# Genome-wide association study of body fat distribution identifies novel loci and sex-specific effects

Mathias Rask-Andersen[*1], Torgny Karlsson[1], Weronica E Ek[1], Åsa Johansson[1]

[1]*Department of Immunology, Genetics and Pathology,* Science for Life Laboratory

*Uppsala University. Box 256, 751 05, Uppsala, Sweden.*

[*]*Corresponding author:* *mathias.rask-andersen@igp.uu.se*

**Body mass and composition are complex traits of clinical interest due to their links to cardiovascular- and metabolic diseases. In this study, we performed genome-wide association studies (GWAS) for the distribution of body fat to the arms, legs and trunk. Proportions of fat, distributed to the different compartments, were calculated for 362,499 individuals from the UK biobank, based on segmental bioimpedance analysis (sBIA) estimates. A total of 85 body fat distribution loci were identified, using data from 116,138 participants, and replicated in an independent set of participants (N = 246,361). Out of these loci, 28 were associated with the proportion of fat in the arms, 43 with the legs and 57 with the trunk. A large degree of overlap was observed between legs and trunk loci (N=33), while arm loci overlapped to a smaller degree with leg and trunk loci (N=4 and 6, respectively). As many as 50 of the loci have not previously been associated with any adiposity-related trait. Within the novel loci we found lipid metabolism-related genes such as *CILP2* and *OSBPL7*, as well as androgen receptor function-related genes such as *ESR1*, *ID4* and *ADAMTS17*. Significant interactions between the top SNP and sex were observed for 38 loci. Our findings provide evidence for multiple loci that affect the distribution of body fat to discrete compartments of the human body, and highlight that genetic effects differ between men and women, in particular for distribution of body fat to the legs and trunk.**

Overweight and obesity have reached epidemic proportions globally [1]. Almost 40% of the world's population are now overweight [2] and 10.8% obese [3]. Overweight and obesity are set to become the world's leading preventable risk factors for disease and early death due to their links to increased risk of developing disease such as type

2 diabetes, cardiovascular disease, and cancer [4]. While body mass index (BMI) serves as a useful proxy for body adiposity, it is unable to fully discriminate between adipose and lean mass. Variation in body fat distribution have been identified by using other proximal measurements that better represent body adiposity, such as waist-to-hip ratio or hip-, and waist circumference. Fat distribution also influences risk for developing cardiovascular disease, e.g. visceral fat stored around the vital organs has been linked to insulin resistance [5], higher risk of developing metabolic syndrome, cardiovascular disease as well as increased mortality [6]. Accumulation of body fat, and distribution of fat to different parts of the body is well known to differ between sexes. For example, after puberty, women accumulate fat in the trunk and limbs to a proportionally greater extent while men accumulate more fat in the trunk [5]. Differences in fat distribution between sexes may be a factor in the lower risk for myocardial infarction and coronary death observed in women, as compared to men [7].

Through genome-wide association studies (GWAS), researchers have identified hundreds of loci to be associated with proximal measurements of body fat mass and distribution, using measurements such as BMI [8–13], waist-to-hip ratio [14,15] or hip-, and waist circumference [15]. The total amount of body fat, estimated by bioelectrical impedance analysis (BIA) or dual energy X-ray absorptiometry (DXA), have also been analyzed in GWAS [16,17]. These methods generate aggregate data for the whole body. As such, they are unable to discriminate adiposity between different compartments of the body. GWAS has been performed for subcutaneous- and visceral adiposity, measured with computed tomography scans (CT), albeit in a relatively limited number of individuals (N=10,577) [18].

Developments in BIA technology has now allowed for cost-efficient segmental body composition scans that estimate of the fat content of the trunk, arms and legs, with high accuracy [19]. In this study, we utilize segmental BIA data on 362,499 participants of the UK Biobank to study the genetic determinants of body fat distribution to the trunk, arms and legs. For this purpose, we performed GWAS on the proportion of body fat distributed to these compartments. We also performed sex-stratified analyses to determine sex-specific effects and interactions.

**Results**

We conducted a two-stage GWAS using data from the interim release of genotype data in UK Biobank as a discovery cohort. Another set of participants for which genotype data were made available as part of the second release was used for replication. After removing non-Caucasians, genetic outliers and related individuals, 116,138 and 246,361 participants remained in the discovery and replication cohort respectively. The proportions of body fat distributed to the arms – arm fat ratio (AFR), the legs – leg fat ratio (LFR) and the trunk – trunk fat ratio (TFR) were calculated by dividing the fat mass per compartment with the total body fat mass in each participant.

Women were found to have higher total fat mass compared to men ($p < 2.2*10^{-16}$, Table 1), as well as higher amount of fat in the arms and legs ($p < 2.2*10^{-16}$, Table 1). Males had on average about a 12% higher proportion of their body fat located in the trunk compared with females (62.2% vs. 50.3%, $p < 2.2*10^{-16}$, Table 1), while women had approximately 12% higher proportion of body fat located in the legs (39.7% vs.

28.1%, $p < 2.2*10^{-16}$, Table 1). While the amount of adipose tissue in the arms was estimated to be higher in women compared to men, the relative amounts of fat in the arms were more similar. Only marginal differences were seen in basic characteristics between the discovery and replication cohorts (Table 1).

**Table 1. General descriptives of UK Biobank participants* included in the analyses.** Data is presented for the discovery and replication cohorts. The cohorts were filtered for unrelated Caucasians. Mean values are presented ± standard deviations.

| | **Men** | | **Women** | |
|---|---|---|---|---|
| | ***Discovery*** | ***Replication*** | ***Discovery*** | ***Replication*** |
| N | 55,006 | 147,374 | 61,132 | 179,191 |
| Age (years) | 57.5 ± 8.1 | 57.0 ± 8.1 | 56.8 ± 7.9 | 56.7 ± 8.0 |
| Height (cm) | 175.7 ± 6.7 | 175.9 ± 6.8 | 162.6 ± 6.2 | 162.6 ± 6.2 |
| Weight (kg) | 86.3 ± 14.5 | 86.1 ± 14.2 | 71.8 ± 14.2 | 71.4 ± 13.9 |
| BMI (kg/m2) | 27.9 ± 4.3 | 27.8 ± 4.2 | 27.2 ± 5.2 | 27.0 ± 5.12 |
| Waist circumference (cm) | 97.4 ± 11.5 | 97.0 ± 11.3 | 85.0 ± 12.7 | 84.5 ± 12.5 |
| Hip circumference (cm) | 103.6 ± 7.7 | 103.5 ± 7.6 | 103.6 ± 10.5 | 103.3 ± 10.3 |
| Waist-to-hip ratio | 0.94 ± 0.1 | 0.94 ± 0.1 | 0.82 ± 0.1 | 0.82 ± 0.1 |
| **Impedance measurements** | | | | |
| Total fat mass (kg) | 22.6 ± 8.4 | 22.4 ± 8.2 | 27.2 ± 10.1 | 26.9 ± 10.0 |
| Leg fat mass (kg) | 3.16 ± 1.3 | 3.12 ± 1.3 | 5.28 ± 1.8 | 5.23 ± 1.7 |
| Arm fat mass (kg) | 1.12 ± 0.5 | 1.11 ± 0.5 | 1.44 ± 0.8 | 1.41 ± 0.8 |
| Trunk fat mass (kg) | 14.0 ± 5.1 | 13.9 ± 5.0 | 13.8 ± 5.3 | 13.6 ± 5.2 |
| **Proportional distribution of body fat** | | | | |
| Leg fat ratio - LFR (%) | 28.0 ± 3.4 | 28.1 ± 3.3 | 39.7 ± 4.0 | 39.7 ± 4.1 |
| Arm fat ratio - AFR (%) | 9.9 ± 1.3 | 9.9 ± 1.3 | 10.1 ± 1.6 | 10.1 ± 1.6 |
| Trunk fat ratio - TFR (%) | 62.2 ± 3.6 | 62.2 ± 3.6 | 50.3 ± 4.1 | 50.3 ± 4.1 |

Interestingly, we did not find a strong correlation between our ratios and other anthropometric traits. A substantial part of the variation in AFR could be explained by

BMI or waist circumference (Table 2). However, anthropomorphic traits did not substantially contribute to explaining the variance in LFR or TFR (Table 2).

**Table 2. The amount of variation in adipose tissue distribution explained by anthropomorphic descriptors.**

| | BMI | Waist circumference | Waist-to-hip ratio | Height |
|---|---|---|---|---|
| **Arm fat ratio (AFR)** | 41.9% | 26.2% | 5.8% | 1.1% |
| **Leg fat ratio (LFR)** | 1.2% | 2.0% | 0.5% | 2.0% |
| **Trunk fat ratio (TFR)** | 0.1% | 0.1% | 0.0% | 2.6% |

*Discovery GWAS for body fat distribution*

In the discovery GWAS, each of the three phenotypes (AFR, LFR and TFR) were analyzed in the whole discovery cohort (sex-combined) and by stratifying by sex (males and females). We first estimated the proportion of variance in body fat distribution to the arms, legs and trunk that were explained by the genotyped SNPs (N=730,616) [20]. We find that approximately 20-24% of the variance can be explained by considering all genotyped SNPs simultaneously in the sex-combined cohort. Interestingly, we also find that a higher degree of variance could be explained by genotyped SNPs in females compared to males (40-42% vs. 18-27%), and that this was consistent across all phenotypes (Table 3).

**Table 3. SNP-heritability estimates for body fat distribution ratios stratified by sex.**

| | N | $h_g^2$ (95% CI) | Number of GWAS regions in the discovery | Number of GWAS regions that replicated |
|---|---|---|---|---|
| **Arm fat ratio (AFR)** | | | | |
| **Sex-combined** | 113,951 | 0.20 (0.19-0.21) | 29 | 19 |
| **Females** | 60,133 | 0.42 (0.40-0.44) | 19 | 14 |
| **Males** | 53,779 | 0.27 (0.24-0.29) | 12 | 7 |
| **Leg fat ratio (LFR)** | | | | |
| **Sex-combined** | 113,912 | 0.24 (0.23-0.25) | 32 | 26 |

| | | | | |
|---|---|---|---|---|
| **Females** | 60,156 | 0.40 (0.38-0.42) | 39 | 33 |
| **Males** | 53,795 | 0.20 (0.18-0.22) | 3 | 1 |
| **Trunk fat ratio (TFR)** | | | | |
| **Sex-combined** | 113,893 | 0.21 (0.20-0.22) | 49 | 32 |
| **Females** | 60,123 | 0.42 (0.40-0.45) | 63 | 49 |
| **Males** | 53,770 | 0.18 (0.15-0.20) | 6 | 1 |

$h_g^2$: SNP heritability estimates from GCTA which represents the fraction of the phenotypic variance that can be explained by the SNPs analyzed.

A total of 25,472,837 imputed SNPs with MAF of at least 0.01% were analyzed in the discovery GWAS. A total of 11,350 associations with $P< 1*10^{-7}$ were identified to be associated with any of the phenotypes: AFR, LFR or TFR in the sex-combined and sex-stratified analyses (Figure 1, S1-2 Fig, S1 supplementary Data). Genomic inflation factors were low for all GWAS ($\lambda$= 1.05-1.14, S3 Fig). The associations corresponded to 5,542 unique SNPs across 111 loci. Many of the regions were associated with multiple phenotypes or associated with one phenotype in more that one of the strata (males, females or sex-combined). Conditional analysis also identified secondary associations with TFR at two loci, near *ACAN* and *ADAMTS17* on chromosome 15q26.1 and 15q26.3, respectively (S1 Table).

*Replication of top GWAS SNPs*

For each phenotype and strata, the most significant SNP in each locus was taken forward for replication. These SNPs were selected independent on whether the locus was associated with multiple phenotypes/strata. Since different phenotypes/strata had different top-SNPs for some loci, a total of 169 unique top-SNPs were selected for replication of 229 associations at the 111 distinct loci. A total of 182 SNP-trait associations distributed across 78 loci replicated (Bonferroni adjusted P< 0.05/229) (S1 Supplementary Data).

Out of the 78 loci that replicated, a total of 28 loci were associated with AFR, 43 with LFR and 57 with TFR in either the sex-combined or sex-stratified analyses. There was substantial overlap between LFR and TFR loci (N= 33). AFR loci overlapped to a smaller degree with LFR (N=4 loci), and TFR (N=6 loci). Three loci in the vicinity of *KRTCAP2/PKLR*, *ADAMTSL3* and *GDF5* were associated with all three phenotypes. As many as 50 of the 78 loci did not overlap with previous GWAS loci for body adiposity-related traits and were considered novel adiposity loci (Table 4). However, 24 of the novel adiposity loci overlap with previous GWAS for height.

**Table 4. Novel body adiposity loci associated with the proportion of adipose tissue distributed to the arms (AFR)**

| | | | | | | Discovery | | | Replication | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Leading SNP** | **Chr** | **Position (bp)** | **Proximal gene(s)** | **MAF1** | **A1** | $N_1$ | $\beta_1$ | $p_1$ | $N_2$ | $\beta_2$ | $p_2$ |
| **Combined** | | | | | | | | | | | |
| **rs4971091** | 1 | 155,143,768 | *KRTCAP/* | 37.82% | T | 112,229 | -0.0235 | $6.31*10^{-8}$ | 240,061 | -0.0128 | $1.64*10^{-5}$ |
| **rs13011472** | 2 | 57,961,602 | *VRK2* | 48.97% | G | 109,797 | -0.0235 | $3.66*10^{-8}$ | 238,574 | -0.0129 | $8.43*10^{-6}$ |
| **rs351855** | 5 | 176,520,243 | *FGFR4* | 29.60% | A | 113,979 | 0.0289 | $3.17*10^{-10}$ | 241,564 | 0.0161 | $3.18*10^{-7}$ |
| **rs56282717** | 7 | 150,657,095 | *KCNH2* | 24.30% | A | 111,762 | -0.0272 | $3.80*10^{-8}$ | 237,649 | -0.0196 | $6.38*10^{-9}$ |
| **rs1138714** | 11 | 825,110 | *PNPLA2* | 43.46% | A | 104,922 | -0.0241 | $4.17*10^{-8}$ | 233,077 | -0.0125 | $2.35*10^{-5}$ |
| **rs1789166** | 11 | 69,482,091 | *ORAOV1* | 35.21% | C | 112,015 | -0.0268 | $1.33*10^{-9}$ | 239,788 | -0.0130 | $1.74*10^{-5}$ |
| **Females** | | | | | | | | | | | |
| **rs61813324** | 1 | 156,049,877 | *MEX3A* | 13.14% | T | 55,824 | 0.0484 | $8.86*10^{-8}$ | 127,869 | 0.0258 | $1.21*10^{-5}$ |
| **rs7562173** | 2 | 46,976,217 | *SOCS5* | 38.35% | C | 59,833 | -0.0341 | $1.00*10^{-8}$ | 131,189 | -0.0175 | $1.28*10^{-5}$ |
| **rs2044387** | 8 | 8,907,950 | *ERI1* | 43.03% | A | 56,089 | 0.0348 | $8.45*10^{-9}$ | 127,578 | 0.0245 | $8.65*10^{-10}$ |
| **rs12546366** | 8 | 10,802,146 | *XKR6* | 45.53% | T | 59,177 | 0.0314 | $7.15*10^{-8}$ | 130,114 | 0.0213 | $6.22*10^{-8}$ |
| **rs1192926** | 11 | 69,476,293 | *ORAOV1* | 36.79% | C | 58,213 | -0.0329 | $6.49*10^{-8}$ | 130,480 | -0.0160 | $8.40*10^{-5}$ |
| **rs8057620** | 16 | 69,884,619 | *WWP2* | 46.12% | T | 58,056 | 0.0350 | $1.64*10^{-9}$ | 130,781 | 0.0245 | $4.14*10^{-10}$ |
| **Males** | | | | | | | | | | | |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **rs56372408** | 2 | 172,416,803 | *CYBRD1* | 24.65% | G | 53,100 | -0.0426 | $2.34*10^{-9}$ | 109,222 | -0.0213 | $1.58*10^{-5}$ |
| **rs6889311** | 5 | 127,431,285 | *SLC12A2* | 24.51% | A | 53,647 | 0.0600 | $2.41*10^{-17}$ | 109,664 | 0.0545 | $5.37*10^{-28}$ |
| **rs11187838** | 10 | 96,026,184 | *PLCE1* | 43.19% | A | 53,752 | 0.0417 | $1.14*10^{-11}$ | 109,788 | 0.0330 | $1.66*10^{-14}$ |
| **rs62066707** | 17 | 43,273,992 | *FMNL1* | 32.75% | A | 50,164 | 0.0369 | $3.65*10^{-8}$ | 105,508 | 0.0176 | $1.57*10^{-4}$ |

Leading SNPs for each locus is presented along with the chromosome and basepair position (hg19/build 37). MAF: minor allele frequency. A1: effect allele. Results from association tests in the discovery ($_1$) and replication cohort ($_2$) are presented. *N*: number of participants with non-missing data for each SNP. $\beta$: estimated effect size (change in rank-transformed AFR) per allele. *p*: p-values from Z-tests for deviance of $\beta$ from zero.

**Table 5. Novel body adiposity loci associated with the proportion of adipose tissue distributed to the legs (LFR)**

| | | | | | | Discovery | | | Replication | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Leading SNP** | **Chr** | **Position (bp)** | **Proximal gene(s)** | **MAF** | **A1** | $N_1$ | $\beta_1$ | $p_1$ | $N_2$ | $\beta_2$ | $p_2$ |
| **Sex-combined** | | | | | | | | | | | |
| **rs180921974** | 1 | 155,268,131 | *PKLR* | 2.3% | G | 113,831 | 0.0868 | $8.44*10^{-10}$ | 241,384 | 0.0908 | $8.31*10^{-21}$ |
| **rs115912456** | 5 | 82,815,158 | *VCAN* | 4.1% | G | 114,007 | 0.0575 | $5.51*10^{-8}$ | 241,624 | 0.0459 | $2.34*10^{-10}$ |
| **rs35344761** | 9 | 78,510,823 | *PCSK5* | 12.0% | A | 111,867 | 0.0359 | $3.89*10^{-8}$ | 238,781 | 0.0336 | $5.54*10^{-14}$ |
| **rs3780327** | 9 | 129,945,847 | *RALGPS1* | 22.1% | A | 111,593 | -0.0310 | $1.40*10^{-9}$ | 237,144 | -0.0165 | $2.42*10^{-6}$ |
| ***rs35826789*** | 11 | 66,913,469 | *KDM2A* | 8.8% | A | 113,976 | 0.0506 | $1.08*10^{-11}$ | 239,941 | 0.0331 | $9.86*10^{-11}$ |
| **rs71420186** | 14 | 50,960,918 | *MAP4K5* | 6.7% | A | 113,034 | -0.0496 | $6.09*10^{-9}$ | 240,081 | -0.0383 | $3.95*10^{-11}$ |
| **rs10153134** | 16 | 90,091,099 | *AFG3L1P* | 36.3% | C | 112,214 | -0.0235 | $8.81*10^{-8}$ | 236,518 | -0.0119 | $9.08*10^{-5}$ |
| **rs2071167** | 17 | 42,287,519 | *UBTF* | 23.4% | T | 114,017 | -0.0280 | $1.37*10^{-8}$ | 236,074 | -0.0132 | $1.27*10^{-4}$ |
| **rs10402308** | 19 | 19,657,500 | *CILP2* | 17.7% | A | 112,922 | 0.0352 | $1.96*10^{-10}$ | 240,376 | 0.0144 | $1.36*10^{-4}$ |
| **Females** | | | | | | | | | | | |
| **rs2273368** | 1 | 113,063,771 | *WNT2B* | 19.6% | T | 59,162 | 0.0397 | $6.70*10^{-8}$ | 130,911 | 0.0297 | $1.36*10^{-9}$ |
| ***rs56310695*** | 4 | 73,535,246 | *ADAMTS3* | 5.8% | A | 58,573 | -0.0674 | $5.29*10^{-8}$ | 129,209 | -0.0573 | $2.35*10^{-12}$ |
| **rs1317415** | 5 | 157,952,404 | *EBF1* | 30.3% | C | 59,403 | -0.0343 | $4.67*10^{-8}$ | 131,265 | -0.0179 | $2.42*10^{-5}$ |
| **rs888762** | 5 | 178,547,313 | *ADAMTS2* | 33.1% | C | 58,318 | -0.0333 | $9.36*10^{-8}$ | 129,236 | -0.0233 | $2.30*10^{-8}$ |
| **rs41271299** | 6 | 19,839,415 | *ID4* | 5.1% | T | 60,196 | -0.0918 | $2.55*10^{-12}$ | 131,669 | -0.0770 | $2.18*10^{-18}$ |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **rs3823974** | 7 | 20,442,796 | *ITGB8* | 41.3% | C | 58,656 | 0.0339 | $1.11*10^{-8}$ | 129,457 | 0.0237 | $3.07*10^{-9}$ |
| **rs10962638** | 9 | 16,846,111 | *BNC2* | 14.8% | A | 57,103 | 0.0473 | $2.80*10^{-8}$ | 127,774 | 0.0212 | $1.27*10^{-4}$ |
| **rs35344761** | 9 | 78,510,823 | *PCSK5* | 12.0% | A | 59,058 | 0.0503 | $2.32*10^{-8}$ | 130,107 | 0.0560 | $2.09´*10^{-20}$ |
| ***rs68049170*** | 10 | 72,432,047 | *ADAMTS14* | 27.3% | A | 58,784 | -0.0367 | $1.99*10^{-8}$ | 129,806 | -0.0233 | $1.20*10^{-7}$ |
| **rs12785906** | 11 | 66,951,966 | *KDM2A* | 5.6% | C | 59,066 | 0.0901 | $1.34*10^{-12}$ | 129,556 | 0.0604 | $9.62*10^{-13}$ |
| **rs1613835** | 12 | 51,205,066 | *ATF1* | 31.3% | T | 59,668 | -0.0360 | $8.41*10^{-9}$ | 131,432 | -0.0213 | $4.08*10^{-7}$ |
| **rs6489111** | 12 | 123,051,018 | *KNTC1* | 35.2% | G | 56,631 | -0.0371 | $3.27*10^{-9}$ | 131,045 | -0.0188 | $4.25*10^{-6}$ |
| **rs1550436** | 15 | 74,221,157 | *LOXL1* | 46.7% | T | 58,904 | 0.0316 | $6.21*10^{-8}$ | 130,110 | 0.0284 | $4.44*10^{-13}$ |
| **rs72755233** | 15 | 100,692,953 | *ADAMTS17* | 11.3% | A | 60,196 | 0.0636 | $2.83*10^{-12}$ | 131,669 | 0.0646 | $7.93*10^{-26}$ |
| **Males** | | | | | | | | | | | |
| **rs1993878** | 5 | 127,476,971 | *SLC12A2* | 24.5 | C | 53,403 | -0.0523 | $1.82*10^{-13}$ | 109,324 | -0.0334 | $2.07*10^{-11}$ |

Leading SNPs for each locus is presented along with the chromosome and basepair position (hg19/build 37). MAF: minor allele frequency. A1: effect allele. Results from association tests in the discovery ($_1$) and replication cohort ($_2$) are presented. *N*: number of participants with non-missing data for each SNP. $\beta$: estimated effect size (change in rank-transformed LFR) per allele. *p*: p-values from Z-tests for deviance of $\beta$ from zero.

**Table 6. Novel body adiposity loci associated with the proportion of adipose tissue distributed to the trunk (TFR)**

| | | | | | | Discovery | | | Replication | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Leading SNP** | **Chr** | **Position (bp)** | **Proximal gene(s)** | **MAF** | **A1** | $N_1$ | $\beta_1$ | $p_1$ | $N_2$ | $\beta_2$ | $p_2$ |
| **Sex-combined** | | | | | | | | | | | |
| **rs4846204** | 1 | 10,308,958 | *KIF1B* | 12.7% | T | 113,962 | 0.0353 | $2.16*10^{-8}$ | 241,509 | 0.0259 | $2.08*10^{-9}$ |
| **rs180921974** | 1 | 155,268,131 | *KRTCAP2* | 2.3% | G | 113,776 | -0.1017 | $6.66*10^{-13}$ | 241,256 | -0.0983 | $4.68*10^{-24}$ |
| **rs17511102** | 2 | 37,960,613 | *CDC42EP3* | 9.1% | T | 113,962 | 0.0421 | $8.13*10^{-9}$ | 241,509 | 0.0198 | $7.87*10^{-5}$ |
| **rs4521268** | 3 | 49,137,904 | *QARS* | 33.2% | G | 113,030 | -0.0247 | $3.26*10^{-8}$ | 240,924 | -0.0212 | $4.06*10^{-12}$ |
| **rs1986599** | 3 | 50,034,637 | *RBM6* | 12.3% | G | 109,220 | -0.0374 | $3.25*10^{-8}$ | 239,307 | -0.0210 | $2.14*10^{-6}$ |
| **rs4694510** | 4 | 73,546,983 | *ADAMTS3* | 6.2% | A | 112,881 | 0.0550 | $3.38*10^{-10}$ | 240,133 | 0.0480 | $1.21*10^{-15}$ |
| **rs115912456** | 5 | 82,815,158 | *VCAN* | 4.1% | G | 113,952 | -0.0694 | $5.32*10^{-11}$ | 241,496 | -0.0519 | $7.77*10^{-13}$ |
| **rs888762** | 5 | 178,547,313 | *ADAMTS2* | 33.1% | C | 110,409 | 0.0265 | $4.65*10^{-9}$ | 237,012 | 0.0198 | $1.41*10^{-10}$ |
| **rs41271299** | 6 | 19,839,415 | *ID4* | 5.2% | T | 113,962 | 0.0547 | $7.25*10^{-9}$ | 241,509 | 0.0469 | $5.10*10^{-13}$ |
| **rs75848127** | 7 | 148,642,553 | *GHET1* | 17.0% | A | 112811 | 0.0315 | $1.86*10^{-8}$ | 240,960 | 0.0205 | $8.16*10^{-8}$ |
| **rs35650604** | 9 | 78,515,195 | *PCSK5* | 13.4% | G | 113,962 | -0.0345 | $1.83*10^{-8}$ | 241,509 | -0.0342 | $7.07*10^{-16}$ |
| **rs12790261** | 11 | 66,988,048 | *KDM2A* | 8.4% | A | 113,962 | -0.0614 | $1.19*10^{-15}$ | 241,509 | -0.0420 | $7.78*10^{-16}$ |
| **rs12905253** | 15 | 74,232,437 | *LOXL1* | 47.0% | A | 113,218 | -0.0232 | $3.37*10^{-8}$ | 240,548 | -0.0223 | $1.12*10^{-14}$ |
| **rs72755233** | 15 | 100,692,953 | *ADAMTS17* | 11.3% | A | 113,962 | -0.0436 | $4.35*10^{-11}$ | 241,509 | -0.0533 | $7.98*10^{-32}$ |
| **rs2074188** | 17 | 45,888,251 | *OSBPL7* | 47.5% | G | 113,055 | 0.0232 | $3.60*10^{-8}$ | 239,440 | 0.0108 | $1.82*10^{-4}$ |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **rs62621197** | 19 | 8,670,147 | *ADAMTS10* | 2.9% | T | 110599 | -0.0690 | $5.65*10^{-8}$ | 239,955 | -0.0968 | $2.77*10^{-31}$ |
| **rs10402308** | 19 | 19,657,500 | *CILP2* | 17.7% | A | 112,869 | -0.0313 | $1.52*10^{-8}$ | 240,251 | -0.0148 | $8.67*10^{-5}$ |
| **Females*** | | | | | | | | | | | |
| **rs2273368** | 1 | 113,063,771 | *WNT2B* | 19.6% | T | 59,132 | -0.0425 | $6.99*10^{-9}$ | 130,850 | -0.0331 | $1.51*10^{-11}$ |
| **rs10916174** | 1 | 227,804,041 | *ZNF678* | 15.8% | A | 59,477 | -0.0433 | $6.20*10^{-8}$ | 130,587 | -0.0367 | $9.34*10^{-12}$ |
| **rs4521268** | 3 | 49,137,904 | *QARS* | 33.3% | G | 59,686 | -0.0327 | $9.25*10^{-8}$ | 131,293 | -0.0245 | $2.81*10^{-9}$ |
| **rs2241069** | 4 | 8,602,798 | *CPZ* | 46.3% | G | 58,819 | 0.0332 | $1.17*10^{-8}$ | 129,272 | 0.0197 | $6.04*10^{-7}$ |
| **rs73825843** | 4 | 73,535,052 | *ADAMTS3* | 6.3% | T | 59,853 | 0.0796 | $3.58*10^{-11}$ | 131,539 | 0.0720 | $2.18*10^{-19}$ |
| **rs888762** | 5 | 178,547,313 | *ADAMTS2* | 33.1% | C | 58,287 | 0.0366 | $4.04*10^{-9}$ | 129,176 | 0.0267 | $1.49*10^{-10}$ |
| **rs41271299** | 6 | 19,839,415 | *ID4* | 5.1% | T | 60,165 | 0.1065 | $4.57*10^{-16}$ | 131,608 | 0.0901 | $1.28*10^{-24}$ |
| **rs2982708** | 6 | 152,356,220 | *ESR1* | 27.5% | C | 59,635 | 0.0354 | $5.22*10^{-8}$ | 131,600 | 0.0250 | $1.21*10^{-8}$ |
| **rs3823974** | 7 | 20,442,796 | *ITGB8* | 41.3% | C | 58,625 | -0.0390 | $4.48*10^{-11}$ | 129,396 | -0.0279 | $2.80*10^{-12}$ |
| **rs4733727** | 8 | 130,731,484 | *GSDMC* | 48.8% | T | 58,394 | -0.0355 | $1.19*10^{-9}$ | 129,730 | 0.0198 | $4.50*10^{-7}$ |
| **rs76937529** | 9 | 78,505,692 | *PCSK5* | 11.5% | T | 56,888 | -0.0547 | $6.57*10^{-9}$ | 126,906 | -0.0611 | $1.21*10^{-22}$ |
| **rs7039458** | 9 | 86,639,999 | *RMI1* | 24.9% | G | 59,585 | 0.0379 | $1.55*10^{-8}$ | 130,985 | 0.0287 | $2.33*10^{-10}$ |
| **rs68049170** | 10 | 72,432,047 | *ADAMTS14* | 27.4% | A | 58,753 | 0.0380 | $6.00*10^{-9}$ | 129,746 | -0.0198 | $2.03*10^{-6}$ |
| **rs12790261** | 11 | 66,988,048 | *KDM2A* | 8.4% | A | 60,165 | -0.0910 | $7.38*10^{-18}$ | 131,608 | -0.0688 | $1.29*10^{-22}$ |
| **rs11614785** | 12 | 50,880,422 | *LARP4* | 33.8% | G | 58,774 | 0.0398 | $9.45*10^{-11}$ | 127,523 | 0.0265 | $2.36*10^{-10}$ |
| **rs6492538** | 13 | 91,993,746 | *MIR17-92HG/GPC5* | 22.1% | A | 59,176 | 0.0392 | $2.39*10^{-8}$ | 129,162 | 0.0212 | $7.45*10^{-6}$ |
| **rs67089639** | 14 | 92,526,240 | *ATXN3* | 24.5% | C | 59,789 | -0.0363 | $7.88*10^{-8}$ | 131,418 | -0.0224 | $7.93*10^{-7}$ |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **rs35874463** | 15 | 67,457,698 | *SMAD3* | 5.8% | G | 60,165 | 0.0687 | $2.72*10^{-8}$ | 131,608 | 0.0647 | $1.20*10^{-14}$ |
| **rs12905253** | 15 | 74,232,437 | *LOXL1* | 47.0% | A | 59,778 | -0.0394 | $1.05E*10^{-11}$ | 131,113 | -0.0322 | $1.84*10^{-16}$ |
| **rs72755233** | 15 | 100,692,953 | *ADAMTS17* | 11.3% | A | 60,165 | -0.0744 | $2.84*10^{-16}$ | 131,608 | -0.0800 | $9.86*10^{-39}$ |
| **rs28394864** | 17 | 47,450,775 | *ZNF652* | 46.2% | A | 60,057 | -0.0357 | $6.33*10^{-10}$ | 129,628 | -0.0237 | $1.95*10^{-9}$ |
| **rs7236575** | 18 | 46,653,380 | *DYM* | 13.7% | A | 59,933 | -0.0463 | $3.54*10^{-8}$ | 131,355 | -0.0220 | $1.13*10^{-4}$ |
| **rs62621197** | 19 | 8,670,147 | *ADAMTS10* | 3.2% | T | 58,413 | -0.1054 | $1.57*10^{-9}$ | 129,021 | -0.1472 | $3.32*10^{-39}$ |
| **rs6038571** | 20 | 6,634,566 | *BMP2* | 48.1% | A | 60,165 | 0.0308 | $9.65*10^{-8}$ | 131,608 | 0.0216 | $2.79*10^{-8}$ |

Leading SNPs for each locus is presented along with the chromosome and position (hg19/build 37). MAF: minor allele frequency. A1: effect allele. Results from association tests in the discovery ($_1$) and replication cohort ($_2$) are presented. *N*: number of participants with non-missing data for each SNP. $\beta$: estimated effect size (change in rank-transformed LFR) per allele. *p*: p-values from Z-tests for deviance of $\beta$ from zero.

*Sex specific effects and SNP x sex interactions*

Clear contrasts could be seen between males and females with regards to the number of associations. As many as 33 loci were associated with LFR in females, but only one in males. For TFR, 49 loci were identified in in females while only one locus in males. For AFR, 14 loci were associated in females compared to seven in males. Only two AFR-associated loci, near the *MC4R* and *FTO*, were observed in both males and females. Altogether, 64 loci appeared to be sex-specific (significant in females but not in males or vice versa), of which 29 showed the same pattern in multiple phenotypes. We further tested all replicated SNPs (N=159) for interaction with sex. Similarly to the GWAS, we first tested for interaction in the discovery cohort using linear regression modeling and validated our findings in the replication cohort using the same covariates.

A total of 58 SNP-sex interactions were observed (Bonferroni adjusted p-values < 0.05/159) in the discovery cohort, of which 43 replicated (Bonferroni adjusted p-values < 0.05/58). However, for 56 SNPs the P-value for the interaction was nominally significant in the replication cohort. The highest number of SNPs with sex-interactions was seen for TFR (N=24), followed by LFR (N=15), and AFR (N=4) (Table 7, S2 Supplementary Data).

**Table 7. Interactions between leading SNPs and sex.** Leading SNPs for each locus is presented along with the chromosome and position (hg19/build 37). $p_{int}$: results from tests of deviance from zero for the interaction term (see mehods). Results from association tests in the discovery and replication cohort are presented. *N*: number of participants with non-missing data for each SNP. $\beta$: estimated effect size (change in rank-transformed LFR) per allele. *p*: p-values from Z-tests for deviance of $\beta$ from zero.

| | | | | | | Discovery | | | | Replication | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | SNP-sex-interaction | | Females | | Males | | Females | | Males | |
| **AFR** | **Chr** | **Position (bp)** | **Proximal gene** | $p_{int}$ **discovery** | $p_{int}$ **replication** | $\beta$ | *p* | $\beta$ | *p* | $\beta$ | *p* | $\beta$ | *p* |
| **rs754537** | 2 | 25,176,277 | *DNAJC27* | $2.66*10^{-7}$ | $2.24*10^{-16}$ | -0.0397 | $6.53*10^{-12}$ | 0.0001 | 0.98 | -0.0439 | $1.79*10^{-29}$ | 0.0005 | 0.91 |
| **rs143384** | 20 | 34,025,756 | *GDF5* | $6.70*10^{-10}$ | $4.71*10^{-11}$ | -0.0137 | $1.96*10^{-2}$ | 0.0354 | $1.00*10^{-8}$ | -0.0120 | $2.59*10^{-3}$ | 0.0247 | $1.25*10^{-8}$ |
| **rs6889311** | 5 | 127,431,285 | *SLC12A2* | $4.98*10^{-7}$ | $7.28*10^{-6}$ | 0.0059 | 0.38 | 0.0602 | $2.51*10^{-17}$ | 0.0207 | $5.42*10^{-6}$ | 0.0545 | $5.38*10^{-28}$ |
| **rs539515** | 1 | 177,889,025 | *FAIM2* | $1.34*10^{-4}$ | $3.27*10^{-4}$ | 0.0547 | $1.41*10^{-14}$ | 0.0182 | $1.63*10^{-2}$ | 0.0486 | $6.02*10^{-24}$ | 0.0257 | $9.69*10^{-7}$ |
| **rs8057620** | 16 | 69,884,619 | *WWP2* | $6.15*10^{-8}$ | $1.27*10^{-3}$ | 0.0350 | $1.64*10^{-9}$ | -0.0081 | 0.19 | 0.0245 | $4.33*10^{-10}$ | 0.0095 | $2.66*10^{-2}$ |
| **rs7562173** | 2 | 46,976,217 | *SOCS5* | $1.41*10^{-6}$ | $1.56*10^{-3}$ | -0.0341 | $1.00*10^{-8}$ | 0.0077 | 0.22 | -0.0175 | $1.31*10^{-5}$ | 0.0018 | 0.67 |
| **rs62066707** | 17 | 43,273,992 | *FMNL1* | $1.88*10^{-4}$ | $8.96*10^{-3}$ | 0.0025 | 0.70 | 0.0369 | $3.65*10^{-8}$ | 0.0026 | 0.53 | 0.0176 | $1.57*10^{-4}$ |
| **rs2044387** | 8 | 8,907,950 | *ERI1* | $5.99*10^{-5}$ | $2.30*10^{-2}$ | 0.0348 | $8.45*10^{-9}$ | -0.0003 | 1.00 | 0.0245 | $8.88*10^{-10}$ | 0.0100 | $2.32*10^{-2}$ |
| **rs56372408** | 2 | 172,416,803 | *CYBRD1* | $1.69*10^{-4}$ | 0.61 | -0.0026 | 0.70 | -0.0426 | $2.34*10^{-9}$ | -0.0157 | $5.32*10^{-4}$ | -0.0213 | $1.58*10^{-5}$ |
| **LFR** | | | | | | | | | | | | | |
| **rs7162542** | 15 | 84,514,290 | *ADAMTSL3* | $5.46*10^{-16}$ | $1.97*10^{-19}$ | 0.0529 | $8.36*10^{-20}$ | -0.0159 | $9.56*10^{-3}$ | 0.0458 | $1.52*10^{-31}$ | -0.0068 | 0.11 |
| **rs2871960** | 3 | 141,121,814 | *ZBTB38* | $1.98*10^{-8}$ | $7.94*10^{-17}$ | -0.0517 | $5.39*10^{-19}$ | -0.0035 | 0.57 | -0.0481 | $2.04*10^{-34}$ | 0.0011 | 0.80 |
| **rs9853018** | 3 | 141,101,961 | *ZBTB38* | $1.26*10^{-8}$ | $2.76*10^{-16}$ | -0.0518 | $4.76*10^{-19}$ | -0.0029 | 0.64 | -0.0477 | $5.41*10^{-34}$ | 0.0006 | 0.89 |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **rs10946808** | 6 | 26,233,387 | *HIST1H1D* | $2.81*10^{-9}$ | $4.66*10^{-14}$ | 0.0558 | $8.79*10^{-18}$ | 0.0014 | 0.84 | 0.0526 | $2.65*10^{-33}$ | 0.0045 | 0.35 |
| **rs4800148** | 18 | 20,724,328 | *RBBP8, CABLES1, C18orf45* | $2.61*10^{-7}$ | $2.03*10^{-10}$ | 0.0403 | $8.32*10^{-9}$ | -0.0113 | 0.13 | 0.0349 | $9.27*10^{-14}$ | -0.0090 | $8.15*10^{-2}$ |
| **rs3817428** | 15 | 89,415,247 | *ACAN* | $1.49*10^{-6}$ | $3.54*10^{-10}$ | 0.0527 | $4.68*10^{-16}$ | 0.0074 | 0.29 | 0.0451 | $1.39*10^{-24}$ | 0.0046 | 0.35 |
| **rs41271299** | 6 | 19,839,415 | *ID4* | $3.15*10^{-6}$ | $4.38*10^{-9}$ | -0.0918 | $2.55*10^{-12}$ | 0.0009 | 0.95 | -0.0770 | $2.05*10^{-18}$ | 0.0026 | 0.79 |
| **rs72755233** | 15 | 100,692,953 | *ADAMTS17* | $3.40*10^{-6}$ | $3.26*10^{-8}$ | 0.0636 | $2.83*10^{-12}$ | 0.0020 | 0.84 | 0.0646 | $7.91*10^{-26}$ | 0.0152 | $2.38*10^{-2}$ |
| **rs2273368** | 1 | 113,063,771 | *WNT2B* | $6.36*10^{-5}$ | $1.41*10^{-7}$ | 0.0397 | $6.70*10^{-8}$ | -0.0024 | 0.75 | 0.0297 | $1.45*10^{-9}$ | -0.0083 | 0.12 |
| **rs3791679** | 2 | 56,096,892 | *EFEMP1* | $1.83*10^{-5}$ | $2.49*10^{-7}$ | 0.0506 | $2.24*10^{-13}$ | 0.0079 | 0.28 | 0.0519 | $8.74*10^{-29}$ | 0.0178 | $5.08*10^{-4}$ |
| **rs1415287** | 1 | 219,742,537 | *LYPLAL1, SLC30A10, ZC3H11B* | $8.25*10^{-5}$ | $1.18*10^{-6}$ | -0.0423 | $3.04*10^{-11}$ | -0.0065 | 0.34 | -0.0397 | $1.21*10^{-20}$ | -0.0103 | $2.68*10^{-2}$ |
| **rs3791675** | 2 | 56,111,309 | *EFEMP1* | $6.45*10^{-6}$ | $2.98*10^{-6}$ | 0.0503 | $1.63*10^{-13}$ | 0.0058 | 0.42 | 0.0488 | $2.99*10^{-26}$ | 0.0184 | $2.73*10^{-4}$ |
| **rs1613835** | 12 | 51,205,066 | *BCDIN3D, FAIM2* | $1.11*10^{-5}$ | $4.13*10^{-6}$ | -0.0360 | $8.41*10^{-9}$ | 0.0043 | 0.52 | -0.0213 | $4.04*10^{-7}$ | 0.0069 | 0.14 |
| **rs12785906** | 11 | 66,951,966 | *KDM2A* | $5.41*10^{-5}$ | $3.85*10^{-5}$ | 0.0901 | $1.34*10^{-12}$ | 0.0137 | 0.31 | 0.0603 | $1.03*10^{-12}$ | 0.0063 | 5.1 |
| **rs994014** | 4 | 82,165,790 | *intergenic* | $3.25*10^{-6}$ | $6.37*10^{-5}$ | -0.0382 | $1.02*10^{-9}$ | 0.0037 | 0.58 | -0.0303 | $9.04*10^{-13}$ | -0.0061 | 0.19 |
| **rs3823974** | 7 | 20,442,796 | *TWISTNB* | $1.72*10^{-4}$ | $1.57*10^{-3}$ | 0.0339 | $1.11*10^{-8}$ | 0.0012 | 0.85 | 0.0236 | $3.25*10^{-9}$ | 0.0053 | 0.22 |
| **rs1993878** | 5 | 127,476,971 | *SLCA2* | $2.06*10^{-8}$ | $1.58*10^{-3}$ | 0.0028 | 0.67 | -0.0523 | $1.82*10^{-13}$ | -0.0122 | $7.40*10^{-3}$ | -0.0334 | $2.07*10^{-11}$ |
| **rs68049170** | 10 | 72,432,047 | *ADAMTS14* | $3.69*10^{-6}$ | $2.51*10^{-3}$ | -0.0367 | $1.99*10^{-8}$ | 0.0081 | 0.24 | -0.0233 | $1.20*10^{-7}$ | -0.0041 | 0.39 |
| **rs1317415** | 5 | 157,952,404 | *EBF1ZZV* | $4.69*10^{-5}$ | 0.22 | -0.0343 | $4.67*10^{-8}$ | 0.0018 | 0.78 | -0.0179 | $2.48*10^{-5}$ | -0.0105 | $2.34*10^{-2}$ |
| | | | | | | | | | | | | | |
| **TFR** | | | | | | | | | | | | | |
| **rs2871960** | 3 | 141,121,814 | *ZBTB38* | $1.39*10^{-11}$ | $2.68*10^{-27}$ | 0.0549 | $2.87*10^{-21}$ | -0.0010 | 0.87 | 0.0558 | $7.05*10^{-46}$ | -0.0064 | 0.13 |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **rs6785012** | 3 | 141,109,348 | *ZBTB38* | $2.48*10^{-11}$ | $7.39*10^{-27}$ | 0.0549 | $4.81*10^{-21}$ | -0.0006 | 0.92 | 0.0563 | $2.06*10^{-46}$ | -0.0056 | 0.19 |
| **rs11856122** | 15 | 84,576,348 | *ADAMTSL 3* | $2.28*10^{-12}$ | $3.58*10^{-20}$ | -0.0579 | $1.06*10^{-23}$ | -0.0006 | 0.93 | -0.0579 | $1.01*10^{-49}$ | -0.0059 | 0.16 |
| **rs11259934** | 15 | 84,580,171 | *ADAMTSL 3* | $1.33*10^{-11}$ | $3.79*10^{-20}$ | -0.0581 | $1.36*10^{-23}$ | -0.0026 | 0.68 | -0.0579 | $1.21*10^{-49}$ | -0.0059 | 0.17 |
| **rs62346126** | 4 | 145,560,166 | *HHIP* | $2.05*10^{-4}$ | $1.61*10^{-13}$ | -0.0515 | $3.89*10^{-12}$ | -0.0123 | 0.12 | -0.0612 | $2.56*10^{-34}$ | -0.0074 | 0.18 |
| **rs9358913** | 6 | 26,239,404 | *HIST1H1D* | $5.32*10^{-11}$ | $3.35*10^{-13}$ | -0.0569 | $5.44*10^{-18}$ | 0.0049 | 0.48 | -0.0529 | $1.23*10^{-32}$ | -0.0061 | 0.21 |
| **rs41271299** | 6 | 19,839,415 | *ID4* | $3.86*10^{-8}$ | $6.82*10^{-13}$ | 0.1065 | $4.57*10^{-16}$ | -0.0008 | 0.95 | 0.0902 | $1.23*10^{-24}$ | -0.0045 | 0.64 |
| **rs143384** | 20 | 34,025,756 | *GDF5* | $4.42*10^{-5}$ | $6.67*10^{-12}$ | 0.0498 | $2.78*10^{-17}$ | 0.0172 | 0.54 | 0.0475 | $5.68*10^{-33}$ | 0.0086 | $4.72*10^{-2}$ |
| **rs3791679** | 2 | 56,096,892 | *EFEMP1* | $4.69*10^{-8}$ | $1.14*10^{-11}$ | -0.0568 | $1.74*10^{-16}$ | -0.0024 | 0.74 | -0.0601 | $4.26*10^{-38}$ | -0.0150 | $3.43*10^{-3}$ |
| **rs72755233** | 15 | 100,692,953 | *ADAMTS17* | $6.25*10^{-7}$ | $1.70*10^{-11}$ | -0.0744 | $2.84*10^{-16}$ | -0.0096 | 0.32 | -0.0800 | $9.85*10^{-39}$ | -0.0215 | $1.42*10^{-3}$ |
| **rs4800148** | 18 | 20,724,328 | *RBBP8, CABLES1, C18orf45* | $1.52*10^{-8}$ | $1.79*10^{-11}$ | -0.0442 | $2.72*10^{-10}$ | 0.0117 | 0.11 | -0.0392 | $6.70*10^{-17}$ | 0.0060 | 0.25 |
| **rs3817428** | 15 | 89,415,247 | *ACAN* | $2.11*10^{-7}$ | $2.83*10^{-9}$ | -0.0608 | $7.49*10^{-21}$ | -0.0120 | $8.17*10^{-2}$ | -0.0513 | $2.46*10^{-31}$ | -0.0126 | $9.25*10^{-3}$ |
| **rs12790261** | 11 | 66,988,048 | *KDM2A* | $8.11*10^{-5}$ | $1.25*10^{-8}$ | -0.0910 | $7.38*10^{-18}$ | -0.0290 | $9.20*10^{-3}$ | -0.0688 | $1.35*10^{-22}$ | -0.0097 | 0.21 |
| **rs2273368** | 1 | 113,063,771 | *WNT2B* | $6.20*10^{-5}$ | $1.54*10^{-8}$ | -0.0425 | $6.99*10^{-9}$ | -0.0010 | 0.90 | -0.0331 | $1.57*10^{-11}$ | 0.0079 | 0.15 |
| **rs798491** | 7 | 2,800,521 | *GNA12* | $1.05*10^{-4}$ | $1.03*10^{-7}$ | -0.0504 | $1.26*10^{-15}$ | -0.0164 | $1.40*10^{-2}$ | -0.0483 | $5.79*10^{-30}$ | -0.0161 | $5.63*10^{-4}$ |
| **rs10916174** | 1 | 227,804,041 | *ZNF678* | $1.07*10^{-4}$ | $1.55*10^{-7}$ | -0.0433 | $6.20*10^{-8}$ | 0.0014 | 0.87 | -0.0367 | $9.64*10^{-12}$ | 0.0065 | 0.27 |
| **rs11614785** | 12 | 50,880,422 | *BCDIN3D, FAIM2* | $5.15*10^{-7}$ | $5.75*10^{-7}$ | 0.0398 | $9.45*10^{-11}$ | -0.0034 | 0.60 | 0.0265 | $2.32*10^{-10}$ | -0.0024 | 0.61 |
| **rs994014** | 4 | 82,165,790 | *PRKG2* | $7.49*10^{-7}$ | $4.64*10^{-6}$ | 0.0416 | $2.72*10^{-11}$ | -0.0027 | 0.68 | 0.0356 | $4.13*10^{-17}$ | 0.0078 | $9.46*10^{-2}$ |
| **rs6038571** | 20 | 6,634,566 | *BMP2* | $3.92*10^{-6}$ | $5.39*10^{-6}$ | 0.0308 | $9.65*10^{-8}$ | -0.0093 | 0.13 | 0.0216 | $2.79*10^{-8}$ | -0.0055 | 0.20 |
| **rs11049361** | 12 | 28,284,841 | *CCDC91* | $2.14*10^{-4}$ | $1.27*10^{-5}$ | -0.0428 | $2.13*10^{-11}$ | -0.0092 | 0.17 | -0.0298 | $5.57*10^{-12}$ | -0.0037 | 0.43 |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **rs991967** | 1 | 218,615,451 | *TGFB2* | $4.99*10^{-5}$ | $1.57*10^{-5}$ | 0.0426 | $2.94*10^{-11}$ | 0.0063 | 0.36 | 0.0244 | $1.64*10^{-8}$ | -0.0016 | 0.74 |
| **rs12905253** | 15 | 74,232,437 | *LOXL1* | $1.51*10^{-5}$ | $7.77*10^{-5}$ | -0.0394 | $1.05*10^{-11}$ | -0.0051 | 0.41 | -0.0322 | $1.87*10^{-16}$ | -0.0105 | $1.40*10^{-2}$ |
| **rs3823974** | 7 | 20,442,796 | *TWISTNB* | $1.02*10^{-5}$ | $1.37*10^{-4}$ | -0.0390 | $4.48*10^{-11}$ | -0.0005 | 0.94 | -0.0279 | $2.91*10^{-12}$ | -0.0058 | 0.19 |
| **rs6492538** | 13 | 91,993,746 | *GPC5* | $5.03*10^{-5}$ | $3.09*10^{-4}$ | 0.0392 | $2.39*10^{-8}$ | -0.0016 | 0.83 | 0.0212 | $7.33*10^{-6}$ | -0.0028 | 0.58 |
| **rs12654493** | 5 | 176,535,209 | *FGFR4* | $1.11*10^{-6}$ | $1.06*10^{-3}$ | 0.0412 | $3.64*10^{-9}$ | -0.0072 | 0.33 | 0.0173 | $2.23*10^{-4}$ | -0.0048 | 0.35 |
| **rs34716573** | 12 | 576,037 | *B4GALNT 3* | $2.21*10^{-4}$ | $2.63*10^{-3}$ | 0.0395 | $1.91*10^{-10}$ | 0.0070 | 0.29 | 0.0174 | $2.41*10^{-5}$ | -0.0015 | 0.75 |
| **rs4733727** | 8 | 130,731,484 | *GSDMC* | $8.38*10^{-5}$ | $6.91*10^{-3}$ | -0.0355 | $1.19*10^{-9}$ | -0.0024 | 0.70 | -0.0198 | $4.46*10^{-7}$ | -0.0033 | 0.44 |
| **rs2982708** | 6 | 152,356,220 | *ESR1* | $4.51*10^{-6}$ | $1.69*10^{-2}$ | 0.0354 | $5.22*10^{-8}$ | -0.0062 | 0.37 | 0.0250 | $1.16*10^{-8}$ | 0.0093 | $5.23*10^{-2}$ |
| **rs68049170** | 10 | 72,432,047 | *ADAMTS1 4* | $4.32*10^{-5}$ | $1.71*10^{-2}$ | 0.0380 | $6.00*10^{-9}$ | -0.0019 | 0.79 | 0.0234 | $1.07*10^{-7}$ | 0.0081 | $9.26*10^{-2}$ |
| **rs79334166** | 4 | 17,859,466 | *LCORL* | $2.16*10^{-5}$ | $2.30*10^{-2}$ | -0.0498 | $7.74*10^{-9}$ | 0.0029 | 0.75 | -0.0272 | $2.84*10^{-6}$ | -0.0088 | 0.16 |

## Discussion

We performed GWAS for body fat distribution, using segmental BIA measurements, and identified 78 loci body fat distribution to the arms (N=28), legs (N=43) and trunk (N=57). As many as 50 of the loci have not been associated with an adiposity related phenotype previously. This is probably due to the low correlation between our derived phenotypes and commonly used variables for adiposity (i.e. BMI). In contrast to previous studies, we have not addressed the total amount of fat but rather the fraction of the total body fat mass that is located in the arms (AFR), the legs (LFR), or the trunk (TFR). While most of the loci were novel, with regards to adiposity, we did see an overlap with previously reported height loci, e.g. loci near *SLC12A2, ADAMTSL3* and *BMP2* [21]. This is surprising since we did see a very limited covariance between height and all our analyzed phenotypes. These results suggest that there is a shared genetic contribution between height and body fat distribution. It follows by logic that genes that are involved in growth can potentially influence several different tissue types such as bone, adipose tissue and muscle.

Interestingly, some of the novel loci overlapped with regions that have previously been associated with lipid-related traits. We found that our lead SNP (rs10402308) at the at the *CLIP2/PBX4*-locus (associated with TFR, and LFR in women) was in strong linkage disequilibrium (LD) ($r > 0.8$) with several SNPs previously associated with triglycerides, and LDL cholesterol [22]. The lead SNP at the TFR-associated locus within *OSBPL7*, rs2074188, was also associated with higher expression of *OSBPL7* in the thyroid [23]. *OSBPL7* encodes oxysterol-binding protein-like protein 7, which is highly expressed in the thyroid, skeletal muscles, GI-tract, kidney and seminal vesicles (www.proteinatlas.org). Oxysterol-binding proteins encompass a

family of lipid-binding proteins involved in lipid trafficking, lipid metabolism and intracellular signaling [24].

Within the novel body loci we also find several genes related to estrogen and androgen signaling. Associations were observed between TFR and variants within the estrogen receptor-encoding gene, *ESR1,* in females. In addition, the TFR-, and LFR-associated SNP at the *ADAMTS17*-locus in females, rs72755233, is a missense mutation [25] which causes a potentially deleterious threonine to isoleucine substitution at position 446 of the ADAMTS17 protein (www.ensembl.org). This gene encodes a secreted metalloproteinase that is inducible in response to estrogen and inhibits breast cancer cell growth [26]. LFR-, and TFR-associated SNPs were also observed near *ID4* in women. *ID4* encodes a helix-loop-helix transcription factor that is highly expressed in the thyroid gland (www.proteinatlas.org) and also regulates androgen receptor function in the prostate [27].

AFR was the phenotype that was most highly correlated with BMI. In agreement with this, the most significant AFR loci were *FTO*, *MC4R*, *TMEM18, SEC16B* and *TFAP2B,* which have previously been associated with BMI and body adiposity-related traits [8–10,28]. Several TFR and LFR-associated loci have also previously been associated with anthropometric traits. In contrast to AFR, most of these loci did not overlap with previously known BMI-loci, but to a larger extent with waist-, and hip-associated loci such as *MTMR11, GDF5, ZBTB38, ADAMTS10, ADAMTS17*. [28]; and with height loci such as *HIST1H1D*, *ADAMTSL3*, *LIN28B* [29].

Comparing men to women showed that genetic effects differ between sexes for a large fraction of the loci. For trunk and legs, many effects were only detected (or significantly higher) in women. In agreement with this, a larger fraction of the variance in fat distribution to different compartments could also be attributed to the SNPs investigated in women, as compared to men. These results are consistent with previous GWAS that have revealed sexual dimorphisms in genetic loci for adiposity-related phenotypes, such as waist-circumference, waist-to-hip ratio, and visceral fat mass [14,30–32]. In our study we find evidence for 43 loci whose effects differed between the males and females, of which one overlapped with a locus (*LYPLAL1)* that has previously been reported to display a different effect between sexes. Our lead SNP (rs1415287) at the *LYPLAL1* locus is in strong LD with rs2820443 and rs4846567 ($R^2$=1.00 and 0.99), which have been associated with stronger effects on WHR and WHR adjusted for BMI in women [14,32].

One possible limitation of our study is the use of segmental BIA measurements for assessments of body adiposity in contrast to using more exact methods such as DXA or MRI. However, the relatively low cost and ease of use has allowed for assessment of body composition in almost the entire UK Biobank cohort, which enables us to perform highly powered association studies. The accuracy in reference to DXA of the body scanner used in UK Biobank, the Tanita BC-418, has previously been assessed in a European sample showing that total fat mass were accurately estimated. However, some biases were present depending on sex and anatomical compartment [33]. This is unlikely to affect our results as we analyzed each compartment separately and also performed sex-stratified analyses in addition to the sex-combined GWAS.

## Conclusions

GWAS of body fat distribution to the arms, legs and trunk revealed 50 novel adiposity loci. Our results indicate that the trunk and legs share genetic determinants of fat distribution, while distribution of fat to the arms is more independent. We also present evidence for 43 SNP-sex interactions that influence adipose tissue distribution. Distribution of adipose tissue between the trunk, legs and arms differ between sexes, which may be due to hormonal differences. Sex hormones primarily affect cellular proliferation, differentiation and fate by binding to and activating nuclear receptors that act as transcription factors. Consequently, the observed interactions with sex are likely to represent genes that are affected by changes in sex hormone levels.

## Methods

*UK Biobank participants*

The first release of genetic data from UK Biobank (N = 152,249) was used as a discovery cohort, and genotype data from an unrelated set of participants from the second genotype batch release (N = 326,565) as a replication cohort. Participants who self-reported as being of British descent (data field 21000) and were classified as Caucasian by principal component analysis (data field 22006) were included in the analysis. Genetic relatedness pairing was provided by the UK Biobank (Data field 22011). In total, 9,385 participants were removed due to relatedness based on kinship data (estimated genetic relationship > 0.044) and individuals with poor call rate (<95%), with high heterozygosity (Data field 22010), or with sex-errors (Data filed 22001) were also removed. After filtering, 116,138 participants were included in the discovery cohort and 246,361 in the replication cohort.

*Genotyping*

Genotyping in the discovery cohort had been performed on two custom-designed microarrays: referred to as UK BiLEVE and Axiom arrays, which genotyped 807,411 and 820,967 SNPs, respectively. Imputation had been performed using UK10K [34] and 1000 genomes phase 3 [35] as reference panels. This dataset included 73,355,667 SNPs. Prior to analysis, we filtered SNPs based on call rate (>0.01%), HWE (P-value > $10^{-20}$), MAF (> 0.0001) and imputation quality (Info >0.3) resulting in 25,472,837 SNPs in the discovery analyses. The second release of data from the UK BioBank contained genotyped and imputed data for 488,366 participants (partly overlapping with the first release). For our replication analyses, we included an independent subset that did not overlap with the discovery cohort (N = 326,565). Genotyping in this subset was performed exclusively on the UK Biobank Axiom Array. This dataset included 47,512,111 SNPs that were filtered based on HWE ($p<10^{-20}$) and MAF (>0.0001).

*Phenotypic measurements*

The phenotypes used in this study derive from impedance measurements produced by the Tanita BC418MA body composition analyzer. Participants were barefoot, wearing light indoor clothing, and measurements were taken with participants in the standing position. Height and weight were entered manually into the analyzer before measurement. The Tanita BC418MA uses eight electrodes: two for each foot and two for each hand. This allows for five impedance measurements: whole body, right leg, left leg, right arm and left arm. Body fat for the whole body and individual body parts had been calculated using a regression formula, that was derived from reference

measurements of body composition by DXA in Japanese and Western subjects, and uses weight, age, height and impedance measurements [36] as input data. Arm and leg fat masses were averaged over both limbs. Arm, leg, and trunk fat masses were then divided by the total body fat mass to obtain the ratios of fat mass for the arms, legs and trunk, i.e. what proportion of the total fat in the body is distributed to each of theses compartments. These variables were analyzed in this study and were named: arm fat ratio (AFR), leg fat ratio (LFR), and trunk fat ratio (TFR).

*Correlation*

Correlations between fat distribution ratios and anthropomorphic traits were assessed by ANOVA of linear regression model fits. BMI, waist circumference, waist-to-hip ratio and height were included as the last term in generalized linear models with adipose tissue distributions (AFR, LFR and TFR) as the response variable. The reduction in residual deviance, i.e. , the reductions in the residual sum of squares as BMI, waist circumference, waist-to-hip ratio and height is is added to the formula, is presented as percentages in Table 1.

*Associations tests*

Ratios (AFR, LFR, and TFR) were adjusted for age and age squared, normalized by rank-transformation separately in males and females using the *rntransform* function included in the GenABEL library [37] in R-Studio v1.0.143 [38]. The GWAS was performed in PLINK v1.90b3n [39] using linear regression models with AFR, LFR, and TFR as the response variables and the SNPs as predictor variables. A batch variable was used as covariate in the discovery analyses to adjust for genotyping array (Axiom and BiLEVE). We also included the first 15 principal components and sex (in

the sex-combined analyses) as covariates. We used a threshold of $p < 10^{-7}$ as threshold for significance in the discovery cohort and one leading SNPs from each loci were taken forward for replication. Individual loci were defined as a region with one or more associated SNPs. The start and stop position of a locus was the position of a SNPs where no additional associated SNP were found (upstream for start position, or downstream for stop position) within 1000kb. For each locus, the leading SNP (lowest P-value) was taken forward for replication. Since imputation and QC had been performed separately for the first and second release of genotype data from UK Biobank, some of the leading SNPs from the discovery analysis were not present (had not passed QC) in the replication cohort. In these cases, a replacement SNP from the same locus, prioritized by the P-value (the second most significant) was taken further. No replacement SNP was taken further for loci that contained only one SNP if that SNP was not available in the replication cohort.

*Conditional analyses*

GWAS were rerun for each phenotype while conditioning for the leading SNPs from the primary analyses to identify individual effects within the same loci. Conditional analyses were performed using PLINK v1.90b3n [39]. The same covariates were used as in the primary analyses.

*Interaction between SNPs and sex*

SNPs were tested for interaction with sex using linear regression modeling in RStudio v1.0.143 [38]. Models included interactions between most covariates (interactions with PCs were not considered) in order to properly control for potential confounders, in accordance with recommendations by Keller [40]. The models we used looked like:

$$P \sim \beta_1 SNP + \beta_2 sex + \beta_3 age + \beta_4 age^2 + \beta_5 (SNP * sex) + \beta_6 (SNP * age) + \beta_7 (SNP * age2) + \beta_8 (sex * age) + \beta_9 (sex * age2) + \beta_{10} (age * age2) + \sum_{i=1}^{15} \beta_{PC,i} PC_i + \varepsilon \qquad (2)$$

where *P* is the phenotype, *SNP* is the genetic variant under interest and *PC*, *sex*, *age* and $age^2$ are covariates and $\beta$ are the estimates. For the interaction analyses; AFR, LFR and TFR were rank-transformed separately for males and females without adjusting for age or $age^2$ to avoid potential interference between covariates that affect the interaction terms. The estimate for *SNP*sex* interaction terms, $\beta_5$, was tested for deviance from zero using a two-sided marginal student's t-test with the null hypothesis $H_0$: $\beta = 0$. The Bonferroni method was used to designate a cutoff for significance of p-values that was adjusted for multiple testing, p-values < $3.1*10^{-4}$ (0.05/159) were considered significant, where 159 was the number of tests performed.

*SNP heritability*

SNP heritabilities, i.e. the heritability explained by the genotyped SNPs investigated were calculated in the discovery cohort using GCTA [20]. We only included genotyped SNPs to avoid confounding due to uncertainties in the imputed data. Additional individuals were excluded from these analyses to ensure that no pairs had an estimated genetic relationship > 0.025. This was done to avoid phenotypic resemblance between relatives resulting from non-genetic effects, e.g. shared environment. Estimates of variance explained by all autosomal SNPs can be biased by genotyping errors and we therefore applied a stricter quality control than for typical GWAS analyses: SNPs with a missing call rates exceeding 5% and 1% minimum

allele frequency. After filtering, 730,616 SNPs remained for these analyses. Ten principal components (principal components) were included as covariates to capture variance due to population stratification. Sex, a batch variable for the two genotyping arrays used in the discovery cohort, as well as age were also included as covariates.

*Cross-reference with previous GWAS*

To identify novel loci, all significant SNPs in the GWAS regions were compared to the SNPs in the NHGRI-EBI catalog of published genome-wide association studies (GWAS Catalog) [41], and leading SNPs from each locus were further investigated using PhenoScanner [42] and HaploReg [43]. We also cross-referenced our identified loci with the findings from a recent GWAS of BMI and body adiposity [44].

## Acknowledgements

We are grateful to the participants and staff of the UK biobank. Access to UK Biobank genetic and phenotypic data was granted under application no. 15152. Computations were performed on the computational cluster at the Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX) under projects b2016021 and b2017066. The work was supported by grants from the Swedish Society for Medical Research (SSMF), the Kjell and Märta Beijers Foundation, Göran Gustafssons Foundation, the Swedish Medical Research Council (Project Number 2015-03327), the Marcus Borgström Foundation, and the Åke Wiberg Foundation.

## References

1. Afshin A, Forouzanfar MH, Reitsma MB, Sur P, Estep K, Lee A. Health

Effects of Overweight and Obesity in 195 Countries over 25 Years The. N Engl J Med. 2017;377: 13–27. doi:10.1056/NEJMoa1614362

2. Ng M, Fleming T, Robinson M, Thomson B, Graetz N, Margono C, et al. Global, regional, and national prevalence of overweight and obesity in children and adults during 1980–2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet. 2014;384: 766–781. doi:10.1016/S0140-6736(14)60460-8

3. Di Cesare M, Bentham J, Stevens GA, Zhou B, Danaei G, Lu Y, et al. Trends in adult body-mass index in 200 countries from 1975 to 2014: A pooled analysis of 1698 population-based measurement studies with 19.2 million participants. Lancet. NCD Risk Factor Collaboration. Open Access article distributed under the terms of CC BY; 2016;387: 1377–1396. doi:10.1016/S0140-6736(16)30054-X

4. Pi-Sunyer FX, Becker DM, Bouchard C, Carleton RA, Colditz GA, Dietz WH, et al. Clinical guidelines on the identification, evaluation, and treatment of overweight and obesity in adults: The Evidence Report. 1998.

5. Bouchard C, Despreis JP, Maurieige P. Genetic and nongenetic determinants of regional fat distribution. Endocr Rev. 1993;14: 72–93. doi:10.1210/edrv-14-1-72

6. Wajchenberg BL. Subcutaneous and Visceral Adipose Tissue□: Endocr Rev. 2000;21: 697–738. doi:0163-769X/00/$03.00/0

7. Tunstall-Pedoe H. Myth and paradox of coronary risk and the menopause. Lancet. 1998;351: 1425–1427. doi:10.1016/S0140-6736(97)11321-6

8. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, et al. Genetic studies of body mass index yield new insights for obesity biology. Nature.

2015;518: 197–206. doi:10.1038/nature14177

9. Willer CJ, Speliotes EK, Loos RJF, Li S, Lindgren CM, Heid IM, et al. Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. Nat Genet. 2009;41: 25–34. doi:10.1038/ng.287
10. Speliotes EK, Willer CJ, Berndt SI, Monda KL, Thorleifsson G, Jackson AU, et al. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. Nat Genet. 2010;42: 937–948. doi:ng.686 [pii]\r10.1038/ng.686
11. Frayling TM, Timpson NJ, Weedon MN, Zeggini E, Freathy RM, Lindgren CM, et al. A Common Variant in the FTO Gene Is Associated with Body Mass Index and Predisposes to Childhood and Adult Obesity. Science (80- ). 2007;316: 889–893.
12. Scuteri A, Sanna S, Chen WM, Uda M, Albai G, Strait J, et al. Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits. PLoS Genet. 2007;3: 1200–1210. doi:10.1371/journal.pgen.0030115
13. Akiyama M, Okada Y, Kanai M, Takahashi A, Momozawa Y, Ikeda M, et al. Genome-wide association study identifies 112 new loci for body mass index in the Japanese population. Nat Genet. Nature Publishing Group; 2017; doi:10.1038/ng.3951
14. Heid IM, Jackson AU, Randall JC, Winkler TW, Qi L, Steinthorsdottir V, et al. Meta-analysis identifies 13 new loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. Nat Genet. 2010;42: 949–60. doi:10.1038/ng.685
15. Wen W, Kato N, Hwang J-Y, Guo X, Tabara Y, Li H, et al. Genome-wide

association studies in East Asians identify new loci for waist-hip ratio and waist circumference. Sci Rep. 2016;6: 17958. doi:10.1038/srep17958

16. Kilpeläinen TO, Zillikens MC, Stančákova A, Finucane FM, Ried JS, Langenberg C, et al. Genetic variation near IRS1 associates with reduced adiposity and an impaired metabolic profile. Nat Genet. Nature Publishing Group; 2011;43: 753–60. doi:10.1038/ng.866
17. Lu Y, Day FR, Gustafsson S, Buchkovich ML, Na J, Bataille V, et al. New loci for body fat percentage reveal link between adiposity and cardiometabolic disease risk. Nat Commun. 2016;7: 10495. doi:10.1038/ncomms10495
18. Yang J, Loos RJF, Powell JE, Medland SE, Speliotes EK, Chasman DI, et al. FTO genotype is associated with phenotypic variability of body mass index. Nature. Nature Publishing Group; 2012;490: 267–72. doi:10.1038/nature11401
19. Pietrobelli A, Rubiano F, St-Onge M-P, Heymsfield SB. New bioimpedance analysis system: improved phenotyping with whole-body analysis. Eur J Clin Nutr. 2004;58: 1479–1484. doi:10.1038/sj.ejcn.1601993
20. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: A tool for genome-wide complex trait analysis. Am J Hum Genet. The American Society of Human Genetics; 2011;88: 76–82. doi:10.1016/j.ajhg.2010.11.011
21. Wood AR, Esko T, Yang J, Vedantam S, Pers TH, Gustafsson S, et al. Defining the role of common variation in the genomic and biological architecture of adult human height. Nat Genet. 2014;46: 1173–86. doi:10.1038/ng.3097
22. Kathiresan S, Melander O, Guiducci C, Surti A, Burtt NP, Rieder MJ, et al. Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. Nat Genet. 2008;40:

189–197. doi:10.1038/ng.75

23. Aguet F, Brown AA, Castel SE, Davis JR, Zappala Z, Abell NS, et al. Local genetic effects on gene expression across 44 human tissues. 2016;

24. Raychaudhuri S, Prinz WA. The Diverse Functions of Oxysterol-Binding Proteins. 2010; doi:10.1146/annurev.cellbio.042308.113334

25. Yates A, Akanni W, Amode MR, Barrell D, Billis K, Carvalho-silva D, et al. Ensembl 2016. 2017;44: 710–716. doi:10.1093/nar/gkv1157

26. Jia Z, Gao S, M'Rabet N, De Geyter C, Zhang H. Sp1 Is Necessary for Gene Activation of Adamts17 by Estrogen. 2014;1839: 1829–1839. doi:10.1002/jcb.24855

27. Komaragiri SK, Bostanthirige DH, Morton DJ, Patel D, Joshi J, Upadhyay S, et al. ID4 promotes AR expression and blocks tumorigenicity of PC3 prostate cancer cells. Biochem Biophys Res Commun. Elsevier Ltd; 2016;478: 60–66. doi:10.1016/j.bbrc.2016.07.092

28. Shungin D, Winkler TW, Croteau-Chonka DC, Ferreira T, Locke AE, Magi R, et al. New genetic loci link adipose and insulin biology to body fat distribution. Nature. 2015;518: 187–196. doi:10.1038/nature14132

29. Wood AR, Esko T, Yang J, Vedantam S, Pers TH, Gustafsson S, et al. Defining the role of common variation in the genomic and biological architecture of adult human height. Nat Genet. 2014;46: 1173–86. doi:10.1038/ng.3097

30. Randall JC, Winkler TW, Kutalik Z, Berndt SI, Jackson AU, Monda KL, et al. Sex-stratified Genome-wide Association Studies Including 270,000 Individuals Show Sexual Dimorphism in Genetic Loci for Anthropometric Traits. PLoS Genet. 2013;9. doi:10.1371/journal.pgen.1003500

31. Fox CS, Liu Y, White CC, Feitosa M, Smith A V., Heard-Costa N, et al. Genome-wide association for abdominal subcutaneous and visceral adipose reveals a novel locus for visceral fat in women. PLoS Genet. 2012;8. doi:10.1371/journal.pgen.1002695

32. Winkler TW, Justice AE, Graff M, Barata L, Feitosa MF, Chu S, et al. The Influence of Age and Sex on Genetic Associations with Adult Body Size and Shape: A Large-Scale Genome-Wide Interaction Study. PLoS Genet. 2015;11: 1–42. doi:10.1371/journal.pgen.1005378

33. Mally K, Trentmann J, Heller M, Dittmar M. Reliability and accuracy of segmental bioelectrical impedance analysis for assessing muscle and fat mass in older Europeans: A comparison with dual-energy X-ray absorptiometry. Eur J Appl Physiol. 2011;111: 1879–1887. doi:10.1007/s00421-010-1795-x

34. Walter K, Min JL, Huang J, Crooks L, Memari Y, McCarthy S, et al. The UK10K project identifies rare variants in health and disease. Nature. 2015;526: 82–90. doi:10.1038/nature14962

35. Auton A, Abecasis GR, Altshuler DM, Durbin RM, Bentley DR, Chakravarti A, et al. A global reference for human genetic variation. Nature. 2015;526: 68–74. doi:10.1038/nature15393

36. Body Composition Analyzer BC-418 Instruction manual. Tokyo;

37. Aulchenko YS, Ripke S, Isaacs A, van Duijn CM. GenABEL: an R library for genome-wide association analysis. Bioinformatics. 2007;23: 1294–1296. doi:10.1093/bioinformatics/btm108

38. Rstudio-Team. RStudio: Integrated Development for R. [Internet]. Boston, MA.: RStudio, Inc.; 2016. Available: http://www.rstudio.com/

39. Weeks JP. plink: An R Package for Linking Mixed-Format Tests Using IRT-

Based Methods. J Stat Softw. 2010;35: 1–33. Available: http://www.jstatsoft.org/v35/i12/

40. Keller MC. Gene × environment interaction studies have not properly controlled for potential confounders: The problem and the (simple) solution. Biol Psychiatry. Elsevier; 2014;75: 18–24. doi:10.1016/j.biopsych.2013.09.006

41. MacArthur J, Bowler E, Cerezo M, Gil L, Hall P, Hastings E, et al. The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). Nucleic Acids Res. 2017;45: D896–D901. doi:10.1093/nar/gkw1133

42. Staley JR, Blackshaw J, Kamat MA, Ellis S, Surendran P, Sun BB, et al. PhenoScanner: A database of human genotype-phenotype associations. Bioinformatics. 2016;32: 3207–3209. doi:10.1093/bioinformatics/btw373

43. Ward LD, Kellis M. HaploReg: A resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. Nucleic Acids Res. 2012;40: 930–934. doi:10.1093/nar/gkr917

44. Graff M, Scott RA, Justice AE, Young KL, Feitosa F, Barata L, et al. Genome-wide physical activity interactions in adiposity — A meta-analysis of 200 , 452 adults. 2017; 1–26.

## Figure legends

**Figure 1. Manhattan plots showing the significance of association between all SNPs and AFR (A), LFR (B) and TFR (C) in the sex-combined discovery analyses.** The red lines show the genome wide significance cut-off ($p < 5*10^{-8}$).

## Supplementary material

**S1 Fig. Manhattan plots showing the significance of association between all SNPs**

**and AFR (A), LFR (B) and TFR (C) in the discovery analyses in females.** The red lines show the genome wide significance cut-off ($p < 5*10^{-8}$).

**S2 Fig. Manhattan plots showing the significance of association between all SNPs and AFR (A), LFR (B) and TFR (C) in the discovery analyses in males.** The red lines show the genome wide significance cut-off ($p < 5*10^{-8}$).

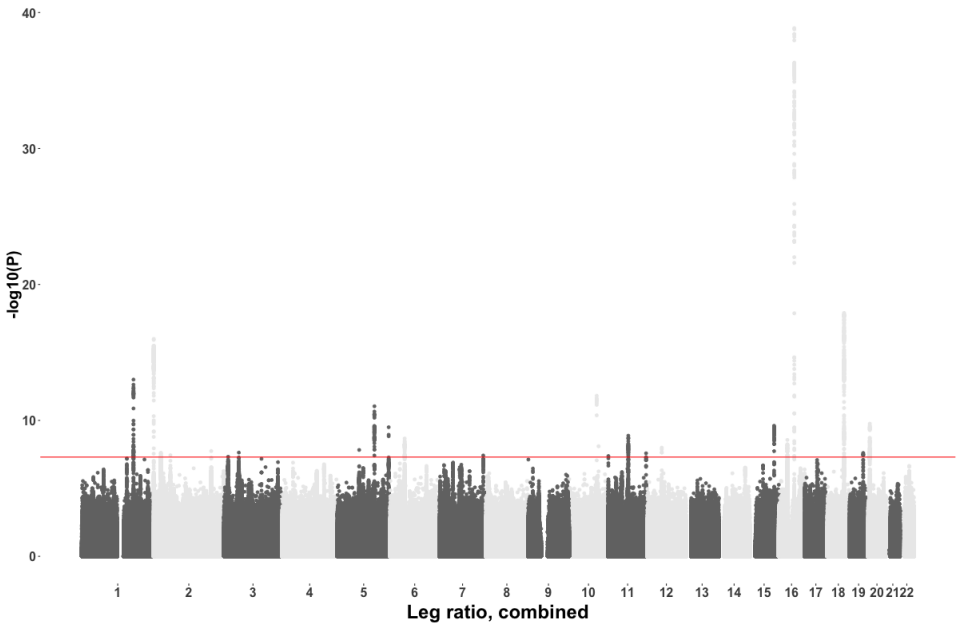
**S3 Fig. Quantile-quantile plots of all SNPs in the discovery analyses for the combined and sex-stratified analyses.** The red lines show the expected distribution of p-values under the null hypotheses.

**S1 Table. Significant GWAS findings when conditioning for the leading SNPs from the initial discovery analyses.**

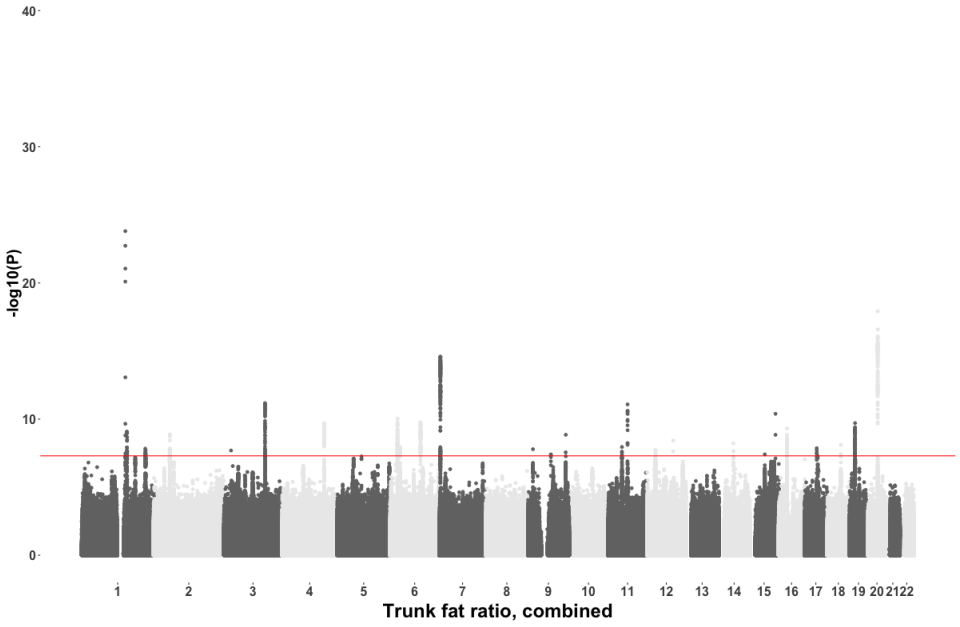
**S1 Supplementary Data. Results from the discovery and replication analyses for leading SNPs associated with AFR, LFR and TFR.** Data is presented for each of the sex-combined and stratified analyses.

**S2 Supplementary Data. Results from tests for SNP-sex interactions for body fat distribution.**

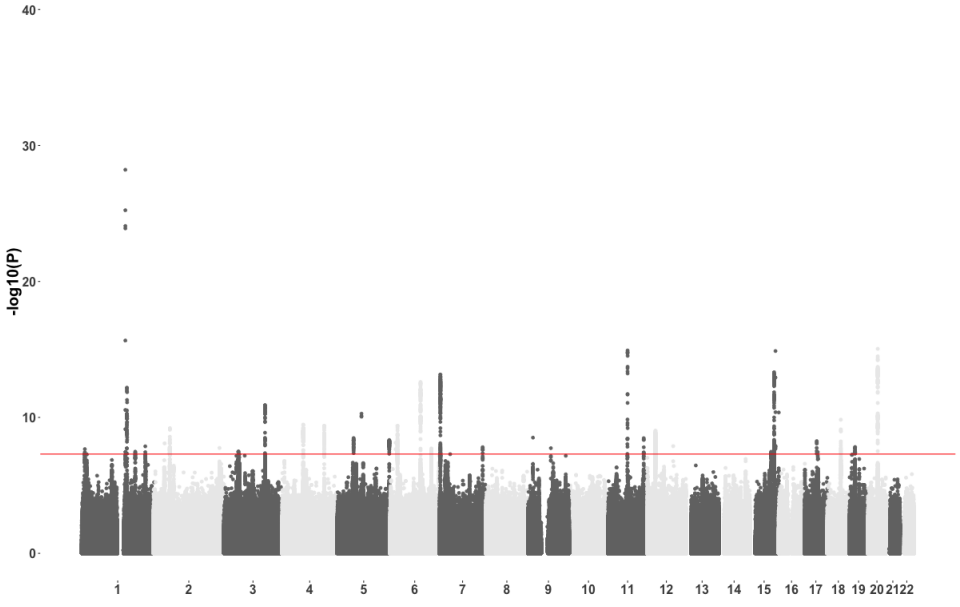
A



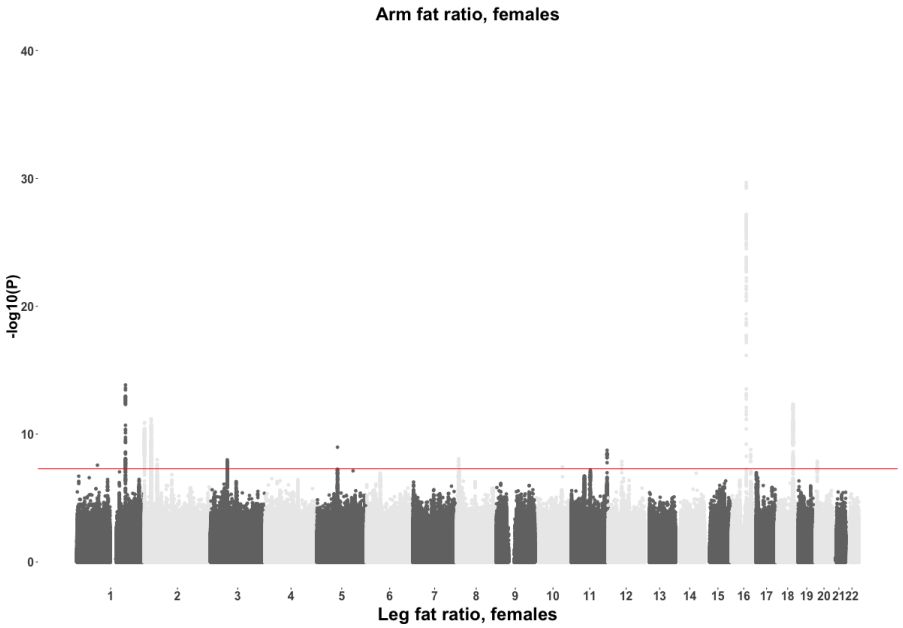

B



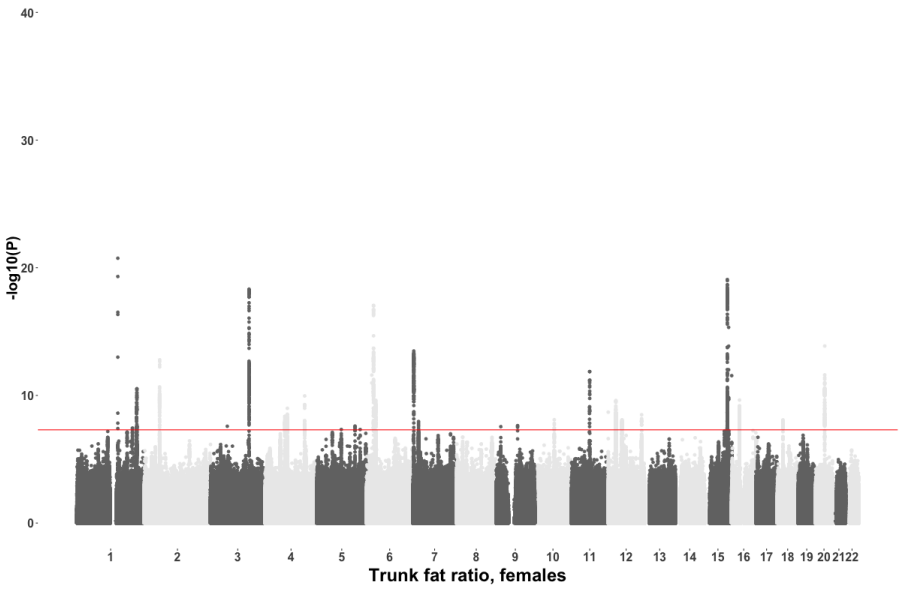

C



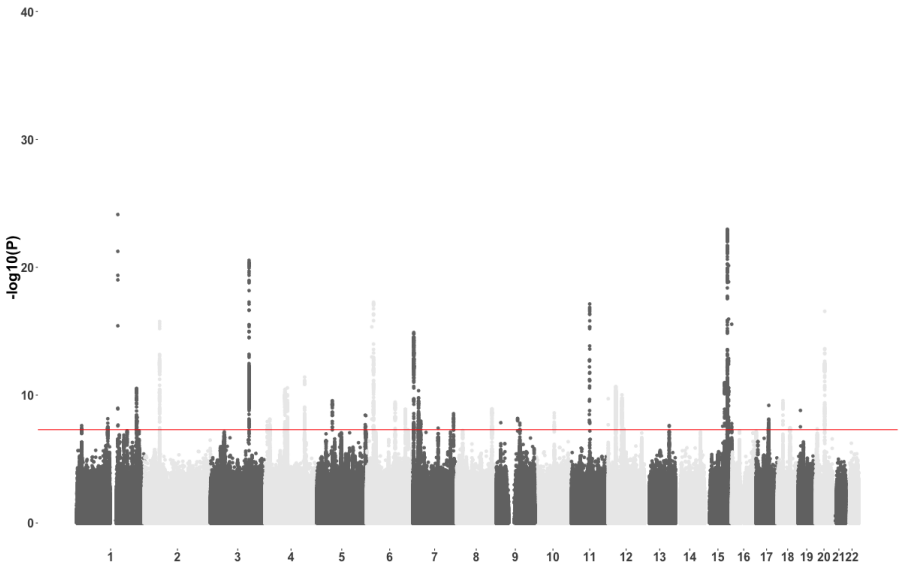
**A**



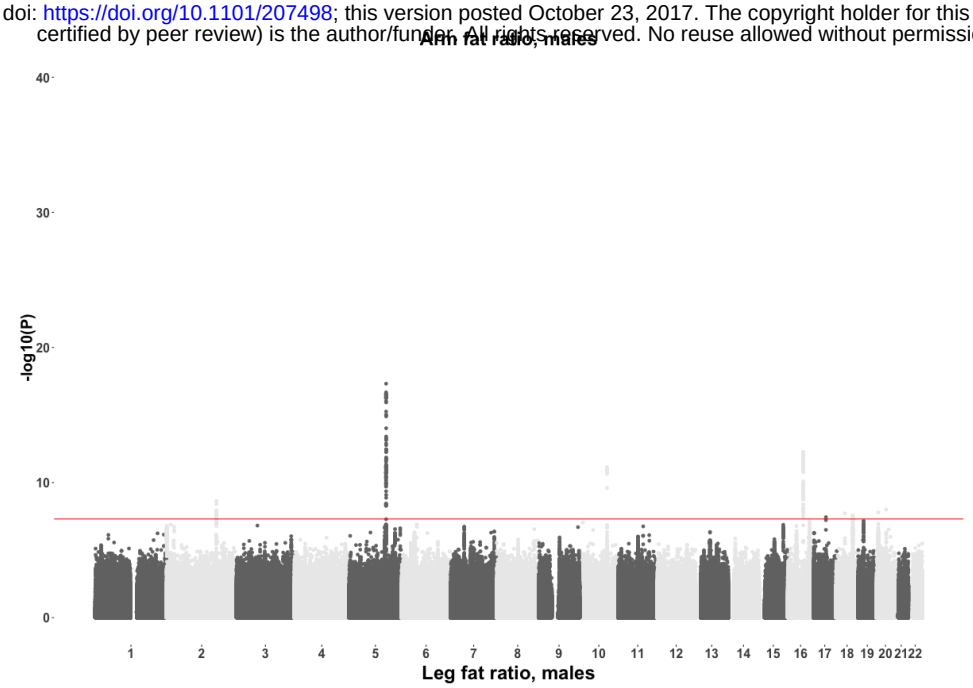

**B**



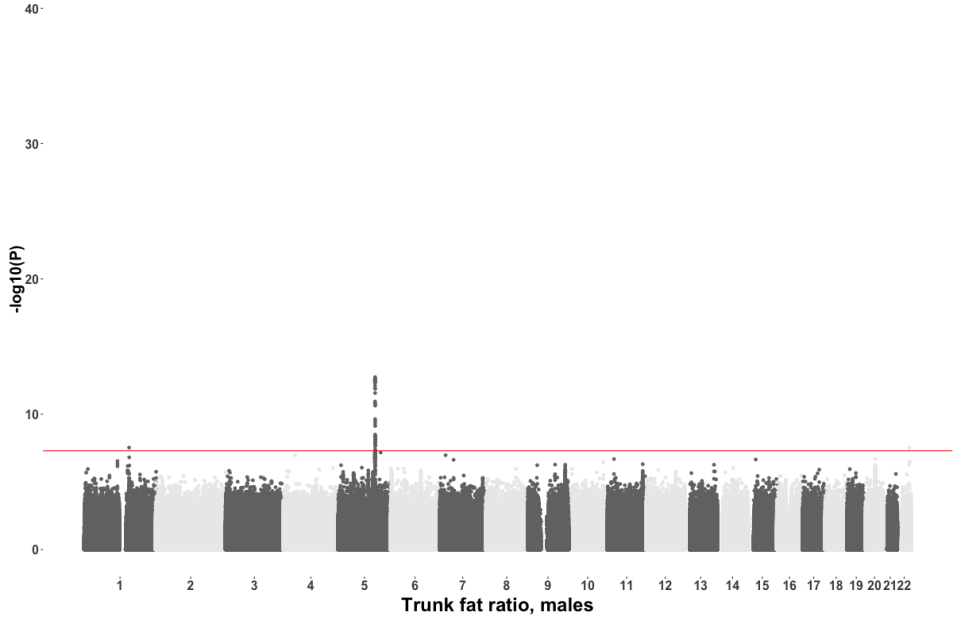

**C**



A




B




C