HHS Public Access

Author manuscript

Nat Genet. Author manuscript; available in PMC 2023 September 01.

Published in final edited form as:

Nat Genet. 2023 March ; 55(3): 461–470. doi:10.1038/s41588-023-01306-0.

Genetics of sexually dimorphic adipose distribution in humans

Grace T. Hansen[1,2,*,#], **Débora R. Sobreira**[1,#], **Zachary T. Weber**[1], **Alexis G. Thornburg**[1], **Ivy Aneas**[1], **Li Zhang**[1], **Noboru J. Sakabe**[1], **Amelia C. Joslin**[1], **Gabriela A. Haddad**[1], **Sophie M. Strobel**[3,4], **Samantha Laber**[3,11], **Farhath Sultana**[5], **Faezeh Sahebdel**[5], **Kohinoor Khan**[5], **Yang I. Li**[1,6], **Melina Claussnitzer**[3,7,8,9,10], **Liang Ye**[5,*], **Ricardo A. Battaglino**[5,*], **Marcelo A. Nóbrega**[1,*]

[1]Department of Human Genetics, University of Chicago

[2]Pritzker School of Medicine, University of Chicago

[3]Broad Institute of MIT and Harvard

[4]Institute of Nutritional Medicine, School of Medicine, Technical University of Munich, Germany

[5]Department of Rehabilitation Medicine, University of Minnesota

[6]Department of Genetic Medicine, University of Chicago

[7]Medical Population genetics Program, The Broad Institute of MIT and Harvard

[8]Massachussetts General Hospital, Harvard Medical School

[9]Diabetes Unit and Center for Genomic Medicine, Massachussetts General Hospital

[10]Department of Medicine, Harvard Medical School

[11]Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, University of Oxford, UK

Abstract

Obesity-associated morbidity is exacerbated by abdominal obesity, which can be measured as waist-to-hip ratio adjusted for body mass index (WHRadjBMI). Here, we identify genes associated

* Correspondence should be addressed to Grace T. Hansen (gthansen@uchicago.edu), Liang Ye (yexxx142@unm.edu), Ricardo A. Battaglino (rbattagl@umn.edu) or Marcelo A Nóbrega (nobrega@uchicago.edu).

#Contributed equally to this work.

Author Contributions

G.T.H, M.A.N, and Y.I.L. conceived of the initial TWAS approach to finding novel genes associated with obesity and WHRadjBMI. Y.I.L. supervised the genetic correlation, ldsc-seg, and TWAS analyses. A.C.J. and G.A.H. performed the MPRA, with assistance and direction from D.R.S. L.Y. and performed the mmΔSnx10^{Adipoq} mouse generation, HFD administration, and DEXA scanning, with supervision from R.A.B. L.Y., F.Sa., and K.K. F.Su. and K.K. performed the shRNA SNX10 shRNA experiments, with supervision and training from R.A.B. A.G.T. and D.R.S. performed the luciferase assays to confirm the regulatory activity of Alu elements. I.A., Z.T.W., L.Z., and D.R.S. performed the CRISPR-mediated deletion of Alu elements. Z.T.W. and N.J.S. analyzed the CRISPR data. S.M.S and S.L. performed the activity-by-contact and Lipocyte Profiler analyses, with supervision from M.C. G.T.H performed the genetic correlation, ldsc-seg, TWAS, MPRA, and motif analys. G.T.H. drafted the manuscript and figures, with editorial assistance from M.A.N.

Competing Interests Statement

The authors of this manuscript have no financial or other interests that may compete with their judgment of this work and have complied with all the ethical requirements of their institutions.

Code Availability

Unless otherwise specified in the Methods section, all code used to generate figures and data included in the study are available at https://github.com/grace-hansen/.

with obesity and WHRadjBMI and characterize allele-sensitive enhancers which are predicted to regulate WHRadjBMI genes in women. We found that several waist-to-hip ratio-associated variants map within primate-specific *Alu* retrotransposons harboring a DNA motif associated with adipocyte differentiation. This suggests that a component of the genetics of adipose distribution in humans may involve co-option of retrotransposons as adipose enhancers. We evaluated the role of the strongest female WHRadjBMI-associated gene, *SNX10*, in adipose biology and determined that it is required for human adipocyte differentiation and function and participates in diet-induced adipose expansion in female mice, but not males. Our data identify genes and regulatory mechanisms that underlie female-specific adipose distribution and mediate genetics of metabolic dysfunction in women.

Several common chronic illnesses, including diabetes and cardiovascular diseases, are precipitated or exacerbated by obesity. However, some obese individuals are more susceptible to metabolic complications than others, with the pattern of excess fat deposition being recognized as a strong modulator of obesity-related cardiometabolic disease[1–3]. Abdominal fat deposition, which represents visceral adipocyte accumulation and ectopic fat deposition on vital organs, is associated with increased insulin resistance and systemic inflammation mediated by the release of free fatty acids and inflammatory cytokines into the bloodstream[4]. Increased immune cell infiltration leads to further insulin resistance, increasing the risk of type 2 diabetes[1]. Waist-to-hip (WHR) circumference ratio, adjusted for BMI (WHRadjBMI) is a proxy for the balance of subcutaneous versus visceral and ectopic fat deposition[5]. A high WHRadjBMI indicates a high burden of abdominal fat and is an independent predictor of cardiovascular disease, type 2 diabetes, non-alcoholic fatty liver disease, renal failure, and stroke[2,4]. WHRadjBMI is a sexually dimorphic trait: women have lower WHRadjBMI than men until menopause, and the lower WHRadjBMI of pre-menopausal women is associated with protection from cardiometabolic disease[2,4].

Obesity and WHRadjBMI are both complex, polygenic genetic traits[5,6], with their genetic heritability distributed over many variants of small phenotypic effect that map primarily to non-coding regions of the genome[5–8]. Despite the identification of these variants, our appreciation of the mechanisms underlying genetic heritability of metabolic health remains limited. For most loci, we have yet to determine which variants are causal, identify which genes mediate the phenotypic effects non-coding risk variants, and understand the role these genes play in human fat distribution and metabolism. To address these questions, we applied a series of experimental and computational approaches to identify genes and variants that may underlie obesity and high WHRadjBMI. We determined the regulatory properties of candidate variants at these loci. We demonstrate the sex-specific effects *in vivo* of a female WHRadjBMI-associated gene, *SNX10*[8,9], illustrating how this gene is involved in adipocyte differentiation and diet-induced adiposity.

Our work advances the understanding of sex-specific fat distribution from variants of small effect that act as adipocyte enhancers to genes whose roles in adipocyte biology may underlie pathologic weight gain in humans.

Results

Tissue specificity and sexual dimorphism of WHRadjBMI and obesity.

To identify specific tissues with enrichment for these traits we performed ldsc-seg[10], which combines genotype and gene expression data from multiple tissues with GWAS data. By identifying the tissue-specific expression patterns of genes nearby risk variants for a trait, ldsc-seg identifies tissues whose gene expression is statistically associated with traits of interest. Using gene expression data from the Genotype-Tissue Expression (GTEx) consortium v8 combined with GWAS data representing obesity and WHRadjBMI from the GIANT consortium[5,6] (obesity n~681,275, WHRadjBMI n=694,649), we find that obesity is strongly enriched in central nervous system tissues, with the strongest enrichment in frontal cortex, as previously seen on other studies[7] (Figure 1a, Supplementary Table 1). In contrast, WHRadjBMI is associated with gene expression in several tissues. Female reproductive tissues and adipose tissues, especially subcutaneous adipose tissue, show the highest enrichment for WHRadjBMI, as previously seen in this trait[11] (Figure 1a, Supplementary Figure 1).

Association between tissue-specific gene expression and human traits is complicated by the heterogeneity of cell types in each tissue. To account for this heterogeneity, we estimated cell type proportions for 63 cell types in all samples from GTEx and 613 human postmortem frontal cortex samples from the CommonMind Consortium using *xCell*[12,13], which compares sample gene expression with gene set enrichments from primary human cell lines to estimate cell type prevalence. The estimated cell type proportions for all samples are available in Supplementary Table 2. We find that estimated cell type proportions are similar between males and females, suggesting that sex-specific differences in phenotypes may not often be driven by changes in cell type proportion (Supplementary Figure 2). Samples from frontal cortex, the tissue most strongly enriched for obesity, show a high estimated proportion of neurons relative to all other cell types. In contrast, subcutaneous adipose tissue showed several cell types with proportions over 5%, including adipocytes and preadipocytes, immune cells such as dendritic cells and natural killer T cells, vascular cells, and stromal cells. (Supplementary Figure 2). While this cell type diversity is a known attribute of adipose tissue[15], the association of adipose bulk tissue with WHR could, in principle, be partially due to gene expression in any of the cell types present at a significant proportion. To determine whether adipocytes in particular are associated with WHRadjBMI, we quantified the estimated adipocyte/preadipocyte proportion (‘fat content’) for each tissue in GTEx. Across all tissues, high estimated fat content was associated with enrichment for waist-to-hip ratio heritability (Figure 1b). Together, these data suggest that WHRadjBMI is associated specifically with gene expression in preadipocytes and adipocytes.

Both obesity and WHRadjBMI are sexually dimorphic traits[16]. Women have lower mean WHR than men and are somewhat more likely to become obese[17–19]. We tested whether genetic risk for obesity and WHRadjBMI is also sexually dimorphic. Genetic correlation analyses showed high correlation between male and female genetic risk for obesity, (r_g=0.923, se=0.0083), and lower correlation between male and female genetic risk for WHRadjBMI (r_g=0.6447, se=0.0262) (Figure 1c). Additionally, we found that total

heritability is similar between women and men for obesity ($p=1.6*10^{-1}$, Z=1.39, female h^2=0.2029, se=0.0064, male h^2=0.2166, se=0.0075), but larger for female WHRadjBMI than male WHRadjBMI ($p=3.4*10^{-4}$, Z=3.58, female h^2=0.1582, se=0.0099, male h^2=0.1169, se=0.0059). These findings suggest that while genetic risk factors for obesity are likely shared between the sexes, there may be female-specific genetic determinants of fat deposition, as suggested by previous studies[5,8].

Identification of genes associated with obesity and WHRadjBMI.

To determine biological mechanisms that may underlie the genetic heritability of obesity and WHRadjBMI, we next sought to identify genes associated with these traits. We performed transcriptome-wide association studies (TWAS), which combine a gene expression panel from a set of individuals, their genotype data, and GWAS summary statistics to identify genes whose expression are associated with genetic risk for a complex trait[21]. Based on the ldsc-seq enrichment data (Figure 1a), we used frontal cortex gene expression (GWAS N~681,275, gene expression N=438 individuals) for the obesity TWAS and subcutaneous adipose gene expression for WHRadjBMI TWAS, using cortex samples from the CommonMind consortium and adipose samples from GTEx. To account for cell type heterogeneity in subcutaneous adipose tissue, we selected the top seven most enriched cell types, collectively representing an average of 47.4% of each sample, and included cell type proportions for these cell types as covariates in the WHRadjBMI TWAS. To identify putative sex-specific mechanisms for WHRadjBMI, female-specific GWAS summary statistics were combined with gene expression samples from females (GWAS N=379,501 females, gene expression N=194 females), and male GWAS data were used in conjunction with male gene expression samples (GWAS N=315,284 males, gene expression N=194 males).

We identified 460 genes whose expression in cortex is associated with obesity risk. The most significant genes include those whose dysfunction in the brain have been implicated in obesity, including *MC4R* and *ADCY3*[22,23], in addition to genes with uncharacterized roles in obesity (Figure 2a). Our WHRadjBMI TWAS identified significantly more genes in female WHRadjBMI than in males, despite equivalent sample sizes, with 91 genes identified for female and 42 genes identified for male (Figure 2b, Supplementary Table 3). The majority of these genes have not been previously highlighted as candidate genes at GWAS loci, with 17 of the 460 obesity TWAS genes previously identified through GWAS gene-finding approaches published in previous literature[7,24], 8 of the 91 female WHRadjBMI TWAS genes, and 4 of the 42 male WHRadjBMI TWAS genes[8,25]. The number of genes found in both obesity TWAS and WHRadjBMI TWAS was significantly less than that expected by chance, an effect likely due to the adjustment of WHR to account for correlation with body mass index (BMI) (obesity and female WHRadjBMI gene overlap= 8 genes, $p<1:10^{-5}$, obesity and male WHRadjBMI overlap=5 genes, $p<1:10^{-5}$, Figure 2a). In contrast, female and male WHRadjBMI TWAS show more genes identified in both analyses than expected by chance, suggesting a shared etiology for some gene effects (female and male WHRadjBMi gene overlap=11 genes, $p=0.226*10^{-2}$, Figure 2a).

Genes associated with WHRadjBMI in both sexes have, on average, a 1.8-fold larger effect size in women (Figure 2b). To evaluate whether sexual dimorphism in WHRadjBMI is the result of sex-specific *cis*-regulatory effects of variants, we compared the effect size of the top eQTL of female TWAS genes in each sex with the effect size of these variants in WHRadjBMI GWAS from each sex. While we observe that eQTLs have comparable effect sizes on gene expression in men and women, they have much larger GWAS effect sizes in women (Figure 2d). This is in concordance with recent observations from GTEx suggesting that there are very few sex-specific eQTLs throughout the genome, despite widespread sexual dimorphism in human traits. We illustrate this effect on S*NX10*, the strongest female-specific WHRadjBMI TWAS gene[5,8,25]. A variant located in an intron of *SNX10*, rs1534696, is a strong eQTL in subcutaneous adipose tissue for *SNX10* in both sexes, and is strongly associated with WHRadjBMI in women, but not in men (Figure 2c,e, female eQTL $p=1.32*10^{-14}$, $\beta=0.68$, male eQTL $p=1.04*10^{-11}$, $\beta=0.64$). We tested the regulatory properties of this region in reporter assays in 3T3L1 pre-adipocytes and show that rs1534696 maps to a genomic region with enhancer properties (Supplementary Figure 14, Supplementary Table 6), as has been previously reported[9], further suggesting that this enhancer regulates *SNX10* expression. While the eQTL effect is preserved in both sexes, *SNX10* expression itself is consistently higher in women (Figure 2c), with a fold change of ~1.3 across rs1534696 genotypes. These data suggest that rs1534696 affects *SNX10* expression in adipocytes in both sexes, and that this activity modulates WHRadjBMI in women, but not in men, perhaps because of higher *SNX10* expression in women. *SNX10* expression does not differ between women of typical pre-menopausal and post-menopausal age (Supplementary Figure 3, p=0.44), suggesting that the female-specific effect is not the result of increased systemic estrogen levels. rs1534696 is also a lead variant for triglyceride levels and high-density lipoprotein (HDL) cholesterol in women, but not in men (Supplementary Figures 4 and 5), suggesting that the sex-specific effect of rs1534696 on WHRadjBMI also modulates metabolic traits closely associated with cardiometabolic disease traits.

WHRadjBMI-associated enhancers map within Alu retrotransposons

Through TWAS, we identified 91 genes whose expression is associated with high WHRadjBMI in women. By identifying enhancers regulating the expression of these genes we can gain insight into the gene regulatory network that controls WHRadjBMI. However, owing to extensive co-inheritance of genetic variants near one another (linkage disequilibrium, LD), it is difficult to differentiate which variants disrupt gene expression, and which are merely in linkage with true causal variants from GWAS summary statistics. Therefore, we performed a massively parallel reporter assay (MPRA) to experimentally determine the regulatory properties of all candidate variants in all loci[26,36]. For each gene we selected the lead GWAS variant, the top eQTL for the TWAS gene, and the *coloc*[37]-identified shared causal variant underlying GWAS and eQTL signal, if present. We also selected all variants in high LD ($r^2 > 0.95$) with the variants identified above, totaling 1,882 variants. This stringent LD threshold, chosen to accommodate limitations with number of synthesized constructs, and likely resulted in us not testing several functional and causal variants in our MPRA.

To functionally interrogate each variant, we synthesized 175-bp DNA fragments centered around both alleles of each variant, totaling 3,764 DNA fragments. We cloned the library of DNA constructs into the pMPRA1 reporter vector[27] and transfected them into 3T3-L1 preadipocytes in four replicates. We identified 426 variants in regions that display enhancer activity in preadipocytes (significant MPRA enhancers, Supplementary Table 4). Of these, 58 harbored variants whose alternate alleles modulate enhancer activity (enhancer-modulating variants) (Figure 3a, Supplementary Figure 7). Significant MPRA enhancer variants were more likely to lie within active chromHMM annotations in adipose nuclei than nonsignificant variants (25% of enhancer variants vs. 20% of nonsignificant variants, Chi-square p=3.53*10–2, χ2=4.43, df=1). Supplementary Figure 8).

To uncover genetic networks potentially coordinating the regulatory properties of these enhancers, we performed two transcription factor motif enrichment analyses. First, we compared the 175-bp sequences centered around each of the 426 significant MPRA enhancer variants to the same-sized sequences corresponding to nonsignificant MPRA variants. While limited by small sample size, this analysis yielded 18 transcription factors whose motifs are enriched in significant MPRA enhancer sequences. Of these, eight are directly involved in adipogenesis, based on previous literature[28–35], and five others are involved in adipocyte-relevant metabolic processes such as cholesterol synthesis, adipocyte browning, and glucose metabolism[30,38–41] (Supplementary Table 5). In the second analysis, we looked for enrichment of transcription factor motifs in the sequence surrounding the 58 enhancer-modulating variants versus nonsignificant variants, allowing for the identification of binding sites beyond the narrowly defined 175-bp MPRA fragments. We also found strong enrichment for factors involved in adipogenesis and adipocyte metabolism, and significant overlap with motifs enriched in significant MPRA enhancers. Six of the 14 motifs are directly involved in adipogenesis[42], and three are involved in adipocyte metabolism (Figure 3b). Almost all the identified adipogenic motifs are characterized by a common core sequence (Figure 3b), AGGTCA, corresponding to half of the DR1 DNA motif which binds *PPARG*, a master regulator of adipogenesis[43].

To further assess the functional importance of MPRA sequences, we tested whether significant MPRA enhancers were more evolutionarily conserved than nonsignificant sequences across mammalian species, with the expectation that sequence conservation reflects functional importance. Surprisingly, we found that significant MPRA enhancer sequences are less conserved than nonsignificant sequences, and sequences containing enhancer-modulating variants are the least conserved sequences in the MPRA (Figure 3c, Student's t-test nonsignificant variants vs. significant MPRA enhancer variants p=7.073*10^{-2}, nonsignificant variants vs. enhancer-modulating variants p=3.173*10^{-2}). Further inspection revealed that this unusual lack of evolutionary conservation is due to a strong enrichment of primate-specific *Alu* repeat elements overlapping significant MPRA enhancers (fold change (FC) over nonsignificant variants=1.29, chi-square χ^2=16.39, df=1, p=5.15*10^{-5}). Notably, 78% of identified enhancer-modulating variants map within *Alu* elements (Figure 3c). In our MPRA data, most *Alu* sequences tested (287 of 479 *Alu* elements containing an MPRA variant) do not drive reporter gene expression in adipocytes. However, a subset of *Alu* elements are associated with very high enhancer activity in pre-adipocytes (Figure 3d) suggesting that some *Alus* may have been co-opted as enhancers

of nearby metabolism-related genes. *Alu* elements do not drive enhancer activity in other MPRA datasets (Supplementary Figure 8), suggesting that these *Alu* elements may act as enhancers specific in the context of adipocyte biology.

Given that only some *Alu* elements stimulate gene expression in preadipocytes, we suspected that the AGGTCA motif enriched in significant MPRA results (Figure 3b) may be part of the sequence of *Alus* with enhancer activity. To test this, we first determined that *Alu* elements and adipogenesis-associated motifs co-occur more frequently than is expected by chance (FC=1.39, chi-square χ^2=41.933, df=1, p=9.44*10^{-11}). Next, we mapped the relative locations of *Alu* elements and motifs and identified that AGGTCA motifs in significant MPRA sequences are consistently located within the first monomer of the *Alu* element (65–80bp from the *Alu* start site) (Figure 3e). The AGGTCA motif is not likely to be the product of a mutation that has emerged from a single *Alu* ancestor, as multiple sequence alignment shows that the *Alu* sequences which drive expression in the MPRA are not more closely related than nonsignificant sequences (Supplementary Figure 10).

To confirm the regulatory activity of *Alu*-containing sequences from the MPRA, we first performed luciferase assays in 12 *Alu*-containing sequences with enhancer activity and six sequences lacking enhancer activity and constructed DNA (mean length= 1000bp) fragments centered on each element, a significant expansion over the 175-bp fragments tested in MPRA. We obtained significant luciferase activity in 8 of these, while none of the six elements that were not enhancers in the MPRA assay showed enhancer activity (Supplementary Figure 11, Supplementary Table 6).

To directly test the regulatory role of *Alu*-containing enhancers we next engineered CRISPR-mediated deletions of three *Alus* tested by MPRA. Two of the regions showed enhancer properties in the MPRA, and one was nonsignificant. Each region was deleted in primary human mesenchymal stem cells followed by clonal expansion of single cells, genotyping and transcriptome analyses. The deletion of a MPRA enhancer harboring rs4955430 led to down-regulation, in an enhancer dosage-sensitive manner, of several genes in cis, including *DAG1*, a gene which forms long-range chromatin interactions in preadipocytes with the deleted enhancer and is a target gene of the WHRadjBMI TWAS (Supplementary Figure 12). Another tested *Alu*-containing MPRA enhancer harbors rs6446275 and is directly upstream of *NICN1*. Deletion of this region resulted in a dramatic down-regulation, also in an enhancer dosage-sensitive manner (Supplementary figure 12) of *NICN1*, suggesting that this *Alu* element has been co-opted as a promoter of *NICN1*, further supported by the chromatin state in human adipose tissue from the Epigenome Roadmap Project[44]. Deletion of all three regions, ranging from 585 to 1,300 bp (Supplementary table 7) resulted in reduced expression of nearby genes in mesenchymal stem cells (Supplementary Figures 12 and 13), suggesting that regulatory elements located within *Alus* can modulate gene expression in adipocyte progenitors, and corroborating our findings that a subset of *Alu* retrotransposons display adipose enhancer properties.

There are limitations to our strategy and data interpretation. We carried out MPRA in an artificial, *in vitro* milieu, at a single time point early in adipocyte differentiation and without any cellular perturbations or stimuli which might better reflect enhancer properties *in vivo*.

Also, while most enhancer-modulating variants map within *Alu* repeats, the number of *Alu*-containing enhancer-modulating variants represent but a fraction of the genes and loci associated with WHRadjBMI. Future studies will determine the extent to which co-option of *Alu* repeats as adipose enhancers participate in the genetic architecture of WHRadjBMI in humans.

Our data suggest that a subset of *Alu* elements in the human genome may have been coopted to build a gene regulatory network in adipocytes and genetic variation in a subset of these enhancers may modulate body fat distribution in women.

Cellular phenotypic effects of an WHRadjBMI-associated variant

Our TWAS and MPRA results suggest that many WHRadjBMI genes may impart their effects on body fat distribution by modulating the development and/or function of adipocytes. The strongest WHRadjBMI TWAS signal we identified was for *SNX10,* a gene that has been repeatedly shown to be associated with WHRadjBMI[5,8,25] but whose roles in adipose tissue remain uncharacterized. The lead variant in this locus, rs1534696, shows strong evidence for being causal for female WHRadjBMI risk. rs1534696 is the top eQTL for *SNX10*, and colocalization analyses suggest that rs1534686 is responsible for both the GWAS and eQTL signals at this locus (Supplementary Figure 14). rs1534696 lies within an enhancer in pre-adipocytes (Supplementary Figure 15) and previous work showed that this enhancer has allele-specific properties in mature adipocytes[9]. To further explore the regulatory effects of each allele of this variant across adipocyte differentiation, we first applied activity-by-contact (ABC), a modeling tool that predicts interaction between enhancers and gene promoters based on chromatin accessibility and H3K27ac histone modifications[45], to pre-adipocytes at several time-points of adipogenesis. ABC analysis predicts that the rs1534696 enhancer interacts with *SNX10* in fully differentiated adipocytes, but not at earlier stages in adipocyte maturation (Figure 4a).

To decipher the consequences of rs1534696 on adipocyte morphological and cellular phenotypes, we next used a recently developed high-content imaging approach, Lipocyte Profiler[46]. In brief, Lipocyte Profiler uses microscopy and image analysis to distinguish differences in the morphological features of cellular organelles according to genotype. Specifically, Lipocyte Profiler uses the fluorescent dyes AGP (actin cytoskeleton, Golgi body, and plasma membrane), Bodipy (lipid droplets and cytoplasmic RNA), Mito (mitochondria) and DNA (nucleic-acid related phenotypes) and performs subsequent image analysis to detect cellular and morphological feature changes in *in vitro* cell cultures according to genotype. We assessed lipocyte profiles across four time-points of adipocyte differentiation (days 0, 3, 8, and 14) in subcutaneous and visceral human adipose-derived mesenchymal stem cells (AMSCs) between different rs1534696 genotypes (Subcutaneous mESCs: 20 CC/CA (14 females), 11 AA genotypes (9 females). Visceral mESCs: 23 CC/CA (17 females), 12 AA genotypes (9 females), Figure 4b). Due to power limitations associated with low number of male samples in the analysis, we were only able to observe genotype-dependent features on the sex-combined sample and did not investigate the interaction between sex and genotype. To obtain sufficient power, we analyzed rs1534696 genotype data as AA vs. CA/CC. We observed significant changes in adipocyte morphological

and cellular features dependent on the rs1534696 genotype only in mature subcutaneous adipocytes (Figure 4c, Supplementary Table 8). Specifically, the majority (50 of 74) of haplotype-driven differential features are lipid storage- associated features describing the structure of intracellular lipids (Figure 4d). This suggests that rs1534696 affects the function of mature adipocytes by dysregulating the cells' ability to store intracellular lipids.

Together, these data suggest that the rs1534696 allele may act by disrupting an enhancer for *SNX10* specifically in mature subcutaneous adipocytes which results in altered lipid handling and/or storage in these cells. This depot-specific adipocyte cellular phenotype driven by rs1534696 is consistent with an organismal phenotype of altered body fat distribution in which subcutaneous white adipose tissue would be affected by rs1534696.

Physiology of SNX10 in adipose biology in humans and mice

To validate the function of the WHRadjBMI TWAS target gene *SNX10* in adipocyte development and function, we first knocked down *SNX10* in human mesenchymal stem cells via the transfection of anti-*SNX10* short hairpin RNA (shRNA). We subsequently differentiated these cells into mature adipocytes. In comparison to wild type cells or a scrambled vector, *SNX10* knockdown resulted in impaired lipid accumulation, and reduced expression of markers of mature adipocytes (Figure 5a, Supplementary Table 9) (*ADIPOQ* ΔCt scramble shRNA=6.18, ΔCt S*NX10* shRNA=8.86, $p=1.7*10^{-4}$), confirming an important role of *SNX10* in adipocyte differentiation and function of mature adipocytes. In line with rs1534696-driven effects specific to mature adipocytes, we found that *SNX10* expression in all GTEx samples correlated with expression of markers of mature adipocytes, but not with markers of progenitor cells or stromal cells (Supplementary Figure 16).

To assess whether *SNX10* has an effect on adipose tissue-related phenotypes at an organismal level, we next developed a mouse knockout (KO) of *Snx10* specific to mature adipocytes, crossing adiponectin promoter-Cre mice with a floxed *Snx10* allele. Mice harboring an adipose-specific deletion of *Snx10* (mmΔSnx10^{Adipoq}), were subjected to a 12-week high fat diet (HFD, D12492, 60kcal% fat). Control male and female littermates of knockout animals displayed diet-induced obesity upon HFD administration, as did male mmΔSnx10^{Adipoq} mice. However, female mmΔSnx10^{Adipoq} mice were resistant to diet-induced obesity, primarily due to reduced fat mass (Figures 5b–d, Supplementary Figures 17–18), (p(FKO weight vs MKO weight)=$9.15*10^{-5}$, t=10.6,df=4,CI=16.2– 26.3) p(FKO fat percent vs MKO fat percent)=$4.00*10^{-4}$, t=7.23,df=4,CI=12.4–25.2). These data suggest that *Snx10* regulates diet-induced expansion of adipose tissue only in females, which in humans may explain its robust association with female-specific WHRadjBMI, despite shared eQTL effects of rs1534696 between males and females.

Discussion

Our work illustrates how sexual dimorphism on cellular and organismal traits can emerge despite shared variant-mediated gene regulation between both sexes. By closely investigating a genetic locus that is associated with fat metabolism in women, we found gene regulation of *SNX10* to be depot-specific, developmental time-point-specific, and sexually

dimorphic, potentially explaining the female-specific association of this gene with changes in fat distribution in humans[5,8,25].

Our data also suggest that human patterns of fat deposition may be, in part, regulated through a subset of co-opted, primate-specific *Alu* retrotransposons, the extension to which remains to be determined as well as the possible phenotypic differences emerging from the genetic variants within these *Alu* retrotransposososns. The involvement of primate-specific elements in body fat distribution could represent one mechanism for rapid evolution of this trait within the human lineage. Humans have a body fat phenotype unique among great apes, characterized by increased body fat, especially in women[48]. If primate-specific *Alu* elements are involved in the evolution of human-specific body fat distribution, then some gene targets of *Alu* elements may be missed by approaches that focus on non-primate model organisms. Recent reports have illustrated examples of repeat elements participating in transcriptional networks involved in mammalian pregnancy and species-specific interferon-mediated immune responses[49,50]. A recent report has shown that pancreatic beta cell enhancers in loci associated with type 2 diabetes and responsive to endoplasmic reticulum (ER) stress are often *Alu* retrotransposons[51]. Interestingly, we found that the *Alu* AGGTCA motif we found enriched in our MPRA is also overrepresented in enhancer-modulating variants and enhancers in their ER-stress MPRA (X-squared = 6.6442, df = 1, p-value = 0.009948, Supplementary figure 18). Future work will determine the extent to which *Alu* repeats participate in gene regulation in adipocytes and shape regulatory landscapes of genes involved with metabolic traits and diseases in humans.

We characterized the adipose-related functions of *SNX10*, the strongest target gene of WHRadjBMI association[5,8,25]. *SNX10* is a sorting nexin that regulates early endosome trafficking and is essential for the function of osteoclasts and gastric zymogenic cells[52,53]. We demonstrate that *SNX10* also plays a role in adipocyte differentiation, lipid accumulation, and expansion of adipose tissue in response to a high-fat diet, suggesting that the disruption of endosome trafficking may also affect the function of mature adipocytes, with effects on human metabolism.

We found that the diet-induced fat expansion was sexually dimorphic, with female mmΔSnx10^{Adipoq} mice being resistant to diet-induced fat expansion and obesity. While the GWAS signal at this locus shows a strong association with WHRadjBMI in women, but not in men, the likely causal variant for *SNX10* expression in human mature adipocytes, rs1534696, affects gene expression in both sexes. These findings illuminate a vexing problem in genetics, understanding how sexually dimorphic traits emerge from effects of regulatory variants on gene expression that is common to both sexes. Sexual dimorphism is present in many human traits, from height and weight to lifespan and disease burden[54,55]. Recently, the GTEx consortium[56] showed that 37% of genes have sex-biased expression, most only in a few tissues. However, this sex-biased expression is not driven by sex-biased effects of allelic variation. Of the almost 500,000 cis-eQTLs identified by GTEx, only 369 (less than 0.1%) had effects that were biased by sex[56].

Our data expand the knowledge about how the genome regulates adipocytes to control fat deposition, determining the impact of caloric excess on human health. By identifying

candidate genes regulating female-specific fat accumulation we highlight new genes and pathways that may lead to metabolic complications of obesity in women.

Methods

This research complies with all ethical regulations. Mouse care and procedures were approved by the University of Minnesota Institutional Animal Care and Use Committee (IACUC, protocol #1904–36987A). All other data using animal models and human data was previously generated. The relevant protocols can be found in the references section.

LDSC-seg:

LDSC-seg analyses were conducted using ldsc v1.0.0 as described by Finucane et al[11]. For each tissue, enrichment for a set of specifically-expressed genes was calculated in comparison to a baseline model measuring non-specific enrichment near genes, histone annotations, open chromatin, and enhancers. The set of specifically expressed genes was used as a functional annotation, and the enrichment of this annotation in linkage to GWAS variants of low p-value, conditioning on the baseline model and the expression of all genes. Tissues included all 53 tissues available from the GTEx v8[14]. Sample sizes were comparable to those found in Finucane et al.

xCell:

xCell analyses were performed using xCell version 1.1.0 as described in Aran et al.[12]. Sample sizes were comparable to those in Aran et al. All samples from GTEx v8 (n=17382, from 938 human subjects representing 53 tissues) were combined with frontal cortex data from the CommonMind Consortium (n=613 tissue samples)[13]. Enrichment analyses for 63 cell types and cellular landscapes were calculated, transformed to proportion estimates, and corrected for spillover between similar cell types. Estimations for similar cell types (β-cells, CD4+ T-cells, CD8+ T-cells, dendritic cells, macrophages, endothelial cells) were averaged, and outliers (smooth muscle) and stem cell types removed prior to visualization.

Genetic correlation:

Heritability estimation and genetic correlation were performed using cross-trait LD score regression in ldsc version 1.0.0 as described by Bulik-Sullivan et al. Summary statistics were gathered from sex-specific GWAS meta-analyses of obesity and WHRadjBMI[5,6]. Heritability was obtained by solving the LD score regression equation for h^2, given GWAS effect sizes and linkage disequilibrium data. Genetic correlation was obtained by normalizing genetic covariance, obtained via the product of the Z-scores for each variant, accounting for linkage disequilibrium and sample size by SNP heritability. Significance testing between sex-specific heritability was performed by calculating the Z-score of the difference between the means and reporting the associated two-tailed p value. Sample sizes were equivalent to other GWAS-based heritability and covariance estimation in Bulik-Sullivan et al.

Transcriptome-wide association studies:

For the TWAS investigating WHRadjBMI, we used the following data: 1) subcutaneous adipose tissue samples from GTEx v8 with matching genotype data, segregated by sex; 2) Sex-specific WHRadjBMI GWAS summary statistics; and 3) xCell estimates of cell type proportion derived from GTEx v8 and CommonMind Consortium data as described above. To allow effect size comparison between male and female WHRadjBMI TWAS, a random subset of male adipose subcutaneous samples was selected from GTEx that was the same size as the total available number of female samples(n=194). Sample sizes were based on power estimates performed by Gusev et al. Estimates for the 6 most prevalent cell types in subcutaneous adipose tissue (preadipocytes, adipocytes, dendritic cells, chondrocytes, epithelial cells, and fibroblasts) were included as covariates, as were participant age and race.

We used FUSION scripts to perform TWAS on gene expression. For WHRadjBMI, we performed both sex-combined and sex-specific analyses. Given sex-specific heritability of WHRadjBMI, we performed further analyses on sex-specific WHRadjBMI data. LD reference panels were made from the relevant subject population using LDAK version 5[57]. To perform TWAS analyses, we first estimated heritability of gene expression as a function of the genotypes of the variants in *cis*, using GCTA-GREML to compute h^2 and a p-value representing the result of a likelihood ratio test. Genes where $p>0.05$ after Bonferroni correction were included in further analyses. Predictive models were calculated estimating the effect of *cis*-variation on gene expression usage. Four different models were used: lasso, elastic net, top single eQTL, and best linear unbiased predictor. The best-performing model was selected via cross-validation, using shuffled gene identifiers to generate a null distribution. The imputed gene expression under the model was then correlated with WHRadjBMI summary statistics for the relevant sex. Bonferroni correction was applied to the resulting set of genes to yield a set of genes that are statistically associated with sex-specific WHRadjBMI under strict criteria.

For the TWAS investigating obesity, we used the following data: 1) frontal cortex samples derived from the CommonMind Consortium, and 2) Sex-combined GWAS summary statistics derived from the same samples as the WHRadjBMI GWAS. Cell type estimation was performed but cell type proportions were not included in obesity cortex TWAS due to the homogeneity of frontal cortex samples (Figure 1). We performed sex-separate and sex-combined TWAS for obesityObesity TWAS was otherwise performed as described for WHRadjBMI. Overlap significance between obesity TWAS, WHR_F TWAS, and WHR_M TWAS was evaluated with a label-swap approach, where sets of equal size to each TWAS-identified gene set were randomly sampled from the list of all significant genes. 100,000 iterations were performed to generate an empirical distribution of overlaps expected by chance.

Colocalization analyses were performed using *coloc*[37], version 3.1 to detect shared causal variants between eQTL and GWAS signal. Sample sizes were similar to those used by Giambartolomei et al. For each significant gene in each TWAS, *coloc* was performed between the GWAS summary statistics (e.g. female, male, or combined) and the sex-combined GTEx-supplied eQTL data for that gene. The locus used was +/−0.5 MB of the

canonical transcription start site (TSS) of the gene. A variant was considered a single shared causal variant if 1) it had the maximum posterior probability of being the single shared causal variant, and 2) the posterior probability of there being a single shared causal variant (PP.H4) was > 0.5. TWAS and *coloc* results for all studies are available in Supplementary Table 3. At the *SNX10* locus, a secondary colocalization analysis was performed to account for multiple female WHRadjBMI GWAS signals located within +/−0.5 MB of the *SNX10* TSS. The locus boundaries of the new locus were chr7:26200000–26400000(hg19).

Sex-specific eQTL analyses (Figure 2d) were performed on rank-normalized expression data from GTEx and matching genotypes using fastQTL version 2.0[58], including the first three principal components from the genotype data as covariates. To ensure a normal distribution, significance testing was performed using two-tailed Student's t-tests on log2tpm values. For comparison with MPRA data, imputation of female WHR GWAS summary statistics to variants identified in the MPRA was performed using 1000 genomes reference genotypes[59] and code provided by the Im lab (https://github.com/hakyimlab/summary-gwas-imputation/wiki, Supplementary Table 4.

UK Biobank (UKBB) GWAS locus zoom:

UKBB GWAS locus zooms were generated from sex-specific GWAS data generated by the Neale lab (http://www.nealelab.is/uk-biobank). Summary statistics were combined with linkage disequilibrium information generated from GTEx v8 genotype data. Points shown represent variants available in both UKBB GTEx v8 and UKBB.

Massively parallel reporter assay (MPRA):

Variant selection: For each gene identified as associated with female WHRadjBMI by TWAS, a collection of tag variants was created. Variants were included if they were either a) the top eQTL for the gene within 1MB of the gene's transcription start site; b) the variant with the lowest p-value for female WHR within 1MB of the gene's transcription start site, or c) the variant predicted to be the single shared causal variant that explains the eQTL or GWAS signal at that locus. This search resulted in 169 core variants.

Next, all variants were selected with close linkage to tag variants, defined as $r^2 > 0.95$. Loci were removed where 500 or more variants were in close linkage to allow for investigation of other loci. This resulted in a final list of 1882 biallelic variants included in the final analysis.

Oligonucleotide assembly: For each variant, a 175-bp sequence centered around the variant was extracted from the GrCh38 (h38) genome assembly. This sequence was edited to include each allele of the variant. Variants representing indels were excluded. Sequences were assembled with a unique 10-bp barcode, primers, and *Kpn1* and *Xba1* binding sites as follows: 3' | F Primer | Enhancer sequence | *Kpn1* binding site | *Xba1* binding site | barcode | R primer | 5'. 19 barcodes were included per allele of each variant.

3T3-L1 cell culture and transfection: 3T3-L1 preadipocyte cells (ATCC, #CL-173) were grown at 37°C in 5% CO2 in DMEM (Gibco # 11995–065) media supplemented with 10% FBS, 1% penicillin-streptomycin solution (10,000 U/ml; Gibco #15140122), 0.8mg/ml

Biotin (Sigma; #B4639) and 0.8mg/ml Panthotenic Acid. For transfection, cells were plated into 6 well plates and transfected with Lipofectamine LTX & Plus reagent (Invitrogen; #15338100) when 30–50% confluent.

MPRA library preparation: MPRA library design and preparation was performed as previously described in Ulirsch et al[26]. Briefly, oligonucleotides representing each allele of selected variants (see 'Variant selection' above) were synthesized on an Agilent 100k oligonucleotide array. Homology arms were added onto the fragments using emulsion PCR, which reduces PCR amplification bias. Oligos were then cloned into a linearized pMPRA1 vector (addgene #49349) using Gibson assembly. This pool of pMPRA1 vector plus inserts was then linearized and a truncated eGFP reporter gene containing a minimal promoter and spacer sequence was then cloned into the vector using T4 DNA ligase (NEB; # B0202S) to create the input DNA library. Once the final constructs were produced, they were transfected into 3T3-L1 preadipocytes (4 technical replicates) as described above. At least 10 million cells were transfected per replicate. Cells were collected 48 hours after transfection and flash frozen in liquid nitrogen. After all replicates were transfected, RNA was extracted using a Qiagen RNeasy mini kit. mRNA was isolated from total RNA using Invitrogen Dynabeads (ThermoFisher #61006). All mRNA was then converted to cDNA using Superscript III reverse transcriptase and then PCR amplified using Illumina multiplexing primers. Two reactions of input DNA (DNA used as transfection material) were also PCR amplified with Illumina multiplexing primers at this point. Libraries were amplified with 10–11 PCR cycles and pooled before being cleaned. Library quality was assessed using the Agilent DNA 1000 bioanalyzer chip, where a single sharp peak of around 250bp is expected. Samples were then sent for paired end NGS sequencing. A 25% PhiX genome spike in was added to each sequencing run due to low complexity.

Significance testing: For each technical replicate of the MPRA, counts per barcode were log2cpm format. Barcodes where log2cpm <= −3 were removed, as were variants where either allele was left with fewer than 7 barcodes after removal. Enhancer activity per barcode was defined as log2cpm(replicate) – log2cpm(input). A nonparametric (Mann-Whitney U) test was performed to identify significant MPRA enhancers, comparing each allele of a variant with all other sequences included in the analysis. Subsequent FDR correction was performed on the set of significant MPRA enhancers to correct for multiple comparisons. Enhancers are considered significant if FDR q<= 0.05 in two or more replicates. Within the set of significant MPRA enhancers, subsequent Mann-Whitney U tests were performed to identify significant enhancer-modulating variants (enhancer-modulating variants), comparing one allele of the significant MPRA enhancer versus the other. Subsequent FDR correction was performed and enhancer-modulating variants with FDR q<=0.05 in two or more replicates were considered to be significant. Re-analysis of Ulirsch et al. data for comparison to our MPRA (Supplementary Figure 8) was limited by low power, as few of their significant variants overlapped *Alu* elements.

Haplotype estimation: To determine which alleles of MPRA variants were co-inherited, haplotype blocks were calculated with HaploView version 4.2, using strong LD spine to identify haplotype blocks with a D' setting of 0.4 to ensure capture of all variants in LD with

the prime variant in the same haplotype block, and to allow MPRA lead variants in LD with one another to be compared.

Enrichment of significant MPRA enhancers in functional chromatin:

15-mark chromHMM marks for adipose nuclei were derived from the Epigenome Roadmap (https://egg2.wustl.edu/roadmap/data/byFileType/chromhmmSegmentations/ChmmModels/coreMarks/jointModel/final/). Active chromHMM marks were defined as active TSS(1_TssA), flanking active transcription start sites (TSS) (2_TssAFlnk), transcription at 3' and 5' gene ends (3_TxFlnk), strong transcription (4_Tx), weak transcription (5_TxWk), genic enhancers (6_EnhG), enhancers (7_Enh), bivalent/poised TSS (10_TssBiv), flanking bivalent TSS/enhancers (11_BivFlnk), and bivalent enhancers (12_EnhBiv). Chi-square testing was performed to see if more MPRA enhancers lay within active chromHMM annotations vs nonsignificant MPRA variants than expected by chance.

MPRA Motif analysis:

Significantly enriched motifs were identified using HOMER version 3.12, using the set of known motifs included in the HOMER motif database. Motif detection was done using the 175-bp sequence surrounding each variant in hg38, using the reference allele of each variant. To correct for strand-specific biases introduced by autonormalization in Homer v3, the analysis was performed on both the sequences and their reverse complements. Motifs were considered enriched if $p<0.05$ and FDR $q<0.05$. Motif enrichment was identified in a) the set of significant MPRA enhancers compared to nonsignificant tested sequences, and b) the set of enhancer-modulating variants compared to nonsignificant tested sequences. The location of the common adipose motif was defined as the span from the earliest start of any of the transcription factor motifs which contributed to this motif (*COUP-TFII, EAR2, ERRA, LXRE, RARA, THRB, ATF1, USF2*) to the latest end of any of these motifs. The ability to detect small effect sizes was limited by small sample size in the motif enrichment analyses.

Luciferase assays:

Construct Design: To further validate candidate variants from the MPRA, 18 variants were tested using a luciferase reporter assay. gBlock data are contained within Supplementary Table 7. ~1 kb genomic sequence fragments containing each variant were constructed (gBlocks Gene Fragments, IDT). Each variant construct was cloned at a 3:1 insert-to-vector ratio into the pGL4.23[luc2/minP] vector (Promega) that contains firefly luciferase. DNA sequencing was used to verify proper cloning. For each luciferase assay, mouse preadipocyte cells (3T3-L1) were transfected with each construct in triplicate. Four independent luciferase assays were conducted. 3T3-L1 cells were also transfected with a negative control construct (scramble) using a DNA sequence (chr9:6,161,550–6,162,024) devoid of any active chromatin epigenetic marks.

Assays: 3T3-L1 cells were seeded in 96-well plates at 5,000 cells per well. 24 hours after plating when cells were at approximately 50% confluency, 3T3-L1 cells were co-transfected with 100 ng of each construct and 10 ng of pGL4.73[hRluc/SV40] *Renilla*

luciferase reporter vector (Promega) to normalize transfection efficiency. Co-transfection of each construct occurred in triplicate using Lipofectamine LTX with Plus Reagent (Thermo Fisher Scientific, #153308030). Firefly and *Renilla* luciferase activity were measured using the Dual-Luciferase Reporter Assay System (Promega, #E1960) 48 hours post-transfection in which cells were washed with PBS, lysed with 1X passive lysis buffer on a rocking platform for 15 minutes at room temperature, and stored at −20 °C until luciferase activity was measured. The Dual-Luciferase assay was conducted four separate times using different DNA Lipofectamine transfection preparations where the cells were transfected, collected, and luciferase activity was measured on different days.

Analysis: Four technical replicates of the luciferase assay were conducted where firefly and *Renilla* luciferase of each construct was measured in triplicate per assay. For all three measurements of each construct and the control, firefly luciferase activity was normalized to *Renilla* luciferase activity (and the average luciferase/*Renilla* ratio was calculated (luciferase/*Renilla* will now be referred to as luciferase activity.). To identify significant increases in luciferase activity for each construct, the average luciferase activity for each construct was compared to the average luciferase activity for the control per technical replicate ($n = 4$) using a one-tailed paired t-test and a significant p-value threshold of $p < 0.05$. Our ability to detect small effects in the luciferase assays is limited by small sample sizes. L ggbreak R package, version 0.1.0 was used to create y-axis breaks[60].

CRISPR-cas9 deletion of *Alu* elements in human mesenchymal stem cell (hMSC):

CRISPR-cas9 deletion: Undifferentiated hMSC cells (Lonza, Walkersville, MD, USA, Document #AA-2501–16 07/11) were grown in mesenchymal stem cell growth medium (MSCGM, #PT-3001), containing mesenchymal stem cell basal medium (MSCBM, # PT-3238), mesenchymal cell growth supplement (MCGS, # PT-4105), L-glutamine, and gentamycin/ampicillin (GA-1000) SingleQuotsTM. Cells were plated at 5.0×10^3 cells/cm^2 in 6-well plates (with cultures allowed to reach 50%–60% confluence prior to transfection. Two guides were designed per region in order to delete three regions containing *Alu* elements. IDT Alt-R® S.p. Cas9 Nuclease V3 (IDT, # 1081059) and Alt-R® CRISPR-Cas9 tracrRNA (IDT, #1072532) were complexed according to the manufacturer's instructions. For the transfection, 50 μM of each guide and cas9 protein were combined and incubated at room temperature for 20 min to form RNPs. One million undifferentiated hMSC were transfected using the AMAXA P1 primary cell 4D nucleofector X (Lonza, # V4XP-1012) kit and pulse FF104. Cells were plated in a 6-well plate. After three days, individual cells were are sorted into 96-well plates. Colonies were genotyped by PCR and expanded. PCR primers for genotyping and guides are in Supplementary Table 8.

RNA-sequencing assay: Total RNA was extracted using the RNeasy Micro Kit (Qiagen, #74004). The total RNA quality was assessed via an Agilent RNA 6000 Pico Kit (Agilent, # 5067–1513). Two nanograms total RNA was used for cDNA synthesis (Clonetech, #634896) and 125pg cDNA for the RNA-sequencing library preparation using the Nextera XT DNA Library Prep Kit (Illumina, #FC-131–1024). Library size and replicate number was based on analyses in Joslin et al. We generated 50–37 bp paired-end sequencing at the University of Chicago Genomics Core facility. We then aligned the data using STAR

v2.7.10a[61]with an index derived from the GRCh38.p13 human genome. Gene level counts were produced using the htseq-count utility in STAR (--quantMode gene). We used edgeR v3.38.1 for differential expression analysis of counts data. We employed pre-filtering and TMM normalization, followed by modeling allelic deletion effect sizes using the quasi-likelihood generalized linear model approach ($p<0.05$)[62,63]. We controlled type-I error using FDR corrections and a 5% threshold on q-values. Differential gene expression was defined relative to other genes in the *cis*-regulatory region (1MB window around the deletion). Experimental batch variables were included in models to account for global differences in gene expression arising from differences in experimental preparation and cell culture. Code for alignment and differential expression analysis can be found at https://github.com/zwebbs/Hansen-et-al-RNA-seq.

Repeat element enrichment:

Repeat element location and class were obtained from the RepeatMasker data available for assembly GrCh38 from UCSC (http://hgdownload.cse.ucsc.edu/goldenpath/hg38/database/). The *Alu* canonical sequence was obtained from RepeatMasker at http://www.repeatmasker.org/AluSubfamilies/humanAluSubfamilies.html. The sequences of all *Alu* elements overlapping an MPRA variant were aligned using Clustal Omega, version 1.2.4, using full distance matrix calculation. The resulting guidetree was visualized as a cladogram (no branch length information) using ggtree, with leaves labeled according to MPRA significance.

Repeat Element Conservation:

Conservation of repeat elements was determined using the PhastCons score generated by UCSC using a 20-way comparison between mammalian species (http://hgdownload.cse.ucsc.edu/goldenPath/hg38/phastCons20way/). Conservation was converted to a per-base score, and the average score across the 175-bp sequences used in the MPRA was calculated and compared between enhancer-modulating variant sequences, significant MPRA enhancer sequences, and nonsignificant sequences. Significance was evaluated using two-tailed Student's wt-tests between the three categories of variants.

Lipocyte Profiler in human adipose-derived mesenchymal stem cells

LipocyteProfiler data were obtained from Laber et al. (ref. 46). In brief, lipocyte profiles of morphological and cellular features were generated from a genotyped cohort of subcutaneous and visceral adipose-derived mesenchymal stem cells from patients undergoing abdominal laparoscopic surgeries. Visceral adipose tissue was derived from the proximity of the angle of His and subcutaneous adipose tissue was obtained from beneath the skin at the site of surgical incision. Pre-adipocytes were differentiated for 14 days. LipocytePainting was performed at four time-points of differentiation (days 0, 3, 8 14) and features were extracted as described in Laber et al[46]. We compared the lipocyte profiles between rs1534696 CC/CA and AA allele carriers using a multi-way analysis of variance (ANOVA) adjusted for BMI, age and batch. P-values were corrected for multiple testing using the false discovery rate approach (FDR). Features with FDR < 10% were classified to be significantly different between risk and non-risk haplotype carriers. Features that represented measurements of the same quantity at different scales

(e.g. Nuclei_Texture_Variance_DNA_20_01,Nuclei_Texture_Variance_DNA_20_01) were collapsed to avoid redundancy, and the largest effect size of the group was retained.

Activity-by-contact modeling of adipocyte enhancer-promoter interactions:

Activity-by-contact (ABC) was performed as previously described[64]. In brief, ABC predictions were modelled based on H3K27ac ChIPseq and ATACseq[65] from primary human adipose-derived mesenchymal stem cells from one female donor at four time-points of differentiation (days 0, 2, 7, 14). Sample sizes were similar to those chosen by Fulco et al.[64]

shRNA-mediated knockdown of *SNX10* during human adipocyte differentiation:

Adipogenesis differentiation in human mesenchymal stem cells (hMSC) was followed as per the manufacturer's instruction (Lonza, Walkersville, MD, USA, Document #AA-2501–16 07/11). hMSC 2.1×10^4 cells/well were plated on a 24 well plate in MSCGM™ at 37°C, 5% CO_2 until confluency. At 100% confluence, the MSCGM™ were replaced with adipogenic induction media supplemented with SingleQuots™ of h-insulin (recombinant), L-glutamine, MCGS, dexamethasone, indomethacin, IBMX (3-isobutyl-l-methyl-xanthine) and GA-1000 (Lonza, Walkersville, MD, USA). hMSC cells were transfected with shRNA control (Sigma-Aldritch #SHC202V) and *SNX10* MISSION shRNA Lentiviral Transduction Particles (Sigma-Aldritch #SHCLNV) 5 hours prior to adding the adipogenic induction media. hMSC adipogenic cultures were incubated with shRNA for 3 days, and f the medium was replaced with Adipogenic Maintenance Medium supplemented with SingleQuots™ of h-insulin (recombinant), L-glutamine, MCGS and GA-1000 (Lonza, Walkersville, MD, USA). The cells again were transfected with *SNX10* MISSION shRNA Lentiviral Transduction Particles 5 hours prior to adding the Adipogenic Maintenance Media and cultured for 3 days. The hMSC cells were maintained in induction/maintenance media for two more cycles without shRNA for optimal adipogenic differentiation. The cells were then maintained in Adipogenic Mainteinance Medium for 7 more days for a total of 21 days differentiation. hMSC adipogenic cultures were harvested for RNA as well as stained for Oil Red O. Two independent experiments were performed. For RT-qPCR, three measurements were taken for each experiment. Our ability to detect evaluate significance in the shRNA knockdown studies is limited by small sample sizes.

Generation of Adipose tissue-specific *Snx10*-deficient mice:

We obtained the *Snx10* targeting vector, PG00216_Z_2_C06, from the European Conditional Mouse Mutagenesis Program (EUCOMM). This vector is a "knockout first" gene trap which inserts a flippase site-flanked Neo selection cassette with an IRES and LacZ reporter into intron 3 and inserts LoxP sites flanking exons 4 and 5 (Snx10Neo-f allele) which. contain the functional domain of *Snx10*, the PX domain required for phospholipid interactions. The vector was electroporated into V6.5 ES cells. Infection was carried out at a multiplicity of infection (MOI) =1. Control cells were infected with scrambled *Snx10* shRNA virus. Both *SNX10* MISSION shRNA Lentiviral Transduction Particles and shRNA Control Transduction Scrambled Particles were purchased from Sigma-Aldrich. Neomycin resistant clones were picked, expanded and screened for correct insertion by Long Range PCR using vector-specific primers and gene-specific primers (Supplementary Table 11). The

resulting clones were injected into host blastocysts to generate mouse chimeras. Two male 129 chimeras were bred with C57BL/6J females. Germ line transmission was confirmed by PCR, using primers flanking the third loxP site (Supplementary Table 11). Snx10Neo-f/+ offspring were bred to homozygosity to generate animals at the expected Mendelian ratio of 25%. PCR genotyping using the preceding primer pair produced bands of 213 bp for wild type (+/+), 273 bp for Snx10Neo-f/Neo-f, and both bands for heterozygotes Snx10Neof/+.

Adipose tissue-specific *Snx10*-deficient mice were generated in two steps. First, the Neo cassette was removed by crossing Snx10Neo-f/+ males with females homozygous for Rosa26-driven FLP recombinase (strain B6.129S4-Gt(ROSA)26 Sortm1(FLP1)Dym/RainJ), resulting in an allele with exons 4 and 5 flanked by loxP sites (Snx10fl/+). Then, homozygous males were crossed with female mice hemizygous for the Adipoq-Cre BAC transgene (B6.FVB-Tg(Adipoq-cre)1Evdr/J Stock No: 028020)[66], in order to breed Adipoq;Snx10fl/+ males which were backcrossed with Snx10fl/fl females. The resulting mice carry an adipose tissue-specific *Snx10* allele with exons 4 and 5 deleted, for adipose tissue-specific knockout. The level of gene expression was detected with TaqMan primers for *SNX10* (assay ID Hs01007226_m1; cat no- 4351372; Applied Biosystems; Thermo Fisher Scientific, Inc.); *ADIPOQ* (assay(Hs00605917_m1; cat no- 4331182; Applied Biosystems; Thermo Fisher Scientific, Inc.); and the endogenous control β-ACTIN gene (assay ID Hs01060665_g1; cat. no. 4331182; Applied Biosystems; Thermo Fisher Scientific, Inc.).

In vivo measurement of fat and lean body mass:

Dual Energy X-ray Absorptiometry (DEXA) analysis was performed using the InAlyzer DEXA (Microphotonics InAlyzer Type S, Serial No. INZ-201021-S002, MEDIKORS). The values for mass, fat mass, lean mass, body fat percentage, and femur length were obtained. Magnetic resonance imaging of body composition estimation was acquired. Data collection and analysis were not performed blind to the conditions of the experiments. Significance testing of was performed with two-sided Student's t-tests. Our ability to detect small effects in the high-fat diet studies is limited by small sample sizes.

Statistics and reproducibility:

Sample sizes for analyses performed in this paper were based on previous research, where available, or from power analyses. Power was considered sufficient if there was a >50% chance of finding a moderate-sized effect (defined as an effect size d >0.5 for t-tests and w> 0.3 for chi-square tests). Preferred sample sizes for motif enrichment analyses based on MPRA data, luciferase assays, shRNA data, and high-fat diet mouse data could not be achieved due to experimental limitations; this limits our ability to detect small effect sizes for these analyses and is noted in the text. Randomization was not required for analyses, except to generate empirical p distributions as described in the Methods. Investigators were not blinded to genotype status when performing DEXA scanning of mmΔSnx10$^{\mathbf{Adipoq}}$ mice, data analysis of luciferase data, or data analysis of CRISPR deletion data. No data were excluded from this study, except as based on quality control analyses where noted in the Methods section. Normality testing was performed for all analyses that assumed a normal

distribution except for analyses of *SNX10* knockdown shRNA and mmΔSnx10^{Adipoq} mice, due to small sample sizes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the Novo Nordisk Foundation (challenge grant NNF18OC0033754 to M.A.N.), Novo Nordisk Foundation grant NNF21SA0072102 to M.C; the National Institutes of Health (grants R01HL128075, R01HL119577 to M.A.N.; grant R01AR064793 to R.B), and the American Heart Association (grant 20PRE35210899 to G.T.H.). We thank Aparecida Candles for support with the manuscript.

Data Availability

Previously published data that are used in this study can be accessed via the GTEx consortium, the CommonMind consortium, and the GIANT consortium(https://gtexportal.org/home/biobank, https://www.nimhgenetics.org/resources/commonmind, https://portals.broadinstitute.org/collaboration/giant/index.php/GIANT_consortium_data_files. Sequencing data generated in this study can be found in the NCBI Sequence Read Archive at BioProject PRJNA847334. Tables that contain data generated in this study, including ldsc_seg enrichment for obesity and WHR, xCell cell type proportion estimations, obesity and WHRadjBMi-associated genes identified in the paper, variants tested in the MPRA, with enhancer activity levels and significance status, transcription factor motif enrichment data, significant Lipocyte Profiler features, *SNX10* shRNA knockdown and *SNX10* mouse data, have been uploaded to Open Science Framework in a repository entitled “Genetics of sexually dimorphic adipose distribution in humans”

References

1. Alexopoulos N, Katritsis D & Raggi P Visceral adipose tissue as a source of inflammation and promoter of atherosclerosis. Atherosclerosis 233, 104–112 (2014). [PubMed: 24529130]
2. Canoy D Distribution of body fat and risk of coronary heart disease in men and women. Curr. Opin. Cardiol. 23, 591–8 (2008). [PubMed: 18830075]
3. Bergman RN et al. Why Visceral Fat is Bad: Mechanisms of the Metabolic Syndrome. Obesity 14, 16S–19S (2006). [PubMed: 16642958]
4. Censin JC et al. Causal relationships between obesity and the leading causes of death in women and men. PLOS Genet. 15, e1008405 (2019). [PubMed: 31647808]
5. Pulit SL et al. Meta-analysis of genome-wide association studies for body fat distribution in 694 649 individuals of European ancestry. Hum. Mol. Genet. 28, 166–174 (2019). [PubMed: 30239722]
6. Yengo L et al. Meta-analysis of genome-wide association studies for height and body mass index in $\sim$700000 individuals of European ancestry. Hum. Mol. Genet. 27, 3641–3649 (2018). [PubMed: 30124842]
7. Locke AE et al. Genetic studies of body mass index yield new insights for obesity biology. Nature 518, 197–206 (2015). [PubMed: 25673413]
8. Shungin D et al. New genetic loci link adipose and insulin biology to body fat distribution. Nature 518, 187–196 (2015). [PubMed: 25673412]
9. Cannon ME et al. Open Chromatin Profiling in Adipose Tissue Marks Genomic Regions with Functional Roles in Cardiometabolic Traits. Genes Genomes Genet. 9, 2521–2533 (2019).

10. Finucane HK et al. Partitioning heritability by functional annotation using genome-wide association summary statistics. Nat. Genet. 47, 1228–1235 (2015). [PubMed: 26414678]
11. Finucane HK et al. Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. Nat. Genet. 50, 621–629 (2018). [PubMed: 29632380]
12. Aran D, Hu Z & Butte AJ xCell: digitally portraying the tissue cellular heterogeneity landscape. Genome Biol. 18, 220–220 (2017). [PubMed: 29141660]
13. Hoffman GE et al. CommonMind Consortium provides transcriptomic and epigenomic data for Schizophrenia and Bipolar Disorder. Sci. Data 6, 180 (2019). [PubMed: 31551426]
14. The Genotype-Tissue Expression (GTEx) project | Nature Genetics. https://www.nature.com/articles/ng.2653.
15. Merrick D et al. Identification of a mesenchymal progenitor cell hierarchy in adipose tissue. Science 364, (2019).
16. Palmer BF & Clegg DJ The sexual dimorphism of obesity. Mol. Cell. Endocrinol. 402, 113–9 (2015). [PubMed: 25578600]
17. Chooi YC, Ding C & Magkos F The epidemiology of obesity. Metabolism. 92, 6–10 (2019). [PubMed: 30253139]
18. Paeratakul S, Lovejoy JC, Ryan DH & Bray GA The relation of gender, race and socioeconomic status to obesity and obesity comorbidities in a sample of US adults. Int. J. Obes. 26, 1205–1210 (2002).
19. Borgeraas H Association of time of obesity onset with comorbidities in treatment-seeking men and women with severe obesity. Obes. Sci. Pract 4, 427–436 (2018). [PubMed: 30338113]
20. Bulik-Sullivan B et al. An atlas of genetic correlations across human diseases and traits. Nat. Genet. 47, 1236–41 (2015). [PubMed: 26414676]
21. Gusev A et al. Integrative approaches for large-scale transcriptome-wide association studies. Nat. Genet. 48, 245–252 (2016). [PubMed: 26854917]
22. Grarup N et al. Loss-of-function variants in ADCY3 increase risk of obesity and type 2 diabetes. Nat. Genet. 50, 172–174 (2018). [PubMed: 29311636]
23. Dubern B et al. Mutational analysis of melanocortin-4 receptor, agouti-related protein, and alpha-melanocyte-stimulating hormone genes in severely obese children. J. Pediatr. 139, 204–209 (2001). [PubMed: 11487744]
24. Fall T & Ingelsson E Genome-wide association studies of obesity and metabolic syndrome - ScienceDirect. https://www.sciencedirect.com/science/article/pii/S0303720712004133.
25. Justice AE et al. Protein-coding variants implicate novel genes related to lipid homeostasis contributing to body-fat distribution. Nat. Genet. 51, 452–469 (2019). [PubMed: 30778226]
26. Ulirsch JC et al. Systematic Functional Dissection of Common Genetic Variation Affecting Red Blood Cell Traits. Cell 165, 1530–1545 (2016). [PubMed: 27259154]
27. Melnikov A, Zhang X, Rogov P, Wang L & Mikkelsen TS Massively parallel reporter assays in cultured mammalian cells. J. Vis. Exp. JoVE (2014).
28. Xu Z, Yu S, Hsu C-H, Eguchi J & Rosen ED The orphan nuclear receptor chicken ovalbumin upstream promoter-transcription factor II is a critical regulator of adipogenesis. Proc. Natl. Acad. Sci. 105, 2421–2426 (2008). [PubMed: 18250317]
29. Small KS et al. Regulatory variants at KLF14 influence type 2 diabetes risk via a female-specific effect on adipocyte size and body composition. Nat. Genet. 50, 572–580 (2018). [PubMed: 29632379]
30. Seo JB et al. Activated Liver X Receptors Stimulate Adipocyte Differentiation through Induction of Peroxisome Proliferator-Activated Receptor γ Expression. Mol. Cell. Biol. 24, 3430–3444 (2004). [PubMed: 15060163]
31. Lu C & Cheng S-Y Thyroid hormone receptors regulate adipogenesis and carcinogenesis via crosstalk signaling with peroxisome proliferator-activated receptors. J. Mol. Endocrinol. 44, 143–154 (2010). [PubMed: 19741045]
32. Guo L, Li X & Tang Q-Q Transcriptional Regulation of Adipocyte Differentiation: A Central Role for CCAAT/Enhancer-binding Protein (C/EBP) β. J. Biol. Chem. 290, 755–761 (2015). [PubMed: 25451943]

33. Cao J, Ma Y, Yao W, Zhang X & Wu D Retinoids Regulate Adipogenesis Involving the TGFβ/SMAD and Wnt/β-Catenin Pathways in Human Bone Marrow Mesenchymal Stem Cells. Int. J. Mol. Sci. 18, (2017).
34. Batchvarova N, Wang XZ & Ron D Inhibition of adipogenesis by the stress-induced protein CHOP (Gadd153). EMBO J. 14, 4654–4661 (1995). [PubMed: 7588595]
35. Fox KE et al. Depletion of cAMP-response element-binding protein/ATF1 inhibits adipogenic conversion of 3T3-L1 cells ectopically expressing CCAAT/enhancer-binding protein (C/EBP) alpha, C/EBP beta, or PPAR gamma 2. J. Biol. Chem. 281, 40341–40353 (2006). [PubMed: 17071615]
36. Joslin AC et al. A functional genomics pipeline identifies pleiotropy and cross-tissue effects within obesity-associated GWAS loci. Nat. Commun. 12, 5253 (2021). [PubMed: 34489471]
37. Giambartolomei C et al. Bayesian Test for Colocalisation between Pairs of Genetic Association Studies Using Summary Statistics. PLoS Genet. 10, e1004383–e1004383 (2014). [PubMed: 24830394]
38. Terrados G et al. Genome-wide localization and expression profiling establish Sp2 as a sequence-specific transcription factor regulating vitally important genes. Nucleic Acids Res. 40, 7844–7857 (2012). [PubMed: 22684502]
39. Huss JM, Torra IP, Staels B, Giguère V & Kelly DP Estrogen-Related Receptor α Directs Peroxisome Proliferator-Activated Receptor α Signaling in the Transcriptional Control of Energy Metabolism in Cardiac and Skeletal Muscle. Mol. Cell. Biol. 24, 9079–9091 (2004). [PubMed: 15456881]
40. Casado M, Vallet VS, Kahn A & Vaulont S Essential Role in Vivo of Upstream Stimulatory Factors for a Normal Dietary Response of the Fatty Acid Synthase Gene in the Liver. J. Biol. Chem. 274, 2009–2013 (1999). [PubMed: 9890958]
41. Yeo CR et al. SGBS cells as a model of human adipocyte browning: A comprehensive comparative study with primary human white subcutaneous adipocytes. Sci. Rep. 7, 4031 (2017). [PubMed: 28642596]
42. Yu K et al. Activating transcription factor 4 regulates adipocyte differentiation via altering the coordinate expression of CCATT/enhancer binding protein β and peroxisome proliferator-activated receptor γ. FEBS J. 281, 2399–2409 (2014). [PubMed: 24673832]
43. Xue JC, Schwarz EJ, Chawla A & Lazar MA Distinct stages in adipogenesis revealed by retinoid inhibition of differentiation after induction of PPARgamma. Mol. Cell. Biol. 16, 1567–1575 (1996). [PubMed: 8657131]
44. Kundaje A et al. Integrative analysis of 111 reference human epigenomes. Nature 518, 317–330 (2015). [PubMed: 25693563]
45. Fulco CP et al. Systematic mapping of functional enhancer–promoter connections with CRISPR interference. Science 354, 769–773 (2016). [PubMed: 27708057]
46. Laber S et al. Discovering cellular programs of intrinsic and extrinsic drivers of metabolic traits using LipocyteProfiler. bioRxiv.
47. Smemo S et al. Obesity-associated variants within FTO form long-range functional connections with IRX3. Nature 507, 371–375 (2014). [PubMed: 24646999]
48. Pontzer H et al. Metabolic acceleration and the evolution of human brain size and life history. Nature 533, 390–392 (2016). [PubMed: 27144364]
49. Lynch VJ et al. Ancient Transposable Elements Transformed the Uterine Regulatory Landscape and Transcriptome during the Evolution of Mammalian Pregnancy. Cell Rep. 10, 551–561 (2015). [PubMed: 25640180]
50. Chuong EB, Elde NC & Feschotte C Regulatory evolution of innate immunity through co-option of endogenous retroviruses. Science 351, 1083–1087 (2016). [PubMed: 26941318]
51. Khetan S et al. Functional characterization of T2D-associated SNP effects on baseline and ER stress-responsive β cell transcriptional activation. Nat. Commun. 12, 5242 (2021). [PubMed: 34475398]
52. Qin B, He M, Chen X & Pei D Sorting Nexin 10 Induces Giant Vacuoles in Mammalian Cells. J. Biol. Chem. 281, 36891–36896 (2006). [PubMed: 17012226]

53. Ye L et al. Osteopetrorickets due to Snx10 Deficiency in Mice Results from Both Failed Osteoclast Activity and Loss of Gastric Acid-Dependent Calcium Absorption. PLOS Genet. 11, e1005057 (2015). [PubMed: 25811986]
54. Teriokhin AT, Budilova EV, Thomas F & Guegan J-F Worldwide variation in life-span sexual dimorphism and sex-specific environmental mortality rates. Hum. Biol. 76, 623–641 (2004). [PubMed: 15754977]
55. Mauvais-Jarvis F et al. Sex and gender: modifiers of health, disease, and medicine. The Lancet 396, 565–582 (2020).
56. Oliva M et al. The impact of sex on gene expression across human tissues. Science 369, (2020).

Methods references

57. Berrandou T-E, Balding D & Speed D LDAK-GBAT: fast and powerful gene-based association testing using summary statistics. medRxiv (2022).
58. Ongen H, Buil A, Brown AA, Dermitzakis ET & Delaneau O Fast and efficient QTL mapper for thousands of molecular phenotypes. Bioinformatics 32, 1479–1485 (2016). [PubMed: 26708335]
59. Siva N 1000 genomes project. Nat. Biotechnol. 26, (2008).
60. Xu S et al. Use ggbreak to Effectively Utilize Plotting Space to Deal With Large Datasets and Outliers. Front. Genet. 12, 2122 (2021).
61. Dobin A et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21 (2013). [PubMed: 23104886]
62. McCarthy DJ, Chen Y & Smyth GK Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. Nucleic Acids Res. 40, 4288–4297 (2012). [PubMed: 22287627]
63. Robinson MD, McCarthy DJ & Smyth GK edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139–140 (2010). [PubMed: 19910308]
64. Fulco CP et al. Activity-by-contact model of enhancer–promoter regulation from thousands of CRISPR perturbations. Nat. Genet. 51, 1664–1669 (2019). [PubMed: 31784727]
65. Sinnott-Armstrong N et al. A regulatory variant at 3q21.1 confers an increased pleiotropic risk for hyperglycemia and altered bone mineral density. Cell Metab. 33, 615–628 (2021). [PubMed: 33513366]
66. Eguchi J et al. Transcriptional Control of Adipose Lipid Handling by IRF4. Cell Metab. 13, 249–259 (2011). [PubMed: 21356515]

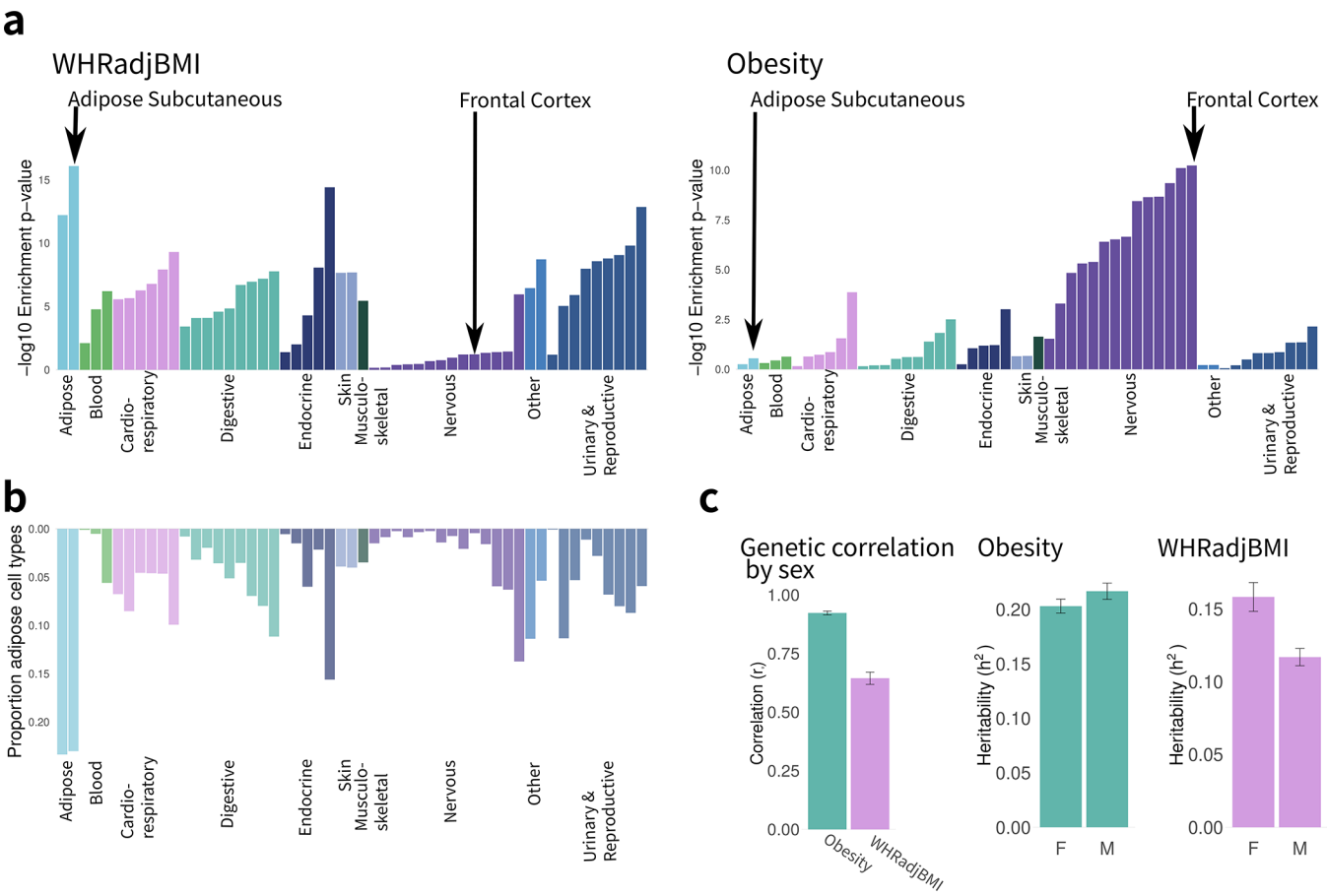

Figure 1: Characteristics of obesity and WHRadjBMI genetic risk.

a) Linkage disequilibrium score regression (LDSC-seg) measuring enrichment of trait near genes with tissue-specific expression. P-values are based on standard error, and represent contribution of tissue-specific expression to trait heritability after correction for baseline annotations. Bars are colored by tissue category. b) xCell-derived cell type enrichment predictions, normalized to generate proportions and averaged across adipose cell types (preadipocyte and adipocyte. Bars are colored by tissue category. c) Genetic correlation and heritability for sex-specific genome-wide association studies (GWAS) conducted for obesity and WHRadjBMI (obesity n~681,275 individuals, WHRadjBMI n=694,649 individuals). Bars are colored by trait. Data are provided as correlation or heritability estimate +/− standard error. Genetic correlation is high between sexes in obesity and lower for WHRadjBMI. The sexual dimorphism of WHRadjBMI is partially driven by increased heritability of WHRadjBMI in women. Welch's two-sided t-test was used to evaluate within-trait heritability between sexes (WHR $p=3.3*10^{-4}$, obesity p=0.16)

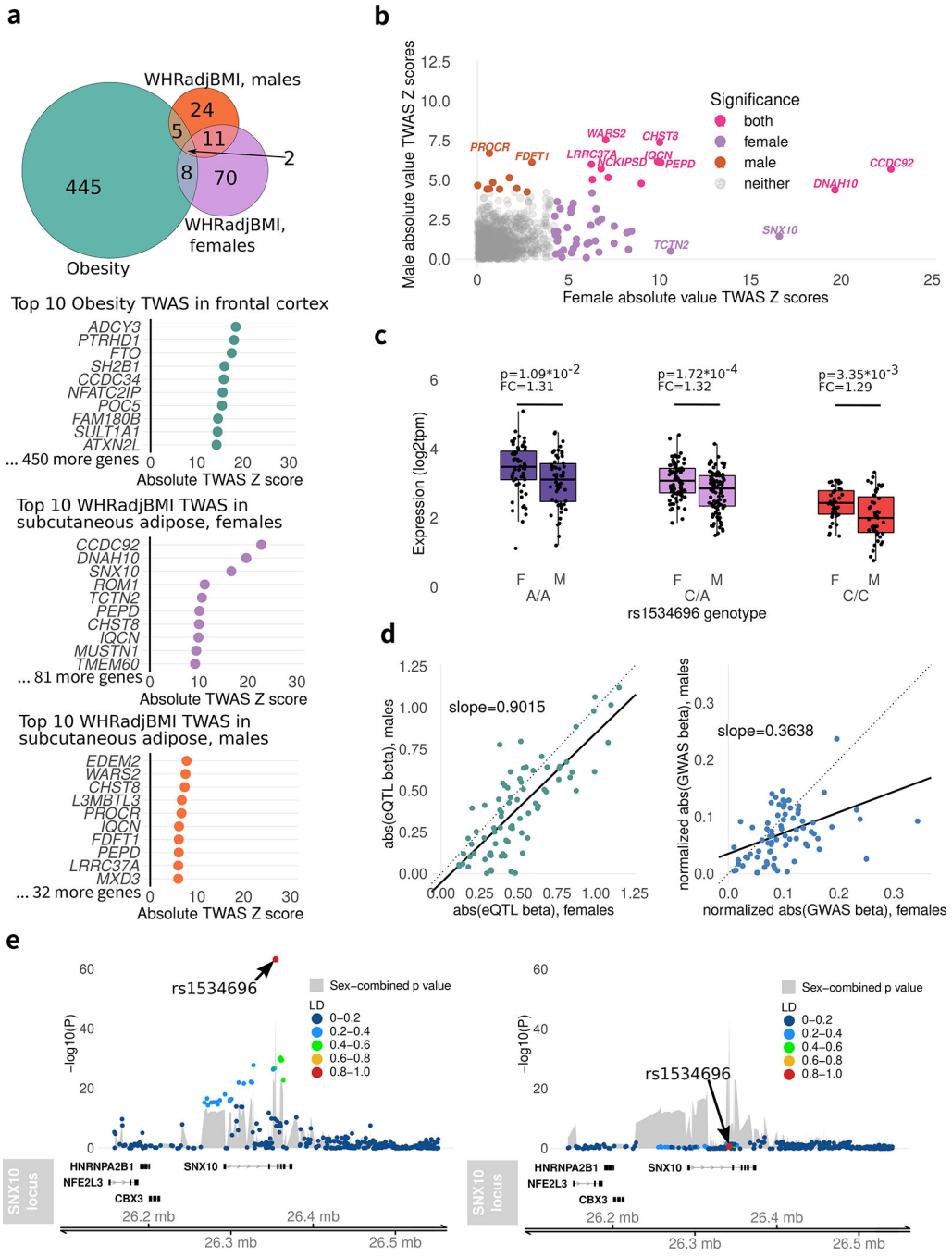

Figure 2: Genes identified by TWAS.

a) Overlap between TWAS-identified gene sets, and top results for each TWAS. Results show some known and some novel genes associated with obesity and WHR. Significance based on permutation distribution with subsequent Bonferroni correction. Data are colored by sex and trait. b) Sexual dimorphism of WHRadjBMI TWAS results. Every gene that was tested in both male and female TWAS is displayed. Data are colored by sex. c) rs1534696 eQTL effect on *SNX10* in women and men. Data are colored by rs1534696 genotype. P-values between sexes at each rs1534696 genotype obtained via Student's two-sided t-test of

log2tpm values. (females: 60 A/A participants, 91 C/A, 42 A/A. Males: 54 A/A participants, 96 C/A, 43 C/C). Box plot whiskers represent minimum (1st percentile) and maximum (99th percentile) of data, box bounds represent first quartile (25th percentile and 75th percentile), and center line represents median of data (50th percentile).d) Sexual dimorphism of eQTL and GWAS for female WHR. Data are colored by data type (eQTL vs GWAS). Each dot represents the top eQTL for a female WHR TWAS gene. GWAS betas are normalized by sex-specific trait heritability (h^2). On the left is the p-value of that variant as a sex-specific eQTL; on the right is the p-value of that variant in WHRadjBMI sex-specific GWAS. Variants show sexual dimorphism in GWAS, but not as eQTLs. e) GWAS locus near *SNX10* in women and men. rs1534696 is sexually dimorphic for WHRadjBMI.

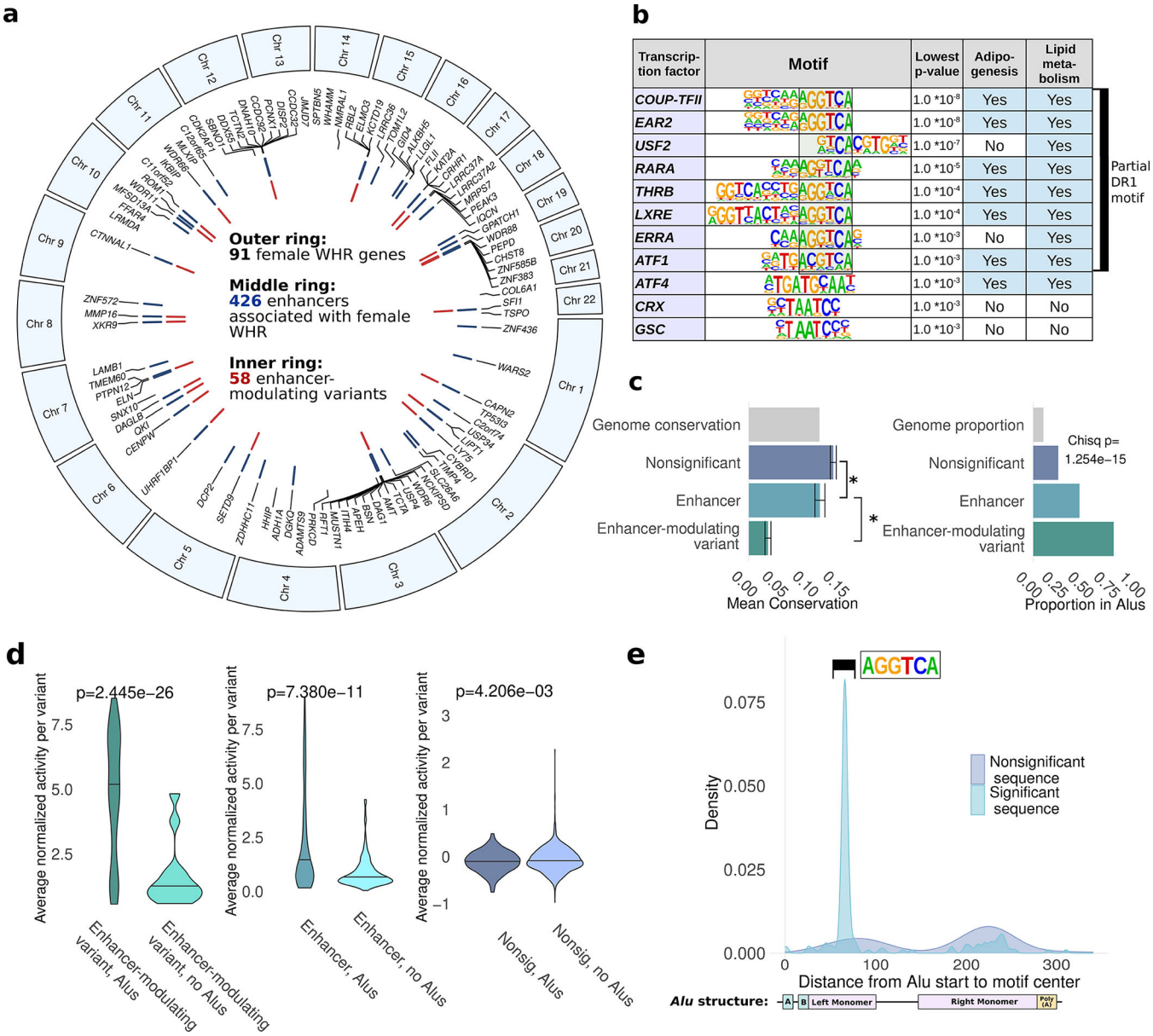

Figure 3: Regulatory network of female WHRadjBMI.

a) MPRA results. On the outside ring are shown female WHRadjBMI TWAS genes (n=91). In the blue ring are variants that had significant enhancer activity (n=426). In the red ring are variants which modulated the activity of significant enhancers (n=58). Nonsignificant variants n=1455. Significance determined by one-sided Mann-Whitney U-tests and subsequent FDR correction. b) Transcription factor motifs enriched in enhancer-modulating variants relative to nonsignificant enhancer sequences. Transcription factors contributing to common adipose motif are highlighted in green. P-values represent probability of finding degree of enrichment by chance, based on HOMER significance testing based on null binomial distribution of motifs in the genome and FDR multiple comparisons correction. c) Conservation of MPRA sequences by significance status and presence of *Alu* repeat elements in MPRA sequences. Data are colored by MPRA

significance status. Conservation significance represents Student's two-sided t-tests between average conservation across MPRA sequences. Data represent mean conservation +/− standard error (enhancer-modulatig variant vs. enhancer $p=1.81*10^{-17}$, enhancer vs. nonsignificant variant=$2.22*10^{-2}$). *Alu* proportion significance calculated via chi-square test comparing observed proportions of MPRA variants in *Alu* elements to that expected by chance in MPRA variants (enhancer-modulating variant overlap = 44/58 sequences, enhancer overlap=148/368 sequences, nonsignificant overlap=287/1433 sequences). d) Average enhancer activity of MPRA sequences with and without *Alu* elements. Data are colored by MPRA significance status. Significance determined by t-tests comparing average normalized activity between variants within or not within *Alu* elements per MPRA significance level. e) Density plot reflecting presence of enriched adipogenesis motifs within first monomer of *Alu* elements in MPRA sequences.

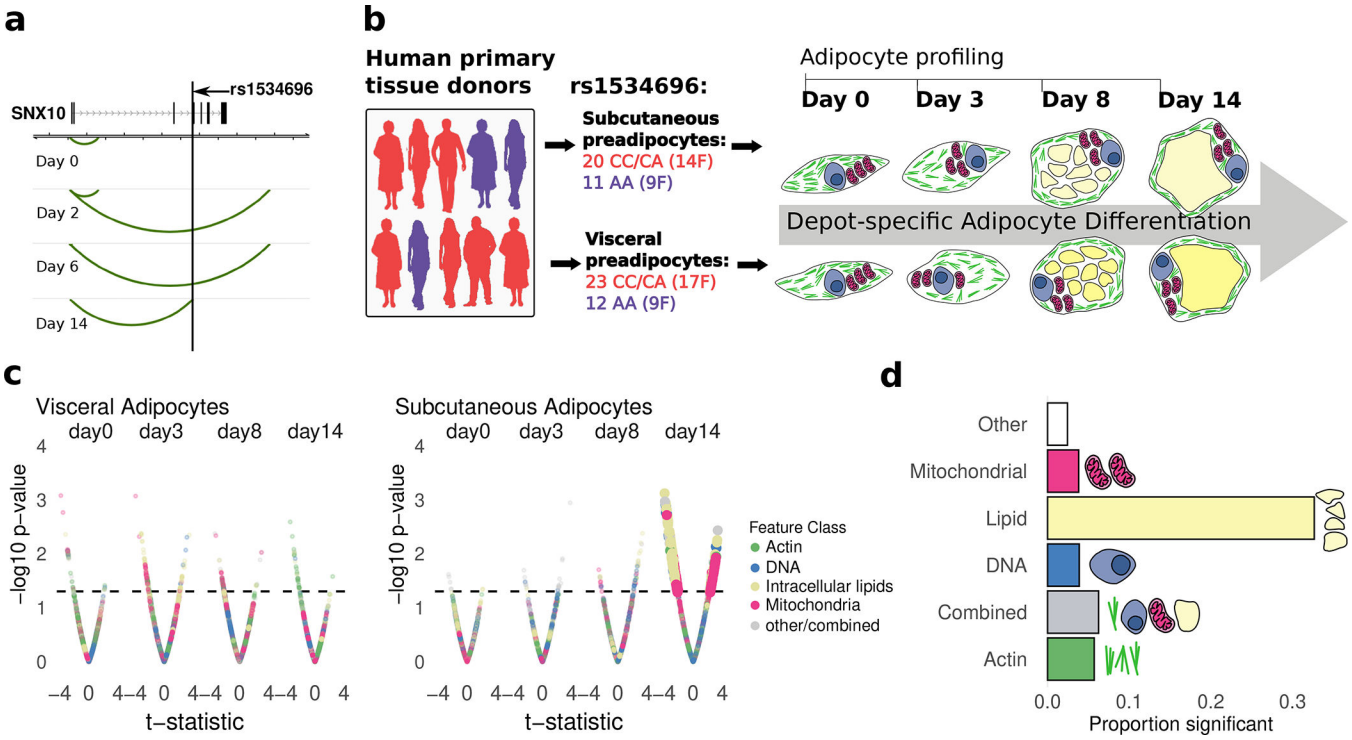

Figure 4: Effects of rs1534696 allele on human primary adipocytes.

a) ABC-predicted interactions between *SNX10* promoter and gene body loci in differentiating primary adipocytes. b) Lipocyte Profiler paradigm. c) Lipocyte Profiler features across primary human adipocyte differentiation. Data are colored by feature class. Larger dots represent features that differ significantly between AA and CC/AA genotype groups at each time point for adipocytes from subcutaneous and visceral depots. Significance is defined as $p<0.05$ and FDR $q<0.05$. Y axis represents p-value of significance of each feature, X-axis represents effect size, represented as t-statistic, of each feature. Dotted line represents p-value significance threshold. d) Lipocyte Profiler features significant in day 14 subcutaneous adipocytes. Data are colored by feature class. X axis shows significant features grouped by cellular organelle staining target. Organelle targets are visualized and consist of 1) mitochondria 2) intracellular lipids, 3) DNA, and 4) actin cytoskeleton and Golgi membrane. Features shown are collapsed across redundant features (e.g. different measures of cell shape). Subcutaneous adipocytes: CC/CA N=20, AA N=11. Visceral adipocytes: CC/CA N=23, AA N=12.

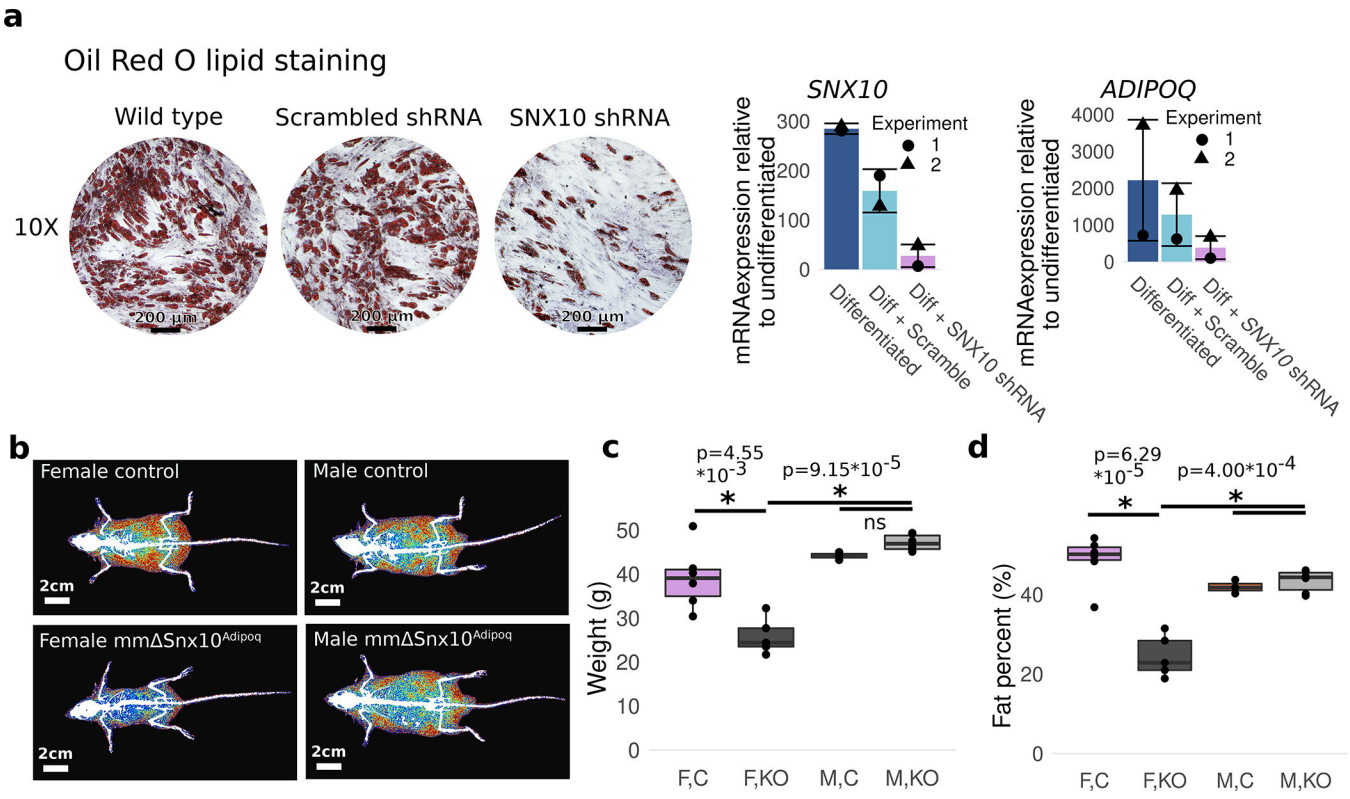

Figure 5: *SNX10*-mediated inhibition of adipogenesis.

a) Oil Red O staining and RT-qPCR quantification representing lipid accumulation after adipocyte differentiation in human mesenchymal stem cells upon *SNX10* shRNA introduction. Oil Red O images are representative of samples from that condition. Data presented as mean expression +/− standard deviation. n=2 technical replicates. Data are colored by shRNA experiment. b) DEXA scans from control mice (top) and female and male mmΔSnx10^{Adipoq} mice. c) Body weight and d) Body fat mass of control and mmΔSnx10^{Adipoq} mice after high-fat diet administration. (F-control N=6 mice, F-KO N=5, M-control N=3, M-KO N=6, pairwise t-test, $p<0.01$). Asterisks represent results of two-sided Student's t-testing. Box plot whiskers represent minimum (1st percentile) and maximum (99th percentile) of data, box bounds represent first quartile (25th percentile and 75th percentile), and center line represents median of data (50th percentile). Data are colored by sex and *SNX10* genotype.