worth noting these conclusions cannot be straightforwardly extrapolated to the general population. Experiments in animals have identified loci/genes associated with thinness/decreased body weight due to reduced food intake/increased energy expenditure/resistance to high fat diet-induced obesity [48,49], mechanisms that we hypothesise may contribute to human thinness. The STILTS cohort, being uncorrelated to anorexia nervosa, is an excellent resource in which to conduct such additional genetic exploration. Further genetic and phenotypic studies focused on persistently thin individuals may provide new insights into the mechanisms regulating human energy balance and may uncover potential anti-obesity drug targets.

Methods

ETHICS STATEMENT

The study was reviewed and approved by the South Cambridgeshire Research Ethics Committee (12/EE/0172). All participants provided written informed consent prior to inclusion.

COHORTS

SCOOP, STILTS and UKHLS cohorts were used for the heritability, genetic correlation, genetic risk score and association analyses with established BMI loci, as well as, used as a discovery cohort in the genome-wide association study (GWAS) and gene-based tests. UK Biobank samples were used for genetic correlation analysis and in the replication stages of the GWAS and gene-based tests. ALSPAC was used as an additional control dataset to UKHLS for comparison against SCOOP in the established BMI loci analysis.

STILTS

The aim was to recruit a new cohort of UK European people who are thin (defined as a body mass index $\leq$ 18kg/m^2) and well. After ethical committee approval (12/EE/0172), we worked with the NIHR Primary Care Research Network (PCRN) to collaborate with 601 GP practices in England. Each practice searched their electronic health records using our inclusion criteria (age 18-65 years, BMI$\leq$18 kg/m^2) and exclusion criteria (medical conditions that could potentially affect weight (chronic renal, liver, gastrointestinal problems, metabolic and psychiatric disease, known eating disorders). A small number of individuals (n=43) with a BMI of 19.0 kg/m^2 were included as they

had a strong family history of thinness. The case notes of each potential participant were reviewed by the GP or a senior nurse with clinical knowledge of the participant to exclude other potential causes of low body weight in discussion with the study team. Through this approach we identified 25,000 individuals who fitted our criteria for inclusion in the study. These individuals were invited to participate in the study; approximately 12% (2,900) replied consenting to take part. We obtained a detailed medical and medication history, screened for eating disorders using a questionnaire (SCOFF) that has been validated against more formal clinical assessment [50]. We excluded all participants who stated that they exercised every day/more than 3 times a week/whose reported activity exceeded 6 metabolic equivalents (METs) for any duration or frequency (http://www.who.int/dietphysicalactivity/physical_activity_intensity/en/). With these rather strict criteria for exercise, we sought to limit the contribution of exercise as a contributor to the thinness of participants in the STILTS cohort. We excluded people who were thin only at a certain point in their lives (often as young adults) to focus on those who were persistently thin/always thin throughout life as we hypothesised that this group would be enriched for genetic factors contributing to their thinness. We asked a specific question to identify these individuals: “have you always been thin?” Only those who answered positively were included. Questionnaires were manually checked by senior clinical staff for these parameters and for reported ethnicity (non-European ancestry excluded). DNA was extracted from salivary samples obtained from these individuals using the Oragene 500 kit according to manufacturer’s instructions (**S1 Table**).

SCOOP

With ethical committee approval (MREC 97/5/21), we have recruited 7,000 individuals with severe early-onset obesity (BMI standard deviation score (SDS) > 3; onset of obesity before the age of 10 years) to the Genetics of Obesity Study (GOOS) [51]. The Severe Childhood Onset Obesity Project (SCOOP) cohort [31] is a sub-cohort of GOOS comprised of ~4,800 British individuals of European ancestry; **S1 Table**). SCOOP individuals likely to have congenital leptin deficiency, a treatable cause of severe obesity, were excluded by measurement of serum leptin, and individuals with mutations in the melanocortin 4 receptor gene (*MC4R*) (the most common genetic form of penetrant obesity) were excluded by prior Sanger sequencing.

UKHLS

Understanding Society (UKHLS) is a longitudinal household study designed to capture economic, social and health information from UK individuals[52]. A subset of 10,484 individuals was selected for genome-wide array genotyping. This cohort was used as a control dataset with SCOOP and STILTS cases (**S1 Table**).

UK BIOBANK (UKBB)

This study includes approximately 487,411 participants with genetic data released (including ~50,000 from the UKBiLEVE cohort [53]) of the total 502,648 individuals from UK BioBank (UKBB). UKBB samples were genotyped on the UK Biobank Axiom array at the Affymetrix Research Services Laboratory in Santa Clara, California, USA and imputed to the Haplotype Reference Consortium (HRC) panel [54]. UKBiLEVE samples were genotyped on the UK BiLEVE array which is a previous

449 version of the UK Biobank Axiom array sharing over 95% of the markers. To date, 487,411 samples
450 with directly genotyped and imputed data are available and data was downloaded using tools
451 provided by UK Biobank. Extensive data from health and lifestyle questionnaires is currently
452 available as well as linked clinical records. BMI, as well as other physical measurements were taken
453 on attendance of recruitment centre. Severely obese participants in the available data were defined
454 as those with BMI ≥ 40 kg/m^2 (N=9,706) and thin individuals were defined as those with BMI ≤ 19
455 kg/m^2 (N=4,538). Given that it has been previously shown that type I error rate for variants with a
456 low minor allele count (MAC) is inadequately controlled for in very unbalanced case-control
457 scenarios[55], we randomly subsampled 35,000 individuals from the original 487,411 genotyped
458 individuals and removed those with BMI≤19 or BMI ≥30, to generate an independent control set.
459 The 25,856 participants remaining after BMI exclusions from the tails, generated a non-extreme set
460 of individuals kept as putative controls (**S2 Fig**). The other 452,411 genotyped samples were kept as
461 the BMI dataset for downstream analyses (**S11 Table, S2 Fig**). An interim release consisting of a
462 subset 152,249 individuals from UKBB was released in May 2015. This interim release was imputed
463 to a combined UK10K and 1000G Phase 3 reference panel and contains several variants which are
464 not currently present in the HRC panel, as such it was used in some of the analyses described.
465
466 **ALSPAC**
467 The Avon Longitudinal Study of Parents and Children (ALSPAC) [27,56], also known as Children of
468 the 90s, is a prospective population-based British birth cohort study. Ethical approval for the study
469 was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics

Committees. Please note that the study website contains details of all the data that is available through a fully searchable data dictionary (http://www.bris.ac.uk/alspac/researchers/data-access/data-dictionary/). Further information about this cohort, including details of the genotyping and imputation procedures, can be found in **S2 Appendix**. This analysis was restricted to a subset of unrelated (identity-by-state < 0.05 [57]) children with genetic data and BMI measured between the age of 12 and 17 years (n=4,964, 48.5% male). The mean age of the children was 14 years and the mean BMI 20.5.

GENOTYPING AND QUALITY CONTROL

SCOOP, STILTS and UKHLS

For the SCOOP cohort, DNA was extracted from whole blood as previously described [31]. For the STILTS cohort, DNA was extracted from saliva using the Oragene saliva DNA kits (online protocol) and quantified using Qubit. All samples from SCOOP, STILTS and UKHLS were typed across 30 SNPs on the Sequenom platform (Sequenom Inc. California, USA) for sample quality control. Of the 3,607 SCOOP and STILTS samples submitted for Sequenom genotyping, 3,280 passed quality controls filters (90.9% pass rate). Of the 10,433 UKHLS samples, 9,965 passed Sequenom sample quality control (95.5% pass rate). Subsequently, UKHLS controls were genotyped on the Illumina HumanCoreExome-12v1-0 Beadchip. The 3,280 SCOOP and STILTS samples, and 48 overlapping UKHLS samples (to test for possible array version effects) were genotyped on the Illumina HumanCoreExome-12v1-1 Beadchip by the Genotyping Facility at the Wellcome Sanger Institute (WSI). Genotype calling was performed centrally for all batches at the WSI using GenCall. Criteria

for excluding samples were as follows: i) concordance against Sequenom genotypes <90%; ii) for each pair of sample duplicates, exclude one with highest missingness; iii) sex inferred from genetic data different from stated sex ; iv) sample call rate <95%; v) sample autosome heterozygosity rate >3 SDS from mean done separately for low (<1%) and high MAF(>1%) bins; vi) magnitude of intensity signal in both channels <90%; and vii) for each pair of related individuals (proportion of IBD (PI_HAT) >0.05), the individual with the lowest call rate was excluded. We performed SNP QC using PLINK v1.07[58]. Criteria for excluding SNPs was: i) Hardy-Weinberg equilibrium (HWE) $p<1x10^{-6}$; ii) Call rate <95% for MAF≥5%, call rate <97% for 1% ≤MAF<5%, and call rate <99% for MAF <1%. SMARTPCA v10210 [59] was used for principal component analysis (PCA). To verify the absence of array version effects we used PCA on the subset of shared controls genotyped on both versions of the array. Cut-offs for samples that diverged from the European cluster were chosen manually after inspecting the PCA plot. SNPs with discordant MAFs in the different versions of the array were excluded. After removal of non-European samples and 13 samples due to cryptic relatedness, 1,456 SCOOP and 1,471 STILTS samples remained for analysis. For UKHLS, 82 samples were removed after applying a strict European filter and 680 related samples were removed after applying a "3^{rd} degree" kinship filter in KING[60]. A total of 9,203 samples remained, of which 6,460 had a BMI >19 and <30 ("controls").

UK BIOBANK

Sample QC was performed using all 487,411 samples. Criteria for excluding samples were as follows: i) supplied and genetically inferred sex mismatches; ii) heterozygosity and missingness

512 outliers according to centrally provided sample QC files; iii) samples not used in kinship estimation
513 by UKBB; iv) individuals that did not identify as "white british" or did not cluster with other "white
514 british" in PCA analysis ; v) samples that withdrew consent and vi) for each pair of related
515 individuals (KING kinship estimate>0.0442), we randomly selected an individual preferentially
516 keeping cases if one related individual is a control. After sample QC, thirteen individuals with
517 underlying health conditions that could influence their BMI were also removed, twelve had BMI<14,
518 and one had BMI>74. In the end, 7,526 obese, 3,532 thin and 20,720 non-extreme controls
519 remained for case-control analyses. In addition, 387,164 samples remained for analysis of BMI as a
520 continuous trait. There is an overlap of 10, 282 samples (~2.6% of the BMI dataset) with obese and
521 thin cases (**S2 Fig**). The same procedure was performed on the interim release of 152,249 UKBB
522 samples to produce a set of 2,799 obese, 1,212 thin, 8,193 controls and 127,672 individuals for the
523 independent BMI dataset. All subsequent analyses on UKBB were also performed on this subset to
524 query variants that are not currently available in the full UKBB release.
525
526 **IMPUTATION AND GENOME-WIDE ASSOCIATION ANALYSES**
527 **SCOOP, STILTS and UKHLS single-variant association analysis**
528 Genotypes from SCOOP, STILTS and UKHLS controls were phased together with SHAPEITv2 [61], and
529 subsequently imputed with IMPUTE2 [62,63] to the merged UK10K and 1000G Phase 3 reference
530 panel [64], containing ~91.3 million autosomal and chromosome X sites, from 6,285 samples. More
531 than 98% of variants with MAF ≥0.5% had an imputation quality score of r^2≥0.4, however variants
532 with MAF <0.1% had a poor imputation quality with only 27% variants with r^2≥0.4 (**S5 Fig**). First-

pass single-variant association tests were done for all variants irrespective of MAF, or imputation quality score (see below). Analyses of 1,456 SCOOP, 1,471 STILTS and 6,460 controls (BMI range 19-30) of European ancestry were based on the frequentist association test, using the EM algorithm, as implemented in SNPTEST v2.5 [65], under an additive model and adjusting for six PCs and sex as covariates.

UKBB BMI dataset single-variant association analysis

For the BMI dataset, we used BOLT-LMM [66] to perform an association analysis with BMI using sex, age, 10 PCs and UKBB genotyping array as covariates.

Heritability estimates and genetic correlation

Summary statistics from the SCOOP vs. UKHLS, STILTS vs. UKHLS, UKBB obese vs controls, UKBB thin vs controls and UKBB BMI analyses were filtered and a subset of 1,197,969 HapMap3 SNPs was kept in each dataset. Using LD score regression [67] we first calculated the heritability of severe childhood obesity (SCOOP vs UKHLS) and persistent thinness (STILTS vs UKHLS). For severe childhood obesity, we estimated a prevalence of 0.15% using the BMI centile equivalent to 3SDS in children [68]. In the case of persistent thinness (BMI<=19), we used a GP based cohort for our prevalence estimates: CALIBER [69]. The CALIBER database consists of 1,173,863 records derived from GP practices. For the heritability analysis, we used a prevalence estimate of 2.8% for BMI<=19 (Claudia Langenberg and Harry Hemingway, personal communication). We also used LD score

553 regression to calculate the genetic correlation of SCOOP with STILTS, SCOOP with UKBB obese,
554 SCOOP with BMI, STILTS with UKBB thin and STILTS with BMI. The genetic correlation between
555 obesity and persistent thinness with anorexia was estimated using the summary statistics from
556 SCOOP vs UKHLS and STILTS vs. UKHLS, and summary statistics available from the Genetic
557 Consortium for Anorexia Nervosa (GCAN) in LD Hub [70]. The same analysis was repeated for UKBB
558 obese vs controls and UKBB thin vs controls. Genetic correlation estimates for BMI vs Overweight,
559 Obesity Class 1, Obesity Class 2 and Obesity Class 3 were also extracted from LD Hub (**S4 Fig**).
560
561 **Comparison with established GIANT BMI associated loci**
562 We obtained the list of 97 established BMI associated loci from the publicly available data from the
563 GIANT consortium [24]. We used this list as we wanted to focus on established common variation in
564 Europeans with accurate effect sizes for simulations. In order to test whether there is evidence of
565 enrichment of nominally significant signals with consistent direction of effect, we performed a
566 binomial test using the subset of signals with nominal significance in the SCOOP vs UKHLS, and
567 STILTS vs UKHLS analyses. Variance explained was calculated using the rms package [71] v4.5.0 in R
568 [72] and Nagelkerke's R^2 is reported. Power calculations were performed using Quanto [73]. To
569 calculate ORs and SE from the ALSPAC BMI summary statistics we used genotype counts from
570 SNPTEST output. We then used a z-test to test for significant differences between the OR calculated
571 using genotype counts of SCOOP and ALSPAC against the SCOOP vs. UKHLS OR.
572
573 **Simulations under an additive model**

We created 10,000 simulations of 1 million individuals for the 97 GIANT BMI loci randomly sampling alleles based on the allele frequency from the sex-combined European dataset reported in Locke *et al*. [24] using an R script. For each simulated genotype, we simulated phenotypes with DISSECT [74] using the effect size in GIANT and then removed all samples from the lower tail where the phenotype was <3SDs to better reproduce the actual BMI distribution. Afterwards we randomly sampled 1,471 individuals from the bottom 2.8% and 1,456 from top 0.15% and compared against a random set of 6,460 controls from the equivalent percentiles to BMI 19-30. Finally, for each of these loci, we calculated the absolute difference between our observed OR and the mean OR from the simulations and counted how many times we saw an equal or larger absolute difference in the simulated data and assigned a p-value. This was done separately for SCOOP vs UKHLS and STILTS vs UKHLS.

Genetic Risk Score

The R package GTX (https://cran.r-project.org/web/packages/gtx/index.html) was used to transpose genotype probabilities into dosages, and a combined dosage score, weighted by the effect size from GIANT, for 97 BMI SNPs [24] was calculated and standardised. We checked whether there was an ordinal relationship between the genetic risk score and BMI category (i.e. thin, normal, or obese) using ordinal logistic regression with the clm function in the ordinal R package. While the assumption of equal variance appears to hold (**S6 Fig**)**,** the proportional odds assumption indicating equal odds between thin, normal, and obese groups is violated for the BMI genetic risk score and some of the principal component covariates (i.e., PC2, PC3, and PC6). As our primary

model, we ran a partial proportional odds model adjusting for PC1, PC4, and PC5 and allowing the BMI genetic score, PC2, PC3, and PC6 to vary between BMI category. To check for consistency, we ran a partial proportional odds model adjusting for the first six PCs and allowing only the BMI genetic score to vary between BMI group and a full proportional odds model allowing all six PCs and the BMI genetic score to vary between BMI group (**S1 Appendix**). Using ANOVA, we formally tested the proportional odds assumption for the BMI genetic risk score. A genetic risk score was created and an ordinal logistic regression was run for each of the 10,000 simulations. We compared the observed test statistic testing whether the odds were the same by BMI category to the 10,000 simulation test statistics. We calculated the p-value as the number of simulations with a test statistic larger than that observed in the real data. A mean genetic risk score was also calculated for each BMI category (obese, thin and controls) across the 10,000 simulations. A t-test was used to test whether the mean observed GRS score in each category was significantly different from the one estimated using the simulations.

Discovery stage GWAS

First pass single-variant association analyses results were used as discovery datasets for the GWAS. After association analysis, we removed variants with MAF<0.5%, an INFO score <0.4, and HWE $p<1x10^{-6}$, as these highlighted regions of the genome that were problematic, including CNV regions with poor imputation quality. Quantile-quantile plots indicated that the genomic inflation was well controlled for in SCOOP-UKHLS (λ=1.06) and STILTS-UKHLS (λ =1.04), and slightly higher for SCOOP-STILTS (λ =1.08**, S7 Fig**). We used LD score regression [67] to correct for inflation not due to

polygenicity. To identify distinct loci, we performed clumping as implemented in PLINK [58] using summary statistics from the association tests and LD information from the imputed data, clumping variants 250kb away from an index variant and with an r^2>0.1. In order to further identify a set of likely independent signals we performed conditional analysis of the lead SNPs in SNPTEST to take into account long-range LD. A total of 135 autosomal variants with $p<1x10^{-5}$ in any of the three case-control analyses were taken forward for replication in UKBB. All case-control results are reported with the lower BMI group as reference.

UKBB association analysis

We tested 1,208,692 SNPs for association under an additive model in SNPTEST using sex, age, 10 PCs and UKBB genotyping array as covariates. Three comparisons were done: obese vs thin, obese vs controls and controls vs thin. Variants with an INFO score <0.4, HWE $p<1x10^{-6}$ were filtered out from the results. Inflation factors were calculated using HapMap markers. The LD score regression intercepts were 1.0074 in obese vs thin, 1.0057 in obese vs controls and 1.009 in thin vs controls. We used all thin individuals, regardless of health status, as our replication cohort to maximize power. However, using ICD10 codes and self-reported illness data (**Tables S12 and S13**) to remove individuals who had a relevant medical diagnosis before date of attendance at UKBB recruitment centre, yielded 2,518 thin individuals and materially equivalent results (**S8 Fig**).

GIANT, EGG and SCOOP 2013 summary statistics

We obtained summary statistics for the GIANT Extremes obesity meta-analysis [20] from http://portals.broadinstitute.org/collaboration/giant/index.php/GIANT_consortium_data_files. Summary statistics for EGG [30] were obtained from http://egg-consortium.org/childhood-obesity.html. We used summary statistics from our previous study of 1,509 early-onset obesity SCOOP cases compared to 5,380 publicly available WTCCC2 controls (SCOOP 2013) [31]. Data for the SCOOP cases is available to download from the European Genome-Phenome Archive (EGA) using accession number EGAD00010000594. The control samples are available to download using accession numbers EGAD00000000021 and EGAD00000000023. These replication studies are largely non-overlapping with our discovery datasets and each-other. When a lead variant was not available in a replication cohort, a proxy ($r^2 \geq 0.8$) was used in the meta-analysis.

Replication meta-analysis

We meta-analysed summary statistics for the 135 variants reaching $p<1x10^{-5}$ in SCOOP/STILTS/UKHLS with the corresponding results from UKBB and study specific replication cohorts (**Tables S5-S7**). For obese vs. thin and obese vs. controls comparisons we used fixed-effects meta-analysis correcting for unknown sample overlap in replication cohorts using METACARPA [75]. For thin vs. controls we used a fixed-effects meta-analysis in METAL [76]. Heterogeneity was assessed using Cochran's Q-test heterogeneity p-value in METAL. A signal was considered to replicate if it met all the following criteria: i) consistent direction of effect; ii) $p<0.05$ in at least one replication cohort; and iii) the meta-analysis p-value reached standard genome-wide significance ($p<5x10^{-8}$). Given that we are querying additional variants on the lower allele frequency spectrum,

one could also use a more strict genome-wide significance threshold taking into account the increased number of tests ($p \leq 1.17 \times 10^{-8}$) [77]. In practice, this only affected one previously established signal (*SULT1A1*, rs3760091) in our obese vs. controls analysis that fell just below this threshold (**S6 Table**). rs4440960 was later removed from final results (SCOOP vs UKHLS and STILTS vs UKHLS) after close examination revealed it was present in a CNV region with poor imputation quality.

Comparison of newly established candidate loci and UKBB independent BMI dataset

We identified eleven signals in SCOOP vs STILTS, nine in SCOOP vs UKHLS and two in UKHLS vs STILTS that were nominally significant in the UKBB BMI dataset GWAS, and directionally consistent. A binomial test was used to check for enrichment of signals with consistent direction of effect (**S9 Table**).

Lookup of previously identified obesity-related signals in our discovery datasets

We took all signals reaching genome-wide significance, or identified for the first time in the GIANT Extremes obesity meta-analysis [20], with either the tails of BMI or obesity classes, and in childhood obesity studies [30,31] and performed look-up of those signals in all three of our discovery analyses (SCOOP vs STILTS, SCOOP vs UKHLS and UKHLS vs STILTS). ORs and p-values from the previous studies and look-up results from our discovery datasets are reported in **S10 Table**.

677 **Data availability**

678 Summary statistics for the discovery analyses will be available to download from EGA
679 (EGAS00001002624). UKHLS data is available for download in EGA with accession code
680 EGAS00001001232.

Appendix A

Table 1

Obese vs. thin						Discovery cohort										Replication cohorts					Combined analysis		
rsID	Nearest gene	Chr.	Position (bp)	EA	NEA	OR (95% CI)	P value	EAF Ob	EAF Th	Cohort	OR (95% CI)	P value	EAF Ob	EAF Th	OR (95% CI)	P value	HetPVal
rs9930333	*FTO*	16	53799977	G	T	1.70(1.52,1.90)	2.30E-20	49.59%	37.46%	UKBB	1.46(1.38,1.55)	3.60E-36	48.26%	38.93%	1.48(1.42,1.54)	8.52E-76	3.34E-02
										GIANT[k]	1.43(1.34,1.54)	8.10E-25					
rs2168711	*MC4R*	18	57848531	C	T	1.66(1.45,1.89)	8.29E-14	28.90%	19.95%	UKBB	1.23(1.15,1.32)	2.19E-09	26.75%	22.90%	1.27(1.21,1.33)	2.02E-21	1.12E-04
										GIANT[k]	1.20(1.10,1.30)	1.80E-05					
rs6748821	*TMEM18*[a]	2	629601	G	A	1.65(1.42,1.91)	9.45E-11	86.69%	79.84%	UKBB	1.27(1.18,1.37)	1.31E-09	85.00%	81.69%	1.32(1.24,1.39)	7.76E-21	2.81E-03
										GIANT[k]	1.26(1.14,1.39)	9.90E-06					
rs506589	*SEC16B*	1	177894287	C	T	1.46(1.27,1.67)	5.42E-08	23.98%	18.07%	UKBB	1.25(1.17,1.35)	5.44E-10	23.11%	19.16%	1.28(1.21,1.35)	3.14E-20	1.21E-01
										GIANT[k]	1.25(1.14,1.37)	2.70E-06					
rs6738433	*ADCY3-DNAJC27*[b]	2	25159501	C	G	1.43(1.28,1.60)	1.71E-10	47.31%	43.92%	UKBB	1.21(1.14,1.28)	2.74E-10	50.70%	45.96%	1.19(1.14,1.24)	3.19E-17	6.25E-03
										GIANT[k]	1.10(1.03,1.17)	5.70E-03					
rs7132908	*FAIM2*	12	50263148	A	G	1.31(1.17,1.47)	2.26E-06	42.45%	36.27%	UKBB	1.18(1.11,1.25)	5.43E-08	41.11%	37.39%	1.20(1.15,1.26)	1.93E-16	2.52E-01
										GIANT[k]	1.20(1.10,1.30)	6.60E-06					
rs62107261	*FAM150B*	2	422144	T	C	2.37(1.75,3.20)	2.07E-08	96.37%	93.38%	UKBB	1.54(1.35,1.76)	3.57E-10	96.28%	94.36%	1.65(1.46,1.87)	1.15E-15	1.07E-02
rs12507026	*GNPDA2*[c]	4	45181334	T	A	1.30(1.17,1.46)	3.69E-06	47.29%	40.92%	UKBB	1.14(1.08,1.21)	8.76E-06	45.30%	41.98%	1.18(1.13,1.23)	5.53E-15	4.06E-02
										GIANT[k]	1.20(1.12,1.28)	3.10E-07					
rs75398113	*SNRPC*	6	34728071	C	A	1.53(1.27,1.85)	8.91E-06	11.95%	8.04%	UKBB	1.24(1.12,1.37)	2.07E-05	10.47%	8.52%	1.30(1.19,1.42)	5.19E-09	5.56E-02
rs13135092	*SLC39A8*	4	103198082	G	A	1.58(1.30,1.93)	4.70E-06	10.50%	7.24%	UKBB	1.25(1.12,1.39)	5.57E-05	9.24%	7.52%	1.32(1.20,1.45)	1.06E-08	3.59E-02
Obese vs. controls																	
rsID	Nearest gene	Chr.	Position (bp)	EA	NEA	OR (95% CI)	P value	EAF Ob	EAF Co	Cohort	OR (95% CI)	P value	EAF Ob	EAF Co	OR (95% CI)	P value	HetPVal

Appendix A

rs9928094	*FTO*	16	53799905	G	A	1.44(1.33,1.57)	1.42E-18	49.50%	41.32%	UKBB	1.30(1.25,1.35)	2.74E-41	48.34%	41.91%	1.32(1.29,1.36)	5.94E-101	4.41E-05
										SCOOP 2013	1.46(1.34,1.60)	4.88E-17					
										EGG	1.21(1.15,1.28)	7.20E-13					
										GIANT[i]	1.43(1.34,1.54)	6.60E-25					
rs35614134	*MC4R*[d]	18	57832856	AC	A	1.31(1.20,1.44)	6.27E-09	29.01%	23.69%	UKBB	1.22(1.16,1.27)	1.25E-18	26.72%	23.15%	1.23(1.20,1.27)	1.57E-43	3.55E-01
										SCOOP 2013	1.32(1.19,1.46)	1.22E-07					
										EGG	1.22(1.15,1.30)	1.27E-10					
										GIANT[i]	1.20(1.10,1.30)	1.70E-05					
rs66906321	*TMEM18*[e]	2	630070	C	T	1.40(1.24,1.57)	2.35E-08	85.78%	81.35%	UKBB	1.17(1.11,1.24)	3.44E-09	84.44%	82.20%	1.25(1.21,1.29)	9.72E-35	1.33E-02
										SCOOP 2013	1.39(1.24,1.57)	6.65E-08					
										EGG	1.28(1.19,1.37)	5.15E-12					
										GIANT[i]	1.27(1.15,1.40)	3.40E-06					
rs7132908	*FAIM2*[f]	12	50263148	A	G	1.22(1.12,1.32)	3.27E-06	42.45%	37.82%	UKBB	1.15(1.10,1.19)	5.37E-12	41.11%	37.71%	1.17(1.14,1.21)	2.38E-31	4.86E-01
										SCOOP 2013	1.23(1.12,1.35)	8.89E-06					
										EGG	1.18(1.11,1.25)	1.24E-08					
										GIANT[i]	1.20(1.10,1.30)	6.60E-06					
rs2384060	*ADCY3-DNAJC27*[g]	2	25135438	G	A	1.23(1.13,1.34)	1.53E-06	43.52%	38.90%	UKBB	1.11(1.07,1.15)	4.89E-08	47.67%	44.93%	1.14(1.11,1.17)	9.39E-23	1.13E-01
										SCOOP 2013	1.09(1.00,1.19)	5.01XE-02					
										EGG	1.18(1.12,1.24)	1.02E-09					
										GIANT[i]	1.12(1.04,1.19)	1.60E-03					
rs11209947	*NEGR1*[h]	1	72808551	A	T	1.30(1.17,1.44)	8.51E-07	76.58%	72.18%	UKBB	1.11(1.05,1.16)	4.53E-05	81.18%	79.76%	1.17(1.13,1.21)	5.17E-20	7.26E-05
										SCOOP 2013	1.46(1.30,1.63)	2.21E-10					
										EGG	1.13(1.06,1.22)	4.60E-04					
										GIANT[i]	1.22(1.11,1.35)	5.60E-05					
rs12735657	*SEC16B*[i]	1	177809133	C	T	1.24(1.13,1.37)	9.72E-06	24.26%	20.46%	UKBB	1.12(1.07,1.17)	1.48E-06	22.87%	20.94%	1.15(1.12,1.19)	7.26E-19	1.79E-01

										SCOOP 2013	1.20(1.07,1.33)	1.18E-03					
										EGG	1.14(1.06,1.21)	1.52E-04					
										GIANT[i]	1.22(1.11,1.34)	1.80E-05					
rs13104545	*GNPDA2*	4	45184907	A	G	1.27(1.15,1.40)	1.61E-06	27.41%	23.45%	UKBB	1.07(1.02,1.11)	5.35E-03	24.36%	23.26%	1.13(1.09,1.17)	1.47E-11	9.39E-05
										EGG	1.13(1.04,1.22)	3.39E-03					
										GIANT[i]	1.34(1.20,1.49)	1.20E-07					
rs112446794	***CEP120***[j]	5	122665465	T	C	1.23(1.13,1.35)	2.08E-06	33.15%	28.69%	UKBB	1.07(1.02,1.11)	2.55E-03	29.47%	28.21%	1.09(1.06,1.13)	3.45E-10	3.33E-02
										SCOOP 2013	1.08(0.98,1.19)	1.38E-01					
										EGG	1.12(1.06,1.18)	1.22E-04					
										GIANT[i]	1.05(0.97,1.13)	2.40E-01					
rs3760091	*SULT1A1*	16	28620800	C	G	1.24(1.14,1.35)	1.56E-06	64.89%	60.23%	UKBB	1.09(1.04,1.14)	1.19E-04	63.49%	61.44%	1.12(1.07,1.16)	2.65E-08	8.49E-03

Acknowledgements

We are indebted to the participants of the STILTS cohort and the patients and families involved in the Genetics of Obesity Study (GOOS) cohort. We thank the staff of the NIHR Primary Care Research Network, the GPs, Physicians and nurses involved in identifying and recruiting participants to STILTS and GOOS. These data are from Understanding Society: The UK Household Longitudinal Study, which is led by the Institute for Social and Economic Research at the University of Essex and funded by the Economic and Social Research Council. The data were collected by NatCen and the genome wide scan data were analysed by the Wellcome Sanger Institute. The Understanding Society DAC have an application system for genetics data and all use of the data should be approved by them. This research has been conducted using the UK Biobank Resource (Application Number 14069). Data on the childhood obesity trait has been contributed by EGG Consortium and has been downloaded from www.egg-consortium.org. We are extremely grateful to all the families who took part in the ALSPAC study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. The data used for the analyses described in this manuscript were obtained from the GTEx Portal on 08/Feb/2018. The authors would like to thank Emma Gray, Michelle Dignam and staff of the WSI Sample Management and Genotyping facilities for their contribution, as well as, Konstantinos Hatzikotoulas and Ioanna Tachmazidou for their assistance in the QC of UK Biobank data. Understanding Society Scientific Group members: Michaela Benzeval[1], Jonathan Burton[1], Nicholas Buck[1], Annette Jäckle[1], Meena

Kumari[1], Heather Laurie[1], Peter Lynn[1], Stephen Pudney[1], Birgitta Rabe[1], Dieter Wolke[2]. 1) Institute for Social and Economic Research, University of Essex, UK; 2) University of Warwick, UK.

References

1. Ogden CL, Carroll MD, Flegal KM (2014) Prevalence of obesity in the United States. JAMA 312: 189-190.
2. Wardle J, Carnell S, Haworth CM, Plomin R (2008) Evidence for a strong genetic influence on childhood adiposity despite the force of the obesogenic environment. Am J Clin Nutr 87: 398-404.
3. Silventoinen K, Magnusson PK, Tynelius P, Kaprio J, Rasmussen F (2008) Heritability of body size and muscle strength in young adulthood: a study of one million Swedish men. Genet Epidemiol 32: 341-349.
4. Allison DB, Kaprio J, Korkeila M, Koskenvuo M, Neale MC, et al. (1996) The heritability of body mass index among an international sample of monozygotic twins reared apart. Int J Obes Relat Metab Disord 20: 501-506.
5. Akiyama M, Okada Y, Kanai M, Takahashi A, Momozawa Y, et al. (2017) Genome-wide association study identifies 112 new loci for body mass index in the Japanese population. Nat Genet 49: 1458-1467.
6. Grarup N, Moltke I, Andersen MK, Dalby M, Vitting-Seerup K, et al. (2018) Loss-of-function variants in ADCY3 increase risk of obesity and type 2 diabetes. Nat Genet 50: 172-174.
7. Justice AE, Winkler TW, Feitosa MF, Graff M, Fisher VA, et al. (2017) Genome-wide meta-analysis of 241,258 adults accounting for smoking behaviour identifies novel loci for obesity traits. Nat Commun 8: 14977.
8. Minster RL, Hawley NL, Su CT, Sun G, Kershaw EE, et al. (2016) A thrifty variant in CREBRF strongly influences body mass index in Samoans. Nat Genet 48: 1049-1054.
9. Pigeyre M, Yazdi FT, Kaur Y, Meyre D (2016) Recent progress in genetics, epigenetics and metagenomics unveils the pathophysiology of human obesity. Clin Sci (Lond) 130: 943-986.
10. Turcot V, Lu Y, Highland HM, Schurmann C, Justice AE, et al. (2018) Protein-altering variants associated with body mass index implicate pathways that control energy intake and expenditure in obesity. Nat Genet 50: 26-41.
11. Ramachandrappa S, Raimondo A, Cali AM, Keogh JM, Henning E, et al. (2013) Rare variants in single-minded 1 (SIM1) are associated with severe obesity. J Clin Invest 123: 3042-3050.
12. Doche ME, Bochukova EG, Su HW, Pearce LR, Keogh JM, et al. (2012) Human SH2B1 mutations are associated with maladaptive behaviors and obesity. J Clin Invest 122: 4732-4736.
13. O'Rahilly S, Farooqi IS (2008) Human obesity: a heritable neurobehavioral disorder that is highly sensitive to environmental conditions. Diabetes 57: 2905-2910.
14. Saeed S, Bonnefond A, Tamanini F, Mirza MU, Manzoor J, et al. (2018) Loss-of-function mutations in ADCY3 cause monogenic severe obesity. Nat Genet 50: 175-179.
15. Bulik CM, Allison DB (2001) The genetic epidemiology of thinness. Obes Rev 2: 107-115.
16. Costanzo PR, Schiffman SS (1989) Thinness--not obesity--has a genetic component. Neurosci Biobehav Rev 13: 55-58.

17. Magnusson PK, Rasmussen F (2002) Familial resemblance of body mass index and familial risk of high and low body mass index. A study of young men in Sweden. Int J Obes Relat Metab Disord 26: 1225-1231.
18. Laskarzewski PM, Khoury P, Morrison JA, Kelly K, Mellies MJ, et al. (1983) Familial obesity and leanness. Int J Obes 7: 505-527.
19. Whitaker KL, Jarvis MJ, Boniface D, Wardle J (2011) The intergenerational transmission of thinness. Arch Pediatr Adolesc Med 165: 900-905.
20. Berndt SI, Gustafsson S, Magi R, Ganna A, Wheeler E, et al. (2013) Genome-wide meta-analysis identifies 11 new loci for anthropometric traits and provides insights into genetic architecture. Nat Genet 45: 501-512.
21. Hinney A, Nguyen TT, Scherag A, Friedel S, Bronner G, et al. (2007) 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.
22. Scannell Bryan M, Argos M, Pierce B, Tong L, Rakibuz-Zaman M, et al. (2014) Genome-wide association studies and heritability estimates of body mass index related phenotypes in Bangladeshi adults. PLoS One 9: e105062.
23. Lee SH, Wray NR, Goddard ME, Visscher PM (2011) Estimating missing heritability for disease from genome-wide association studies. Am J Hum Genet 88: 294-305.
24. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, et al. (2015) Genetic studies of body mass index yield new insights for obesity biology. Nature 518: 197-206.
25. Yan X, Wang Z, Schmidt V, Gauert A, Willnow TE, et al. (2018) Cadm2 regulates body weight and energy homeostasis in mice. Mol Metab 8: 180-188.
26. Klimentidis YC, Raichlen DA, Bea J, Garcia DO, Wineinger NE, et al. (2018) Genome-wide association study of habitual physical activity in over 377,000 UK Biobank participants identifies multiple variants including CADM2 and APOE. Int J Obes (Lond).
27. Boyd A, Golding J, Macleod J, Lawlor DA, Fraser A, et al. (2013) Cohort Profile: the 'children of the 90s'--the index offspring of the Avon Longitudinal Study of Parents and Children. Int J Epidemiol 42: 111-127.
28. Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, et al. (2015) An atlas of genetic correlations across human diseases and traits. Nat Genet 47: 1236-1241.
29. Winkler TW, Kutalik Z, Gorski M, Lottaz C, Kronenberg F, et al. (2015) EasyStrata: evaluation and visualization of stratified genome-wide association meta-analysis data. Bioinformatics 31: 259-261.
30. Bradfield JP, Taal HR, Timpson NJ, Scherag A, Lecoeur C, et al. (2012) A genome-wide association meta-analysis identifies new childhood obesity loci. Nat Genet 44: 526-531.
31. Wheeler E, Huang N, Bochukova EG, Keogh JM, Lindsay S, et al. (2013) Genome-wide SNP and CNV analysis identifies common and low-frequency variants associated with severe early-onset obesity. Nat Genet 45: 513-517.
32. Tachmazidou I, Suveges D, Min JL, Ritchie GRS, Steinberg J, et al. (2017) Whole-Genome Sequencing Coupled to Imputation Discovers Genetic Signals for Anthropometric Traits. Am J Hum Genet 100: 865-884.
33. Hinney A, Volckmar AL, Knoll N (2013) Melanocortin-4 receptor in energy homeostasis and obesity pathogenesis. Prog Mol Biol Transl Sci 114: 147-191.
34. Geller F, Reichwald K, Dempfle A, Illig T, Vollmert C, et al. (2004) Melanocortin-4 receptor gene variant I103 is negatively associated with obesity. Am J Hum Genet 74: 572-581.
35. Mitchell JA, Hakonarson H, Rebbeck TR, Grant SF (2013) Obesity-susceptibility loci and the tails of the pediatric BMI distribution. Obesity (Silver Spring) 21: 1256-1260.

36. Beyerlein A, von Kries R, Ness AR, Ong KK (2011) Genetic markers of obesity risk: stronger associations with body composition in overweight compared to normal-weight children. PLoS One 6: e19057.
37. Chan Y, Holmen OL, Dauber A, Vatten L, Havulinna AS, et al. (2011) Common variants show predicted polygenic effects on height in the tails of the distribution, except in extremely short individuals. PLoS Genet 7: e1002439.
38. Young AI, Wauthier F, Donnelly P (2016) Multiple novel gene-by-environment interactions modify the effect of FTO variants on body mass index. Nat Commun 7: 12724.
39. Winkler TW, Justice AE, Graff M, Barata L, Feitosa MF, et al. (2015) The Influence of Age and Sex on Genetic Associations with Adult Body Size and Shape: A Large-Scale Genome-Wide Interaction Study. PLoS Genet 11: e1005378.
40. Qi Q, Kilpelainen TO, Downer MK, Tanaka T, Smith CE, et al. (2014) FTO genetic variants, dietary intake and body mass index: insights from 177,330 individuals. Hum Mol Genet 23: 6961-6972.
41. Bjornland T, Langaas M, Grill V, Mostad IL (2017) Assessing gene-environment interaction effects of FTO, MC4R and lifestyle factors on obesity using an extreme phenotype sampling design: Results from the HUNT study. PLoS One 12: e0175071.
42. Wood AR, Tyrrell J, Beaumont R, Jones SE, Tuke MA, et al. (2016) Variants in the FTO and CDKAL1 loci have recessive effects on risk of obesity and type 2 diabetes, respectively. Diabetologia 59: 1214-1221.
43. Shungin D, Winkler TW, Croteau-Chonka DC, Ferreira T, Locke AE, et al. (2015) New genetic loci link adipose and insulin biology to body fat distribution. Nature 518: 187-196.
44. Lango Allen H, Estrada K, Lettre G, Berndt SI, Weedon MN, et al. (2010) Hundreds of variants clustered in genomic loci and biological pathways affect human height. Nature 467: 832-838.
45. Wen W, Kato N, Hwang JY, Guo X, Tabara Y, et al. (2016) Genome-wide association studies in East Asians identify new loci for waist-hip ratio and waist circumference. Sci Rep 6: 17958.
46. Shaheen R, Schmidts M, Faqeih E, Hashem A, Lausch E, et al. (2015) A founder CEP120 mutation in Jeune asphyxiating thoracic dystrophy expands the role of centriolar proteins in skeletal ciliopathies. Hum Mol Genet 24: 1410-1419.
47. Roosing S, Romani M, Isrie M, Rosti RO, Micalizzi A, et al. (2016) Mutations in CEP120 cause Joubert syndrome as well as complex ciliopathy phenotypes. J Med Genet 53: 608-615.
48. Morton NM, Nelson YB, Michailidou Z, Di Rollo EM, Ramage L, et al. (2011) A stratified transcriptomics analysis of polygenic fat and lean mouse adipose tissues identifies novel candidate obesity genes. PLoS One 6: e23944.
49. Simoncic M, Horvat S, Stevenson PL, Bunger L, Holmes MC, et al. (2008) Divergent physical activity and novel alternative responses to high fat feeding in polygenic fat and lean mice. Behav Genet 38: 292-300.
50. Morgan JF, Reid F, Lacey JH (1999) The SCOFF questionnaire: assessment of a new screening tool for eating disorders. BMJ 319: 1467-1468.
51. Bochukova EG, Huang N, Keogh J, Henning E, Purmann C, et al. (2010) Large, rare chromosomal deletions associated with severe early-onset obesity. Nature 463: 666-670.
52. University of Essex. Institute for Social and Economic Research and NatCen Social Research Understanding Society: Waves 1-5, 2009-2014 [computer file]. 7th Edition. Colchester, Essex: UK Data Archive [distributor] November 2015 SN: 6614.
53. Wain LV, Shrine N, Miller S, Jackson VE, Ntalla I, et al. (2015) Novel insights into the genetics of smoking behaviour, lung function, and chronic obstructive pulmonary disease (UK BiLEVE): a genetic association study in UK Biobank. Lancet Respir Med 3: 769-781.

54. McCarthy S, Das S, Kretzschmar W, Delaneau O, Wood AR, et al. (2016) A reference panel of 64,976 haplotypes for genotype imputation. Nat Genet 48: 1279-1283.
55. Ma C, Blackwell T, Boehnke M, Scott LJ, Go TDi (2013) Recommended joint and meta-analysis strategies for case-control association testing of single low-count variants. Genet Epidemiol 37: 539-550.
56. Fraser A, Macdonald-Wallis C, Tilling K, Boyd A, Golding J, et al. (2013) Cohort Profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. Int J Epidemiol 42: 97-110.
57. Yang J, Lee SH, Goddard ME, Visscher PM (2011) GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet 88: 76-82.
58. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81: 559-575.
59. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, et al. (2006) Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 38: 904-909.
60. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, et al. (2010) Robust relationship inference in genome-wide association studies. Bioinformatics 26: 2867-2873.
61. Delaneau O, Marchini J, Zagury JF (2011) A linear complexity phasing method for thousands of genomes. Nat Methods 9: 179-181.
62. Howie BN, Donnelly P, Marchini J (2009) A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genet 5: e1000529.
63. Howie B, Marchini J, Stephens M (2011) Genotype imputation with thousands of genomes. G3 (Bethesda) 1: 457-470.
64. Huang J, Howie B, McCarthy S, Memari Y, Walter K, et al. (2015) Improved imputation of low-frequency and rare variants using the UK10K haplotype reference panel. Nat Commun 6: 8111.
65. Marchini J, Howie B, Myers S, McVean G, Donnelly P (2007) A new multipoint method for genome-wide association studies by imputation of genotypes. Nat Genet 39: 906-913.
66. Loh PR, Tucker G, Bulik-Sullivan BK, Vilhjalmsson BJ, Finucane HK, et al. (2015) Efficient Bayesian mixed-model analysis increases association power in large cohorts. Nat Genet 47: 284-290.
67. Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang J, et al. (2015) LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet 47: 291-295.
68. Felix JF, Bradfield JP, Monnereau C, van der Valk RJ, Stergiakouli E, et al. (2016) Genome-wide association analysis identifies three new susceptibility loci for childhood body mass index. Hum Mol Genet 25: 389-403.
69. Denaxas SC, George J, Herrett E, Shah AD, Kalra D, et al. (2012) Data resource profile: cardiovascular disease research using linked bespoke studies and electronic health records (CALIBER). Int J Epidemiol 41: 1625-1638.
70. Zheng J, Erzurumluoglu AM, Elsworth BL, Kemp JP, Howe L, et al. (2017) LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. Bioinformatics 33: 272-279.
71. Harrell FE rms: R functions for biostatistical/epidemiologic modeling, testing, estimation, validation, graphics, prediction, and typesetting by storing enhanced model design attributes in the fit, 2013. Implements methods in *Regression Modeling Strategies*, New York:Springer, 2001.
72. R Development Core Team (2011) R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.
73. Gauderman W, Morrison J (2006) QUANTO 1.1: A computer program for power and sample size calculations for genetic-epidemiology studies, http://hydra.usc.edu/gxe.
74. Canela-Xandri O, Law A, Gray A, Woolliams JA, Tenesa A (2015) A new tool called DISSECT for analysing large genomic data sets using a Big Data approach. Nat Commun 6: 10162.

75. Southam L, Gilly A, Suveges D, Farmaki AE, Schwartzentruber J, et al. (2017) Whole genome sequencing and imputation in isolated populations identify genetic associations with medically-relevant complex traits. Nat Commun 8: 15606.
76. Willer CJ, Li Y, Abecasis GR (2010) METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 26: 2190-2191.
77. Xu C, Tachmazidou I, Walter K, Ciampi A, Zeggini E, et al. (2014) Estimating genome-wide significance for whole-genome sequencing studies. Genet Epidemiol 38: 281-290.

Supporting information captions

S1 Appendix. **Assessing equal vs. unequal effects for the genetic risk score.**

S2 Appendix. **The Avon Longitudinal Study of Parents and Children.**

S1 Fig. **Mean GRS for SCOOP and STILTS compared to simulations**. Histogram represents mean GRS scores for each BMI category across 10,000 simulations. Vertical red line highlights the observed value in real data. p=p-value of difference**.**

S2 Fig. **Summary of the UKBB sample sets after QC.** Venn Diagram showing sample numbers and overlap between UKBB sample sets used in genetic correlation (BMI dataset) and GWAS replication (obese, controls, thin) analyses.

S3 Fig. Manhattan plot of SCOOP vs STILTS. Manhattan plot produced in EasyStrata, red line indicates genome-wide significance threshold at p=5x10-08. Orange line indicates discovery significance threshold at p=1x10-05. Black labels highlight known BMI/obesity loci that were taken forward for replication and yellow peaks indicate those that met genome-wide significance after replication.

S4 Fig. Genetic correlation of traits and BMI. Genetic correlation estimates and 95% CI for severe early-onset childhood obesity (SCOOP), healthy persistent thinness (STILTS), Obesity Class 3, Obesity Class 2, Obesity Class 1 and Overweight. Dotted lines represent complete genetic correlation.

S5 Fig. Quality of UK10K+1000G imputed genotypes. Percentage of variants with INFO score (r^2)>0.4, as derived from the IMPUTE2 imputation algorithm, stratified by minor allele frequency across all samples (SCOOP, STILTS and UKHLS).

S6 Fig. Box and density plots of risk score weighted by effect size for 97 BMI associated SNPs from GIANT. A weighted genetic risk score for each individual was obtained by summing genotype dosages multiplied by the effect (beta) estimates from GIANT for each of the 97 SNPs. To check the equal variance assumption, we used a box plot (left) and density plot (right). Density plot: Green = STILTS; Blue = UKHLS; Red = SCOOP.

S7 Fig. Quantile-quantile plots of three discovery analysis cohorts. Q-Q plots of LD Score Regression-corrected p-values for the three analysis cohorts used for the discovery analysis, produced in EasyStrata. Red=SCOOP vs. STILTS; Black=SCOOP vs. UKHLS, Blue=STILTS vs. UKHLS. Variants passing QC and with MAF >=0.5% are shown. LD Score regression intercept (λLD) values before correction are shown for each analysis.

S8 Fig. Quantile-quantile plots for UKBB case-control analysis with different exclusion criteria for thin individuals. Q-Q plot using all thin individuals as cases (Full UKBB) and removing individuals based on ICD10 and self-reported data (ICD10+self-reported filter). Correlation for –log10 p-values is shown (r=0.7462).

S1 Table. Summary of discovery sample sets.

S2 Table. 97 BMI SNPs from the GIANT consortium study and their summary statistics in our three analysis cohorts.

S3 Table. Nominally significant loci for non-additive effect in extremes.

S4 Table. Difference in SCOOP OR when using ALSPAC as control dataset vs. UKHLS.

S5 Table. Discovery, replication and meta-analysis results for 32 SNPs meeting P<10-5 in discovery association results of SCOOP vs STILTS analysis.

S6 Table. Discovery, replication and meta-analysis results for 66 SNPs meeting P<10-5 in discovery association results of SCOOP vs UKHLS analysis.

S7 Table. Discovery, replication and meta-analysis results for 37 SNPs meeting P<10-5 in discovery association results of UKHLS vs STILTS analysis.

S8 Table: Reciprocal analysis of previously established signals and lead signals in this study.

S9 Table. Consistency of the direction of effect in candidate loci meeting p<1x10-5 in the discovery stages with BMI dataset GWAS.

S10 Table. Published loci from GIANT, EGG and SCOOP 2013 not reaching genome-wide significance in our study

S11 Table. Summary of UKBB sample sets.

S12 Table. ICD10 codes used to exclude thin individuals in UKBB

S13 Table. Self-reported illness codes used to exclude thin individuals in UKBB

S2 Table. 97 BMI SNPs from the GIANT consortium study and their summary statistics in our three analysis cohorts.

rsID	Chr.	Position (bp)*	Nearest gene(s)	GIANT						SCOOP vs. STILTS							SCOOP vs. UKHLS							UKHLS vs. STILTS						
				Effect	Other	EAF	Beta	SE	P value	Effect	Other	EAF	OR	95% CI	P value	Direction	Effect	Other	EAF	OR	95% CI	P value	Direction	Effect	Other	EAF	OR	95% CI	P value	Direction
rs1558902	16	53803574	*FTO*	A	T	0.415	0.082	0.003	7.51E-153	A	T	0.419	1.689	(1.509,1.89)	7.31E-20	+	A	T	0.410	1.429	(1.317,1.55)	1.25E-17	+	A	T	0.387	1.174	(1.077,1.28)	2.78E-04	+
rs6567160	18	57829135	*MC4R*	C	T	0.236	0.056	0.004	3.93E-53	C	T	0.244	1.643	(1.44,1.875)	1.77E-13	+	C	T	0.246	1.308	(1.194,1.433)	7.91E-09	+	C	T	0.229	1.257	(1.134,1.394)	1.38E-05	+
rs13021737	2	632348	*TMEM18*	G	A	0.828	0.060	0.004	1.113E-50	G	A	0.832	1.586	(1.365,1.844)	1.81E-09	+	G	A	0.832	1.356	(1.206,1.526)	3.89E-07	+	G	A	0.820	1.211	(1.088,1.347)	4.44E-04	+
rs10938397	4	45182527	*GNPDA2*	G	A	0.434	0.040	0.003	3.205E-38	G	A	0.442	1.302	(1.164,1.457)	4.19E-06	+	G	A	0.440	1.186	(1.093,1.287)	4.50E-05	+	G	A	0.429	1.084	(0.996,1.179)	6.24E-02	+
rs543874	1	177889480	*SEC16B*	G	A	0.193	0.048	0.004	2.618E-35	G	A	0.210	1.452	(1.267,1.664)	8.57E-08	+	G	A	0.213	1.201	(1.09,1.324)	2.22E-04	+	G	A	0.202	1.177	(1.056,1.311)	3.11E-03	+
rs2207139	6	50845490	*TFAP2B*	G	A	0.177	0.045	0.004	4.126E-29	G	A	0.171	1.292	(1.116,1.496)	5.91E-04	+	G	A	0.170	1.175	(1.057,1.306)	2.70E-03	+	G	A	0.163	1.116	(0.994,1.253)	6.21E-02	+
rs11030104	11	27684517	*BDNF*	A	G	0.792	0.041	0.004	5.557E-28	A	G	0.796	1.320	(1.151,1.513)	7.31E-05	+	A	G	0.799	1.143	(1.029,1.269)	1.27E-02	+	A	G	0.792	1.122	(1.015,1.24)	2.43E-02	+
rs3101336	1	72751185	*NEGR1*	C	T	0.613	0.033	0.003	2.661E-26	C	T	0.613	1.245	(1.111,1.395)	1.58E-04	+	C	T	0.605	1.195	(1.098,1.3)	3.66E-05	+	C	T	0.595	1.056	(0.97,1.15)	2.07E-01	+
rs7138803	12	50247468	*BCDIN3D*	A	G	0.384	0.032	0.003	8.153E-24	A	G	0.379	1.286	(1.149,1.441)	1.31E-05	+	A	G	0.371	1.216	(1.118,1.322)	4.68E-06	+	A	G	0.360	1.035	(0.948,1.13)	4.47E-01	+
rs10182181	2	25150296	*ADCY3*	G	A	0.462	0.031	0.003	8.777E-24	G	A	0.490	1.415	(1.268,1.58)	6.34E-10	+	G	A	0.496	1.202	(1.108,1.304)	9.30E-06	+	G	A	0.480	1.186	(1.09,1.29)	6.81E-05	+
rs3888190	16	28889486	*ATP2A1*	A	C	0.403	0.031	0.003	3.14E-23	A	C	0.407	1.136	(1.017,1.271)	2.46E-02	+	A	C	0.401	1.129	(1.04,1.226)	3.87E-03	+	A	C	0.395	1.035	(0.95,1.126)	4.34E-01	+
rs1516725	3	185824004	*ETV5*	C	T	0.872	0.045	0.005	1.886E-22	C	T	0.859	1.303	(1.112,1.526)	1.05E-03	+	C	T	0.863	1.158	(1.024,1.309)	1.89E-02	+	C	T	0.856	1.180	(1.051,1.325)	5.03E-03	+
rs12446632	16	19935389	*GPRC5B*	G	A	0.865	0.040	0.005	1.477E-18	G	A	0.849	1.279	(1.094,1.495)	2.03E-03	+	G	A	0.856	1.097	(0.975,1.235)	1.24E-01	+	G	A	0.850	1.194	(1.066,1.337)	2.20E-03	+
rs2287019	19	46202172	*QPCTL*	C	T	0.804	0.036	0.004	4.585E-18	C	T	0.814	1.038	(0.898,1.2)	6.14E-01	+	C	T	0.816	1.024	(0.919,1.142)	6.61E-01	+	C	T	0.814	1.047	(0.939,1.168)	4.06E-01	+
rs16951275	15	68077168	*MAP2K5*	T	C	0.784	0.031	0.004	1.911E-17	T	C	0.779	1.173	(1.028,1.339)	1.79E-02	+	T	C	0.776	1.133	(1.025,1.251)	1.43E-02	+	T	C	0.770	1.056	(0.957,1.164)	2.80E-01	+
rs3817334	11	47650993	*MTCH2*	T	C	0.407	0.026	0.003	5.145E-17	T	C	0.408	1.202	(1.073,1.347)	1.46E-03	+	T	C	0.414	1.093	(1.006,1.187)	3.52E-02	+	T	C	0.405	1.098	(1.008,1.196)	3.29E-02	+
rs2112347	5	75015242	*POC5*	T	G	0.629	0.026	0.003	6.191E-17	T	G	0.633	1.087	(0.969,1.219)	1.55E-01	+	T	G	0.637	1.032	(0.949,1.124)	4.60E-01	+	T	G	0.634	1.065	(0.977,1.16)	1.51E-01	+
rs12566985	1	75002193	*FPGT-TNNI3K*	G	A	0.446	0.024	0.003	3.282E-15	G	A	0.448	1.273	(1.138,1.424)	2.48E-05	+	G	A	0.439	1.206	(1.11,1.311)	1.04E-05	+	G	A	0.429	1.038	(0.952,1.132)	3.96E-01	+
rs3810291	19	47569003	*ZC3H4*	A	G	0.666	0.028	0.004	4.812E-15	A	G	0.670	1.198	(1.067,1.345)	2.19E-03	+	A	G	0.671	1.134	(1.039,1.237)	4.69E-03	+	A	G	0.663	1.072	(0.983,1.17)	1.15E-01	+
rs7141420	14	79899454	*NRXN3*	T	C	0.527	0.024	0.003	1.23E-14	T	C	0.518	1.151	(1.03,1.287)	1.33E-02	+	T	C	0.513	1.112	(1.025,1.208)	1.11E-02	+	T	C	0.507	1.003	(0.922,1.091)	9.48E-01	+
rs13078960	3	85807590	*CADM2*	G	T	0.196	0.030	0.004	1.737E-14	G	T	0.192	1.170	(1.018,1.344)	2.74E-02	+	G	T	0.206	0.994	(0.899,1.099)	9.08E-01	-	G	T	0.201	1.198	(1.076,1.334)	9.49E-04	+
rs10968576	9	28414339	*LINGO2*	G	A	0.320	0.025	0.003	6.607E-14	G	A	0.322	1.026	(0.91,1.155)	6.78E-01	+	G	A	0.316	1.049	(0.96,1.145)	2.88E-01	+	G	A	0.316	0.964	(0.881,1.055)	4.22E-01	-
rs17024393	1	110154688	*GNAT2*	C	T	0.040	0.066	0.009	7.029E-14	C	T	0.029	1.802	(1.284,2.529)	6.58E-04	+	C	T	0.026	1.568	(1.246,1.973)	1.26E-04	+	C	T	0.023	1.099	(0.824,1.465)	5.20E-01	+
rs657452	1	49589847	*AGBL4*	A	G	0.394	0.023	0.003	5.482E-13	A	G	0.381	1.103	(0.981,1.24)	1.01E-01	+	A	G	0.384	1.035	(0.95,1.126)	4.31E-01	+	A	G	0.380	1.037	(0.95,1.132)	4.14E-01	+
rs12429545	13	54102206	*OLFM4*	A	G	0.133	0.033	0.005	1.094E-12	A	G	0.134	1.076	(0.916,1.264)	3.72E-01	+	A	G	0.130	1.119	(0.992,1.261)	6.73E-02	+	A	G	0.127	0.982	(0.866,1.115)	7.82E-01	-
rs12286929	11	115022404	*CADM1*	G	A	0.523	0.022	0.003	1.31E-12	G	A	0.529	1.143	(1.024,1.276)	1.74E-02	+	G	A	0.530	1.069	(0.985,1.16)	1.11E-01	+	G	A	0.524	1.063	(0.978,1.155)	1.49E-01	+
rs13107325	4	103188709	*SLC39A8*	T	C	0.072	0.048	0.007	1.825E-12	T	C	0.081	1.605	(1.309,1.967)	5.14E-06	+	T	C	0.081	1.284	(1.116,1.477)	4.84E-04	+	T	C	0.075	1.203	(1.019,1.42)	2.89E-02	+
rs11165643	1	96924097	*PTBP2*	T	C	0.583	0.022	0.003	2.07E-12	T	C	0.588	1.022	(0.913,1.144)	7.03E-01	+	T	C	0.591	1.017	(0.936,1.105)	6.97E-01	+	T	C	0.589	1.030	(0.947,1.12)	4.94E-01	+
rs7903146	10	114758349	*TCF7L2*	C	T	0.713	0.023	0.003	1.112E-11	C	T	0.717	1.015	(0.899,1.147)	8.07E-01	+	C	T	0.712	1.049	(0.959,1.149)	2.94E-01	+	C	T	0.711	0.956	(0.872,1.048)	3.34E-01	-
rs10132280	14	25928179	*STXBP6*	C	A	0.682	0.023	0.003	1.141E-11	C	A	0.704	0.968	(0.855,1.095)	6.02E-01	-	C	A	0.703	0.991	(0.906,1.084)	8.48E-01	-	C	A	0.704	0.984	(0.897,1.079)	7.25E-01	-
rs17405819	8	76806584	*HNF4G*	T	C	0.700	0.022	0.003	2.07E-11	T	C	0.705	1.269	(1.125,1.433)	1.09E-04	+	T	C	0.706	1.124	(1.026,1.23)	1.19E-02	+	T	C	0.699	1.089	(0.995,1.191)	6.30E-02	+
rs6091540	20	51087862	*ZFP64*	C	T	0.723	0.019	0.004	2.154E-11	C	T	0.711	1.145	(1.015,1.292)	2.76E-02	+	C	T	0.711	1.067	(0.975,1.168)	1.60E-01	+	C	T	0.707	1.035	(0.945,1.133)	4.62E-01	+
rs1016287	2	59305625	*LINC01122*	T	C	0.287	0.023	0.003	2.253E-11	T	C	0.311	1.007	(0.895,1.131)	9.13E-01	+	T	C	0.298	1.070	(0.98,1.168)	1.29E-01	+	T	C	0.298	0.928	(0.849,1.016)	1.05E-01	-
rs4256980	11	8673939	*TRIM66*	G	C	0.646	0.021	0.003	2.9E-11	G	C	0.650	1.031	(0.919,1.157)	6.04E-01	+	G	C	0.654	1.013	(0.93,1.104)	7.68E-01	+	G	C	0.652	1.039	(0.952,1.134)	3.94E-01	+
rs17094222	10	102395440	*HIF1AN*	C	T	0.211	0.025	0.004	5.942E-11	C	T	0.210	1.133	(0.985,1.303)	8.06E-02	+	C	T	0.212	1.043	(0.942,1.155)	4.15E-01	+	C	T	0.209	1.045	(0.939,1.163)	4.25E-01	+
rs12401738	1	78446761	*FUBP1*	A	G	0.352	0.021	0.003	1.145E-10	A	G	0.378	0.988	(0.881,1.106)	8.30E-01	-	A	G	0.377	1.007	(0.925,1.096)	8.75E-01	+	A	G	0.377	0.998	(0.916,1.089)	9.71E-01	-
rs7599312	2	213413231	*ERBB4*	G	A	0.724	0.022	0.003	1.173E-10	G	A	0.732	1.083	(0.957,1.225)	2.08E-01	+	G	A	0.731	1.047	(0.955,1.148)	3.32E-01	+	G	A	0.729	0.985	(0.897,1.081)	7.45E-01	-
rs2365389	3	61236462	*FHIT*	C	T	0.582	0.020	0.003	1.629E-10	C	T	0.595	1.185	(1.058,1.328)	3.43E-03	+	C	T	0.594	1.092	(1.004,1.187)	3.94E-02	+	C	T	0.589	1.060	(0.973,1.154)	1.80E-01	+
rs205262	6	34563164	*C6orf106*	G	A	0.273	0.022	0.004	1.753E-10	G	A	0.275	1.186	(1.049,1.341)	6.38E-03	+	G	A	0.269	1.163	(1.062,1.274)	1.14E-03	+	G	A	0.262	1.051	(0.955,1.156)	3.12E-01	+
rs2820292	1	201784287	*NAV1*	C	A	0.555	0.020	0.003	1.834E-10	C	A	0.562	1.084	(0.969,1.212)	1.57E-01	+	C	A	0.567	1.031	(0.949,1.119)	4.74E-01	+	C	A	0.563	1.095	(1.007,1.191)	3.47E-02	+
rs12885454	14	29738038	*PRKD1*	C	A	0.642	0.021	0.003	1.943E-10	C	A	0.631	1.084	(0.967,1.214)	1.65E-01	+	C	A	0.633	1.084	(0.995,1.126)	4.35E-01	+	C	A	0.630	1.033	(0.948,1.125)	4.61E-01	+
rs9641123	7	93197732	*CALCR*	C	G	0.429	0.019	0.004	2.077E-10	C	G	0.411	1.161	(1.036,1.302)	1.03E-02	+	C	G	0.410	1.096	(1.008,1.193)	3.19E-02	+	C	G	0.404	1.037	(0.951,1.131)	4.09E-01	+
rs12016871	13	28017782	*MTIF3*	T	C	0.203	0.030	0.005	2.291E-10	T	C	0.188	1.089	(0.947,1.252)	2.32E-01	+	T	C	0.179	1.152	(1.039,1.278)	7.09E-03	+	T	C	0.176	0.962	(0.862,1.073)	4.84E-01	-
rs16851483	3	141275436	*RASA2*	T	G	0.066	0.048	0.008	3.548E-10	T	G	0.066	1.428	(1.137,1.795)	2.23E-03	+	T	G	0.067	1.202	(1.027,1.406)	2.17E-02	+	T	G	0.064	1.170	(0.977,1.403)	8.83E-02	+
rs1167827	7	75163169	*HIP1*	G	A	0.553	0.020	0.003	6.333E-10	G	A	0.569	1.167	(1.046,1.302)	5.65E-03	+	G	A	0.572	1.075	(0.99,1.167)	8.67E-02	+	G	A	0.566	1.075	(0.989,1.169)	8.89E-02	+
rs758747	16	3627358	*NLRC3*	T	C	0.265	0.023	0.004	7.473E-10	T	C	0.270	1.138	(1.007,1.287)	3.88E-02	+	T	C	0.280	1.010	(0.923,1.106)	8.29E-01	+	T	C	0.275	1.089	(0.991,1.197)	7.65E-02	+
rs1928295	9	120378483	*TLR4*	T	C	0.548	0.019	0.003	7.91E-10	T	C	0.576	1.116	(0.998,1.249)	5.38E-02	+	T	C	0.568	1.103	(1.015,1.197)	2.00E-02	+	T	C	0.563	0.990	(0.911,1.076)	8.13E-01	-
rs9925964	16	31129895	*KAT8*	A	G	0.620	0.019	0.003	8.108E-10	A	G	0.642	1.036	(0.924,1.161)	5.47E-01	+	A	G	0.646	1.010	(0.927,1.1)	8.27E-01	+	A	G	0.644	1.025	(0.939,1.119)	5.77E-01	+
rs11126666	2	26928811	*KCNK3*	A	G	0.283	0.021	0.003	1.332E-09	A	G	0.255	1.001	(0.883,1.134)	9.89E-01	+	A	G	0.254	0.992	(0.903,1.089)	8.60E-01	-	A	G	0.255	1.004	(0.913,1.104)	9.33E-01	+
rs2650492	16	28333411	*SBK1*	A	G	0.303	0.021	0.004	1.915E-09	A	G	0.306	1.211	(1.076,1.363)	1.48E-03	+	A	G	0.299	1.175	(1.077,1.281)	2.93E-04	+	A	G	0.292	1.056	(0.964,1.158)	2.42E-01	+
rs6804842	3	25106437	*RARB*	G	A	0.575	0.019	0.003	2.476E-09	G	A	0.580	1.118	(1,1.25)	4.99E-02	+	G	A	0.580	1.083	(0.996,1.177)	6.17E-02	+	G	A	0.574	1.037	(0.953,1.128)	3.96E-01	+
rs12940622	17	78615571	*RPTOR*	G	A	0.575	0.018	0.003	2.494E-09	G	A	0.560	1.185	(1.061,1.324)	2.62E-03	+	G	A	0.559	1.121	(1.031,1.218)	7.20E-03	+	G	A	0.551	1.068	(0.981,1.161)	1.28E-01	+
rs7164727	15	73093991	*LOC100287559*	T	C	0.686	0.018	0.003	3.915E-09	T	C	0.676	0.940	(0.835,1.058)	3.04E-01	-	T	C	0.669	0.992	(0.909,1.082)	8.56E-01	-	T	C	0.672	0.949	(0.867,1.04)	2.61E-01	-
rs11847697	14	30515112	*PRKD1*	T	C	0.042	0.049	0.008	3.99E-09	T	C	0.045	1.420	(1.093,1.846)	8.68E-03	+	T	C	0.045	1.252	(1.041,1.507)	1.72E-02	+	T	C	0.042	1.244	(1,1.548)	5.05E-02	+
rs4740619	9	15634326	*C9orf93*	T	C	0.542	0.018	0.003	4.564E-09	T	C	0.537	1.183	(1.059,1.321)	2.94E-03	+	T	C	0.548	1.054	(0.971,1.143)	2.10E-01	+	T	C	0.540	1.124	(1.034,1.221)	5.88E-03	+
rs492400	2	219349752	*USP37*	C	T	0.423	0.016	0.003	6.784E-09	C	T	0.426	1.097	(0.983,1.225)	9.79E-02	+	C	T	0.431	1.033	(0.951,1.121)	4.42E-01	+	C	T	0.426	1.063	(0.976,1.158)	1.58E-01	+
rs13191362	6	163033350	*PARK2*	A	G	0.879	0.028	0.005	7.339E-09	A	G	0.878	1.164	(0.981,1.38)	8.18E-02	+	A	G	0.878	1.073	(0.945,1.219)	2.78E-01	+	A	G	0.876	1.069	(0.943,1.211)	2.98E-01	+
rs3736485	15	51748610	*DMXL2*	A	G	0.454	0.018	0.003	7.412E-09	A	G	0.472	1.021	(0.914,1.14)	7.18E-01	+	A	G	0.460	1.084	(0.998,1.177)	5.48E-02	+	A	G	0.458	0.961	(0.884,1.045)	3.54E-01	-
rs17001654	4	77129568	*SCARB2*	G	C	0.153	0.031	0.005	7.76E-09	G	C	0.153	0.991	(0.849,1.158)	9.12E-01	-	G	C	0.149	1.012	(0.901,1.137)	8.37E-01	+	G	C	0.150	0.950	(0.844,1.069)	3.94E-01	-
rs11191560	10	104869038	*NT5C2*	C	T	0.089	0.031	0.005	8.446E-09	C	T	0.082	1.255	(1.025,1.538)	2.83E-02	+	C	T	0.076	1.236	(1.069,1.429)	4.23E-03	+	C	T	0.073	1.000	(0.852,1.174)	9.98E-01	+
rs2080454	16	49062590	*CBLN1*	C	A	0.405	0.017	0.003	8.604E-09	C	A	0.385	1.011	(0.903,1.131)	8.50E-01	+	C	A	0.386	0.989	(0.91,1.075)	7.96E-01	-	C	A	0.387	0.998	(0.917,1.087)	9.70E-01	-
rs7715256	5	153537893	*GALNT10*	G	T	0.421	0.016	0.003	8.851E-09	G	T	0.423	1.180	(1.055,1.318)	3.60E-03	+	G	T	0.428	1.067	(0.982,1.158)	1.26E-01	+	G	T	0.421	1.087	(0.998,1.185)	5.56E-02	+
rs2176040	2	227092802	*LOC646736*	A	G	0.365	0.014	0.003	9.989E-09	A	G	0.364	1.025	(0.913,1.151)	6.75E-01	+	A	G	0.350	1.080	(0.993,1.175)	7.37E-02	+	A	G	0.350	0.930	(0.853,1.013)	9.77E-02	-
rs1528435	2	181550962	*UBE2E3*	T	C	0.631	0.018	0.003	1.196E-08	T	C	0.636	1.013	(0.904,1.136)	8.21E-01	+	T	C	0.624	1.087	(0.999,1.183)	5.27E-02	+	T	C	0.622	0.944	(0.866,1.029)	1.90E-01	-
rs2075650	19	45395619	*TOMM40*	A	G	0.848	0.026	0.005	1.247E-08	G	A	0.147	0.968	(0.828,1.131)	6.82E-01	-	A	G	0.853	1.002	(0.893,1.123)	9.77E-01	+	A	G	0.853	1.007	(0.896,1.132)	9.04E-01	+
rs1000940	17	5283252	*RABEP1*	G	A	0.320	0.019	0.003	1.284E-08	G	A	0.307	1.155	(1.026,1.302)	1.75E-02	+	G	A	0.304	1.115	(1.022,1.218)	1.47E-02	+	G	A	0.298	1.061	(0.968,1.164)	2.04E-01	+
rs11583200	1	50559820	*ELAVL4*	C	T	0.396	0.018	0.003	1.479E-08	C	T	0.376	1.083	(0.965,1.214)	1.75E-01	+	C	T	0.381	1.024	(0.943,1.113)	5.70E-01	+	C	T	0.377	1.058	(0.971,1.153)	2.01E-01	+
rs7239883	18	40147671	*LOC284260*	G	A	0.393	0.016	0.003	1.511E-08	G	A	0.380	1.029	(0.917,1.154)	6.25E-01	+	G	A	0.377	1.029	(0.946,1.119)	5.08E-01	+	G	A	0.376	1.004	(0.921,1.093)	9.33E-01	+
rs2836754	21	40291740	*ETS2*	C	T	0.612	0.016	0.003	1.605E-08	C	T	0.643	1.134	(1.011,1.272)	3.22E-02	+	C	T	0.651	1.053	(0.966,1.147)	2.42E-01	+	C	T	0.645	1.120	(1.027,1.221)	1.03E-02	+
rs9400239	6	108977663	*FOXO3*	C	T	0.688	0.019	0.003	1.613E-08	C	T	0.706	1.214	(1.074,1.371)	1.85E-03	+	C	T	0.709	1.118	(1.019,1.226)	1.84E-02	+	C	T	0.702	1.099	(1.003,1.205)	4.31E-02	+
rs10733682	9	129460914	*LMX1B*	A	G	0.478	0.017	0.003	1.83E-08	A	G	0.479	1.084	(0.968,1.212)	1.61E-01	+	A	G	0.475	1.071	(0.986,1.165)	1.06E-01	+	A	G	0.471	1.014	(0.93,1.105)	7.55E-01	+
rs11688816	2	63053048	*EHBP1*	G	A	0.525	0.017	0.003	1.893E-08	G	A	0.542	1.033	(0.924,1.154)	5.69E-01	+	G	A	0.544	0.988	(0.91,1.074)	7.83E-01	-	G	A	0.544	1.001	(0.92,1.09)	9.76E-01	+
rs11057405	12	122781897	*CLIP1*	G	A	0.901	0.031	0.006	2.019E-08	G	A	0.898	1.070	(0.892,1.284)	4.67E-01	+	G	A	0.891	1.129	(0.985,1.293)	8.09E-02	+	G	A	0.890	0.929	(0.81,1.065)	2.89E-01	-
rs9914578	17	2005136	*SMG6*	G	C	0.211	0.020	0.004	2.072E-08	G	C	0.208	1.044	(0.91,1.197)	5.39E-01	+	G	C	0.207	1.042	(0.943,1.152)	4.17E-01	+	G	C	0.205	1.030	(0.929,1.143)	5.74E-01	+
rs977747	1	47684677	*TAL1*	T	G	0.391	0.017	0.003	2.182E-08	T	G	0.412	1.099	(0.983,1.23)	9.84E-02	+	T	G	0.416	1.043	(0.961,1.133)	3.13E-01	+	T	G	0.411	1.050	(0.965,1.143)	2.54E-01	+

Appendix A

rs2121279	2	143043285	*LRP1B*	T	C	0.152	0.025	0.004	2.313E-08	T	C	0.133	1.063	(0.904,1.252)	4.59E-01	+	T	C	0.127	1.118	(0.991,1.26)	6.99E-02	+	T	C	0.126	0.958	(0.845,1.085)	4.99E-01	-
rs29941	19	34309532	*KCTD15*	G	A	0.669	0.018	0.003	2.407E-08	G	A	0.677	1.187	(1.055,1.336)	4.37E-03	+	G	A	0.671	1.132	(1.037,1.236)	5.77E-03	+	G	A	0.665	1.027	(0.94,1.123)	5.56E-01	+
rs11727676	4	145659064	*HHIP*	T	C	0.910	0.036	0.006	2.55E-08	C	T	0.092	0.999	(0.813,1.228)	9.94E-01	-	T	C	0.904	1.056	(0.907,1.23)	4.80E-01	+	T	C	0.905	0.899	(0.766,1.054)	1.90E-01	-
rs3849570	3	81792112	*GBE1*	A	C	0.359	0.019	0.003	2.601E-08	A	C	0.348	1.040	(0.926,1.169)	5.05E-01	+	A	C	0.346	1.021	(0.937,1.113)	6.37E-01	+	A	C	0.346	0.988	(0.905,1.078)	7.80E-01	-
rs9374842	6	120185665	*LOC285762*	T	C	0.748	0.019	0.004	2.673E-08	T	C	0.775	1.222	(1.073,1.393)	2.54E-03	+	T	C	0.772	1.160	(1.05,1.281)	3.41E-03	+	T	C	0.766	1.058	(0.96,1.166)	2.53E-01	+
rs6477694	9	111932342	*EPB41L4B*	C	T	0.365	0.017	0.003	2.673E-08	C	T	0.356	1.161	(1.035,1.303)	1.07E-02	+	C	T	0.353	1.101	(1.011,1.198)	2.73E-02	+	C	T	0.347	1.043	(0.955,1.139)	3.53E-01	+
rs4787491	16	30015337	*INO80E*	G	A	0.509	0.016	0.003	2.696E-08	G	A	0.538	1.014	(0.908,1.132)	8.08E-01	+	G	A	0.536	1.006	(0.927,1.092)	8.87E-01	+	G	A	0.537	0.981	(0.902,1.067)	6.56E-01	-
rs1441264	13	79580919	*MIR548A2*	A	G	0.609	0.018	0.003	2.959E-08	A	G	0.590	1.082	(0.963,1.215)	1.86E-01	+	A	G	0.590	1.051	(0.963,1.146)	2.68E-01	+	A	G	0.587	1.049	(0.961,1.146)	2.86E-01	+
rs7899106	10	87410904	*GRID1*	G	A	0.052	0.040	0.007	2.96E-08	G	A	0.056	1.269	(0.998,1.613)	5.17E-02	+	G	A	0.051	1.240	(1.043,1.475)	1.48E-02	+	G	A	0.050	0.949	(0.786,1.147)	5.90E-01	-
rs2176598	11	43864278	*HSD17B12*	T	C	0.251	0.020	0.004	2.971E-08	T	C	0.237	1.055	(0.926,1.201)	4.19E-01	+	T	C	0.247	0.957	(0.871,1.053)	3.68E-01	-	T	C	0.246	1.076	(0.976,1.187)	1.41E-01	+
rs2245368	7	76608143	*PMS2L11*	C	T	0.180	0.032	0.006	3.187E-08	C	T	0.178	1.190	(1.025,1.382)	2.27E-02	+	C	T	0.167	1.225	(1.098,1.366)	2.73E-04	+	C	T	0.162	0.984	(0.875,1.105)	7.82E-01	-
rs17203016	2	208255518	*CREB1*	G	A	0.197	0.021	0.004	3.406E-08	G	A	0.213	1.128	(0.987,1.289)	7.77E-02	+	G	A	0.206	1.133	(1.026,1.25)	1.32E-02	+	G	A	0.202	0.982	(0.886,1.088)	7.28E-01	-
rs17724992	19	18454825	*PGPEP1*	A	G	0.746	0.019	0.004	3.415E-08	A	G	0.744	1.196	(1.055,1.356)	5.09E-03	+	A	G	0.741	1.155	(1.05,1.271)	2.99E-03	+	A	G	0.734	1.042	(0.949,1.144)	3.90E-01	+
rs7243357	18	56883319	*GRP*	T	G	0.812	0.022	0.004	3.857E-08	T	G	0.825	1.182	(1.02,1.368)	2.56E-02	+	T	G	0.826	1.106	(0.989,1.236)	7.66E-02	+	T	G	0.821	1.090	(0.978,1.214)	1.19E-01	+
rs16907751	8	81375457	*ZBTB10*	C	T	0.916	0.035	0.007	3.888E-08	C	T	0.906	0.966	(0.797,1.171)	7.28E-01	-	C	T	0.908	0.953	(0.828,1.097)	5.01E-01	-	C	T	0.909	1.013	(0.876,1.171)	8.63E-01	+
rs1808579	18	21104888	*C18orf8*	C	T	0.534	0.017	0.003	4.169E-08	C	T	0.522	1.096	(0.981,1.224)	1.05E-01	+	C	T	0.517	1.079	(0.994,1.172)	6.90E-02	+	C	T	0.513	1.026	(0.943,1.115)	5.53E-01	+
rs13201877	6	137675541	*IFNGR1*	G	A	0.142	0.023	0.005	4.285E-08	G	A	0.141	1.181	(1.006,1.385)	4.18E-02	+	G	A	0.141	1.091	(0.971,1.225)	1.43E-01	+	G	A	0.138	1.056	(0.932,1.196)	3.95E-01	+
rs2033732	8	85079709	*RALYL*	C	T	0.747	0.019	0.004	4.889E-08	C	T	0.743	1.008	(0.89,1.142)	8.95E-01	+	C	T	0.744	0.982	(0.895,1.078)	7.08E-01	-	C	T	0.744	1.015	(0.923,1.117)	7.62E-01	+
rs9540493	13	66205704	*MIR548X2*	A	G	0.456	0.017	0.003	4.971E-08	A	G	0.460	1.130	(1.005,1.27)	4.13E-02	+	A	G	0.454	1.120	(1.028,1.222)	9.92E-03	+	A	G	0.449	1.004	(0.92,1.096)	9.28E-01	+
rs1460676	2	164567689	*FIGN*	C	T	0.173	0.020	0.004	4.978E-08	C	T	0.158	1.022	(0.879,1.187)	7.81E-01	+	C	T	0.155	1.044	(0.934,1.168)	4.46E-01	+	C	T	0.154	0.983	(0.876,1.103)	7.66E-01	-
rs6465468	7	95169514	*ASB4*	T	G	0.304	0.017	0.004	4.98E-08	T	G	0.308	1.005	(0.887,1.139)	9.36E-01	+	T	G	0.301	1.047	(0.955,1.149)	3.24E-01	+	T	G	0.301	0.955	(0.868,1.049)	3.36E-01	-
rs751414**	6	40350030	TDRG1	T	G	0.258	0.018	0.004	1.58E-05	T	G	0.283	1.16813	(1.033,1.32)	1.29E-02	+	T	G	0.287	1.04676	(0.957,1.145)	3.18E-01	+	T	G	0.283	1.08231	(0.986,1.188)	9.67E-02	+

*GRCh37/hg19 coordinates
**Proxy for rs2033529
Effect = Effect allele (BMI increasing allele); Other = Other allele; EAF = Effect allele frequency

S4 Table. Difference in SCOOP OR when using ALSPAC as control dataset vs. UKHLS

SNP	Locus	OR.UKHLS	OR.ALSPAC	P.Diff
rs1558902	FTO	1.4287329	1.3427721	2.94E-01
rs6567160	MC4R	1.3080991	1.3604779	5.52E-01
rs13021737	TMEM18	1.3563998	1.2974696	6.00E-01
rs10938397	GNPDA2	1.1857281	1.1860919	9.96E-01
rs543874	SEC16B	1.2010834	1.2045657	9.67E-01
rs2207139	TFAP2B	1.1750903	1.1546588	8.18E-01
rs11030104	BDNF	1.1428476	1.1088972	6.90E-01
rs3101336	NEGR1	1.1948946	1.2385984	5.57E-01
rs7138803	BCDIN3D	1.215858	1.2146898	9.87E-01
rs10182181	ADCY3	1.2020002	1.2265576	7.31E-01
rs3888190	ATP2A1	1.1293237	1.0144525	7.22E-02
rs1516725	ETV5	1.158149	1.026153	1.74E-01
rs12446632	GPRC5B	1.0971063	1.0185721	3.86E-01
rs2287019	QPCTL	1.0244956	1.0421619	8.25E-01
rs16951275	MAP2K5	1.1325514	1.092782	6.20E-01
rs3817334	MTCH2	1.0927407	1.1358904	5.16E-01
rs2112347	POC5	1.0324305	1.004322	6.53E-01
rs12566985	FPGT-TNNI3K	1.2061603	1.1713434	6.23E-01
rs3810291	ZC3H4	1.1339907	1.0873902	5.08E-01
rs7141420	NRXN3	1.1124525	1.1058898	9.20E-01
rs13078960	CADM2	0.99411	1.031164	6.16E-01
rs10968576	LINGO2	1.048838	1.0523973	9.57E-01
rs17024393	GNAT2	1.5681554	1.5545372	9.58E-01
rs657452	AGBL4	1.0346724	1.0741845	5.39E-01
rs12429545	OLFM4	1.1186482	1.1316867	8.93E-01
rs12286929	CADM1	1.0687761	1.0658373	9.63E-01
rs13107325	SLC39A8	1.2837186	1.3563332	5.90E-01
rs11165643	PTBP2	1.0166116	1.0013239	8.01E-01
rs7903146	TCF7L2	1.049512	1.1068024	4.16E-01
rs10132280	STXBP6	0.9912485	0.9586591	6.05E-01
rs17405819	HNF4G	1.1236114	1.0863413	6.08E-01
rs6091540	ZFP64	1.067074	1.1034806	6.08E-01
rs1016287	LINC01122	1.0702148	1.0895905	7.78E-01
rs4256980	TRIM66	1.0129686	0.9606069	3.92E-01
rs17094222	HIF1AN	1.0431979	1.0554176	8.73E-01
rs12401738	FUBP1	1.0068534	0.9709964	5.52E-01
rs7599312	ERBB4	1.0466985	0.9901823	4.06E-01
rs2365389	FHIT	1.0918292	1.1451163	4.30E-01
rs205262	C6orf106	1.1634375	1.0784589	2.46E-01
rs2820292	NAV1	1.0305774	0.9731171	3.36E-01
rs12885454	PRKD1	1.0343118	0.9851811	4.27E-01
rs9641123	CALCR	1.0963951	1.0743197	7.35E-01
rs9581854	MTIF3	1.1523104	1.0643572	2.87E-01
rs16851483	RASA2	1.2018139	1.2290979	8.43E-01
rs11167827	HIP1	1.0745054	1.0968666	7.30E-01
rs758747	NLRC3	1.0100159	1.0528825	5.26E-01
rs1928295	TLR4	1.1026364	1.0470854	3.86E-01
rs9925964	KAT8	1.009594	1.0531275	4.94E-01
rs11126666	KCNK3	0.9916191	1.0011154	8.88E-01
rs2650492	SBK1	1.1745464	1.1002881	3.00E-01
rs6804842	RARB	1.0826074	1.0744722	9.00E-01
rs12940622	RPTOR	1.1210032	1.0859058	5.96E-01
rs7164727	LOC100287559	0.9919261	0.9667406	6.83E-01
rs11847697	PRKD1	1.2522288	1.1977594	7.40E-01
rs4740619	C9orf93	1.053687	1.0122709	4.98E-01
rs492400	USP37	1.0326143	1.0502736	7.75E-01
rs13191362	PARK2	1.0730103	1.1335706	5.49E-01
rs3736485	DMXL2	1.0839278	1.0843441	9.95E-01
rs17001654	SCARB2	1.0123387	0.9533294	4.70E-01
rs11191560	NT5C2	1.2358896	1.1978983	7.66E-01
rs2080454	CBLN1	0.9890632	0.9957939	9.11E-01
rs7715256	GALNT10	1.0667076	1.0659983	9.91E-01
rs2176040	LOC646736	1.0800962	1.0561891	7.15E-01
rs1528435	UBE2E3	1.0871075	1.0592421	6.71E-01
rs2075650	TOMM40	1.0017207	0.9436388	4.72E-01
rs1000940	RABEP1	1.1155025	1.1561709	5.73E-01
rs11583200	ELAVL4	1.0244068	1.0473277	7.14E-01
rs7239883	LOC284260	1.0288501	0.9941999	5.73E-01
rs2836754	ETS2	1.0526894	1.0507287	9.76E-01
rs9400239	FOXO3	1.1176888	1.0694002	5.06E-01
rs10733682	LMX1B	1.0712934	1.0545872	7.92E-01
rs11688816	EHBP1	0.988461	0.9758464	8.29E-01
rs11057405	CLIP1	1.1285696	1.0499994	4.62E-01
rs9914578	SMG6	1.0423268	1.0775537	6.46E-01
rs977747	TAL1	1.0432587	1.0166069	6.64E-01
rs2121279	LRP1B	1.1174631	1.0960597	8.23E-01
rs29941	KCTD15	1.1320951	1.0502216	2.39E-01
rs11727676	HHIP	1.0563982	1.0477135	9.38E-01
rs3849570	GBE1	1.0208987	0.9989013	7.25E-01
rs9374842	LOC285762	1.159775	1.1754326	8.52E-01
rs6477694	EPB41L4B	1.100504	1.0672617	6.17E-01
rs4787491	INO80E	1.0059218	1.0361099	6.17E-01
rs1441264	MIR548A2	1.0505803	1.0275484	7.19E-01
rs7899106	GRID1	1.240441	1.3269198	5.94E-01
rs2176598	HSD17B12	0.9573421	0.9531657	9.49E-01
rs2245368	PMS2L11	1.2247062	0.8928163	3.81E-05
rs17203016	CREB1	1.1327383	1.1718323	6.35E-01
rs17724992	PGPEP1	1.1553103	1.156706	9.86E-01
rs7243357	GRP	1.1056075	1.0651079	6.41E-01
rs16907751	ZBTB10	0.9527799	1.0182715	5.11E-01
rs1808579	C18orf8	1.0793057	1.0641407	8.11E-01
rs13201877	IFNGR1	1.0907763	1.0761121	8.71E-01
rs2033732	RALYL	0.9823586	0.9258369	3.80E-01
rs9540493	MIR548X2	1.1204384	1.0826008	5.72E-01
rs1460676	FIGN	1.0443083	1.0803401	6.74E-01
rs6465468	ASB4	1.0475326	0.949917	1.33E-01
rs2033529	TDRG1	0.9553109	0.9790073	7.05E-01

OR.UKHLS= OR when using UKHLS as control group
OR.ALSPAC= OR when using age-matched ALSPAC as control group
P.Diff=p value for difference

Appendix A

S5 Table. Discovery, replication and meta-analysis results for 32 SNPs meeting P<10-5 in discovery association results of SCOOP vs STILTS analysis.

						SCOOP				UKBB					GIANT BMI Tails				Combined analysis		
rsID	Nearest gene	Chr.	Position (bp)	EA	NEA	OR (95% CI)	P value	EAF Obese	EAF Thin	proxy rsID	OR (95% CI)	P value	EAF Obese	EAF Thin	proxy rsID	r2	OR (95% CI)	P value	OR (95% CI)	P value	HetPVal
rs9930333	FTO	16	53799977	G	T	1.70(1.52,1.90)	2.30E-20	49.59%	37.46%		1.46(1.38,1.55)	3.60E-36	48.26%	38.93%			1.43(1.34,1.54)	8.10E-25	1.48(1.42,1.54)	8.52E-76	3.34E-02
rs2168711	MC4R	18	57848531	C	T	1.66(1.45,1.89)	8.29E-14	28.90%	19.95%		1.23(1.15,1.32)	2.19E-09	26.75%	22.90%			1.20(1.10,1.30)	1.80E-05	1.27(1.21,1.33)	2.02E-21	1.12E-04
rs6748821	TMEM18	2	629601	G	A	1.65(1.42,1.91)	9.45E-11	86.69%	79.84%		1.27(1.18,1.37)	1.31E-09	85.00%	81.69%	rs12995480	0.998	1.26(1.14,1.39)	9.90E-06	1.32(1.24,1.39)	7.76E-21	2.81E-03
rs506589	SEC16B	1	177894287	C	T	1.46(1.27,1.67)	5.42E-08	23.98%	18.07%		1.25(1.17,1.35)	5.44E-10	23.11%	19.16%			1.25(1.14,1.37)	2.70E-06	1.28(1.21,1.35)	3.14E-20	1.21E-01
rs6738433	ADCY3-DNAJC27	2	25159501	C	G	1.43(1.28,1.60)	1.71E-10	47.31%	43.92%		1.21(1.14,1.28)	2.74E-10	50.70%	45.96%	rs2384054	0.968	1.10(1.03,1.17)	5.70E-03	1.19(1.14,1.24)	3.19E-17	6.25E-03
rs7132908	FAIM2	12	50263148	A	G	1.31(1.17,1.47)	2.26E-06	42.45%	36.27%		1.18(1.11,1.25)	5.43E-08	41.11%	37.39%			1.20(1.10,1.30)	6.60E-06	1.20(1.15,1.26)	1.93E-16	2.52E-01
rs62107261	FAM150B	2	422144	T	C	2.37(1.75,3.20)	2.07E-08	96.37%	93.38%		1.54(1.35,1.76)	3.57E-10	96.28%	94.36%	NA	NA	NA	NA	1.65(1.46,1.87)	1.15E-15	1.07E-02
rs12507026	GNPDA2	4	45181334	T	A	1.30(1.17,1.46)	3.69E-06	47.29%	40.92%		1.14(1.08,1.21)	8.76E-06	45.30%	41.98%	rs12641981	0.998	1.20(1.12,1.28)	3.10E-07	1.18(1.13,1.23)	5.53E-15	4.06E-02
rs75398113	SNRPC	6	34728071	C	G	1.53(1.27,1.85)	8.91E-06	11.95%	8.04%		1.24(1.12,1.37)	2.07E-05	10.47%	8.52%	NA	NA	NA	NA	1.30(1.19,1.42)	5.19E-09	5.56E-02
rs13135092	SLC39A8	4	103198082	G	A	1.58(1.30,1.93)	4.70E-06	10.50%	7.24%		1.25(1.12,1.39)	5.57E-05	9.24%	7.52%	NA	NA	NA	NA	1.32(1.20,1.45)	1.06E-08	3.59E-02
rs57988840	TFAP2B	6	50817748	T	A	1.69(1.39,2.05)	1.27E-07	92.53%	88.81%		1.13(1.02,1.24)	1.65E-02	91.05%	90.04%	rs7769978	1	1.20(1.03,1.39)	1.90E-02	1.22(1.13,1.31)	3.86E-07	2.87E-04
rs4447506	PIK3C3	18	39510074	G	A	1.32(1.17,1.48)	4.21E-06	41.83%	36.39%		1.07(1.01,1.14)	2.60E-02	39.34%	37.71%			1.10(1.02,1.18)	7.80E-03	1.11(1.06,1.16)	1.46E-06	7.85E-03
rs375252497*	SEMA3B	3	50310286	AAATAATAATAAT	A	1.35(1.20,1.53)	1.74E-06	37.22%	31.78%		1.13(1.02,1.26)	2.50E-02	34.30%	31.95%	NA	NA	NA	NA	1.22(1.13,1.32)	1.49E-06	3.05E-02
rs7927262	HIPK3	11	33384447	C	T	1.41(1.24,1.60)	1.81E-07	41.78%	35.78%		1.08(0.89,1.32)	4.43E-01	97.52%	97.37%	NA	NA	NA	NA	1.31(1.17,1.45)	1.58E-06	2.87E-02
rs654240	CCND1	11	69448373	T	C	1.35(1.21,1.52)	2.99E-07	43.85%	37.39%		1.05(0.99,1.12)	9.25E-02	41.43%	40.23%			1.08(1.00,1.16)	5.30E-02	1.10(1.05,1.15)	2.10E-05	6.81E-04
rs35403928*	PRDM6	5	122416569	GT	G	1.39(1.23,1.56)	6.79E-08	39.85%	32.94%		1.05(0.95,1.16)	3.61E-01	37.64%	36.49%	NA	NA	NA	NA	1.18(1.09,1.28)	2.46E-05	4.77E-04
rs516579	MTCL1	18	8801634	G	T	1.40(1.22,1.61)	2.07E-06	82.14%	77.25%		1.03(0.96,1.11)	4.52E-01	80.35%	80.05%	rs518561	0.998	1.15(1.04,1.27)	6.40E-03	1.11(1.05,1.18)	9.70E-05	1.11E-04
rs397859802*	FLJ26850	19	50556007	C	CA	1.92(1.45,2.53)	4.49E-06	6.02%	3.44%		1.11(0.86,1.44)	4.28E-01	4.25%	3.78%	NA	NA	NA	NA	1.43(1.18,1.73)	2.12E-04	4.77E-03
rs2784243	PKHD1	6	51454640	T	C	1.30(1.16,1.45)	5.99E-06	61.89%	56.06%		1.07(1.01,1.13)	2.90E-02	58.99%	57.34%	rs2784187	0.988	1.02(0.95,1.10)	5.40E-01	1.08(1.04,1.13)	3.14E-04	2.55E-03
rs11792928	LMX1B	9	129401550	T	C	1.36(1.20,1.53)	1.32E-06	32.13%	26.91%		1.05(0.99,1.12)	1.17E-01	29.94%	29.01%			1.03(0.95,1.11)	5.00E-01	1.08(1.03,1.13)	8.05E-04	5.19E-04
rs6711131*	BAZ2B	2	160407777	A	G	1.31(1.17,1.47)	4.30E-06	65.12%	58.62%		1.02(0.92,1.13)	6.81E-01	63.33%	63.04%	NA	NA	NA	NA	1.14(1.05,1.23)	8.90E-04	1.33E-03
rs73145387	ABI3BP	3	100813661	C	G	2.48(1.67,3.69)	7.36E-06	98.00%	96.42%		1.15(0.96,1.37)	1.29E-01	97.55%	97.19%	NA	NA	NA	NA	1.31(1.11,1.54)	1.29E-03	5.19E-04
rs599291	SLC44A5	1	75691616	T	C	1.31(1.17,1.47)	2.35E-06	47.71%	41.63%		1.02(0.96,1.08)	4.95E-01	44.55%	44.01%			1.04(0.97,1.11)	2.20E-01	1.06(1.02,1.11)	3.44E-03	4.01E-04
rs11185396	LOC100129138	1	104754536	C	T	1.50(1.26,1.80)	8.13E-06	12.78%	9.21%		1.06(0.97,1.17)	2.13E-01	10.37%	9.65%			1.01(0.89,1.14)	9.20E-01	1.10(1.03,1.18)	6.95E-03	8.13E-04
rs2836760	LOC400867	21	40300052	T	G	1.65(1.33,2.03)	3.28E-06	10.33%	7.12%		1.03(0.93,1.14)	5.92E-01	9.14%	8.91%			1.07(0.93,1.23)	3.50E-01	1.11(1.03,1.20)	9.44E-03	3.30E-04
rs11159277	SPTLC2	14	78932957	A	T	1.35(1.20,1.53)	1.56E-06	71.04%	66.32%		1.03(0.95,1.08)	6.53E-01	68.83%	68.55%	NA	NA	NA	NA	1.08(1.02,1.14)	9.74E-03	4.58E-05
rs10546790	CDH22	20	44910100	C	CAT	1.34(1.19,1.52)	1.91E-06	72.94%	66.87%	rs2425853	1.03(0.97,1.10)	3.42E-01	70.11%	69.59%	rs2425853	0.998	1.00(0.93,1.08)	9.90E-01	1.06(1.01,1.11)	1.35E-02	1.57E-04
rs11319985*	CNTN6	3	1377810	T	TA	1.29(1.15,1.45)	9.85E-06	61.56%	56.63%		0.97(0.88,1.07)	5.75E-01	57.91%	58.39%	NA	NA	NA	NA	1.10(1.02,1.18)	1.38E-02	2.03E-04
rs4790399	RAP1GAP2	17	2883199	C	T	1.33(1.18,1.51)	6.95E-06	74.28%	69.50%		1.02(0.96,1.09)	5.43E-01	71.22%	70.85%			0.99(0.91,1.08)	8.30E-01	1.05(1.00,1.10)	4.40E-02	2.50E-04
rs536093	PDE10A	6	165945641	T	C	1.38(1.22,1.58)	1.01E-06	27.05%	21.65%		0.97(0.90,1.03)	3.17E-01	24.39%	25.03%			1.06(0.97,1.15)	2.00E-01	1.05(1.00,1.10)	6.84E-02	9.29E-06
rs936249	CACNA1B	9	140971315	T	C	2.41(1.66,3.49)	3.81E-06	6.31%	4.63%		1.01(0.88,1.15)	9.30E-01	4.78%	4.77%	NA	NA	NA	NA	1.11(0.98,1.27)	9.53E-02	1.62E-05
rs1692144	GJA5	1	147281349	C	T	1.37(1.19,1.57)	8.19E-06	81.52%	77.06%		1.04(0.97,1.12)	2.90E-01	79.54%	79.06%			0.92(0.84,1.01)	7.00E-02	1.04(0.99,1.10)	1.29E-01	1.68E-05

***Interim release used in UKBB for these signals. Nobese=2,799. Nthin=1,212**

EA= Effect allele (BMI increasing allele); NEA= Non-effect allele; OR = Odds ratio; 95% CI = 95% confidence interval for the odds ratio; EAF = effect allele frequency. HetPVal= Heterozygosity p-value

Positions mapped to hg19

Blue line: Conventional genome-wide significant threshold (p<5E-08) in combined analysis.

S6 Table. Discovery, replication and meta-analysis results for 66 SNPs meeting P<10-5 in discovery association results of SCOOP vs UKHLS analysis.

						SCOOP				UKBB					SCOOP 2013				EGG				GIANT obesity class III				Combined analysis		
rsID	Nearest gene	Chr.	Position (bp)	EA	NEA	OR (95% CI)	P value	EAF Obese	EAF Non-extremes	proxy rsID	OR (95% CI)	P value	EAF Obese	EAF Non-extremes	proxy rsID	r2	OR (95% CI)	P value	proxy rsID	r2	OR (95% CI)	P value	proxy rsID	r2	OR (95% CI)	P value	OR (95% CI)	P value	HetPVal
rs9928094	FTO	16	53799905	G	A	1.44(1.33,1.57)	1.42E-18	49.50%	41.32%		1.30(1.25,1.35)	2.74E-41	48.34%	41.91%			1.46(1.34,1.60)	4.88E-17			1.21(1.15,1.28)	7.20E-13			1.43(1.34,1.54)	6.60E-25	1.32(1.29,1.36)	5.94E-101	4.41E-05
rs35614134	MC4R	18	57832856	AC	A	1.31(1.20,1.44)	6.27E-09	29.01%	23.69%	rs663129	1.22(1.16,1.27)	1.25E-18	26.72%	23.15%	rs663129	0.99814	1.32(1.19,1.46)	1.22E-07	rs663129	0.99814	1.22(1.15,1.30)	1.27E-10	rs663129	0.99814	1.20(1.10,1.30)	1.70E-05	1.23(1.20,1.27)	1.57E-43	3.55E-01
rs66906321	TMEM18	2	630070	C	T	1.40(1.24,1.57)	2.35E-08	85.78%	81.35%		1.17(1.11,1.24)	3.44E-09	84.44%	82.20%	rs13007080	0.95655	1.39(1.24,1.57)	6.65E-08	rs13007080	0.956548	1.28(1.19,1.37)	5.15E-12	rs13007080	0.956548	1.27(1.15,1.40)	3.40E-06	1.25(1.21,1.29)	9.72E-35	1.33E-02
rs7132908	FAIM2	12	50263148	A	G	1.22(1.12,1.32)	3.27E-06	42.45%	37.82%		1.15(1.10,1.19)	5.37E-12	41.11%	37.71%	rs7138803	0.90165	1.23(1.12,1.35)	8.89E-06			1.18(1.11,1.25)	1.24E-08			1.20(1.10,1.30)	6.60E-06	1.17(1.14,1.21)	2.38E-31	4.86E-01
rs2384060	ADCY3-DNAJC27	2	25135438	G	A	1.23(1.13,1.34)	1.53E-06	43.52%	38.90%	rs6722587	1.11(1.07,1.15)	4.89E-08	47.67%	44.93%	rs6722587	0.90829	1.09(1.00,1.19)	5.01E-02	rs6722587	0.908287	1.18(1.12,1.24)	1.02E-09	rs6722587	0.908287	1.12(1.04,1.19)	1.60E-03	1.14(1.11,1.17)	9.39E-23	1.13E-01
rs11209947	NEGR1	1	72808551	A	T	1.30(1.17,1.44)	8.51E-07	76.58%	72.18%	rs1460940	1.11(1.05,1.16)	4.53E-05	81.18%	79.76%	rs1460940	0.80203	1.46(1.30,1.63)	2.21E-10	rs1460940	0.802029	1.13(1.06,1.22)	4.60E-04	rs1460940	0.802029	1.22(1.11,1.35)	5.60E-05	1.17(1.13,1.21)	5.17E-20	7.26E-05
rs12735657	SEC16B	1	177809133	C	T	1.24(1.13,1.37)	9.72E-06	24.26%	20.46%		1.12(1.07,1.17)	1.48E-06	22.87%	20.94%	rs4132288	0.9914	1.20(1.07,1.33)	1.18E-03	rs4132288	0.991399	1.14(1.06,1.21)	1.52E-04	rs4132288	0.991399	1.22(1.11,1.34)	1.80E-05	1.15(1.12,1.19)	7.26E-19	1.79E-01
rs13104545	GNPDA2	4	45184907	A	G	1.27(1.15,1.40)	1.61E-06	27.41%	23.45%		1.07(1.02,1.11)	5.35E-03	24.36%	23.26%	NA	NA	NA	NA			1.13(1.04,1.22)	3.39E-03			1.34(1.20,1.49)	1.20E-07	1.13(1.09,1.17)	1.47E-11	9.39E-05
rs112446794	CEP120	5	122665465	T	C	1.23(1.13,1.35)	2.08E-06	33.15%	28.69%		1.07(1.02,1.11)	2.55E-03	29.47%	28.21%	rs1366333	0.95876	1.08(0.98,1.19)	1.38E-01	rs1366333	0.958762	1.12(1.06,1.18)	1.22E-04	rs1366333	0.958762	1.05(0.97,1.13)	2.40E-01	1.09(1.06,1.13)	3.45E-10	3.33E-02
rs3760091	SULT1A1	16	28620800	C	G	1.24(1.14,1.35)	1.56E-06	64.89%	60.23%		1.09(1.04,1.14)	1.19E-04	63.49%	61.44%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.12(1.07,1.16)	2.65E-08	8.49E-03
rs115474151	SLC7A14	3	170272967	A	T	1.36(1.20,1.55)	2.63E-06	12.44%	9.53%		1.27(0.86,1.87)	2.36E-01	0.26%	0.20%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.35(1.20,1.53)	1.40E-06	7.24E-01
rs8096590	LINC01541	18	69231235	A	G	1.24(1.14,1.35)	1.26E-06	34.39%	29.80%		1.03(0.99,1.07)	1.68E-01	31.28%	30.75%	rs7228473	0.98094	1.08(0.99,1.19)	9.70E-02	rs7228473	0.980939	1.02(0.97,1.08)	4.30E-01	rs7228473	0.980939	1.14(1.05,1.23)	1.40E-03	1.07(1.04,1.10)	7.38E-06	7.32E-04
rs201388971*	CDKAL1	6	20514945	T	TTG	1.33(1.17,1.50)	7.17E-06	85.90%	82.52%		1.10(1.01,1.20)	3.58E-02	84.15%	82.90%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.17(1.09,1.25)	1.77E-05	1.36E-02
rs141442356	FAT1	4	187723286	T	TG	1.23(1.13,1.35)	6.26E-06	42.33%	38.18%		0.97(0.76,1.24)	8.27E-01	99.39%	99.42%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.20(1.10,1.31)	2.98E-05	7.57E-02
rs10456655	PKHD1	6	51817204	G	C	1.30(1.16,1.44)	2.13E-06	20.35%	16.84%		1.06(1.01,1.12)	2.00E-02	17.45%	16.53%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.10(1.05,1.16)	3.34E-05	1.10E-03
rs7752955	TFAP2D	6	50402660	T	C	1.28(1.16,1.40)	6.46E-07	76.66%	72.16%		1.04(1.00,1.09)	6.33E-02	73.38%	72.61%			1.07(0.97,1.18)	1.89E-01			1.06(1.00,1.12)	6.05E-02			1.00(0.92,1.09)	9.50E-01	1.06(1.03,1.10)	3.90E-05	2.37E-03
rs898708	PNOC	8	28119194	C	T	1.26(1.15,1.38)	9.94E-07	73.10%	68.50%		1.02(0.98,1.07)	2.77E-01	69.52%	69.11%			1.06(0.96,1.17)	2.51E-01			1.04(0.98,1.10)	1.53E-01			1.08(0.99,1.16)	7.10E-02	1.06(1.03,1.09)	9.27E-05	2.38E-03
rs34208875	CASC17	17	69195603	C	T	2.95(1.90,4.59)	1.55E-06	1.33%	0.59%		1.18(0.90,1.55)	2.29E-01	0.52%	0.43%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.53(1.21,1.93)	3.78E-04	5.57E-04
rs540249707	METTL7B	12	56072632	G	A	3.60(2.13,6.08)	1.77E-06	1.03%	0.41%		1.17(0.78,1.75)	4.46E-01	0.23%	0.21%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.78(1.29,2.46)	4.25E-04	9.06E-04
rs10944524	MIR4643	6	91898609	T	C	1.28(1.15,1.43)	5.88E-06	18.12%	14.73%		1.05(0.99,1.10)	9.69E-02	15.62%	15.13%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.09(1.04,1.14)	4.94E-04	8.25E-04
rs60581051	ANKS1B	12	100147984	T	C	1.30(1.16,1.45)	7.81E-06	16.53%	13.40%		1.05(0.99,1.11)	8.71E-02	14.04%	13.55%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.09(1.04,1.15)	5.02E-04	1.01E-03
rs147725108	LINC00354	13	112579895	C	G	1.89(1.44,2.48)	4.66E-06	3.10%	1.92%		1.11(0.93,1.32)	2.61E-01	1.20%	1.07%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.30(1.12,1.51)	5.74E-04	1.27E-03
rs142450848	CARD18	11	105111566	C	T	2.34(1.69,3.25)	3.54E-07	2.08%	0.97%		1.10(0.92,1.33)	3.00E-01	1.07%	0.96%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.33(1.13,1.57)	5.89E-04	9.62E-05
rs5873242*	RANBP17	5	170321490	A	T	1.25(1.14,1.38)	2.73E-06	72.23%	67.87%		1.03(0.96,1.10)	4.25E-01	68.78%	68.25%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.10(1.04,1.17)	6.63E-04	8.90E-04
rs67529790*	LINC00977	8	130055514	AT	A	1.26(1.14,1.39)	2.78E-06	42.00%	37.87%		1.02(0.95,1.10)	5.13E-01	38.60%	38.10%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.10(1.04,1.17)	8.05E-04	8.21E-04
rs138019013	ALS2CL	3	46737713	T	G	3.92(2.29,6.69)	5.75E-07	1.01%	0.40%		0.97(0.58,1.60)	8.90E-01	0.14%	0.14%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.87(1.29,2.70)	8.55E-04	1.86E-04
rs28581396*	SLC38A11	2	165876796	C	T	2.14(1.54,2.98)	6.32E-06	2.01%	1.06%		1.07(0.80,1.42)	6.50E-01	1.42%	1.34%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.44(1.16,1.79)	9.60E-04	1.84E-03
rs144209184	SCD5	4	83646583	C	T	2.40(1.64,3.51)	7.09E-06	1.70%	0.88%		1.14(0.90,1.44)	2.74E-01	0.68%	0.59%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.40(1.15,1.71)	1.01E-03	1.15E-03
rs116931808*	CSMD1	8	4035481	C	G	2.29(1.59,3.29)	7.45E-06	1.60%	0.73%		1.01(0.70,1.47)	9.45E-01	0.72%	0.72%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.54(1.19,2.00)	1.10E-03	2.13E-03
rs11563327	HOXA1	7	27071888	C	T	1.24(1.13,1.36)	9.41E-06	74.64%	70.42%		1.03(0.99,1.08)	1.16E-01	71.80%	71.12%	NA	NA	NA	NA			1.03(0.97,1.09)	3.30E-01			1.03(0.95,1.11)	5.30E-01	1.05(1.02,1.08)	1.21E-03	4.77E-03
rs540119135	PTPRR	12	71094332	G	A	3.17(2.03,4.95)	4.21E-07	1.24%	0.45%		1.05(0.75,1.46)	7.90E-01	0.33%	0.33%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.55(1.19,2.03)	1.22E-03	9.42E-05
rs66671632	ARC	8	143680772	T	C	1.31(1.17,1.47)	2.62E-06	16.48%	13.14%		1.04(0.98,1.10)	2.32E-01	12.61%	12.26%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.09(1.03,1.15)	1.35E-03	2.70E-04
rs141214244*	RAB11FIP2	10	119652851	T	GTGTG	1.32(1.18,1.49)	3.45E-06	21.01%	17.88%		1.02(0.94,1.12)	5.92E-01	19.37%	18.93%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.12(1.04,1.20)	1.51E-03	5.94E-04
rs2237402	POU6F2	7	39449768	G	A	1.23(1.12,1.34)	6.31E-06	69.78%	65.38%		1.04(1.00,1.08)	7.72E-02	66.68%	65.95%	rs2237403	0.99506	1.02(0.93,1.12)	7.31E-01	rs2237403	0.995062	1.01(0.95,1.06)	7.75E-01	rs2237403	0.995062	1.04(0.96,1.11)	3.30E-01	1.04(1.02,1.07)	1.78E-03	5.14E-03
rs147345620	GPC6	13	94017255	G	A	3.02(1.89,4.82)	3.46E-06	1.25%	0.60%		1.12(0.85,1.47)	4.20E-01	0.51%	0.45%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.44(1.14,1.83)	2.37E-03	3.17E-04
rs11589523	FAM46C	1	118270980	T	C	1.22(1.12,1.32)	6.46E-06	36.66%	32.31%		1.02(0.98,1.07)	2.64E-01	33.42%	32.89%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.06(1.02,1.10)	3.26E-03	3.27E-04
rs556954774*	LOC100130992	10	22549699	C	CT	1.33(1.17,1.51)	9.52E-06	16.87%	14.13%		1.01(0.91,1.12)	8.35E-01	14.57%	14.53%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.12(1.04,1.22)	3.54E-03	8.29E-04
rs34515326*	MIR99AHG	21	17987843	CA	C	1.95(1.46,2.60)	6.33E-06	3.12%	2.02%		1.00(0.79,1.26)	9.72E-01	2.17%	2.16%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.30(1.09,1.57)	4.58E-03	4.39E-04
rs139198909	ZFHX3	16	72929809	C	A	2.50(1.68,3.73)	7.18E-06	1.50%	0.74%		1.08(0.85,1.36)	5.37E-01	0.68%	0.64%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.34(1.09,1.64)	5.10E-03	3.66E-04
rs191312158*	CCNY	10	35871176	C	T	2.76(1.80,4.23)	3.30E-06	1.42%	0.74%		0.96(0.69,1.35)	8.27E-01	0.98%	1.04%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.44(1.11,1.88)	6.68E-03	1.54E-04
6:153400217_AT_A	RGS17	6	153400217	AT	A	1.24(1.13,1.36)	6.04E-06	62.00%	57.94%		1.02(0.97,1.07)	4.13E-01	74.18%	73.83%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.06(1.02,1.10)	6.69E-03	1.98E-04
rs191634319	SOS1	2	39318135	C	G	3.54(2.13,5.88)	1.09E-06	1.04%	0.41%		1.05(0.78,1.40)	7.66E-01	0.43%	0.43%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.42(1.10,1.83)	7.11E-03	4.52E-05
rs62179502	SPATS2L	2	201146942	G	A	2.03(1.51,2.73)	2.31E-06	98.06%	96.28%		1.06(0.95,1.19)	3.12E-01	97.12%	96.98%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.15(1.04,1.28)	8.37E-03	5.18E-05
rs543354609	MDS2	1	23930694	C	T	2.79(1.78,4.37)	7.11E-06	1.20%	0.52%		1.04(0.80,1.36)	7.73E-01	0.51%	0.50%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.35(1.07,1.70)	1.04E-02	2.15E-04
rs144759478	PTH2R	2	209323283	T	C	1.64(1.34,2.02)	2.06E-06	5.13%	3.42%		1.02(0.92,1.14)	7.04E-01	3.29%	3.23%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.13(1.03,1.25)	1.10E-02	5.66E-05
rs6496033	LOC440311	15	95245915	A	C	1.47(1.25,1.73)	4.82E-06	8.81%	6.65%		1.00(0.93,1.08)	9.73E-01	7.37%	7.35%	NA	NA	NA	NA			1.07(0.96,1.19)	2.48E-01			1.19(0.96,1.47)	1.10E-01	1.07(1.02,1.13)	1.16E-02	3.80E-04
rs1571570	PBX3	9	128748249	C	G	1.41(1.22,1.63)	3.08E-06	9.45%	6.91%		1.08(1.00,1.16)	4.00E-02	7.33%	6.82%	NA	NA	NA	NA			0.97(0.89,1.06)	5.17E-01			0.96(0.83,1.12)	6.30E-01	1.07(1.01,1.12)	1.28E-02	1.40E-04
rs28700201*	TP53TG3D	16	32275881	A	G	1.33(1.18,1.50)	2.04E-06	49.33%	45.88%		0.96(0.88,1.06)	4.72E-01	54.34%	54.74%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.10(1.02,1.18)	1.36E-02	3.76E-05
rs571984391*	LMX1B	9	129414279	TG	T	1.22(1.12,1.33)	6.58E-06	44.94%	40.48%		0.99(0.93,1.06)	7.89E-01	42.42%	42.68%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.07(1.01,1.12)	1.54E-02	1.39E-04
rs524349	UNC13C	15	54689366	A	G	1.29(1.17,1.41)	8.01E-08	73.26%	67.99%		1.01(0.97,1.05)	7.10E-01	69.42%	69.20%	rs563881	0.9026	1.00(0.91,1.11)	9.41E-01	rs563881	0.902595	1.02(0.96,1.08)	5.48E-01	rs563881	0.902595	1.02(0.95,1.11)	5.80E-01	1.04(1.01,1.07)	1.64E-02	8.42E-05
rs144435735	LINC00682	4	41881253	A	G	1.98(1.49,2.61)	1.96E-06	3.26%	2.06%		0.95(0.80,1.13)	5.51E-01	1.25%	1.33%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.16(1.00,1.34)	4.81E-02	1.23E-05
rs73676140	NXPH1	7	8729124	G	A	1.59(1.29,1.95)	9.74E-06	4.75%	3.16%		0.99(0.88,1.10)	7.95E-01	3.07%	3.12%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.10(1.00,1.21)	6.06E-02	5.89E-05
rs139813768	SNX5	20	17828396	A	G	2.47(1.66,3.67)	8.32E-06	1.57%	0.83%		0.94(0.75,1.19)	6.21E-01	0.65%	0.68%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.21(0.99,1.48)	6.46E-02	4.32E-05
rs6806545	LOC647323	3	193660056	T	C	1.21(1.11,1.32)	9.08E-06	41.15%	36.68%		0.99(0.95,1.03)	6.64E-01	37.72%	37.84%	NA	NA	NA	NA			1.03(0.98,1.09)	2.74E-01			1.00(0.94,1.07)	9.20E-01	1.02(1.00,1.05)	8.16E-02	4.24E-04
rs146883791	DCHS2	4	155121474	A	ATAAG	1.74(1.37,2.21)	5.48E-06	3.42%	2.02%	rs7680244	0.94(0.83,1.07)	3.43E-01	2.26%	2.41%	rs7680244	0.99439	1.43(1.04,1.96)	2.95E-02	rs7680244	0.994386	0.99(0.80,1.23)	9.25E-01	NA	NA	NA	NA	1.09(0.99,1.20)	8.79E-02	3.30E-05
rs9863931	ADAMTS9-AS2	3	65014778	T	C	1.31(1.17,1.47)	3.89E-06	15.79%	12.47%		1.01(0.96,1.07)	6.52E-01	12.40%	12.17%	NA	NA	NA	NA			0.98(0.91,1.06)	6.73E-01			0.99(0.88,1.12)	9.20E-01	1.04(0.99,1.08)	9.31E-02	2.84E-04
rs7069309	CAMK1D	10	12351089	T	C	1.20(1.11,1.30)	5.59E-06	49.24%	45.94%		1.01(0.97,1.05)	5.94E-01	50.11%	49.66%	NA	NA	NA	NA			0.99(0.94,1.05)	7.17E-01			0.99(0.92,1.07)	8.10E-01	1.02(1.00,1.05)	9.91E-02	3.72E-04
rs79325679	LARGE	22	33929898	T	C	2.07(1.55,2.76)	7.71E-07	2.50%	1.26%		0.93(0.80,1.09)	3.70E-01	1.47%	1.59%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.12(0.97,1.28)	1.16E-01	1.79E-06
rs11872477		18	2037477	T	C	1.26(1.14,1.39)	5.92E-06	37.93%	34.04%		0.99(0.94,1.03)	6.07E-01	32.52%	32.80%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.03(0.99,1.08)	1.44E-01	1.55E-05
rs190999252	CCSER1	4	92847036	A	G	2.07(1.51,2.84)	7.12E-06	2.25%	1.27%		0.93(0.79,1.11)	4.34E-01	1.26%	1.34%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.12(0.96,1.30)	1.52E-01	1.49E-05
rs75809547	PTBP2	1	97331151	C	T	2.61(1.74,3.93)	3.85E-06	1.46%	0.65%		0.86(0.68,1.09)	2.21E-01	0.62%	0.72%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.14(0.93,1.41)	2.00E-01	4.10E-06
rs147842378	AGO4	1	36276202	C	G	2.01(1.49,2.71)	5.08E-06	3.16%	2.10%		0.91(0.78,1.07)	2.67E-01	1.48%	1.63%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.08(0.94,1.25)	2.54E-01	5.19E-06
11:23234822_ATGG_A	MIR8054	11	23234822	ATGG	A	1.22(1.13,1.33)	2.68E-06	43.06%	38.46%	rs2403819	0.99(0.96,1.03)	7.41E-01	39.17%	39.33%	rs2403819	0.97568	1.01(0.92,1.11)	8.49E-01	rs2403819	0.975677	0.96(0.91,1.01)	1.04E-01	rs2403819	0.975677	1.03(0.96,1.10)	4.00E-01	1.01(0.98,1.04)	4.19E-01	5.19E-05
rs12261064	GPR26	10	125429706	A	G	1.55(1.30,1.86)	1.48E-06	6.21%	4.14%		0.90(0.82,0.99)	2.71E-02	4.32%	4.72%			0.86(0.69,1.05)	1.44E-01			0.92(0.82,1.04)	2.00E-01			1.05(0.86,1.28)	6.20E-01	1.02(0.96,1.09)	4.68E-01	2.40E-06
rs1188854	ECT2L	6	139196524	G	A	1.39(1.21,1.61)	6.57E-06	10.13%	7.69%		0.97(0.90,1.03)	3.14E-01	8.20%	8.43%	NA	NA	NA	NA			1.00(0.91,1.10)	9.58E-01			0.94(0.81,1.11)	4.80E-01	1.02(0.97,1.07)	5.48E-01	8.42E-05
rs3104056	COL13A1	10	71510039	A	G	1.53(1.27,1.84)	6.31E-06	5.64%	3.78%		0.85(0.76,0.94)	1.16E-03	3.55%	4.09%			1.05(0.83,1.32)	6.83E-01			0.98(0.84,1.14)	8.09E-01			1.20(0.93,1.54)	1.60E-01	1.00(0.94,1.08)	9.08E-01	1.17E-06

***Interim release used in UKBB for these signals. Nobese=2,799. Ncontrols=8,193**

EA= Effect allele (BMI increasing allele); NEA= Non-effect allele; OR = Odds ratio; 95% CI = 95% confidence interval for the odds ratio; EAF = effect allele frequency. HetPVal= Heterozygosity p-value

Positions mapped to hg19

S7 Table. Discovery, replication and meta-analysis results for 37 SNPs meeting P<10-5 in discovery association results of UKHLS vs STILTS analysis.

						STILTS				UKBB				Combined analysis		
rsID	Nearest gene	Chr.	Position (bp)	EA	NEA	OR (95% CI)	P value	EAF Non-extremes	EAF Thin	OR (95% CI)	P value	EAF Non-extremes	EAF Thin	OR (95% CI)	P value	HetPVal
rs13262703	PI15	8	75819902	A	T	4.15 (2.42,7.11)	2.32E-07	99.62%	98.88%	1.69(0.98,2.91)	5.68E-02	99.84%	99.74%	2.66(1.81,3.89)	5.46E-07	2.15E-02
rs558258836*	LOC102724874	8	78716821	T	A	4.04 (2.25,7.26)	3.07E-06	99.60%	99.04%	1.69 (0.89,3.2)	1.07E-01	99.50%	99.28%	2.71 (1.76,4.18)	5.99E-06	4.90E-02
rs2123163	CADM2	3	85243797	T	G	1.68 (1.35,2.1)	4.90E-06	6.40%	4.20%	1.14(1.00,1.29)	4.22E-02	5.12%	4.60%	1.25(1.12,1.39)	6.18E-05	2.76E-03
rs150756788	SLC2A7	1	9050295	G	T	2.09 (1.52,2.87)	4.96E-06	98.21%	97.25%	1.20(0.91,1.60)	2.00E-01	99.26%	99.11%	1.54(1.25,1.90)	6.41E-05	1.07E-02
rs545797179*	FOXN2	2	48546924	AT	A	2.79 (1.8,4.32)	4.31E-06	99.24%	98.48%	1.15 (0.65,2.05)	6.23E-01	99.18%	99.07%	2.02 (1.42,2.86)	7.77E-05	1.65E-02
rs117638949*	PIGZ	3	196692722	T	A	3.5 (2.27,5.4)	1.50E-08	99.50%	98.55%	0.54 (0.27,1.09)	8.60E-02	99.30%	99.62%	2.09 (1.44,3.02)	9.25E-05	8.97E-06
rs576762972*	CACNA1C	12	2244717	T	C	2.17 (1.55,3.05)	7.23E-06	98.99%	98.03%	1.16 (0.78,1.72)	4.69E-01	98.87%	98.73%	1.66 (1.29,2.15)	1.05E-04	1.79E-02
rs138454709*	COL8A2	1	36592131	A	G	2.58 (1.72,3.88)	5.29E-06	99.03%	98.33%	1.11 (0.65,1.9)	7.04E-01	99.04%	99.00%	1.89 (1.37,2.62)	1.17E-04	1.41E-02
rs75937976	C3orf38	3	88321976	G	C	2.95 (2.02,4.32)	2.43E-08	99.20%	98.25%	1.10(0.84,1.44)	4.96E-01	99.13%	99.05%	1.53(1.23,1.91)	1.52E-04	3.33E-05
rs190051670	PHF2	9	96460947	C	T	2.55 (1.73,3.76)	2.11E-06	99.25%	98.35%	1.19(0.91,1.56)	2.00E-01	99.18%	99.05%	1.53(1.23,1.91)	1.68E-04	1.59E-03
rs56152157	EDIL3	5	83171742	G	A	1.21 (1.11,1.31)	6.91E-06	47.95%	42.99%	1.04(0.99,1.10)	1.21E-01	47.37%	46.44%	1.09(1.04,1.14)	2.11E-04	2.85E-03
rs139226692*	ASAH1	8	17928720	C	CA	2.82 (1.78,4.46)	9.46E-06	99.56%	98.81%	1.11 (0.63,1.92)	7.24E-01	99.37%	99.34%	1.93 (1.35,2.75)	2.73E-04	1.08E-02
rs112958625*	KNDC1	10	134969737	G	A	2.8 (1.81,4.33)	3.61E-06	99.00%	98.39%	1.01 (0.59,1.72)	9.73E-01	98.93%	98.94%	1.86 (1.33,2.62)	3.02E-04	3.75E-03
rs68090520*	ZMAT3	3	178717361	C	A	1.24 (1.13,1.36)	4.04E-06	54.37%	50.43%	1.02 (0.93,1.11)	7.45E-01	53.49%	53.13%	1.12 (1.05,1.2)	4.39E-04	2.72E-03
rs17544568	ONECUT1	15	53321119	G	A	2.04 (1.54,2.7)	6.53E-07	97.94%	96.67%	1.09(0.93,1.29)	2.78E-01	97.58%	97.40%	1.28(1.11,1.47)	5.80E-04	1.74E-04
rs143866745*	LOC101927495	11	61356693	C	T	1.31 (1.17,1.46)	1.26E-06	60.23%	56.57%	0.89 (0.51,1.55)	6.88E-01	99.16%	99.21%	1.84 (1.29,2.61)	6.65E-04	9.33E-04
rs184273748*	PTPRU	1	29562801	G	A	2.53 (1.71,3.73)	3.04E-06	99.12%	98.24%	0.65 (0.33,1.26)	2.00E-01	99.17%	99.35%	1.79 (1.28,2.5)	7.11E-04	5.41E-04
rs115861768	MIR4426	16	60885992	C	T	3.27 (1.93,5.52)	9.57E-06	99.53%	98.98%	1.14(0.73,1.76)	5.68E-01	99.65%	99.64%	1.76(1.25,2.46)	1.04E-03	2.46E-03
rs191980904*	UQCRC2	16	21946517	C	T	2.98 (1.85,4.79)	6.69E-06	99.56%	98.93%	0.84 (0.46,1.55)	5.84E-01	99.36%	99.44%	1.85 (1.27,2.7)	1.28E-03	1.39E-03
rs11665052	MC4R	18	57908675	G	A	1.31 (1.18,1.44)	1.40E-07	27.11%	22.61%	1.02(0.96,1.08)	5.63E-01	26.22%	25.78%	1.09(1.03,1.14)	1.48E-03	2.22E-05
2:25411587_C_CT*	POMC	2	25411587	C	CT	1.36 (1.21,1.51)	6.52E-08	83.75%	79.76%	1.01(0.95,1.08)	7.22E-01	82.18%	82.10%	1.10(1.04,1.16)	1.76E-03	9.91E-06
rs137887309	CDH23	10	73211425	G	A	2.66 (1.74,4.07)	6.24E-06	99.48%	98.66%	1.03(0.71,1.50)	8.67E-01	99.54%	99.50%	1.56(1.18,2.06)	1.94E-03	9.94E-04
rs11757467	EYA4	6	133808153	A	T	1.71 (1.35,2.17)	8.55E-06	97.57%	96.11%	1.05(0.90,1.24)	5.26E-01	97.48%	97.33%	1.23(1.08,1.40)	2.43E-03	9.08E-04
rs148209625	ZNF664-FAM101A	12	124681051	C	T	2.2 (1.58,3.07)	2.97E-06	99.02%	97.95%	1.04(0.82,1.32)	7.63E-01	98.89%	98.85%	1.35(1.11,1.64)	2.74E-03	3.19E-04
rs71515311*	TMEM72-AS1	10	45116672	A	ATAT	1.25 (1.13,1.38)	8.55E-06	70.65%	66.71%	0.99 (0.89,1.09)	7.72E-01	69.59%	70.05%	1.11 (1.04,1.19)	2.84E-03	9.00E-04
rs11557769	ACTN1	14	69341653	T	A	1.95 (1.5,2.55)	8.61E-07	98.33%	96.94%	0.91(0.70,1.17)	4.48E-01	98.83%	98.98%	1.31(1.09,1.57)	4.41E-03	4.45E-05
rs142441937	KLF15	3	126030681	G	A	2.53 (1.69,3.79)	6.78E-06	99.16%	98.44%	1.02(0.78,1.34)	8.75E-01	99.08%	99.11%	1.36(1.08,1.70)	8.09E-03	2.71E-04
rs117944743	ZNF93	19	20060211	C	G	2.06 (1.5,2.82)	7.33E-06	98.81%	97.84%	1.01(0.82,1.25)	9.34E-01	98.49%	98.47%	1.26(1.05,1.50)	1.06E-02	2.28E-04
rs142425331	CHCHD3	7	132583478	G	A	1.6 (1.31,1.96)	4.36E-06	96.58%	94.82%	0.98(0.84,1.13)	7.44E-01	96.61%	96.71%	1.16(1.03,1.30)	1.59E-02	8.80E-05
rs138251346	LOC101929452	2	7279064	A	G	2.99 (1.9,4.7)	2.23E-06	99.27%	98.62%	0.77(0.50,1.19)	2.42E-01	99.53%	99.64%	1.46(1.07,2.00)	1.66E-02	2.20E-05
rs1435711	ADAMTS20	12	43429113	G	A	1.32 (1.18,1.49)	2.11E-06	86.34%	83.30%	0.99(0.92,1.07)	8.17E-01	85.74%	85.93%	1.08(1.01,1.15)	1.76E-02	3.92E-05
rs553440779	KCNJ3	2	155835504	T	C	2.67 (1.74,4.09)	6.98E-06	99.27%	98.53%	0.72(0.45,1.15)	1.66E-01	99.61%	99.70%	1.46(1.07,2.00)	1.78E-02	4.88E-05
rs77709566	INTU	4	128466995	A	G	2 (1.5,2.66)	2.01E-06	98.42%	97.20%	0.97(0.81,1.17)	7.51E-01	97.98%	98.03%	1.20(1.03,1.40)	2.02E-02	3.20E-05
rs514529	LRP5	11	68090836	T	A	1.23 (1.13,1.34)	1.09E-06	53.61%	51.60%	0.99(0.94,1.05)	7.62E-01	52.12%	52.27%	1.05(1.01,1.10)	2.04E-02	1.68E-05
rs200275909*	ADAMTS20	12	43954570	A	AT	3.21 (1.93,5.35)	7.29E-06	99.42%	98.76%	0.88 (0.78,1)	5.84E-02	85.23%	86.54%	1.1 (1.01,1.2)	2.38E-02	4.69E-06
rs73085383	ZNF343	20	2503465	C	T	2.13 (1.56,2.92)	2.05E-06	98.63%	97.60%	0.85(0.66,1.10)	2.28E-01	98.79%	98.97%	1.23(1.01,1.50)	3.92E-02	8.79E-06
rs527595266	ADAMTS16	5	5341419	C	G	2.91 (1.81,4.68)	9.95E-06	99.42%	98.79%	0.93(0.68,1.27)	6.52E-01	99.30%	99.30%	1.31(1.01,1.71)	4.00E-02	8.21E-05

***Interim release used in UKBB for these signals. Nthin=1,212. Ncontrols=8,193**

****rs4665779 was used as a proxy in UKBB**

EA= Effect allele (BMI increasing allele); NEA= Non-effect allele; OR = Odds ratio; 95% CI = 95% confidence interval for the odds ratio; EAF = effect allele frequency. HetPVal= Heterozygosity p-value

Positions mapped to hg19

Red line: Strict genome-wide significant threshold (p<1.17E-08) in combined analysis. Blue line: Conventional genome-wide significant threshold (p<5E-08) in combined analysis.

S10 Table. Published loci from GIANT, EGG and SCOOP 2013 not reaching genome-wide significance in our study

Known BMI loci with meta p <5E-8 in GIANT BMI tails study but not in this study (obese vs thin)

rsID	Gene	OR GIANT BMI tails Stage 1	P GIANT BMI tails Stage 1	OR SCOOP/STILTS	P SCOOP/STILTS	OR SCOOP/UKHLS	P SCOOP/UKHLS	OR UKHLS/STILTS	P UKHLS/STILTS
rs2568958	NEGR1	1.17 (1.12,1.23)	6.80E-10	1.25 (1.11,1.39)	1.00E-04	1.19(1.09,1.29)	5.65E-05	1.06(0.97,1.16)	1.73E-01
rs987237	TFAP2B	1.20 (1.12,1.28)	4.30E-07	1.31 (1.14,1.50)	2.00E-04	1.17(1.05,1.29)	3.25E-03	1.14(1.01,1.27)	2.72E-02
rs2030323	BDNF	1.21 (1.13,1.30)	5.20E-08	1.31 (1.13,1.50)	7.46E-05	1.15(1.03,1.27)	1.03E-02	1.10(1.00,1.22)	4.92E-02
rs1516725	ETV5	1.30 (1.19, 1.42)	2.10E-08	1.30 (1.11,1.52)	8.00E-04	1.16(1.02,1.31)	1.89E-02	1.18(1.05,1.33)	5.03E-03

Loci identified in S.I. Berndt, et al. (2013)

rsID	Gene	OR GIANT Stage 1	P GIANT Stage 1	OR SCOOP/STILTS	P SCOOP/STILTS	OR SCOOP/UKHLS	P SCOOP/UKHLS	OR UKHLS/STILTS	P UKHLS/STILTS	Reported Trait
rs7989336	HS6ST3	1.12	5.88E-09	1.13(1.01,1.26)	3.17E-02	1.03(0.95,1.12)	4.42E-01	1.09(1.00,1.19)	4.15E-02	Obesity class 2
rs17381664	ZZZ3	1.11	7.61E-08	1.00(0.89,1.12)	9.86E-01	0.98(0.91,1.07)	6.99E-01	1.03(0.95,1.12)	4.82E-01	Obesity class 2
rs17024258	GNAT2	1.23	1.41E-06	1.80(1.29,2.53)	6.27E-04	1.57(1.25,1.97)	1.18E-04	1.10(0.82,1.46)	5.32E-01	Obesity class 1
rs4735692	HNF4G	1.07	5.03E-08	1.08(0.97,1.21)	1.57E-01	1.00(0.92,1.09)	9.87E-01	1.07(0.98,1.16)	1.27E-01	Obesity class 1
rs13041126	MRPS33P4	1.07	3.05E-07	1.14(1.01,1.28)	3.88E-02	1.07(0.97,1.17)	1.71E-01	1.03(0.94,1.13)	5.43E-01	Obesity class 1
rs2531995	ADCY9	1.06	3.17E-06	1.14(1.01,1.28)	3.22E-02	1.06(0.97,1.16)	2.06E-01	1.08(0.98,1.18)	1.04E-01	Obesity class 1
rs4735692	HNF4G	1.05	6.13E-09	1.08(0.97,1.21)	1.57E-01	1.00(0.92,1.09)	9.87E-01	1.07(0.98,1.16)	1.27E-01	Overweight
rs7503807	RPTOR	1.04	4.20E-06	1.18(1.06,1.32)	2.90E-03	1.11(1.03,1.21)	1.04E-02	1.07(0.98,1.16)	1.24E-01	Overweight

Loci indentified in J.P. Bradfield, H.R. Taal, et al. (2012)

rsID	Gene	OR EGG Stage 1	P EGG Stage 1	OR SCOOP/STILTS	P SCOOP/STILTS	OR SCOOP/UKHLS	P SCOOP/UKHLS	OR UKHLS/STILTS	P UKHLS/STILTS
rs9568856	OLFM4	1.21	6.58E−7	1.09(0.93,1.28)	2.71E-01	1.14(1.01,1.28)	2.99E-02	0.97(0.86,1.10)	6.41E-01
rs9299	HOXB5	1.14	9.12E−7	1.18(1.05,1.32)	6.46E-03	1.03(0.94,1.12)	5.68E-01	1.09(1.00,1.19)	5.34E-02

Loci identified in E. Wheeler, et al. (2013)

rsID	Gene	OR SCOOP 2013 Stage 1	P SCOOP 2013 Stage 1	OR SCOOP/STILTS	P SCOOP/STILTS	OR SCOOP/UKHLS	P SCOOP/UKHLS	OR UKHLS/STILTS	P UKHLS/STILTS
rs1993709	NEGR1	1.46	1.98E-12	1.30(1.13,1.50)	2.54E-04	1.29(1.16,1.44)	4.45E-06	1.03(0.93,1.14)	6.16E-01
rs1957894	PRKCH	1.64	1.01E-08	1.25(1.03,1.51)	2.61E-02	1.17(1.02,1.35)	2.40E-02	1.01(0.87,1.18)	8.79E-01
rs11208659	LEPR	1.63	1.16E-10	1.22(1.01,1.48)	4.33E-02	1.28(1.12,1.48)	4.35E-04	0.95(0.81,1.10)	4.90E-01
rs564343	PACS1	1.25	5.81E-08	1.01(0.90,1.13)	9.18E-01	1.04(0.95,1.13)	4.12E-01	0.96(0.89,1.05)	4.12E-01
rs11109072	RMST	1.79	1.48E-07	0.87(0.63,1.20)	3.83E-01	0.95(0.74,1.21)	6.74E-01	0.97(0.76,1.24)	8.13E-01

S12 Table. ICD10 codes used to exclude thin individuals in UKBB

Code	ICD10 Code Description
A071	A07.1 Giardiasis [lambliasis]
A150	A15.0 Tuberculosis of lung, confirmed by sputum microscopy with or without culture
A151	A15.1 Tuberculosis of lung, confirmed by culture only
A152	A15.2 Tuberculosis of lung, confirmed histologically
A159	A15.9 Respiratory tuberculosis unspecified, confirmed bacteriologically and histologically
A162	A16.2 Tuberculosis of lung, without mention of bacteriological or histological confirmation
A169	A16.9 Respiratory tuberculosis unspecified, without mention of bacteriological or histological confirmation
B181	B18.1 Chronic viral hepatitis B without delta-agent
B182	B18.2 Chronic viral hepatitis C
B203	B20.3 HIV disease resulting in other viral infections
B204	B20.4 HIV disease resulting in candidiasis
B238	B23.8 HIV disease resulting in other specified conditions
B24	B24 Unspecified human immunodeficiency virus [HIV] disease
C01	C01 Malignant neoplasm of base of tongue
C099	C09.9 Tonsil, unspecified
C108	C10.8 Overlapping lesion of oropharynx
C109	C10.9 Oropharynx, unspecified
C155	C15.5 Lower third of oesophagus
C159	C15.9 Oesophagus, unspecified
C169	C16.9 Stomach, unspecified
C172	C17.2 Ileum
C180	C18.0 Caecum
C182	C18.2 Ascending colon
C184	C18.4 Transverse colon
C187	C18.7 Sigmoid colon
C189	C18.9 Colon, unspecified
C20	C20 Malignant neoplasm of rectum
C210	C21.0 Anus, unspecified
C211	C21.1 Anal canal
C220	C22.0 Liver cell carcinoma
C221	C22.1 Intrahepatic bile duct carcinoma
C250	C25.0 Head of pancreas
C258	C25.8 Overlapping lesion of pancreas
C259	C25.9 Pancreas, unspecified
C341	C34.1 Upper lobe, bronchus or lung
C342	C34.2 Middle lobe, bronchus or lung
C343	C34.3 Lower lobe, bronchus or lung
C349	C34.9 Bronchus or lung, unspecified
C411	C41.1 Mandible
C435	C43.5 Malignant melanoma of trunk
C436	C43.6 Malignant melanoma of upper limb, including shoulder
C437	C43.7 Malignant melanoma of lower limb, including hip
C439	C43.9 Malignant melanoma of skin, unspecified
C441	C44.1 Skin of eyelid, including canthus
C442	C44.2 Skin of ear and external auricular canal
C443	C44.3 Skin of other and unspecified parts of face
C444	C44.4 Skin of scalp and neck
C445	C44.5 Skin of trunk
C446	C44.6 Skin of upper limb, including shoulder
C447	C44.7 Skin of lower limb. including hip
C482	C48.2 Peritoneum, unspecified
C503	C50.3 Lower-inner quadrant of breast
C504	C50.4 Upper-outer quadrant of breast
C505	C50.5 Lower-outer quadrant of breast
C509	C50.9 Breast, unspecified
C541	C54.1 Endometrium
C56	C56 Malignant neoplasm of ovary
C61	C61 Malignant neoplasm of prostate
C64	C64 Malignant neoplasm of kidney, except renal pelvis
C66	C66 Malignant neoplasm of ureter
C679	C67.9 Bladder, unspecified
C719	C71.9 Brain, unspecified
C73	C73 Malignant neoplasm of thyroid gland
C770	C77.0 Lymph nodes of head, face and neck
C771	C77.1 Intrathoracic lymph nodes
C772	C77.2 Intra-abdominal lymph nodes
C773	C77.3 Axillary and upper limb lymph nodes
C779	C77.9 Lymph node, unspecified
C780	C78.0 Secondary malignant neoplasm of lung
C786	C78.6 Secondary malignant neoplasm of retroperitoneum and peritoneum
C787	C78.7 Secondary malignant neoplasm of liver
C788	C78.8 Secondary malignant neoplasm of other and unspecified digestive organs
C793	C79.3 Secondary malignant neoplasm of brain and cerebral meninges
C795	C79.5 Secondary malignant neoplasm of bone and bone marrow
C80	C80 Malignant neoplasm without specification of site
C829	C82.9 Follicular non-Hodgkin's lymphoma, unspecified
C844	C84.4 Peripheral T-cell lymphoma
C859	C85.9 Non-Hodgkin's lymphoma, unspecified type
C880	C88.0 Waldenstr m's macroglobulinaemia
C910	C91.0 Acute lymphoblastic leukaemia
C920	C92.0 Acute myeloid leukaemia
D001	D00.1 Oesophagus
D013	D01.3 Anus and anal canal
D033	D03.3 Melanoma in situ of other and unspecified parts of face
D371	D37.1 Stomach
D374	D37.4 Colon
D375	D37.5 Rectum
D377	D37.7 Other digestive organs
D381	D38.1 Trachea, bronchus and lung
D391	D39.1 Ovary
D400	D40.0 Prostate
D430	D43.0 Brain, supratentorial
D432	D43.2 Brain, unspecified
D443	D44.3 Pituitary gland
D45	D45 Polycythaemia vera
D469	D46.9 Myelodysplastic syndrome, unspecified
D471	D47.1 Chronic myeloproliferative disease
D477	D47.7 Other specified neoplasms of uncertain or unknown behaviour of lymphoid, haematopoietic and related tissue
D481	D48.1 Connective and other soft tissue
D485	D48.5 Skin
D486	D48.6 Breast
D500	D50.0 Iron deficiency anaemia secondary to blood loss (chronic)
D508	D50.8 Other iron deficiency anaemias
D509	D50.9 Iron deficiency anaemia, unspecified
D510	D51.0 Vitamin B12 deficiency anaemia due to intrinsic factor deficiency
D519	D51.9 Vitamin B12 deficiency anaemia, unspecified
D529	D52.9 Folate deficiency anaemia, unspecified
D539	D53.9 Nutritional anaemia, unspecified
E039	E03.9 Hypothyroidism, unspecified
E041	E04.1 Non-toxic single thyroid nodule
E042	E04.2 Non-toxic multinodular goitre
E049	E04.9 Non-toxic goitre, unspecified
E050	E05.0 Thyrotoxicosis with diffuse goitre
E052	E05.2 Thyrotoxicosis with toxic multinodular goitre
E059	E05.9 Thyrotoxicosis, unspecified
E063	E06.3 Autoimmune thyroiditis
E079	E07.9 Disorder of thyroid, unspecified
E101	E10.1 With ketoacidosis
E103	E10.3 With ophthalmic complications
E104	E10.4 With neurological complications
E105	E10.5 With peripheral circulatory complications
E109	E10.9 Without complications
E110	E11.0 With coma
E112	E11.2 With renal complications
E113	E11.3 With ophthalmic complications
E114	E11.4 With neurological complications
E115	E11.5 With perioheral circulatory comolications
E118	E11.8 With unspecified complications
E162	E16.2 Hypoglycaemia, unspecified
E210	E21.0 Primary hyoerparathyroidism
E211	E21.1 Secondary hyperparathyroidism, not elsewhere classified
E212	E21.2 Other hyperparathyroidism
E213	E21.3 Hyoerparathyrodism, unspecified
E222	E22.2 Syndrome of inappropriate secretion of antidiuretic hormone
E229	E22.9 Hyperfunction of pituitary gland, unspecified
E230	E23.0 Hypopituitarism
E232	E23.2 Diabetes insipidus
E236	E23.6 Other disorders of pituitary gland
E237	E23.7 Disorder of pituitary gland, unspecified
E271	E27.1 Primary adrenocortical insufficiency
E272	E27.2 Addisonian crisis
E274	E27.4 Other and unspecified adrenocortical insufficiency
E279	E27.9 Disorder of adrenal gland, unspecified
E348	E34.8 Other specified endocrine disorders
E349	E34.9 Endocrine disorder, unspecified
E46	E46 Unspecified protein-energy malnutrition
E538	E53.8 Deficiency of other specified B group vitamins
E559	E55.9 Vitamin D deficiency, unspecified
E804	E80.4 Gilbert's syndrome
E831	E83.1 Disorders of iron metabolism
E833	E83.3 Disorders of phosphorus metabolism
E834	E83.4 Disorders of magnesium metabolism
E835	E83.5 Disorders of calcium metabolism
E853	E85.3 Secondary systemic amyloidosis
E854	E85.4 Organ-limited amyloidosis
E859	E85.9 Amvloidosis, unspecified
E86	E86 Volume depletion
E870	E87.0 Hyperosmolality and hypernatraemia
E871	E87.1 Hypo-osmolality and hyponatraemia
E872	E87.2 Acidosis
E873	E87.3 Alkalosis
E875	E87.5 Hyperkalaemia
E876	E87.6 Hypokalaemia
E878	E87.8 Other disorders of electrolyte and fluid balance. not elsewhere classified
E880	E88.0 Disorders of plasma-protein metabolism, not elsewhere classified
E881	E88.1 Lipodystrophy, not elsewhere classified
E890	E89.0 Postoracedural hvoothvroidism
E891	E89.1 Postprocedural hypoinsulinaemia
F009	F00.9 Dementia in Alzheimer's disease, unspecified
F019	F01.9 Vascular dementia. unspecified
F03	F03 Unspecified dementia
F059	F05.9 Delirium, unspecified
F067	F06.7 Mild cognitive disorder
F069	F06.9 Unspecified mental disorder due to brain damage and dysfunction and to physical disease
F072	F07.2 Postconcussional syndrome
F09	F09 Unspecified organic or symptomatic mental disorder
F100	F10.0 Acute intoxication
F101	F10.1 Harmful use
F102	F10.2 Dependence syndrome
F103	F10.3 Withdrawal state
F104	F10.4 Withdrawal state with delirium

F105	F10.5 Psychotic disorder
F106	F10.6 Amnesic syndrome
F109	F10.9 Unspecified mental and behavioural disorder
F110	F11.0 Acute intoxication
F111	F11.1 Harmful use
F112	F11.2 Dependence syndrome
F115	F11.5 Psychotic disorder
F121	F12.1 Harmful use
F122	F12.2 Dependence syndrome
F170	F17.0 Acute intoxication
F171	F17.1 Harmful use
F172	F17.2 Dependence syndrome
F173	F17.3 Withdrawal state
F191	F19.1 Harmful use
F193	F19.3 Withdrawal state
F200	F20.0 Paranoid schizophrenia
F206	F20.6 Simple schizophrenia
F208	F20.8 Other schizophrenia
F209	F20.9 Schizophrenia, unspecified
F220	F22.0 Delusional disorder
F230	F23.0 Acute polymorphic psychotic disorder without symptoms of schizophrenia
F231	F23.1 Acute polymorphic psychotic disorder with symptoms of schizophrenia
F239	F23.9 Acute and transient psychotic disorder, unspecified
F258	F25.8 Other schizoaffective disorders
F29	F29 Unspecified nonorganic psychosis
F300	F30.0 Hypomania
F309	F30.9 Manic episode, unspecified
F310	F31.0 Bipolar affective disorder, current episode hypomanic
F312	F31.2 Bipolar affective disorder, current episode manic with psychotic symptoms
F315	F31.5 Bipolar affective disorder, current episode severe depression with psychotic symptoms
F317	F31.7 Bipolar affective disorder, currently in remission
F319	F31.9 Bipolar affective disorder, unspecified
F320	F32.0 Mild depressive episode
F321	F32.1 Moderate depressive episode
F322	F32.2 Severe deoressive episode without osychotic symotoms
F323	F32.3 Severe depressive episode with psychotic symptoms
F328	F32.8 Other depressive episodes
F329	F32.9 Deoressive episode. unsoecified
F330	F33.0 Recurrent depressive disorder, current episode mild
F331	F33.1 Recurrent depressive disorder, current episode moderate
F332	F33.2 Recurrent depressive disorder. current episode severe without osychotic symotoms
F333	F33.3 Recurrent depressive disorder, current episode severe with psychotic symptoms
F334	F33.4 Recurrent depressive disorder, currently in remission
F339	F33.9 Recurrent depressive disorder, unspecified
F341	F34.1 Dysthymia
F380	F38.0 Other single mood [affective] disorders
F402	F40.2 Specific (isolated) phobias
F410	F41.0 Panic disorder [episodic paroxysmal anxiety]
F411	F41.1 Generalised anxiety disorder
F412	F41.2 Mixed anxiety and depressive disorder
F419	F41.9 Anxiety disorder, unspecified
F420	F42.0 Predominantlv obsessional thouehts or ruminations
F428	F42.8 Other obsessive-compulsive disorders
F429	F42.9 Obsessive-compulsive disorder, unspecified
F430	F43.0 Acute stress reaction
F431	F43.1 Posttraumatic stress disorder
F432	F43.2 Adjustment disorders
F439	F43.9 Reaction to severe stress. unspecified
F458	F45.8 Other somatoform disorders
F500	F50.0 Anorexia nervosa
F501	F50.1 Atvpical anorexia nervosa
F502	F50.2 Bulimia nervosa
F508	F50.8 Other eating disorders
F509	F50.9 Eating disorder, unspecified
F522	F52.2 Failure of genital response
F530	F53.0 Mild mental and behavioural disorders associated with the puerperium, not elsewhere classified
F55	F55 Abuse of non-dependence-producing substances
F603	F60.3 Emotionally unstable personality disorder
F605	F60.5 Anankastic personality disorder
F606	F60.6 Anxious [avoidant] personality disorder
F607	F60.7 Dependent personality disorder
F609	F60.9 Personality disorder. unspecified
F633	F63.3 Trichotillomania
F640	F64.0 Transsexualism
F801	F80.1 Expressive laneuaee disorder
F819	F81.9 Developmental disorder of scholastic skills, unspecified
F911	F91.1 Unsocialised conduct disorder
F99	F99 Mental disorder. not otherwise soecified
G309	G30.9 Alzheimer's disease, unspecified
G35	G35 Multiple sclerosis
J440	J44.0 Chronic obstructive oulmonary disease with acute lower resoiratorv infection
J441	J44.1 Chronic obstructive pulmonary disease with acute exacerbation, unspecified
J448	J44.8 Othersoecified chronic obstructive pulmonary disease
J449	J44.9 Chronic obstructive pulmonary disease, unspecified
K500	K50.0 Crohn's disease of small intestine
K501	K50.1 Crohn's disease of large intestine
K508	K50.8 Other Crohn's disease
K509	K50.9 Crohn's disease, unspecified
K510	K51.0 Ulcerative (chronic) enterocolitis
K512	K51.2 Ulcerative (chronic) proctitis
K513	K51.3 Ulcerative (chronic) rectosigmoiditis
K518	K51.8 Other ulcerative colitis
K519	K51.9 Ulcerative colitis, unspecified
K521	K52.1 Toxic gastro-enteritis and colitis
K528	K52.8 Other soecified non-infective gastro-enteritis and colitis
K529	K52.9 Non-infective gastro-enteritis and colitis, unspecified
K580	K58.0 Irritable bowel syndrome with diarrhoea
K589	K58.9 Irritable bowel svndrome without diarrhoea
K700	K70.0 Alcoholic fatty liver
K703	K70.3 Alcoholic cirrhosis of liver
K709	K70.9 Alcoholic liver disease. unspecified
K720	K72.0 Acute and subacute hepatic failure
K729	K72.9 Hepatic failure, unspecified
K740	K74.0 Hepatic fibrosis
K743	K74.3 Primary biliary cirrhosis
K744	K74.4 Secondary biliary cirrhosis
K745	K74.5 Biliary cirrhosis, unspecified
K746	K74.6 Other and unspecified cirrhosis of liver
K750	K75.0 Abscess of liver
K760	K76.0 Fatty (change of) liver, not elsewhere classified
K766	K76.6 Portal hypertension
K767	K76.7 Hepatorenal syndrome
K768	K76.8 Other specified diseases of liver
K769	K76.9 Liver disease, unspecified
K770	K77.0 Liver disorders in infectious and parasitic diseases classified elsewhere
K900	K90.0 Coeliac disease
K904	K90.4 Malabsorption due to intolerance, not elsewhere classified
K909	K90.9 Intestinal malabsorption. unspecified
K910	K91.0 Vomiting following gastro-intestinal surgery
K911	K91.1 Postgastric surgery syndromes
K912	K91.2 Postsurgical malabsorption, not elsewhere classified
K913	K91.3 Postoperative intestinal obstruction
K914	K91.4 Colostomy and enterostomy malfunction
K918	K91.8 Other postprocedural disorders of digestive system, not elsewhere classified
K920	K92.0 Haematemesis
K921	K92.1 Melaena
K922	K92.2 Gastro-intestinal haemorrhage, unspecified
K928	K92.8 Other specified diseases of digestive system
K929	K92.9 Disease of digestive svstem. unsoecified
N180	N18.0 End-stage renal disease
N185	N18.5 Chronic kidney disease, stage 5
N188	N18.8 Other chronic renal failure
N189	N18.9 Chronic renal failure, unspecified
N19	N19 Unspecified renal failure
Q02	Q02 Microcephaly
Q874	Q87.4 Marfan's syndrome
R630	R63.0 Anorexia
R633	R63.3 Feeding difficulties and mismanagement
R634	R63.4 Abnormal weight loss
R64	R64 Cachexia
Y835	Y83.5 Amputation of limb(s)
Z511	Z51.1 Chemotherapy session for neoplasm
Z512	Z51.2 Other chemotherapy
Z800	Z80.0 Family history of malignant neoplasm of digestive organs
Z850	Z85.0 Personal history of malignant neoplasm of digestive organs
Z851	Z85.1 Personal history of malignant neoplasm of trachea, bronchus and lung
Z853	Z85.3 Personal history of malignant neoplasm of breast
Z854	Z85.4 Personal history of malignant neoplasm of genital organs
Z855	Z85.5 Personal history of malienant neoplasm of urinary tract
Z856	Z85.6 Personal history of leukaemia
Z857	Z85.7 Personal history of other malignant neoplasms of lymphoid, haematopoietic and related tissues
Z858	Z85.8 Personal history of malienant neoplasms of other oreans and svstems
Z860	Z86.0 Personal history of other neoplasms
Z864	Z86.4 Personal history of psychoactive substance abuse
Z865	Z86.5 Personal history of other mental and behavioural disorders
Z895	Z89.5 Acquired absence of leg at or below knee
Z896	Z89.6 Acquired absence of leg above knee
Z899	Z89.9 Acquired absence of limb. unspecified
Z901	Z90.1 Acquired absence of breast(s)
Z902	Z90.2 Acquired absence of lung [part of]
Z903	Z90.3 Acquired absence of part of stomach
Z904	Z90.4 Acquired absence of other parts of digestive tract
Z905	Z90.5 Acquired absence of kidney
Z906	Z90.6 Acquired absence of other parts of urinary tract
Z907	Z90.7 Acquired absence of genital organ(s)
Z992	Z99.2 Dependence on renal dialysis
Z993	Z99.3 Dependence on wheelchair

S13 Table. Self-reported illness codes used to exclude thin individuals in UKBB

	Psychiatric
1286	depression
1287	anxiety/panic attacks
1288	nervous breakdown
1289	schizophrenia
1290	deliberate self-harm/suicide attempt
1291	mania/bipolar disorder/manic depression
1469	post-traumatic stress disorder
1470	anorexia/bulimia/other eating disorder
1614	stress
1615	obsessive compulsive disorder (ocd)
1616	insomnia
1408	alcohol dependency
1409	opioid dependency
1410	other substance abuse/dependency
1531	post-natal depression
	Liver
1136	liver/biliary/pancreas problem
1155	hepatitis
1158	liver failure/cirrhosis
1159	bile duct disease
1161	gall bladder disease
1164	pancreatic disease
1507	haemochromatosis
1508	jaundice (unknown cause)
1156	infective/viral hepatitis
1157	non-infective hepatitis
1578	hepatitis a
1579	hepatitis b
1580	hepatitis c
1581	hepatitis d
1582	hepatitis e
1506	primary biliary cirrhosis
1604	alcoholic liver disease / alcoholic cirrhosis
1160	bile duct obstruction/ascending cholangitis
1475	sclerosing cholangitis
1165	pancreatitis
	Cardiac
1076	heart failure/pulmonary odema
	Renal
1192	renal/kidney failure
1193	renal failure requiring dialysis
1194	renal failure not requiring dialysis
1405	other renal/kidney problem
1196	urinary tract infection/kidney infection
1515	pyelonephritis
1427	polycystic kidney
1519	kidney nephropathy
1608	nephritis
1520	iga nephropathy
1607	diabetic nephropathy
1609	glomerulnephritis
	Gut
1154	irritable bowel syndrome
1456	malabsorption/coeliac disease
1457	duodenal ulcer
1459	colitis/not chrons or ulcerative colitis
1461	inflammatory bowel disease
1502	appendicitis
1503	anal problem
1599	constipation
1600	bowel / intestinal perforation
1601	bowel / intestinal infarction
1602	bowel / intestinal obstruction
1603	rectal prolapse
1462	crohns disease
1463	ulcerative colitis
	Abdominal
1400	peptic ulcer
	Endocrine
1224	thyroid problem (not cancer)
1229	parathyroid gland problem (not cancer)
1232	disorder of adrenal gland
1237	disorder of pituitary gland
1239	cushings syndrome
1432	carcinoid syndrome
1682	benign insulinoma
1221	gestational diabetes
1222	type 1 diabetes
1225	hyperthyroidism/thyrotoxicosis
1226	hypothyroidism/myxoedema
1228	thyroid radioablation therapy
1428	thyroiditis
1522	grave's disease
1610	thyroid goitre
1230	parathyroid hyperplasia/adenoma
1611	hyperparathyroidism
1233	adrenal tumour
1234	adrenocortical insufficiency/addison's disease
1235	hyperaldosteronism/conn's syndrome
1236	phaeochromocytoma
1238	pituitary adenoma/tumour
1429	acromegaly
1430	hypopituitarism
1431	hyperprolactinaemia
	COPD
1112	COPD
	Infections
1439	hiv/aids
1567	infectious mononucleosis / glandular fever / epstein barr virus (ebv)
1440	tuberculosis (tb)
1575	herpes simplex

Cancer (responded yes to "Have you ever been diagnosed with cancer?")

Appendix B

Supplementary Tables 1 and 2 are too large to print. They are located here:

Supplementary Table 1

https://docs.google.com/spreadsheets/d/1HYbX5qI81pvMjAM7bn8yWlN34OGtwudpDJLzLVUbu5A/edit?usp=sharing

Supplementary Table 2

https://docs.google.com/spreadsheets/d/19s_C6eb7uX4etbaTQ0M-XUvYYhOeTutiyIJwM1XwJ4A/edit?usp=sharing

Supplementary Table 4: Gene set analyses results

Gene set id	Trait	Meta-p	Meta-p (no APO)	WES p	N WES	WGS p	N WGS	Description	Source
C0020445	lhdlfc_	2.31E-10	0.02813214	1.01E-05	35	7.62E-06	21	Hypercholesterolemia Familial	DisGeneNet
C0020476	lhdlfc_	1.58E-11	0.000932652	2.39E-06	14	7.77E-07	7	Hyperlipoproteinemias	DisGeneNet
C0020476	hdld	1.81E-10	0.000279994	0.000496	14	1.80E-08	7	Hyperlipoproteinemias	DisGeneNet
C0020476	lhdlc_	2.90E-08	0.00385449	2.23E-05	14	0.000201	7	Hyperlipoproteinemias	DisGeneNet
C0020476	lhdlpl_	2.15E-06	0.002200132	0.000977	14	0.000793	7	Hyperlipoproteinemias	DisGeneNet
C0342881	idltg	2.02E-11	0.015485781	2.03E-09	11	0.002838	8	Familial hypercholesterolemia - homozygous	DisGeneNet
C0342881	xsvldlp	3.79E-10	0.014275635	4.03E-07	11	0.00085	8	Familial hypercholesterolemia - homozygous	DisGeneNet
C0342881	ldltg	7.64E-10	0.006844523	9.76E-09	11	0.004302	8	Familial hypercholesterolemia - homozygous	DisGeneNet
C0342881	xsvldltg	1.08E-09	0.023413237	1.84E-07	11	0.006007	8	Familial hypercholesterolemia - homozygous	DisGeneNet
C0342881	lldltg	3.58E-09	0.005062039	8.20E-08	11	0.003857	8	Familial hypercholesterolemia - homozygous	DisGeneNet
C0342881	apob	7.72E-09	0.005089742	2.38E-07	11	0.002934	8	Familial hypercholesterolemia - homozygous	DisGeneNet
C0342881	svldlfc	3.18E-08	0.012250296	2.71E-05	11	0.002389	8	Familial hypercholesterolemia - homozygous	DisGeneNet
C0342881	mldltg	7.07E-08	0.013478956	5.24E-08	11	0.029378	8	Familial hypercholesterolemia - homozygous	DisGeneNet
C0342881	sldltg	8.59E-08	0.016697804	5.88E-08	11	0.026173	8	Familial hypercholesterolemia - homozygous	DisGeneNet
C0342881	mufa	1.10E-07	0.018070242	0.00013	11	0.007047	8	Familial hypercholesterolemia - homozygous	DisGeneNet
C0342881	idll	1.75E-07	0.010999563	3.69E-06	11	0.003782	8	Familial hypercholesterolemia - homozygous	DisGeneNet
C0342881	apobapoa1	2.15E-07	0.004237918	1.04E-06	11	0.00795	8	Familial hypercholesterolemia - homozygous	DisGeneNet
C0342881	lldlp	2.43E-07	0.009922028	4.65E-07	11	0.012224	8	Familial hypercholesterolemia - homozygous	DisGeneNet
C0342881	svldlpl	2.48E-07	0.0089107	4.13E-05	11	0.002879	8	Familial hypercholesterolemia - homozygous	DisGeneNet
C0342881	lldll	2.84E-07	0.010485712	8.43E-07	11	0.013297	8	Familial hypercholesterolemia - homozygous	DisGeneNet
C0342881	xsvldlpl	3.71E-07	0.03201298	1.63E-07	11	0.004467	8	Familial hypercholesterolemia - homozygous	DisGeneNet
C0342881	idlp	3.89E-07	0.009724476	1.56E-06	11	0.002886	8	Familial hypercholesterolemia - homozygous	DisGeneNet
C0342881	idlpl	4.91E-07	0.012464279	3.13E-06	11	0.008312	8	Familial hypercholesterolemia - homozygous	DisGeneNet
C0342881	ldlc	6.95E-07	0.013848465	1.34E-06	11	0.026768	8	Familial hypercholesterolemia - homozygous	DisGeneNet
C0342881	lldlpl	7.04E-07	0.013332371	3.51E-06	11	0.018528	8	Familial hypercholesterolemia - homozygous	DisGeneNet
C0342881	lldlce	7.20E-07	0.01120345	2.29E-06	11	0.018631	8	Familial hypercholesterolemia - homozygous	DisGeneNet
C0342881	mldlpl	7.91E-07	0.012623335	1.12E-06	11	0.030335	8	Familial hypercholesterolemia - homozygous	DisGeneNet
C0342881	totfa	9.12E-07	0.020070097	2.66E-05	11	0.006704	8	Familial hypercholesterolemia - homozygous	DisGeneNet
C0342881	lldlc	9.44E-07	0.01233823	3.16E-06	11	0.018568	8	Familial hypercholesterolemia - homozygous	DisGeneNet
C0342881	mldlp	9.49E-07	0.012045521	2.99E-07	11	0.043184	8	Familial hypercholesterolemia - homozygous	DisGeneNet
C0342881	mldll	1.10E-06	0.011701482	4.54E-07	11	0.047026	8	Familial hypercholesterolemia - homozygous	DisGeneNet
C0342881	mldlfc	1.64E-06	0.03871723	2.76E-06	11	0.044593	8	Familial hypercholesterolemia - homozygous	DisGeneNet
C0342883	lhdlfc_	9.97E-14	0.001782186	6.12E-07	9	1.04E-09	4	Cholesteryl Ester Transfer Protein Deficiency	DisGeneNet
C0342883	tgpg	9.85E-10	0.016207152	5.21E-05	9	2.13E-06	4	Cholesteryl Ester Transfer Protein Deficiency	DisGeneNet
C0542037	lhdlfc_	3.57E-13	0.003632137	6.12E-07	9	1.74E-09	3	Hypotriglyceridaemia	DisGeneNet
C0542037	tgpg	3.23E-09	0.01845352	5.21E-05	9	2.21E-06	3	Hypotriglyceridaemia	DisGeneNet
C0745103	idltg	1.90E-10	0.008406138	1.83E-08	21	0.010046	17	Hyperlipoproteinemia Type IIa	DisGeneNet
C0745103	xsvldltg	2.03E-10	0.001916666	3.13E-07	21	0.008834	17	Hyperlipoproteinemia Type IIa	DisGeneNet
C0745103	svldlfc	1.22E-09	0.001385636	3.49E-05	21	0.001432	17	Hyperlipoproteinemia Type IIa	DisGeneNet
C0745103	xsvldlp	3.75E-09	0.014609129	3.02E-06	21	0.00142	17	Hyperlipoproteinemia Type IIa	DisGeneNet
C0745103	svldlpl	4.15E-09	0.000606073	3.76E-05	21	0.001927	17	Hyperlipoproteinemia Type IIa	DisGeneNet
C0745103	svldll	1.06E-08	0.001568385	7.76E-05	21	0.002428	17	Hyperlipoproteinemia Type IIa	DisGeneNet
C0745103	svldlp	1.49E-08	0.001319162	8.22E-05	21	0.003239	17	Hyperlipoproteinemia Type IIa	DisGeneNet
C0745103	mufa	1.26E-07	0.00369098	0.000211	21	0.003996	17	Hyperlipoproteinemia Type IIa	DisGeneNet
C0745103	ldltg	2.00E-07	0.020495788	2.63E-07	21	0.013265	17	Hyperlipoproteinemia Type IIa	DisGeneNet
C0745103	lldltg	4.58E-07	0.020209296	1.50E-06	21	0.01609	17	Hyperlipoproteinemia Type IIa	DisGeneNet
C0745103	sldltg	5.19E-07	0.01743014	2.02E-06	21	0.032317	17	Hyperlipoproteinemia Type IIa	DisGeneNet
C0745103	apob	1.19E-06	0.006937804	2.87E-06	21	0.001865	17	Hyperlipoproteinemia Type IIa	DisGeneNet
C0745103	apobapoa1	1.34E-06	0.004883344	1.88E-05	21	0.010105	17	Hyperlipoproteinemia Type IIa	DisGeneNet
C0745103	mvldlfc	1.68E-06	0.000836725	0.000477	21	0.019124	17	Hyperlipoproteinemia Type IIa	DisGeneNet
C0745103	totfa	1.71E-06	0.006830084	6.15E-05	21	0.006047	17	Hyperlipoproteinemia Type IIa	DisGeneNet
C1848486	xsvldlpl	5.53E-08	0.004265067	6.53E-07	11	0.005985	9	Premature arteriosclerosis	DisGeneNet
C1848486	sldltg	2.10E-07	0.026356402	9.08E-08	11	0.036179	9	Premature arteriosclerosis	DisGeneNet
C1848486	mldltg	8.28E-07	0.02980488	2.03E-07	11	0.044683	9	Premature arteriosclerosis	DisGeneNet
C4280503	xsvldlpl	5.53E-08	0.004265067	6.53E-07	11	0.005985	9	Premature hardening of arteries	DisGeneNet
C4280503	sldltg	2.10E-07	0.026356402	9.08E-08	11	0.036179	9	Premature hardening of arteries	DisGeneNet
C4280503	mldltg	8.28E-07	0.02980488	2.03E-07	11	0.044683	9	Premature hardening of arteries	DisGeneNet
R-HSA-204174	idlpl	7.85E-07	7.85E-07	0.005939	12	0.000503	4	Regulation of pyruvate dehydrogenase (PDH) complex	Reactome
R-HSA-204174	mldlpl	1.01E-06	1.01E-06	0.002671	12	0.000594	4	Regulation of pyruvate dehydrogenase (PDH) complex	Reactome
R-HSA-204174	estc	1.09E-06	1.09E-06	0.004754	12	0.001175	4	Regulation of pyruvate dehydrogenase (PDH) complex	Reactome
R-HSA-204174	idlp	1.17E-06	1.17E-06	0.003992	12	0.000593	4	Regulation of pyruvate dehydrogenase (PDH) complex	Reactome
R-HSA-204174	lldlp	1.20E-06	1.20E-06	0.004822	12	0.000258	4	Regulation of pyruvate dehydrogenase (PDH) complex	Reactome
R-HSA-204174	lldlpl	1.21E-06	1.21E-06	0.004853	12	0.000423	4	Regulation of pyruvate dehydrogenase (PDH) complex	Reactome
R-HSA-204174	idll	1.21E-06	1.21E-06	0.004313	12	0.000574	4	Regulation of pyruvate dehydrogenase (PDH) complex	Reactome
R-HSA-204174	serumc	1.24E-06	1.24E-06	0.005999	12	0.001071	4	Regulation of pyruvate dehydrogenase (PDH) complex	Reactome
R-HSA-204174	lldll	1.35E-06	1.35E-06	0.005082	12	0.000275	4	Regulation of pyruvate dehydrogenase (PDH) complex	Reactome
R-HSA-204174	idlc	1.40E-06	1.40E-06	0.00475	12	0.001019	4	Regulation of pyruvate dehydrogenase (PDH) complex	Reactome
R-HSA-204174	lldlfc	1.46E-06	1.46E-06	0.00681	12	0.0003	4	Regulation of pyruvate dehydrogenase (PDH) complex	Reactome
R-HSA-204174	lldlc	1.87E-06	1.87E-06	0.006489	12	0.000275	4	Regulation of pyruvate dehydrogenase (PDH) complex	Reactome
R-HSA-204174	mldlp	1.96E-06	1.96E-06	0.006409	12	0.000132	4	Regulation of pyruvate dehydrogenase (PDH) complex	Reactome
R-HSA-204174	lldlce	2.01E-06	2.01E-06	0.006486	12	0.000277	4	Regulation of pyruvate dehydrogenase (PDH) complex	Reactome
R-HSA-204174	sldll	2.13E-06	2.13E-06	0.006413	12	0.000115	4	Regulation of pyruvate dehydrogenase (PDH) complex	Reactome
R-HSA-204174	sldlp	2.13E-06	2.13E-06	0.005994	12	0.000113	4	Regulation of pyruvate dehydrogenase (PDH) complex	Reactome
R-HSA-204174	mldll	2.13E-06	2.13E-06	0.006416	12	0.000164	4	Regulation of pyruvate dehydrogenase (PDH) complex	Reactome
R-HSA-204174	ldlc	2.17E-06	2.17E-06	0.007809	12	0.000177	4	Regulation of pyruvate dehydrogenase (PDH) complex	Reactome
R-HSA-204174	apob	2.20E-06	2.20E-06	0.00504	12	0.000803	4	Regulation of pyruvate dehydrogenase (PDH) complex	Reactome
R-HSA-204174	idlfc	2.22E-06	2.22E-06	0.009798	12	0.000399	4	Regulation of pyruvate dehydrogenase (PDH) complex	Reactome
R-HSA-8866423	xsvldlp	1.49E-12	0.027026999	2.06E-09	8	0.000246	7	VLDL assembly	Reactome
R-HSA-8866423	xsvldll	6.57E-12	0.029658511	3.13E-09	8	0.000254	7	VLDL assembly	Reactome
R-HSA-8866423	xsvldlpl	3.27E-10	0.047296943	4.87E-10	8	0.000529	7	VLDL assembly	Reactome
R-HSA-8866423	idlp	5.94E-10	0.012821521	2.79E-09	8	0.000385	7	VLDL assembly	Reactome
R-HSA-8866423	apob	9.00E-10	0.035805827	1.23E-09	8	0.001105	7	VLDL assembly	Reactome
R-HSA-8866423	lldlpl_	1.21E-09	0.003361758	2.31E-11	8	0.006697	7	VLDL assembly	Reactome
R-HSA-8866423	idll	2.82E-09	0.014169646	1.95E-08	8	0.000547	7	VLDL assembly	Reactome
R-HSA-8866423	ldlc	2.02E-08	0.015814492	6.36E-09	8	0.003754	7	VLDL assembly	Reactome

R-HSA-8866423	lldlp	2.09E-08	0.010925413	1.78E-09	8	0.001674	7	VLDL assembly	Reactome
R-HSA-8866423	remnantc	6.95E-08	0.005468158	4.44E-09	8	0.00083	7	VLDL assembly	Reactome
R-HSA-8866423	lldlfc	1.75E-07	0.012845409	7.16E-08	8	0.002439	7	VLDL assembly	Reactome
R-HSA-8866423	idlpl	1.84E-07	0.011330913	2.70E-08	8	0.000806	7	VLDL assembly	Reactome
R-HSA-8866423	xsvldlfc	1.96E-07	0.037068974	1.64E-07	8	0.002008	7	VLDL assembly	Reactome
R-HSA-8866423	lldlpl	2.05E-07	0.009682997	1.12E-08	8	0.00295	7	VLDL assembly	Reactome
R-HSA-8866423	lldlce	2.22E-07	0.012777309	8.81E-09	8	0.002613	7	VLDL assembly	Reactome
R-HSA-8866423	lldlc	2.28E-07	0.012348304	1.40E-08	8	0.002473	7	VLDL assembly	Reactome
R-HSA-8866423	lldll	2.53E-07	0.010991255	3.93E-09	8	0.00178	7	VLDL assembly	Reactome
R-HSA-8866423	idlfc	2.64E-07	0.020705061	2.07E-07	8	0.001976	7	VLDL assembly	Reactome
R-HSA-8866423	xsvldlc	2.81E-07	0.010931813	7.53E-06	8	0.000766	7	VLDL assembly	Reactome
R-HSA-8866423	idlc	3.52E-07	0.018406604	6.83E-07	8	0.001729	7	VLDL assembly	Reactome
R-HSA-8866423	serumc	4.74E-07	0.023383675	6.02E-07	8	0.008607	7	VLDL assembly	Reactome
R-HSA-8866423	idlce	5.22E-07	0.00201215	1.58E-06	8	0.001804	7	VLDL assembly	Reactome
R-HSA-8866423	mldlp	5.32E-07	0.019315992	6.25E-09	8	0.008059	7	VLDL assembly	Reactome
R-HSA-8866423	mldll	5.50E-07	0.017018952	9.19E-09	8	0.008598	7	VLDL assembly	Reactome
R-HSA-8866423	estc	5.80E-07	0.024024992	4.43E-07	8	0.012954	7	VLDL assembly	Reactome
R-HSA-8866423	freec	6.04E-07	0.027416347	6.56E-06	8	0.004008	7	VLDL assembly	Reactome
R-HSA-8866423	idlpl_	6.58E-07	0.039687097	7.17E-07	8	0.01239	7	VLDL assembly	Reactome
R-HSA-8866423	mldlpl	7.13E-07	0.015426761	1.99E-08	8	0.010748	7	VLDL assembly	Reactome
R-HSA-8866423	xsvldlce	7.44E-07	0.009844208	4.89E-05	8	0.000835	7	VLDL assembly	Reactome
R-HSA-8866423	sldlc	7.51E-07	0.024307244	4.84E-09	8	0.017042	7	VLDL assembly	Reactome
R-HSA-8866423	sldlp	7.67E-07	0.027289638	2.54E-09	8	0.015185	7	VLDL assembly	Reactome
R-HSA-8866423	pufa	7.71E-07	0.08454695	1.50E-06	8	0.008925	7	VLDL assembly	Reactome
R-HSA-8866423	vldlc	8.93E-07	0.052364901	1.41E-05	8	0.002975	7	VLDL assembly	Reactome
R-HSA-8866423	sldlce	9.02E-07	0.007812149	2.95E-09	8	0.01715	7	VLDL assembly	Reactome
R-HSA-8866423	sldll	9.22E-07	0.027649486	1.60E-09	8	0.015745	7	VLDL assembly	Reactome
R-HSA-8866423	mldlc	1.18E-06	0.019562719	3.67E-08	8	0.012777	7	VLDL assembly	Reactome
R-HSA-8866423	mldlce	1.27E-06	0.021261762	4.76E-08	8	0.012747	7	VLDL assembly	Reactome
R-HSA-8866423	svldlce	1.44E-06	0.024795814	8.30E-05	8	0.001773	7	VLDL assembly	Reactome
R-HSA-8866423	mldlfc	2.02E-06	0.005231542	3.73E-08	8	0.016205	7	VLDL assembly	Reactome
R-HSA-8866423	sldlfc	2.14E-06	0.015190169	6.26E-09	8	0.027399	7	VLDL assembly	Reactome
R-HSA-8963888	xsvldlp	2.49E-14	0.206996778	2.15E-10	10	2.02E-05	11	Chylomicron assembly	Reactome
R-HSA-8963888	svldlc	3.38E-14	0.378917505	1.71E-09	10	2.65E-05	11	Chylomicron assembly	Reactome
R-HSA-8963888	xsvldll	2.89E-13	0.204758667	3.87E-10	10	5.27E-05	11	Chylomicron assembly	Reactome
R-HSA-8963888	apobapoa1	8.43E-13	0.167387417	1.72E-09	10	3.83E-06	11	Chylomicron assembly	Reactome
R-HSA-8963888	vldlc	2.12E-11	0.280931433	3.23E-09	10	4.19E-05	11	Chylomicron assembly	Reactome
R-HSA-8963888	lldlfc_	7.49E-11	0.195504351	2.15E-10	10	0.000573	11	Chylomicron assembly	Reactome
R-HSA-8963888	svldlce	2.15E-10	0.147784098	1.40E-08	10	0.000173	11	Chylomicron assembly	Reactome
R-HSA-8963888	mvldlce	2.15E-10	0.281624878	1.40E-07	10	0.000117	11	Chylomicron assembly	Reactome
R-HSA-8963888	remnantc	5.86E-10	0.085210798	2.92E-08	10	0.000799	11	Chylomicron assembly	Reactome
R-HSA-8963888	ldltg	9.59E-10	0.396710914	6.92E-08	10	0.000441	11	Chylomicron assembly	Reactome
R-HSA-8963888	lldltg	3.23E-09	0.294471306	3.09E-07	10	0.000602	11	Chylomicron assembly	Reactome
R-HSA-8963888	xsvldlfc	3.29E-09	0.081395683	2.97E-08	10	0.002742	11	Chylomicron assembly	Reactome
R-HSA-8963888	mufa	2.93E-08	0.388099899	7.78E-06	10	0.002762	11	Chylomicron assembly	Reactome
R-HSA-8963888	xsvldlce	2.94E-08	0.069144815	2.02E-06	10	0.001391	11	Chylomicron assembly	Reactome
R-HSA-8963888	xsvldlc	3.01E-08	0.06186095	4.02E-07	10	0.001125	11	Chylomicron assembly	Reactome
R-HSA-8963888	idlpl_	3.01E-08	0.098183672	7.18E-08	10	0.005842	11	Chylomicron assembly	Reactome
R-HSA-8963888	apob	6.39E-08	0.24090487	6.48E-09	10	0.001438	11	Chylomicron assembly	Reactome
R-HSA-8963888	xsvldlpl	6.41E-08	0.303509995	6.64E-09	10	0.000858	11	Chylomicron assembly	Reactome
R-HSA-8963888	lhdltg_	2.38E-07	0.248013219	6.88E-05	10	0.000773	11	Chylomicron assembly	Reactome
R-HSA-8963888	lldlpl_	4.25E-07	0.046875529	2.37E-07	10	0.02747	11	Chylomicron assembly	Reactome
R-HSA-8963888	xlvldlpl	6.11E-07	0.195788822	0.000226	10	0.001796	11	Chylomicron assembly	Reactome
R-HSA-8963888	xlvldll	6.12E-07	0.215946625	0.000218	10	0.001764	11	Chylomicron assembly	Reactome
R-HSA-8963888	xlvldlp	6.27E-07	0.336990987	0.000219	10	0.001784	11	Chylomicron assembly	Reactome
R-HSA-8963888	xsvldltg_	6.28E-07	0.089073028	0.00775	10	1.70E-05	11	Chylomicron assembly	Reactome
R-HSA-8963888	xlvldlce	1.14E-06	0.309746179	0.000159	10	0.001166	11	Chylomicron assembly	Reactome
R-HSA-8963888	xxlvldltg	1.16E-06	0.176848043	0.000474	10	0.002696	11	Chylomicron assembly	Reactome
R-HSA-8963888	xlvldltg	1.18E-06	0.321278515	0.000222	10	0.002044	11	Chylomicron assembly	Reactome
R-HSA-8963888	xlvldlc	1.20E-06	0.309130119	0.000172	10	0.001447	11	Chylomicron assembly	Reactome
R-HSA-8963888	xxlvldlce	1.21E-06	0.210451215	0.000326	10	0.001414	11	Chylomicron assembly	Reactome
R-HSA-8963888	xxlvldll	1.29E-06	0.192140471	0.000347	10	0.002374	11	Chylomicron assembly	Reactome
R-HSA-8963888	lvldlpl_	1.43E-06	0.551285259	8.06E-06	9	0.014211	11	Chylomicron assembly	Reactome
R-HSA-8963888	xxlvldlp	2.07E-06	0.186176084	0.000354	10	0.00242	11	Chylomicron assembly	Reactome
R-HSA-8963888	totfa	2.11E-06	0.511678701	8.44E-07	10	0.045133	11	Chylomicron assembly	Reactome
R-HSA-8963888	xxlvldlc	2.14E-06	0.219048098	0.000325	10	0.001807	11	Chylomicron assembly	Reactome
R-HSA-8963888	xlvldlfc	2.19E-06	0.314011612	0.000228	10	0.002124	11	Chylomicron assembly	Reactome
R-HSA-8963898	xsvldltg	9.97E-10	0.237293907	2.89E-08	23	0.003889	19	Plasma lipoprotein assembly	Reactome
R-HSA-8963898	svldlpl	6.28E-07	0.819781918	1.17E-06	23	0.002978	19	Plasma lipoprotein assembly	Reactome
R-HSA-8963898	svldlfc	6.79E-07	1	1.11E-06	23	0.004052	19	Plasma lipoprotein assembly	Reactome
R-HSA-8963898	svldlp	1.19E-06	1	1.35E-06	23	0.004011	19	Plasma lipoprotein assembly	Reactome
R-HSA-8963898	svldll	1.25E-06	1	1.13E-06	23	0.004472	19	Plasma lipoprotein assembly	Reactome
R-HSA-8963901	hdld	9.72E-10	0.001414545	0.000108	12	8.68E-06	12	Chylomicron remodeling	Reactome
R-HSA-8963901	xlhdlfc	3.04E-09	0.004188796	0.000336	12	6.07E-05	12	Chylomicron remodeling	Reactome
R-HSA-8963901	lhdlc_	1.01E-08	0.003841981	4.60E-05	12	0.000594	12	Chylomicron remodeling	Reactome
R-HSA-8963901	xlhdlpl	1.13E-08	0.007480561	0.000162	12	4.45E-05	12	Chylomicron remodeling	Reactome
R-HSA-8963901	xlhdlc	1.76E-07	0.011331821	0.002411	12	0.000666	12	Chylomicron remodeling	Reactome
R-HSA-8964058	tgpg	5.88E-10	0.006630914	1.81E-05	17	2.46E-06	8	HDL remodeling	Reactome

Meta-p= Meta-analysis p-value
Meta-p (no APO) = Meta-analysis p-value after removing APO genes from gene sets (APOB and APOC3)
WES p = p-value in WES dataset
N WES = number of variants tested in WES dataset
WGS p = p-value in WGS dataset
N WGS = number of variants tested in WGS dataset

Supplementary Table 9: Detailed results for gene sets with enriched rare variation in tails of lipoprotein traits

S-VLDL-C lower tail outliers. Hyperlipidemia gene set.

gene	snp	dataset	MAC	residuals in all carriers
AGL	rs200459772	WES	5	2.36762834154852,-0.334045067074641,0.431527558983269,-0.838811852821138,-3.05000388882286
APOB	2:21236148	WES	1	-2.661258903
APC	rs150973053	WES	1	-3.089584993
APC	rs201830995	WES	3	-2.87066740721444,-0.787318922230483,0.420463200843388
CYP19A1	rs141305220	WES	2	-3.494055746710 22,-1.2437172570647
CYP19A1	rs200111039	WES	9	-2.975904533000663,0.253051068847167,0.795701074251656,1.01065228811834,-0.403431340606028,-0.144560598282279,-3.08958499345356,0.741693060794646,-0.4324749949308
NPHS1	rs368988883	WES	1	-3.374778926
GCG	2:163003928	WGS	1	-3.123944186
APC	5:112173509	WGS	2	-3.543944498814 21,-0.121786221385006
APC	5:112174919	WGS	2	-3.543944498814 21,-0.121786221385006
APC	5:112178070	WGS	2	-3.543944498814 21,-0.121786221385006
APC	5:112179437	WGS	2	-3.543944498814 21,-0.121786221385006
NOS3	7:150698995	WGS	2	-0.18836307152566,-2.7037021818663
NOS3	7:150706632	WGS	5	-0.304682642445497,-0.26183599116571,-0.48084454181164,-2.83884053615798,-0.759193154129386
CETP	rs150236668	WGS	2	-1.08925353711354,-3.543944498814 21
NPHS1	19:36342715	WGS	3	0.920578019402659,-0.398163133632229,-2.93246487402699

XS-VLDL-P lower tail outliers. Hyperlipidemia gene set.

gene	snp	dataset	MAC	effects
APOB	2:21236148	WES	1	-3.436640493
APC	rs150973053	WES	1	-3.202863174
APC	rs201830995	WES	3	-2.86524374013287,-0.168052075323293,0.077312983454771
NOS3	rs141170595	WES	7	-1.18611115166881,-2.955898255405 99,-0.246085238062439,1.13215214546922,0.154253491587311,-0.217108986788457,-1.74689283105004
CYP19A1	rs141305220	WES	2	-3.205389847208 28,-0.876451886676179
CYP19A1	rs200111039	WES	9	-2.640986452125 51,0.545406135777777,1.01054251354388,0.763927578916 17,-0.245365417517183,-0.682701582154253,-3.20286317422441,0.569449665319146,-0.251861115970539
NPHS1	rs368988883	WES	1	-3.318749511
NPHS2	1:179520511	WGS	2	0.213855282424323,-2.73710673267041
NPHS2	1:179530462	WGS	6	-2.85736273031488,0.500033274189366,0.129175297645043,0.476908535573381,-0.94191940828643,0.175183524144263
APOB	2:21255263	WGS	1	-2.965806851
GCG	2:163003928	WGS	1	-3.430062283
APC	5:112173509	WGS	2	-2.949070995374 61,-0.259525678305062
APC	5:112174919	WGS	2	-2.949070995374 61,-0.259525678305062
APC	5:112178070	WGS	2	-2.949070995374 61,-0.259525678305062
APC	5:112179437	WGS	2	-2.949070995374 61,-0.259525678305062
NOS3	7:150698995	WGS	2	-0.466020371752611,-2.83066719904639
CETP	rs150236668	WGS	2	-0.633213631434203,-2.94907099537461

S-VLDL-CE lower tail outliers. Hyperlipidemia gene set.

gene	snp	dataset	MAC	effects
AGL	rs200459772	WES	5	2.36762834154852,-0.334045067074641,0.431527558983269,-0.838811852821138,-3.05000388882286
APOB	2:21236148	WES	1	-2.661258903
APC	rs150973053	WES	1	-3.089584993
APC	rs201830995	WES	3	-2.87066740721444,-0.787318922230483,0.420463200843388
CYP19A1	rs141305220	WES	2	-3.494055746710 22,-1.2437172570647
CYP19A1	rs200111039	WGS	9	-2.975904533000663,0.253051068847167,0.795701074251656,1.01065228811834,-0.403431340606028,-0.144560598282279,-3.08958499345356,0.741693060794646,-0.4324749949308
NPHS1	rs368988883	WES	1	-3.374778926
GCG	2:163003928	WGS	1	-3.123944186
APC	5:112173509	WGS	2	-3.543944498814 21,-0.121786221385006
APC	5:112174919	WGS	2	-3.543944498814 21,-0.121786221385006
APC	5:112178070	WGS	2	-3.543944498814 21,-0.121786221385006
APC	5:112179437	WGS	2	-3.543944498814 21,-0.121786221385006
NOS3	7:150698995	WGS	2	-0.18836307152566,-2.7037021818663
NOS3	7:150706632	WGS	5	-0.304682642445497,-0.26183599116571,-0.48084454181164,-2.83884053615798,-0.759193154129386
CETP	rs150236668	WGS	2	-1.08925353711354,-3.543944498814 21
NPHS1	19:36342715	WGS	3	0.920578019402659,-0.398163133632229,-2.93246487402699

S-HDL-P lower tail outliers. HDL remodeling gene set

gene	snp	dataset	MAC	effects
CETP	rs140547417	WES	10	0.578651608406939,0.610798008574449,0.292679415486239,0.395395459347943,-1.113868535475629,-2.932639377408 99,-0.0998285578608295,-0.0864903418646204,-0.318903381965163,0.775064146714445
LIPG	18:47107925	WES	1	-3.234598237
APOE	rs199768005	WES	7	-0.540244700238687,1.92520088605348,0.92978260411206,-3.02709326825206,-0.789305787720864,0.121976706457689,-1.34111543948004
ABCG1	rs148226451	WES	1	-2.932639377
APOA1	11:11670686:	WGS	1	-3.003505735
APOA1	rs199759119	WGS	7	-1.005363495326 65,-2.86126280384725,-0.582922966059555,-0.693164051709012,1.64763326906752,-2.39922951075938,-1.79280325835833
CETP	rs142750310	WGS	1	-3.046409293

Supplementary Table 10: Sensitity analyses for rare variant enrichment in tails analysis using different percentile cutoffs to define tails of the phenotypic distribution

.5% Percentile upper tails

trait	p.wes	p.wgs	meta-p	Gene set
lldlc	0.00432	0.03209	0.0007737	LDL_clearance
vldld	0.02887	0.00607	0.0009188	VLDL_clearance

.5% Percentile lower tails

trait	p.wes	p.wgs	meta-p	Gene set
svldlce	0.02992	0.01634	0.0022477	Hyperlipidemia
svldlfc	0.01287	0.00676	0.0004448	Hyperlipidemia
xsvldlp	0.02992	0.0024	0.0004422	Hyperlipidemia
idltg	0.00032	0.01528	4.02E-05	LDL_remodeling
idltg	1.00E-05	0.01621	2.97E-06	VLDL_assembly

1 Percentile lower tails

trait	p.wes	p.wgs	meta-p	Gene set
mhdltg	0.04487	0.00976	0.0021777	Hyperlipidemia

p.wes: permutation p-value in WES
p.wgs: permutation p-value in WGS
meta-p: p-value after meta-analysis using Stouffer's method
Highlighted in yellow are gene sets that are significant after meta-analysis using Stouffer's method and after adjusting for multiple traits (p<=0.00037).