

1 **Meta-analysis of genome-wide association studies for body fat distribution in 694,649 individuals of European**  
2 **ancestry**

3  
4 Sara L. Pulit<sup>1,2,3\*</sup>, Charli Stoneman<sup>4\*</sup>, Andrew P. Morris<sup>5,6</sup>, Andrew R. Wood<sup>4</sup>, Craig A. Glastonbury<sup>1</sup>, Jessica  
5 Tyrrell<sup>4</sup>, Loïc Yengo<sup>7</sup>, Teresa Ferreira<sup>1</sup>, Eirini Marouli<sup>8</sup>, Yingjie Ji<sup>4</sup>, Jian Yang<sup>7,9</sup>, Samuel Jones<sup>4</sup>, Robin Beaumont<sup>4</sup>,  
6 Damien C. Croteau-Chonka<sup>10</sup>, Thomas W. Winkler<sup>11</sup>, GIANT Consortium, Andrew T. Hattersley<sup>4</sup>, Ruth J. F.  
7 Loos<sup>12</sup>, Joel N. Hirschhorn<sup>13,14,15,16</sup>, Peter M. Visscher<sup>7,9</sup>, Timothy M. Frayling<sup>4#</sup>, Hanieh Yaghootkar<sup>4#</sup>, Cecilia M.  
8 Lindgren<sup>1,3,6#</sup>

- 9  
10  
11 1. Big Data Institute, Li Ka Shing Center for Health Information and Discovery, Oxford University, Oxford,  
12 UK  
13 2. Department of Genetics, Center for Molecular Medicine, University Medical Center Utrecht, Utrecht, The  
14 Netherlands  
15 3. Program in Medical and Population Genetics, Broad Institute, Boston, MA, USA  
16 4. University of Exeter Medical School, University of Exeter, Royal Devon and Exeter NHS Trust, Exeter,  
17 UK  
18 5. Biostatistics Department, University of Liverpool, Liverpool, UK  
19 6. Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK  
20 7. Institute for Molecular Bioscience, University of Queensland, Brisbane, Australia  
21 8. William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary  
22 University of London, London, UK  
23 9. Queensland Brain Institute, The University of Queensland, Brisbane, Queensland, Australia  
24 10. Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and  
25 Harvard Medical School, Boston, MA, USA  
26 11. Department of Genetic Epidemiology, University of Regensburg, Regensburg, Germany  
27 12. The Charles Bronfman Institute for Personalized Medicine, The Mindich Child Health and Development  
28 Institute, the Icahn School of Medicine at Mount Sinai, New York, USA  
29 13. Broad Institute of MIT and Harvard, Cambridge, MA, USA  
30 14. Department of Genetics, Harvard Medical School, Boston, MA, USA  
31 15. Division of Endocrinology and Center for Basic and Translational Obesity Research, Boston Children's  
32 Hospital, Boston, MA, USA  
33 16. Department of Pediatrics, Harvard Medical School, Boston, MA, USA

34  
35  
36

1 **Corresponding author**

2

3 Cecilia M Lindgren

4 Address: The Big Data Institute

5 Li Ka Shing Centre for Health Information and Discovery

6 Old Rd, Oxford OX3 7FZ, UK

7 Phone: +44 01865 287591

8 E-mail: [cecilia.lindgren@bdi.ox.ac.uk](mailto:cecilia.lindgren@bdi.ox.ac.uk)

9

10

## 1 Abstract

2 **One in four adults worldwide are either overweight or obese. Epidemiological studies indicate that the**  
3 **location and distribution of excess fat, rather than general adiposity, is most informative for predicting risk of**  
4 **obesity sequelae, including cardiometabolic disease and cancer. We performed a genome-wide association**  
5 **study meta-analysis of body fat distribution, measured by waist-to-hip ratio adjusted for BMI**  
6 **(WHRadjBMI), and identified 463 signals in 346 loci. Heritability and variant effects were generally stronger**  
7 **in women than men, and we found approximately one-third of all signals to be sexually dimorphic. The 5% of**  
8 **individuals carrying the most WHRadjBMI-increasing alleles were 1.62 times more likely than the bottom**  
9 **5% to have a WHR above the thresholds used for metabolic syndrome. These data, made publicly available,**  
10 **will inform the biology of body fat distribution and its relationship with disease.**

11

## 12 Introduction

13 One in four adults worldwide are either overweight or obese (1,2) and are at increased risk of metabolic disease.  
14 While higher adiposity increases morbidity and mortality (1,3), epidemiological studies indicate that the location  
15 and distribution of excess fat within particular depots is more informative than general adiposity for predicting  
16 disease risk. Independent of their overall body mass index (BMI), individuals with higher central adiposity have  
17 increased risk of cardiometabolic diseases, including type 2 diabetes (T2D) and stroke (4,5); in contrast, individuals  
18 with higher gluteal adiposity have lower risk of such outcomes.(5) Previous studies indicate that fat distribution, as  
19 assessed by waist-to-hip ratio (WHR), is a trait with a strong heritable component, independent of overall adiposity  
20 (measured by BMI), with twin-based heritability estimates ranging between 30-60% (5,6) and narrow-sense  
21 heritability estimates have been estimated at ~50% in women and ~20% in men (5). The most recent genome-wide  
22 association study in 224,459 samples implicated 49 loci associated with WHR adjusted for BMI (5), and recent  
23 Mendelian randomisation studies using known WHR-associated genetic variants showed putative causal effects of  
24 higher WHR on T2D and coronary artery disease independently of BMI (7).

25

## 26 Results

27 With the goal of pinpointing genetic variants associated to body shape and fat distribution and motivated by the  
28 recent release of genetic data from half a million individuals (8), we performed a meta-analysis of WHR adjusted for  
29 BMI (WHRadjBMI). WHRadjBMI is an easily-measured fat distribution phenotype that correlates well with  
30 imaging-based fat distribution measures (9). We performed genome-wide association studies (GWAS) of  
31 WHRadjBMI in the UK Biobank data set (8), a collection of 484,563 samples with densely-imputed genotype data,  
32 using a linear mixed model (10) to account for relatedness and ancestral heterogeneity. We then combined the  
33 results with publicly-available GWAS data generated by the GIANT consortium for the same phenotype (**Table 1**  
34 **and Methods**) (5), resulting in a meta-analysis of 694,649 samples (**Table 1**) and ~27.4M SNPs (**Methods**). As a  
35 sensitivity analysis and to evaluate the robustness of our results, we also performed a GWAS of WHR unadjusted  
36 for BMI (**Table 1**).

37

1 We identified 346 loci (300 novel) containing 463 independent signals associated with WHRadjBMI ( $p < 5 \times 10^{-9}$ , to  
2 account for the denser imputation data (11); **Methods, Supplementary Table 1** and **Supplementary Fig 1**). The  
3 Linkage Disequilibrium (LD) Score Regression (12) intercept (1.035) of the meta-analysis results indicated that the  
4 observed enrichment in genomic signal was due to polygenicity and not confounding (**Supplementary Table 2**). Of  
5 the 300 novel signals, 234 (78%,  $p_{\text{binomial}} < 1 \times 10^{-7}$ ) were directionally-consistent in an independent dataset with a  
6 relatively small sample size ( $N = 7,721$ ) and signals were consistent in several sensitivity checks (**Supplementary**  
7 **Tables 3-5**, and **Supplementary Fig 2-3**). Combined, these variants explained  $\sim 3.9\%$  of the variance in  
8 WHRadjBMI in the independent study (**Methods** and **Table 1**). We constructed a weighted polygenic risk score  
9 using the 346 index SNPs discovered in the combined meta-analysis and tested this score in the same independent  
10 study. The 5% of individuals carrying the most WHRadjBMI-raising alleles were 1.62 times more likely to meet the  
11 WHR threshold used to define metabolic syndrome (13) than the 5% carrying the fewest (consistent with the results  
12 obtained from unweighted polygenic score; **Methods**). The WHRadjBMI of people in the top 5% of the PRS was  
13 1.05 and 1.06 times greater in men and women, respectively, compared to those in the bottom 5% of the PRS.

14  
15 To investigate the potential for collider bias resulting from conditioning WHR on BMI, we investigated the behavior  
16 of WHRadjBMI-associated SNPs in GWAS of WHR (without adjustment for BMI) and BMI alone. We found that  
17 the majority of WHRadjBMI signals identified have genuine effect on body shape, and that any bias caused by  
18 adjusting WHR for a correlated covariate (14, 15) (that is, BMI) was minimal. Of the 346 index variants, 311  
19 associated with stronger standard deviation effect sizes for WHR (unadjusted) than with standard deviation effect  
20 sizes for BMI (**Supplementary Table 3** and **Supplementary Fig 4**). This observation also indicates that the WHR  
21 association is unlikely to be secondary to the known effect of higher BMI resulting in higher WHR. Furthermore,  
22 the common SNP associated with the largest known effect on BMI, that in the *FTO* gene (16), was not associated  
23 with WHRadjBMI (rs1421085,  $p = 0.40$ ) despite a very strong association with WHR ( $p = 4 \times 10^{-118}$ ). Finally,  
24 carrying each additional (weighted) WHRadjBMI-raising allele was associated with an increase in WHRadjBMI of  
25 0.0199 SD ( $p = 6 \times 10^{-62}$ ; adjusted  $R^2 = 4\%$ ), an increase in WHR of 0.0111 SD ( $p = 3 \times 10^{-20}$ ; adjusted  $R^2 = 0.12\%$ )  
26 and a decrease in BMI of 0.0038 SD ( $p = 1.4 \times 10^{-3}$ ; adjusted  $R^2 = 0.13\%$ ) in our independent dataset, consistent  
27 with the results obtained from an unweighted polygenic score (**Methods**).

28  
29 Given the sex-dimorphism of fat distribution in humans, previously shown to have a genetic basis (5, 17), we next  
30 performed meta-analyses of WHRadjBMI in women and men separately (**Table 1** and **Supplementary Fig 5**). We  
31 found SNP-based heritability ( $h_g^2$ ) of WHRadjBMI, estimated using the restricted maximum likelihood method  
32 implemented in BOLT-REML (10) (**Methods**), to be stronger in women ( $h_g^2 = 25.6\%$ ) compared to men ( $h_g^2 = 16.7\%$ ,  
33  $p_{\text{difference}} = 9 \times 10^{-85}$ ; **Table 1, Supplementary Table 6**, and **Equation 2**). In addition to the heritability dimorphism,  
34 and in keeping with previous studies (5), we found signatures of sex-dimorphism amongst associated loci: a total of  
35 266 loci associated with WHRadjBMI in women, compared to 91 loci in men ( $p < 5 \times 10^{-9}$ ). Genome-wide, SNP  
36 effects on WHRadjBMI were strongly correlated between men and women (LD Score  $r_g = 0.514$  (s.e. = 0.019),  $p =$   
37  $3.43 \times 10^{-159}$ ), but the consistency between the effect size of 266 female index SNPs on WHRadjBMI in women and

1 men (adjusted  $R^2 = 51\%$ ) was greater than the consistency between the effect size of 91 male index SNPs on  
2 WHRadjBMI in men and women (adjusted  $R^2 = 9\%$ ). Of all associated index SNPs ( $p < 5 \times 10^{-9}$  in the combined or  
3 sex-specific analyses), 105 SNPs were sex-dimorphic ( $p_{\text{diff}} < 3.3 \times 10^{-5}$ ; (17) and **Methods**). Variants discovered in  
4 the combined sex analysis will be enriched for those with similar effects in each sex, while variants discovered in  
5 sex-specific analyses will be enriched for those with differing effects between sexes. In the absence of any sex-  
6 specific effects, we would only expect a slight shift towards stronger associations in women due to the larger  
7 available sample size in that analysis. However, we observed that of the 105 sex-dimorphic signals, 97 (92.4%)  
8 showed stronger effects in women compared to men (**Figure 1, Supplementary Fig 6, and Methods**). Scanning  
9 genome-wide for sex-dimorphic SNPs ( $p_{\text{diff}} < 5 \times 10^{-9}$ ), regardless of their association p-values in the sex-specific  
10 analyses, we identified 61 sex-dimorphic SNPs after LD-based clumping ( $r^2 < 0.05$ ). Of these, 19 (31.1%)  
11 overlapped with the sex-dimorphic and genome-wide significant loci, and 54 (88.5%) had stronger effect in women  
12 than in men (**Supplementary Information**).

13  
14 Previous studies have shown that in addition to redistributing body fat, some WHRadjBMI variants are also  
15 associated with total body fat percentage (BF%) (5,18–20). Of relevance to the biology of adipose tissue storage  
16 capacity, these studies have shown that these pleiotropic associations can occur in both directions: some alleles  
17 associated with higher WHRadjBMI are associated with higher total BF%, whilst others are associated with lower  
18 BF% (5,18–20). To test the hypothesis that alleles associated with higher WHRadjBMI could have pleiotropic  
19 effects on total BF%, and that these effects could occur in both directions, we next investigated whether 346 index  
20 variants associated with WHRadjBMI also associated with BF%. Of the 59/346 variants associated with BF% in  
21 443,001 European-ancestry UK Biobank individuals ( $p < 0.05/346 = 1.44 \times 10^{-4}$ ), 25 SNPs associated with higher  
22 WHR and higher BF%, whilst 34 SNPs associated with higher WHR but lower BF% (**Figure 2**). These findings  
23 indicate that WHR-increasing alleles do not strictly influence BF% in one direction but rather can associate with  
24 either higher or lower BF%, yielding biological insight beyond the known epidemiological correlation between BF%  
25 and WHR. Additionally, a large proportion (29%) of WHRadjBMI index SNPs with a stronger effect in women had  
26 a BF% phenotype in men: 28 of the 97 female-specific WHRadjBMI SNPs were associated with BF% in men and  
27 25 were associated with BF% in women ( $p < 0.05/105 = 4.8 \times 10^{-4}$ , **Supplementary Fig 7**). These variants appear to  
28 alter total BF% in men and women to a similar extent but distribute body fat between the upper and lower body to a  
29 much greater extent in women (**Supplementary Table 7-9 and Supplementary Fig 7**). Finally, we tested the index  
30 SNPs from each of the meta-analyses (combined and sex-specific) in a recent GWAS of CT and MRI image-based  
31 measures of ectopic and subcutaneous fat depots (21). Adjusting for the three sample groups and the 8 depots  
32 examined in the imaging-based GWAS ( $p < 0.05/24 = 2.1 \times 10^{-3}$ ), the alleles associated with higher WHRadjBMI  
33 were collectively associated with lower measures of subcutaneous fat, and higher measures of visceral fat, including  
34 pericardial and visceral adipose tissue (**Supplementary Fig 8**).

35

36 **Discussion**

1 In a meta-analysis of nearly 700,000 individuals, we have increased the number of loci associated to WHRadjBMI  
2 by more than seven-fold. Of all the detected signals, 105 are sex-dimorphic, consistent with previous findings (5).  
3 While we have performed the largest meta-analysis of a measure of body-fat distribution to date, a number of  
4 limitations remain. First, the substantially larger number of signals with a stronger effect in women compared to  
5 men may be influenced by the reduction in power (proportional to the product of sample size and SNP heritability)  
6 in the men-only analysis (**Table 1**) compared the women-only analysis. Despite the power difference in the sex-  
7 specific analyses, we would not expect the difference to result in 92% of signals conferring a stronger effect in  
8 women. Second, our replication sample was too small (~1% of the discovery) to formally replicate individual SNP  
9 associations, but the fact that 78% of the 300 previously unknown index associations showed consistent direction of  
10 effect suggests a low false positive rate. Finally, our meta-analysis focused only on European-ancestry samples.  
11 Given the very different body-fat distributions between people of European and non-European ancestry, and their  
12 very different risks of adiposity-related disease, studies in non-Europeans are urgently needed (22,23).  
13  
14 In summary, the genetic variants and loci identified by this meta-analysis will likely provide starting points for  
15 further understanding the biology of body fat distribution and its relationship with disease.  
16  
17

## 1 **Materials and Methods**

2

### 3 I. Data and code availability

4

5 Code and data related to this project, including summary-level data from the meta-analyses, can be found online at  
6 <https://github.com/lindgengroup/fatdistnGWAS>.

7

### 8 II. Phenotypes

9

10 To generate phenotypes for the waist-to-hip ratio (WHR) and waist-to-hip ratio adjusted for body mass index  
11 (WHRadjBMI) analyses in the UK Biobank data (**Supplementary Table 10**), we followed a phenotype conversion  
12 consistent with that performed in previous efforts investigating WHR and WHRadjBMI by the GIANT consortium  
13 (5,24).

14

15 Using phenotype information from UK Biobank, we divided waist circumference by hip circumference to calculate  
16 the WHR measure, and then regressed the WHR measure on sex, age at assessment, age at assessment squared, and  
17 assessment centre. To generate the WHRadjBMI phenotype, we followed the same procedure and included body  
18 mass index (BMI) as an additional independent variable in the regression. We performed rank inverse normalization  
19 on the resulting residuals from the regression (**Supplementary Fig 9**) and used these normalized residuals as the  
20 tested phenotype in downstream genome-wide association testing. To generate phenotypes for the sex-specific  
21 analyses, we followed this same procedure but ran the regressions in sex-specific groups.

22

### 23 III. Genome-wide association analyses

#### 24 *The UK Biobank data*

25

26 We conducted genome-wide association testing in the second release (June 2017) version of the UK Biobank  
27 data(8); this release did not contain the corrected imputation at non-Haplotype Reference Consortium (HRC (25))  
28 sites and we therefore subset all of the SNP data down to HRC SNPs only. The UK Biobank applied quality control  
29 to samples and genotypes and imputed the resulting genotype data using sequencing-based imputation reference  
30 panels. We performed all of our genome-wide association testing and downstream analyses on the publicly-available  
31 imputation data (released in bgen format).

32

33 We excluded samples as suggested by the UK Biobank upon release of the data (**Supplementary Table 11**). Sample  
34 exclusions included samples with genotype but no imputation information, samples with missingness > 5%, samples  
35 with mismatching phenotypic and genotypic sex, and samples that have withdrawn consent since the initiation of the  
36 project.

37

1 *LD scores and genetic relationship matrix for BOLT-LMM*

2

3 We implemented all genome-wide association studies (GWAS) in BOLT-LMM (10), which performs association  
4 testing using a linear mixed model. To run, BOLT-LMM requires three primary components: the (imputed)  
5 genotypic data for association testing; a reference panel of Linkage Disequilibrium (LD) scores per SNP, calculated  
6 using LD Score Regression (12); and genotype data used to approximate a genetic relationship matrix (GRM),  
7 which is the best method available in this sample size to account for all forms of relatedness, ancestral heterogeneity  
8 in the samples, and other (potentially hidden) structure in the data.

9

10 We performed sensitivity testing (**Supplementary Information, Supplementary Tables 12-13 and**  
11 **Supplementary Fig 10**) using three LD Score reference datasets and four SNP-sets to construct the GRM. For our  
12 final GWAS, we used LD scores calculated from a randomly-selected, 9,748 unrelated UK Biobank samples (~2%  
13 of the full UK Biobank sample set; **Supplementary Information**) and a GRM constructed using: imputed SNPs  
14 with imputation info score > 0.8; MAF > 1%; Hardy Weinberg P-value >  $1 \times 10^{-8}$ ; genotype missingness < 1%, after  
15 converting imputed dosages to best-guess genotypes; LD pruned at a threshold ( $r^2$ ) of 0.2; and excluding the major  
16 histocompatibility complex, the lactase locus, and the inversions on chromosomes 8 and 17 (**Supplementary**  
17 **Information**).

18

19 *Association testing*

20

21 For genome-wide association testing, we used BOLT-LMM to run a linear mixed model (LMM). We tested SNPs  
22 with imputation quality (info) > 0.3, minor allele frequency (MAF) > 0.1% (equivalent to ~50 copies of the minor  
23 allele in the full sample), and only those single-nucleotide variants (SNVs) and single-nucleotide polymorphisms  
24 (SNPs) represented in the Haplotype Reference Consortium (25) imputation reference panel. We used only the  
25 standard LMM implementation (i.e., infinitesimal model, using --lmm) in BOLT-LMM (**Supplementary Fig 11-**  
26 **12**); we did not run association testing using a non-infinitesimal model. The only covariate used in the LMM was the  
27 SNP array used to genotype sample; we included no other covariates.

28

29 After association testing, we looked at known SNPs already reported in WHR, WHRadjBMI, and BMI (5, 24). At  
30 the previously-described loci, we checked correlation of frequency, beta, standard error, and  $-\log_{10}(\text{p-value})$  between  
31 our UK Biobank GWAS and the previous GWAS results (**Supplementary Fig 13**). Additionally, we estimated  
32 genomic inflation (lambda) and the LD Score Intercept to check if the P-values were well calibrated  
33 (**Supplementary Table 2**); calculations were performed using the LD Score software (<https://github.com/bulik/ldsc>)  
34 (12).

35

36 IV. Meta-analysis of results from UK Biobank and GIANT

37 *Data preparation and quality control*



1  
2 We downloaded summary-level results from previous meta-analyses of WHR and WHRadjBMI  
3 ([https://portals.broadinstitute.org/collaboration/giant/index.php/GIANT\\_consortium\\_data\\_files](https://portals.broadinstitute.org/collaboration/giant/index.php/GIANT_consortium_data_files) and **Supplementary**  
4 **Information**) performed by the GIANT consortium (5). Marker names in both the GIANT data and UK Biobank  
5 were lifted over to their dbSNP151 identifier. We additionally renamed markers as “rsID:A1:A2” (where A1 was the  
6 tested allele in UK Biobank) to avoid ambiguity at multiallelic SNPs in the UK Biobank data. As the GIANT data  
7 was imputed with HapMap 2 (26,27) data (hg18), we additionally lifted chromosomal positions to hg19 for this data.  
8 SNPs with a frequency difference > 15% between GIANT and UK Biobank were removed from the data  
9 (**Supplementary Fig 14**).

10  
11 *Meta-analysis and downstream quality control*

12  
13 We performed inverse variance-weighted fixed effects meta-analysis in METAL (28). To estimate LD score  
14 intercepts and genomic inflation (lambda) for the meta-analysis results, we first estimated LD scores from the same  
15 samples used to estimate the LD score reference for BOLT-LMM. LD scores were only estimated at high-quality  
16 SNPs (using the same criteria as used for SNPs included in the GRM in BOLT-LMM, but without applying a MAF  
17 threshold; **Supplementary Information**). We then calculated LD Score Regression intercepts and lambda with the  
18 LDSC software (12).

19  
20 As an additional quality control check, we reran all of our GWAS using two different subsets of the UK Biobank  
21 samples: (1) the unrelated samples only, and (2) the unrelated white British samples only. These subsamples were  
22 selected to test if our initial UK Biobank-wide GWAS was confounded by either relatedness or ancestral  
23 heterogeneity. After running these GWAS, we meta-analyzed the results with the existing GIANT summary-level  
24 data and checked the concordance of our signals (**Supplementary Fig 2-3**).

25  
26 V. Identification of index and secondary signals  
27 *Linkage disequilibrium clumping*

28  
29 To identify genomic loci (i.e., genomic windows) containing independent association signals, we first constructed a  
30 reference dataset of best-guess genotypes from 20,275 unrelated UK Biobank samples (equivalent to 5% of the  
31 unrelated sample). We converted imputed dosages of SNPs with info score > 0.3 and MAF > 0.001% to best-guess  
32 genotypes using PLINK (version 1.9), (29,30) and a conversion threshold (--hard-call-threshold) of 0.1  
33 (**Supplementary Information**). SNPs with missingness > 5% after conversion or Hardy-Weinberg equilibrium  $p <$   
34  $1 \times 10^{-7}$  were removed.

35  
36 We then used the PLINK ‘clumping’ algorithm to select top-associated SNPs ( $p < 5 \times 10^{-9}$ ) and identify all SNPs in  
37 LD ( $r^2 > 0.05$ ) with the top associated SNP and  $\pm 5$ Mb away. We determined the genomic span of each LD-based

1 clump and added 1kb up- and downstream as buffer to the region. If any of these windows overlapped, we merged  
2 them together into a single (larger) locus. As a sensitivity analysis, we ran clumping also using a smaller genomic  
3 window to calculate LD ( $\pm 2\text{Mb}$ ); the results were effectively unchanged, as  $<5$  loci appeared independent using the  
4  $\pm 2\text{Mb}$  window but were found to correlate using  $\pm 5\text{Mb}$  windows. Therefore, we report loci using the  $\pm 5\text{Mb}$  window.

5

#### 6 *Proximal conditional and joint testing*

7

8 To identify index and secondary signals within each of the clumping-based loci, we ran proximal joint and  
9 conditional analysis as implemented in the Genome-wide Complex Trait Analysis (GCTA) software (31). We ran  
10 this model (--cojo-slt) using the summary-level data within each locus, the LD reference panel constructed from  
11 UK Biobank data and also used for the locus 'clumping,' and setting genome-wide significance with  $p < 5 \times 10^{-9}$ .

12

#### 13 VI. Validation in an independent dataset

14

15 We used an independent dataset EXTEND (7,721 individuals of European descent collected from South West  
16 England, **Supplementary Table 14**) to validate our findings. We extracted the index SNPs from the HRC imputed  
17 genotypes. To generate the WHRadjBMI variable, we regressed WHR on BMI, age, age-squared, sex and principal  
18 components 1-5. We then performed rank based inverse normalization on the resulting residuals. We validated the  
19 findings in 3 steps:

20

21 (1) *Directional consistency.* We checked for directional consistency between the effect of index SNPs on  
22 WHRadjBMI from the main meta-analysis and EXTEND. We performed linear regression of WHRadjBMI on each  
23 individual SNP. We ensured all alleles were aligned to the WHRadjBMI increasing allele in the original meta-  
24 analysis. We compared directions between all 346 index SNPs and then split these into novel and known signals to  
25 determine the number of novel signals showing consistent directionality.

26

27 (2) *Variance explained.* We evaluated the proportion of variance explained by including all the index SNPs into a  
28 linear regression model and calculated the adjusted  $R^2$ . We performed these analyses using the `lm()` function in R.

29

30 (3) *Polygenic scores.* We created a weighted polygenic score based on the 346 index SNPs associated with  
31 WHRadjBMI. The weighted polygenic risk score (PRS) was calculated by summing the dosage of the  
32 WHRadjBMI-increasing alleles (weighted by the effect size on WHRadjBMI from the meta-analysis). We then  
33 performed linear regression to test the association between WHRadjBMI and the PRS in our independent dataset.

34

35 We sought to determine how likely the 5% of individuals carrying the most WHRadjBMI-increasing alleles were to  
36 meet the World Health Organization (WHO) WHR threshold used to diagnose metabolic syndrome (along with  
37 lipids and type 2 diabetes status) (13) compared to the 5% carrying the least. We used the WHR reference levels of

1 > 0.9 in men and > 0.85 in women to define cases and WHR < 0.9 in men and < 0.85 in women to define controls  
2 (13). We excluded all individuals with missing data leaving a sample size of 7,513. We took 5% of individuals  
3 (7,513 x 0.05 = 376) from the two ends of weighted PRS and coded them as 1 or 2 respectively. We tested for the  
4 likelihood of the top 5% meeting the WHR threshold to diagnose metabolic syndrome (WHO criteria) compared to  
5 the bottom 5% using a binomial logistic regression model adjusting for age, age-squared, sex and principal  
6 components 1-5.

## 7 8 VII. Collider bias analysis

9  
10 Given that we had conditioned WHR on the BMI phenotype for analysis (and BMI and WHR are correlated;  $r =$   
11 0.433 in the UK Biobank data; **Supplementary Fig 15**), we tested all index signals found in the WHRadjBMI  
12 analysis for evidence of collider bias (15, 32). To do this, we ran meta-analyses of BMI and WHR using the UK  
13 Biobank samples and pre-existing summary-level data from GIANT (5, 24) (**Supplementary Methods**). We  
14 performed these meta-analyses using identical methods to the meta-analysis of WHRadjBMI.

15  
16 Then, for each index SNP from the WHRadjBMI meta-analyses (combined as well as sex-specific) we extracted the  
17 association results from the BMI and WHR meta-analyses (**Supplementary Fig 4**). WHRadjBMI-associated SNPs  
18 with a stronger association for BMI than WHR show evidence of collider bias or pleiotropy. We additionally looked  
19 at the effect size and direction of effect in BMI and WHR, but whether the effects are from collider bias or  
20 pleiotropy cannot be determined from this data.

## 21 22 VIII. Identification of sex-dimorphic signals

23  
24 We estimated correlation between WHRadjBMI in females and in males using bivariate LD Score Regression  
25 analysis (12,33).

26  
27 We performed sex-specific GWAS in UK Biobank and meta-analyzed the results with publicly-available sex-  
28 specific data from the GIANT consortium. We identified the primary and secondary signals from these meta-  
29 analyses using methods identical to those performed in the combined analysis. We tested each primary and  
30 secondary signal for a sex-dimorphic effect by estimating the t-statistic:

31  
32

$$t = \frac{\beta_{females} - \beta_{males}}{\sqrt{se_{females}^2 + se_{males}^2 - 2r * se_{females} * se_{males}}} \quad (1)$$

33  
34 where  $se$  is the standard error and  $r$  is the genome-wide Spearman rank correlation coefficient between SNP effects  
35 in females and males. We estimated the t-statistic and the resulting so-called  $p_{diff}$  (p-value from a t-distribution with  
36 one degree of freedom (17)) as implemented in the EasyStrata software (34).

1  
2 We tested a total of 2,162 different index SNPs for sex-dimorphism; we tested all of the secondary signals as well,  
3 but these signals are by definition in linkage disequilibrium with the index SNPs (and therefore not independent).  
4 Given that we tested for sex-dimorphism at index SNPs in not only WHRadjBMI but WHR and BMI as well, we  
5 performed a test at 1,502 distinct genomic loci. Therefore, we set significance for sex-dimorphism at a Bonferroni-  
6 corrected  $p = 0.05/1,502 = 3.3 \times 10^{-5}$ .

7  
8 SNPs were determined to have a stronger effect in women if they fell into one of the following categories (abs,  
9 absolute value):

- 10  
11 (a)  $\beta_{\text{females}} \leq 0$  and  $\beta_{\text{males}} \leq 0$  and  $\text{abs}(\beta_{\text{females}}) > \text{abs}(\beta_{\text{males}})$   
12 (b)  $\beta_{\text{females}} \geq 0$  and  $\beta_{\text{males}} \geq 0$  and  $\text{abs}(\beta_{\text{females}}) > \text{abs}(\beta_{\text{males}})$   
13 (c)  $\beta_{\text{females}} \leq 0$  and  $\beta_{\text{males}} \geq 0$  and  $p_{\text{females}} < p_{\text{males}}$  and  $\text{abs}(\beta_{\text{females}}) > \text{abs}(\beta_{\text{males}})$ , or  
14 (d)  $\beta_{\text{females}} \geq 0$  and  $\beta_{\text{males}} \leq 0$  and  $p_{\text{females}} < p_{\text{males}}$  and  $\text{abs}(\beta_{\text{females}}) > \text{abs}(\beta_{\text{males}})$

15  
16 IX. Heritability calculations

17 *SNP-based heritability calculations*

18  
19 We implemented all heritability calculations in BOLT-LMM.(10) We used the same genetic relationship matrix  
20 (GRM) to estimate SNP-based heritability as we did to run our GWAS (see *Genome-wide association analyses*).  
21 This GRM included 790,000 SNPs. Heritability was estimated using only the UK Biobank samples, for which we  
22 had individual level data; these estimates are likely more accurate than those resulting from only summary-level  
23 data. We used Restricted Maximum Likelihood Estimation, implemented as --reml in BOLT.

24  
25 To test the impact of including lower-frequency SNPs in the heritability estimates, we constructed an additional  
26 GRM identically as we had for association testing but including no minor allele frequency threshold. This GRM  
27 included ~1.7M SNPs. Heritability analyses were calculated identically using this GRM and --reml in BOLT.

28  
29 To calculate whether heritability estimates in men and women were sex-dimorphic, we used the following equation  
30 to generate a z-score:

31  
32 
$$z = \frac{h_{\text{females}}^2 - h_{\text{males}}^2}{\sqrt{\text{variance}_{\text{females}} + \text{variance}_{\text{males}}}} \quad (2)$$

33  
34 We then converted the z-scores to P-values using the following formula in the statistical programming language and  
35 software suite R (version 3.4):

36

1 
$$p = 2 * pnorm(-abs(z))$$
 (3)

2

3 X. Comparison of WHRadjBMI-associated SNPs in other fat distribution phenotypes

4 *Comparison with body fat percentage*

5

6 Similarly to Shungin et al (5), we carried out analysis on the 346 index SNPs and their association with BF% and  
7 WHR. We obtained association statistics for the 346 SNPs on BF% and WHR from a GWAS of 443,001 unrelated,  
8 European-ancestry UK Biobank individuals. We aligned all results to the WHR increasing allele and used a  
9 Bonferroni-corrected P-value ( $0.05/346 = 1.44 \times 10^{-4}$ ) to determine if a SNP was associated with BF% (**Figure 2**).  
10 To determine whether sex-specific WHRadjBMI index SNPs have an adiposity phenotype, we took the 97 (female-  
11 specific) and 8 (male-specific) SNPs and independently compared their effects on WHRadjBMI and BF% in men  
12 and women. To identify which sex-dimorphic SNPs were strongly associated with BF% in men and women  
13 separately, we used a Bonferroni-corrected P-value of  $0.05/105$  ( $4.8 \times 10^{-4}$ ) (**Supplementary Fig 7** and  
14 **Supplementary Table 9**). We obtained Pearson's  $r$  correlations using the `cor()` function in R for each comparison.

15

16 *Comparison with genome-wide analysis of depot-specific traits*

17

18 Recently, Chu et al (21) performed a genome-wide association study of subcutaneous and ectopic fat depots, as  
19 measured by CT and MRI imaging, in a multi-ancestry sample. Since the meta-analysis results are publicly-available  
20 (<https://grasp.nhlbi.nih.gov/FullResults.aspx> and **Supplementary Information** for further details), we took the  
21 index SNPs from our WHRadjBMI meta-analyses (combined sample as well as sex-specific), checked for allele  
22 consistency, aligned effects to the reference allele, and tested for associations with the imaging based measures of  
23 subcutaneous and ectopic fat. We repeated these analyses in men and women separately. The depots investigated in  
24 the imaging-based GWAS were: pericardial tissue (PAT), PAT adjusted for height and weight (PATadjHtWt),  
25 subcutaneous adipose tissue (SAT), SAT Hounsfield units as measured by MRI (SATHU), visceral adipose tissue  
26 (VAT), VAT Hounsfield units (VATHU), ratio of VAT to SAT (VAT/SAT), and VAT adjusted for BMI  
27 (VATadjBMI).

28

29 We calculated Pearson's  $r$  correlations between z-scores in WHRadjBMI (calculated by dividing the SNP beta by  
30 the standard error) and SNP z-scores reported in Chu et al (21). We evaluated significance of the correlation by  
31 performing a t-test (implemented as `cor.test()` in R). Correlations were considered significant if P-value  $< 0.05/3$   
32 sample groups/9 phenotypes =  $1.9 \times 10^{-3}$ .

33

34

1 **Acknowledgements**

2  
3 This research was conducted using the UK Biobank Resource under Application Numbers 11867 and 9072.  
4 EXTEND data were provided by the Peninsula Research Bank, part of the NIHR Exeter Clinical Research Facility.  
5

6 C.M.L is supported by the Li Ka Shing Foundation, WT-SSI/John Fell funds and by the NIHR Biomedical Research  
7 Centre, Oxford, by Widenlife and NIH (CRR00070 CR00.01).  
8

9 S.L.P. is supported by a Veni Fellowship 016.186.071 (ZonMW) from the Dutch Organization for Scientific  
10 Research (Nederlandse Organisatie voor Wetenschappelijk Onderzoek, NWO).  
11

12 H.Y. is funded by Diabetes UK RD Lawrence fellowship (grant: 17/0005594).  
13

14 A.R.W. and T.M.F. are supported by the European Research Council grant: 323195:GLUCOSEGENES-FP7-  
15 IDEAS-ERC. R.B. is funded by the Wellcome Trust and Royal Society grant: 104150/Z/14/Z.  
16

17 J.T. is funded by the ERDF and a Diabetes Research and Wellness Foundation Fellowship.  
18

19 S.E.J. is funded by the Medical Research Council (grant: MR/M005070/1).  
20

21 P.M.V. and J.Y. are funded by Australian National Health and Medical Research Council (1078037 and 1113400).  
22

23 J.Y. is supported by the Sylvia & Charles Viertel Charitable Foundation.  
24

25 D.C.C.-C. is supported by a grant from the U.S. National Institutes of Health (K01 HL127265).  
26

27 A.T.H. is a Wellcome Trust senior investigator and NIHR Senior Investigator.  
28  
29  
30

31 **Authorship contributions**

32  
33 Data collection and analysis: S.L.P., C.S., A.P.M., A.R.W.,  
34 Data interpretation: S.L.P., C.S., C.M.L., T.M.F., H.Y., A.P.M., C.G.  
35 Study supervision: H.Y., T.F., S.L.P., C.M.L.  
36 First draft of the manuscript: S.L.P., C.M.L.  
37 Critical revisions of the manuscript: all co-authors  
38

1 **Conflict of Interest Statement**

2

3 The authors declare no conflict of interest.

4

5

1 **References**

- 2
- 3 1. GBD 2015 Obesity Collaborators, Afshin,A., Forouzanfar,M.H., Reitsma,M.B., Sur,P., Estep,K., Lee,A.,  
4 Marczak,L., Mokdad,A.H., Moradi-Lakeh,M., *et al.* (2017) Health Effects of Overweight and Obesity in 195  
5 Countries over 25 Years. *N. Engl. J. Med.*, **377**, 13–27.
- 6 2. WHO | Obesity and overweight (2018) *WHO | Obesity and overweight*.  
7 <http://www.who.int/mediacentre/factsheets/fs311/en/> Accessed February 22, 2018.
- 8 3. Heymsfield,S.B. and Wadden,T.A. (2017) Mechanisms, Pathophysiology, and Management of Obesity. *N. Engl.*  
9 *J. Med.*, **376**, 254–266.
- 10 4. Wang,Y., Rimm,E.B., Stampfer,M.J., Willett,W.C. and Hu,F.B. (2005) Comparison of abdominal adiposity and  
11 overall obesity in predicting risk of type 2 diabetes among men1–3. *Am. J. Clin. Nutr.*, **81**, 555–563.
- 12 5. Shungin,D., Winkler,T.W., Croteau-Chonka,D.C., Ferreira,T., Locke,A.E., Mägi,R., Strawbridge,R.J., Pers,T.H.,  
13 Fischer,K., Justice,A.E., *et al.* (2015) New genetic loci link adipose and insulin biology to body fat distribution.  
14 *Nature*, **518**, 187–196.
- 15 6. Rose,K.M., Newman,B., Mayer-Davis,E.J. and Selby,J.V. (1998) Genetic and behavioral determinants of waist-  
16 hip ratio and waist circumference in women twins. *Obes. Res.*, **6**, 383–392.
- 17 7. Emdin,C.A., Khera,A.V., Natarajan,P., Klarin,D., Zekavat,S.M., Hsiao,A.J. and Kathiresan,S. (2017) Genetic  
18 Association of Waist-to-Hip Ratio With Cardiometabolic Traits, Type 2 Diabetes, and Coronary Heart Disease.  
19 *JAMA*, **317**, 626–634.
- 20 8. Sudlow,C., Gallacher,J., Allen,N., Beral,V., Burton,P., Danesh,J., Downey,P., Elliott,P., Green,J., Landray,M., *et*  
21 *al.* (2015) UK biobank: an open access resource for identifying the causes of a wide range of complex diseases  
22 of middle and old age. *PLoS Med.*, **12**, e1001779.
- 23 9. Pulit,S.L., Karaderi,T. and Lindgren,C.M. (2017) Sexual dimorphisms in genetic loci linked to body fat  
24 distribution. *Biosci. Rep.*, **37**.
- 25 10. Loh,P.-R., Tucker,G., Bulik-Sullivan,B.K., Vilhjálmsson,B.J., Finucane,H.K., Salem,R.M., Chasman,D.I.,  
26 Ridker,P.M., Neale,B.M., Berger,B., *et al.* (2015) Efficient Bayesian mixed-model analysis increases  
27 association power in large cohorts. *Nat. Genet.*, **47**, 284–290.
- 28 11. Pulit,S.L., de With,S.A.J. and de Bakker,P.I.W. (2017) Resetting the bar: Statistical significance in whole-  
29 genome sequencing-based association studies of global populations. *Genet. Epidemiol.*, **41**, 145–151.
- 30 12. Bulik-Sullivan,B.K., Loh,P.-R., Finucane,H.K., Ripke,S., Yang,J., Consortium,S.W.G. of T.P.G., Patterson,N.,  
31 Daly,M.J., Price,A.L. and Neale,B.M. (2015) LD Score regression distinguishes confounding from polygenicity  
32 in genome-wide association studies. *Nat. Genet.*, **47**, 291–295.
- 33 13. Huang,P.L. (2009) A comprehensive definition for metabolic syndrome. *Dis. Model. Mech.*, **2**, 231–237.
- 34 14. Cole,S.R., Platt,R.W., Schisterman,E.F., Chu,H., Westreich,D., Richardson,D. and Poole,C. (2010) Illustrating  
35 bias due to conditioning on a collider. *Int. J. Epidemiol.*, **39**, 417–420.
- 36 15. Day,F.R., Loh,P.-R., Scott,R.A., Ong,K.K. and Perry,J.R.B. (2016) A Robust Example of Collider Bias in a  
37 Genetic Association Study. *Am. J. Hum. Genet.*, **98**, 392–393.
- 38 16. Frayling,T.M., Timpson,N.J., Weedon,M.N., Zeggini,E., Freathy,R.M., Lindgren,C.M., Perry,J.R.B.,  
39 Elliott,K.S., Lango,H., Rayner,N.W., *et al.* (2007) A common variant in the FTO gene is associated with body  
40 mass index and predisposes to childhood and adult obesity. *Science*, **316**, 889–894.
- 41 17. Randall,J.C., Winkler,T.W., Kutalik,Z., Berndt,S.I., Jackson,A.U., Monda,K.L., Kilpeläinen,T.O., Esko,T.,



- 1 Mägi,R., Li,S., *et al.* (2013) Sex-stratified genome-wide association studies including 270,000 individuals show  
2 sexual dimorphism in genetic loci for anthropometric traits. *PLoS Genet.*, **9**, e1003500.
- 3 18. Lotta,L.A., Gulati,P., Day,F.R., Payne,F., Ongen,H., van de Bunt,M., Gaulton,K.J., Eicher,J.D., Sharp,S.J.,  
4 Luan,J. 'an, *et al.* (2017) Integrative genomic analysis implicates limited peripheral adipose storage capacity in  
5 the pathogenesis of human insulin resistance. *Nat. Genet.*, **49**, 17–26.
- 6 19. Yaghoobkar,H., Lotta,L.A., Tyrrell,J., Smit,R.A.J., Jones,S.E., Donnelly,L., Beaumont,R., Campbell,A.,  
7 Tuke,M.A., Hayward,C., *et al.* (2016) Genetic Evidence for a Link Between Favorable Adiposity and Lower  
8 Risk of Type 2 Diabetes, Hypertension, and Heart Disease. *Diabetes*, **65**, 2448–2460.
- 9 20. Scott,R.A., Fall,T., Pasko,D., Barker,A., Sharp,S.J., Arriola,L., Balkau,B., Barricarte,A., Barroso,I., Boeing,H.,  
10 *et al.* (2014) Common genetic variants highlight the role of insulin resistance and body fat distribution in type 2  
11 diabetes, independent of obesity. *Diabetes*, **63**, 4378–4387.
- 12 21. Chu,A.Y., Deng,X., Fisher,V.A., Drong,A., Zhang,Y., Feitosa,M.F., Liu,C.-T., Weeks,O., Choh,A.C., Duan,Q.,  
13 *et al.* (2017) Multiethnic genome-wide meta-analysis of ectopic fat depots identifies loci associated with  
14 adipocyte development and differentiation. *Nat. Genet.*, **49**, 125–130.
- 15 22. Pulit,S.L., Voight,B.F. and de Bakker,P.I.W. (2010) Multiethnic genetic association studies improve power for  
16 locus discovery. *PLoS One*, **5**, e12600.
- 17 23. Petrovski,S. and Goldstein,D.B. (2016) Unequal representation of genetic variation across ancestry groups  
18 creates healthcare inequality in the application of precision medicine. *Genome Biol.*, **17**, 157.
- 19 24. Locke,A.E., Kahali,B., Berndt,S.I., Justice,A.E., Pers,T.H., Day,F.R., Powell,C., Vedantam,S.,  
20 Buchkovich,M.L., Yang,J., *et al.* (2015) Genetic studies of body mass index yield new insights for obesity  
21 biology. *Nature*, **518**, 197–206.
- 22 25. McCarthy,S., Das,S., Kretzschmar,W., Delaneau,O., Wood,A.R., Teumer,A., Kang,H.M., Fuchsberger,C.,  
23 Danecek,P., Sharp,K., *et al.* (2016) A reference panel of 64,976 haplotypes for genotype imputation. *Nat.*  
24 *Genet.*, **48**, 1279–1283.
- 25 26. Frazer,K.A., Ballinger,D.G., Cox,D.R., Hinds,D.A., Stuve,L.L., Gibbs,R.A., Belmont,J.W., Boudreau,A.,  
26 Hardenbol,P., Leal,S.M., *et al.* (2007) A second generation human haplotype map of over 3.1 million SNPs.  
27 *Nature*, **449**, 851–861.
- 28 27. The International HapMap Consortium., Gibbs,R.A., Belmont,J.W., Hardenbol,P., Willis,T.D., Yu,F., Yang,H.,  
29 Ch'ang,L.-Y., Huang,W., Liu,B., *et al.* (2003) The International HapMap Project. *Nature*, **426**, 789–796.
- 30 28. Willer,C.J., Li,Y. and Abecasis,G.R. (2010) METAL: Fast and efficient meta-analysis of genomewide  
31 association scans. *Bioinformatics*, **26**, 2190–2191.
- 32 29. Chang,C.C., Chow,C.C., Tellier,L.C., Vattikuti,S., Purcell,S.M. and Lee,J.J. (2015) Second-generation PLINK:  
33 rising to the challenge of larger and richer datasets. *Gigascience*, **4**, 1–16.
- 34 30. Purcell,S., Neale,B., Todd-Brown,K., Thomas,L., Ferreira,M. a. R., Bender,D., Maller,J., Sklar,P., de  
35 Bakker,P.I.W., Daly,M.J., *et al.* (2007) PLINK: a tool set for whole-genome association and population-based  
36 linkage analyses. *Am. J. Hum. Genet.*, **81**, 559–575.
- 37 31. Yang,J., Lee,S.H., Goddard,M.E. and Visscher,P.M. (2011) GCTA: A tool for genome-wide complex trait  
38 analysis. *Am. J. Hum. Genet.*, **88**, 76–82.
- 39 32. Munafò,M.R., Tilling,K., Taylor,A.E., Evans,D.M. and Davey Smith,G. (2017) Collider scope: when selection  
40 bias can substantially influence observed associations. *Int. J. Epidemiol.*, 10.1093/ije/dyx206.
- 41 33. Bulik-Sullivan,B., Finucane,H.K., Anttila,V., Gusev,A., Day,F.R., Loh,P.-R., ReproGen Consortium, Psychiatric  
42 Genomics Consortium, Genetic Consortium for Anorexia Nervosa of the Wellcome Trust Case Control

- 1 Consortium 3, Duncan,L., *et al.* (2015) An atlas of genetic correlations across human diseases and traits. *Nat.*  
2 *Genet.*, **47**, 1236–1241.
- 3 34. Winkler,T.W., Kutalik,Z., Gorski,M., Lottaz,C., Kronenberg,F. and Heid,I.M. (2015) EasyStrata: evaluation and  
4 visualization of stratified genome-wide association meta-analysis data. *Bioinformatics*, **31**, 259–261.
- 5
- 6
- 7

1 **Tables and Figures**

2

3 **Table 1 | Large-scale meta-analysis in body fat distribution.** We performed a meta-analysis of fat distribution as  
 4 measured by WHRadjBMI in up to 694,649 individuals. We performed analyses of WHR as a sensitivity measure.  
 5 Our analyses increase the number of WHRadjBMI-associated loci ( $p < 5 \times 10^{-9}$ , to account for SNP density in UK  
 6 Biobank) to 346 loci. SNP-based heritability ( $h_g^2$ ) results, estimated using the restricted maximum likelihood method  
 7 implemented (10), and top-associated loci indicate patterns of sex-dimorphism. The top-associated index SNPs  
 8 explain 3.9% of the overall phenotypic variance (i.e., adjusted  $R^2$ ) in fat distribution (calculated in an independent  
 9 dataset,  $N = 7,721$ ).

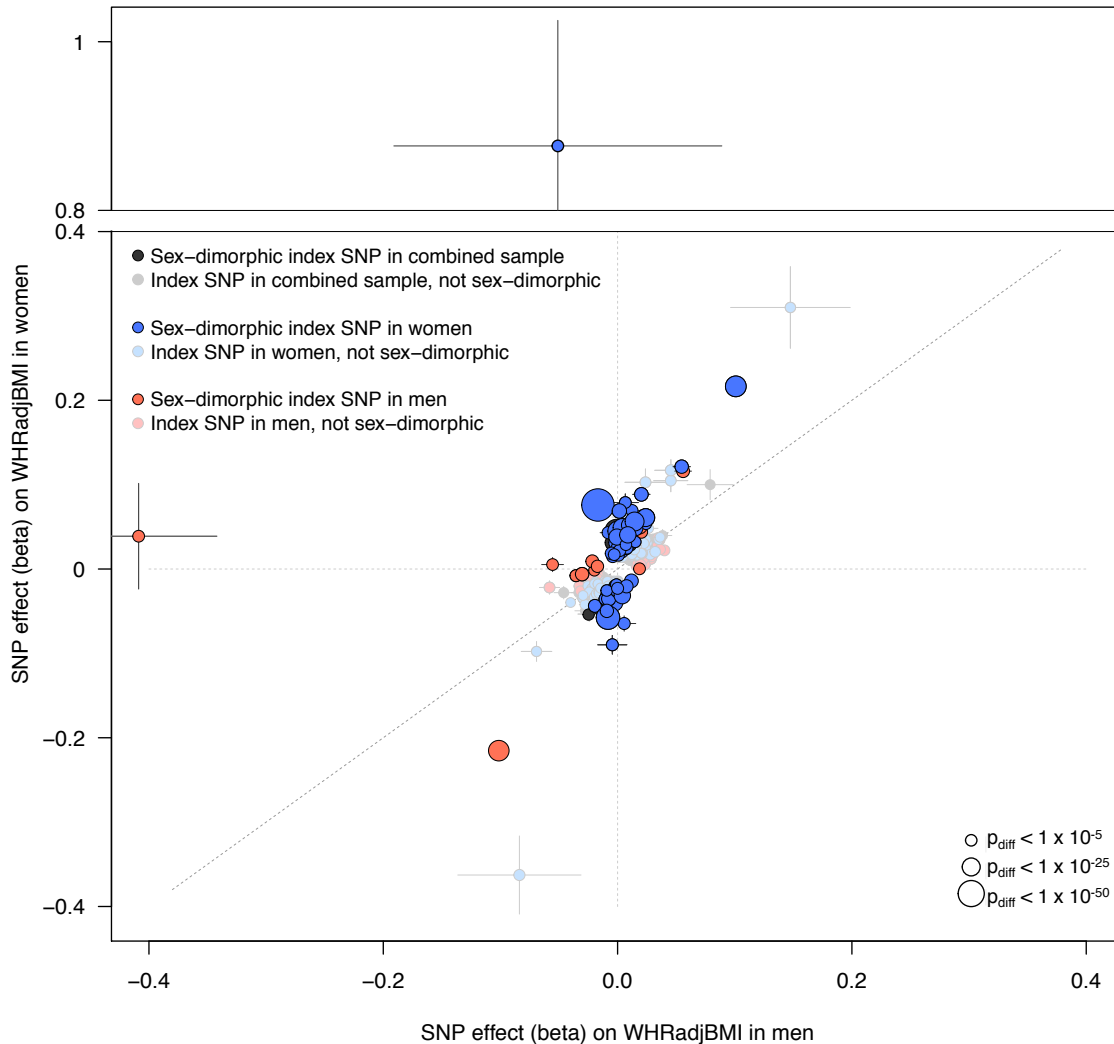
10

11

Phenotype	Sex	Sample sizes			Associated loci $p < 5 \times 10^{-9}$		Dimorphic index SNPs (% of total)	$h_g^2$ (se)	Variance explained
		UKBB	GIANT	Meta	Loci	Independent signals			
WHRadjBMI	Combined	484,563	210,086	694,649	346	463	53 (15.3)	0.174 (0.002)	3.9%
	Women	262,759	116,742	379,501	266	363	77 (28.9)	0.256 (0.003)	3.6%
	Men	221,804	93,480	315,284	91	102	13 (14.3)	0.167 (0.003)	1.0%
WHR	Combined	485,486	212,248	697,734	316	382	37 (11.7)	0.194 (0.002)	3.0%
	Women	263,148	118,004	381,152	203	261	64 (31.5)	0.254 (0.003)	4.0%
	Men	222,338	94,434	316,772	79	82	10 (12.7)	0.208 (0.003)	0.3%

12

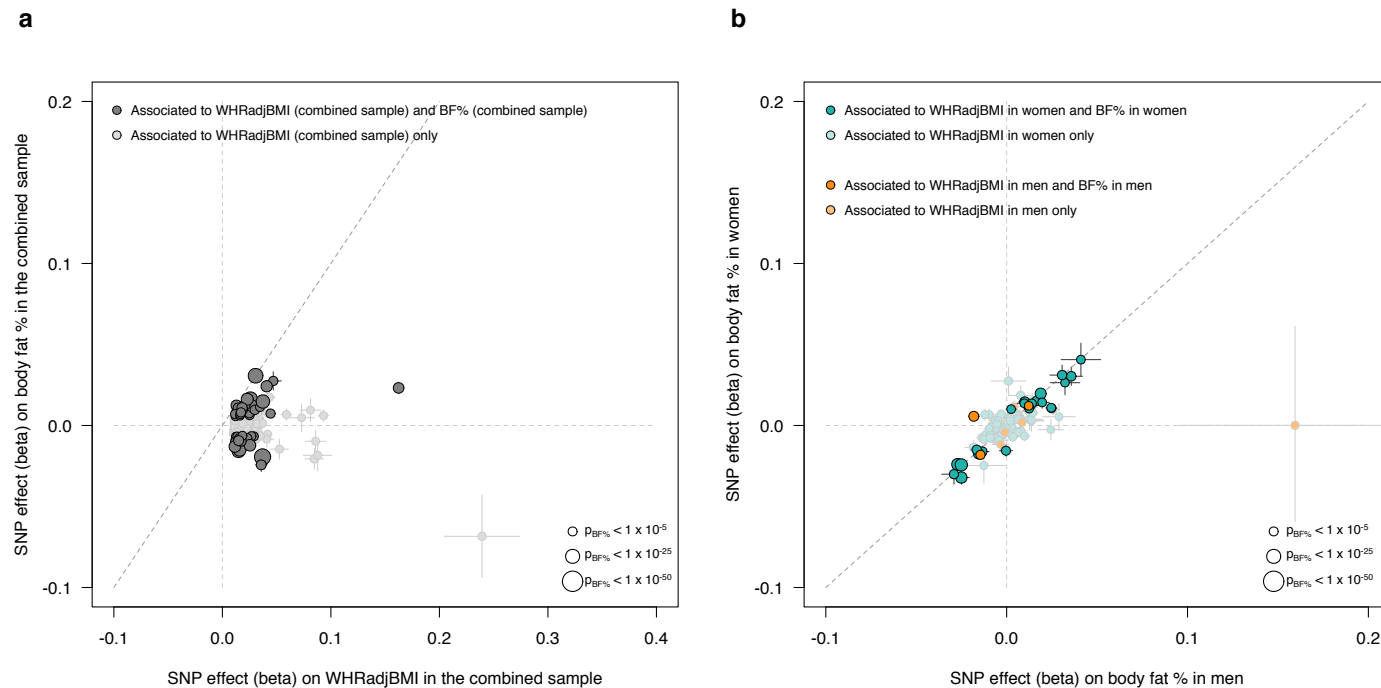
1 **Figure 1 | Sex-dimorphic association signals in fat distribution.** For each associated locus from either the  
2 combined or sex-specific meta-analyses, we tested the index SNP for sex-dimorphism. We plot here all index SNPs  
3 from each of the three meta-analyses (combined, women only, and men only). SNPs that are significantly sex-  
4 dimorphic ( $p_{\text{diff}} < 3.3 \times 10^{-5}$ ) are represented by boldly-colored circles, while index SNPs that are not sex-dimorphic  
5 are plotted with faded colors. Despite the expectation that SNPs identified in the combined sample (men and  
6 women, grey points) will be biased away from sex-dimorphism, and index SNPs identified in the sex-specific  
7 sample will be biased towards sex-dimorphism (due to winner's curse), we observed stronger effects in women  
8 across all SNPs. Of the index SNPs from the men-only analysis (orange points), 14% showed evidence of sex-  
9 dimorphism. In contrast, ~29% of the index SNPs from the women-only analysis (blue points) show evidence of  
10 dimorphism. Over all sex-dimorphic SNPs, 92.4% show a stronger effect in women compared to men. Points are  
11 sized by the  $-\log_{10}(p_{\text{diff}})$  of the sex-dimorphism test. Horizontal bars indicate standard error in men; vertical bars  
12 indicate standard error in women.



13  
14

1 **Figure 2 | Effects of WHRadjBMI-associated SNPs on body fat percentage.** (a) We investigated the impact of the 346 WHRadjBMI index SNPs (discovered  
2 in the combined analysis) on body fat percentage (BF%) in 449,001 UK Biobank individuals. Of the 346 SNPs, 59 (17.1%) are associated with BF% ( $p <$   
3  $0.05/346 = 1.44 \times 10^{-4}$ , dark grey points). We oriented the effects of the SNPs to the WHRadjBMI-increasing effect, and found that 34 of the 59 BF%-associated  
4 SNPs associate with increased BF%, while 25/59 associate with decreased BF%, indicating that WHRadjBMI-associated SNPs can effect BF% in both  
5 directions. (b) Given the sex-dimorphic signature observed in WHRadjBMI-associated SNPs and the increased number of SNPs with stronger effects on  
6 WHRadjBMI in women, we investigated the effect of the 105 sex-dimorphic index SNPs identified from the three meta-analyses (in the combined sample, in  
7 women only, in men only) on BF% in men or women separately. Of the 105 dimorphic SNPs, 97 were female specific (aquamarine points) and conferred a  
8 stronger effect on WHRadjBMI (on average) compared to the 8 male-specific SNPs (orange points). We plot the 105 sex-dimorphic SNPs by their effect on BF%  
9 in men (x-axis) and in women (y-axis). Of the 105 SNPs, 56 associate with BF% ( $p < 0.05/05 = 4.8 \times 10^{-3}$ ). Despite the fact that these SNPs confer different  
10 effects on WHRadjBMI within sex-specific groups, we found that they confer relatively similar effects in BF% in sex-specific groups. All points are scaled  
11 size to their strength of association in BF%.

12  
13



1	<b>Abbreviations</b>	
2		
3	BMI	Body mass index
4	WHR	Waist-to-hip ratio
5	WHRadjBMI	Waist-to-hip ratio, adjusted for body mass index
6	GWAS	Genome-wide association study
7	BF%	Body fat percentage
8	LD	Linkage disequilibrium
9	SNP	Single nucleotide polymorphism
10	UKBB	UK Biobank
11	T2D	Type 2 diabetes
12	GRS	Genetic risk score
13	CT	Computerized tomography
14	MR	Magnetic resonance imaging
15	BOLT-LMM	BOLT Linear Mixed Model
16	BOLT-REML	BOLT restricted maximum likelihood
17	PCA	Principal component analysis
18	SAT	Subcutaneous adipose tissue
19	VAT	Visceral adipose tissue
20	PAT	Pericardial adipose tissue
21	GRM	Genetic relationship matrix
22		
23		