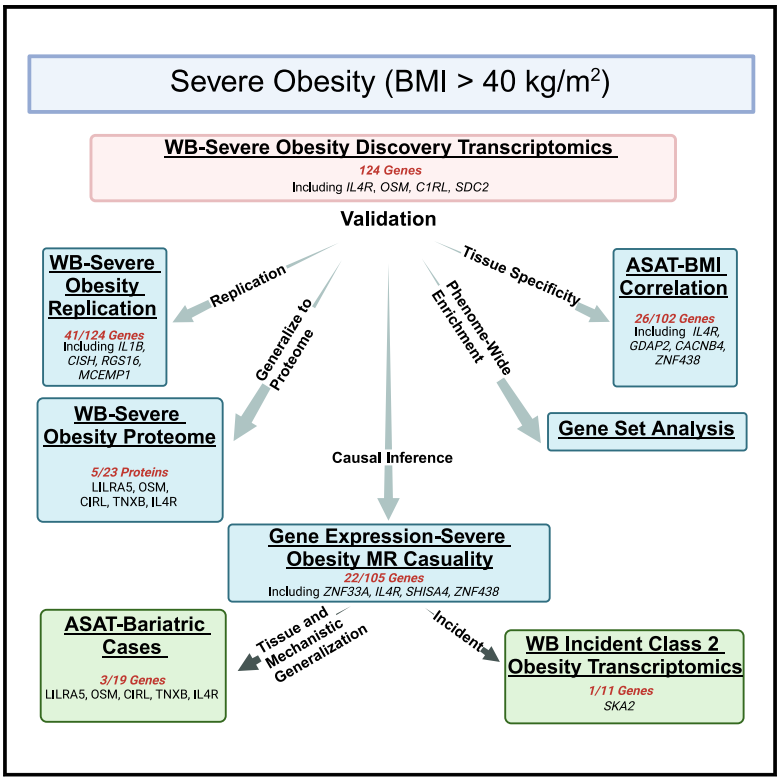
Cell Genomics

Article

Multiomics reveal key inflammatory drivers of severe obesity: *IL4R*, *LILRA5,* and *OSM*

Graphical abstract



Highlights

- 124 genes differentially expressed (DE) in Hispanic people with severe obesity
- *ZNF438*, *LINC01503*, *MSL3*, *ARRB1*, *CACNB4*, *FCGRT*, and *NCF1B* generalized to adipose tissue
- *AKT1*, *SCAP*, and *POLR2E* were DE in adipose before and 1 year after bariatric surgery

Authors

Hung-Hsin Chen, Heather M. Highland, Elizabeth G. Frankel, ..., Joseph B. McCormick, Kari E. North, Jennifer E. Below

Correspondence

kari_north@unc.edu (K.E.N.), jennifer.e.below@vanderbilt.edu (J.E.B.)

In brief

Chen et al. analyzed gene expression in Hispanic Latino individuals with severe obesity (BMI $\geq$40 kg/m^2) in comparison to controls, following up their 124 findings with subcutaneous adipose tissue expression, plasma proteomics, and longitudinal analyses. Their findings include inflammatory drivers and highlight potential mechanistic relationships between severe obesity and known correlates of obesity, including obstructive sleep apnea and hyperaldosteronism.

Multiomics of Hispanic Latino individuals with severe obesity highlights inflammatory drivers and potential mechanistic relationships between severe obesity and its correlates.

Chen et al., 2025, Cell Genomics *5*, 100784
March 12, 2025
https://doi.org/10.1016/j.xgen.2025.100784

CellPress

Cell Genomics

CellPress
OPEN ACCESS

Article

Multiomics reveal key inflammatory drivers of severe obesity: *IL4R*, *LILRA5,* and *OSM*

Hung-Hsin Chen,[1,2,3,24] Heather M. Highland,[4,24] Elizabeth G. Frankel,[1,2] Alyssa C. Scartozzi,[1,2] Xinruo Zhang,[4] Rashedeh Roshani,[1,2] Priya Sharma,[4] Asha Kar,[5,6] Victoria L. Buchanan,[4] Hannah G. Polikowsky,[1,2] Lauren E. Petty,[1,2] Jungkyun Seo,[4,7] Mohammad Yaser Anwar,[4] Daeeun Kim,[4] Mariaelisa Graff,[4] Kristin L. Young,[4] Wanying Zhu,[1,2] Kalypso Karastergiou,[8,9] Douglas M. Shaw,[1,2] Anne E. Justice,[10] Lindsay Fernández-Rhodes,[11] Mohanraj Krishnan,[4] Absalon Gutierrez,[12] Peter J. McCormick,[13] Carlos A. Aguilar-Salinas,[14,15] Maria Teresa Tusié-Luna,[16] Linda Liliana Muñoz-Hernandez,[17] Miguel Herrera-Hernandez,[18] Miryoung Lee,[19] Eric R. Gamazon,[1,2,20] Nancy J. Cox,[1,2] Päivi Pajukanta,[5,6,21] Susan K. Fried,[8,9] Penny Gordon-Larsen,[22] Ravi V. Shah,[23] Susan P. Fisher-Hoch,[19] Joseph B. McCormick,[19] Kari E. North,[4,24,*] and Jennifer E. Below[1,2,24,25,*]

[1]Department of Medicine, Division of Genetic Medicine, Vanderbilt University School of Medicine, Nashville, TN, USA
[2]Vanderbilt Genetics Institute, Vanderbilt University Medical Center, Nashville, TN, USA
[3]Academia Sinica, Institute of Biomedical Sciences, Taipei, Taiwan
[4]Department of Epidemiology, Gillings School of Global Public Health, The University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
[5]Department of Human Genetics, David Geffen School of Medicine at University of California, Los Angeles, Los Angeles, CA, USA
[6]Bioinformatics Interdepartmental Program, University of California, Los Angeles, Los Angeles, CA, USA
[7]Department of MetaBiohealth, Sungkyunkwan University, Suwon, Republic of Korea
[8]Obesity Research Center, Boston University School of Medicine, Boston, MA, USA
[9]Diabetes Obesity and Metabolism Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA
[10]Department of Population Health Services, Geisinger Health, Danville, PA, USA
[11]Department of Biobehavioral Health, The Pennsylvania State University, University Park, PA, USA
[12]Department of Internal Medicine, Division of Endocrinology, Diabetes, and Metabolism, Houston, TX, USA
[13]Centre for Endocrinology, William Harvey Research Institute, Barts and the London School of Medicine, Queen Mary University of London, London, UK
[14]Unidad de Investigación de Enfermedades Metabólicas and Research Direction of the Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, México City, México
[15]Tecnológico de Monterrey, Escuela de Medicina y Ciencias de la Salud, México City, México
[16]Unidad de Biología Molecular y Medicina Genómica, Instituto de Investigaciones Biomédicas UNAM Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, México City, México
[17]Unidad de Investigación de Enfermedades Metabólicas del Instituto Nacional de Ciencias, Médicas, y Nutrición Salvador Zubirán, México City, México
[18]Surgery Direction of the Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, México City, México
[19]Department of Epidemiology, The University of Texas Health Science Center at Houston School of Public Health, Brownsville Regional Campus, Brownsville, TX, USA
[20]MRC Epidemiology Unit, University of Cambridge, Cambridge, UK
[21]Institute for Precision Health at University of California, Los Angeles, Los Angeles, CA, USA
[22]Department of Nutrition, Gillings School of Global Public Health, The University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
[23]Vanderbilt Translational and Clinical Research Center, Cardiovascular Division, Vanderbilt University Medical Center, Nashville, TN, USA
[24]These authors contributed equally
[25]Lead contact
*Correspondence: kari_north@unc.edu (K.E.N.), jennifer.e.below@vanderbilt.edu (J.E.B.)
https://doi.org/10.1016/j.xgen.2025.100784

SUMMARY

Polygenic severe obesity (body mass index [BMI] $\geq$40 kg/m^2) has increased, especially in Hispanic/Latino populations, yet we know little about the underlying mechanistic pathways. We analyzed whole-blood multiomics data to identify genes differentially regulated in severe obesity in Mexican Americans from the Cameron County Hispanic Cohort. Our RNA sequencing analysis identified 124 genes significantly differentially expressed between severe obesity cases (BMI $\geq$40 kg/m^2) and controls (BMI <25 kg/m^2); 33% replicated in an independent sample from the same population. Our integrative approach identified inflammatory genes, including *IL4R*, *ZNF438*, and *LILRA5*. Several genes displayed transcriptomic effects on severe obesity in subcutaneous adipose tissue. We further showed that the genetic regulation of these genes is associated with several traits in a large biobank, including bone fractures, obstructive sleep apnea, and hyperaldosteronism, illuminating potential risk mechanisms. Our findings furnish a molecular architecture of the severe obesity phenotype across multiple molecular domains.

INTRODUCTION

Obesity is a major driver of the population burden of a variety of diseases (e.g., type 2 diabetes [T2D], kidney disease, cancers, maternal birth complications, asthma, liver disease) and disproportionately impacts historically marginalized populations, such as Hispanic/Latinos (HL), as compared to non-Hispanic Whites.[1–9] Severe obesity (SO; body mass index [BMI] $\geq$40 kg/m^2, approximately >100 lb overweight) is an emerging and critical concern, as it affects every organ system and is concomitant with nearly a decade or more of life expectancy reduction.[2,10,11] The prevalence of SO has doubled among US adults over the past 2 decades to 9.2%; women and HL have experienced the greatest increases.[10,12] However, SO has been an exclusion criterion in many epidemiological and clinical studies and is often unidentified in electronic health records and mortality data, limiting knowledge about its complex etiology, mechanistic pathways, and intervention targets, particularly in understudied populations.[13–15]

Obesity is a heterogeneous disease with involvement of the central nervous system, including the pituitary gland and hypothalamus (e.g., the leptin melanocortin system), hippocampus, and limbic system (reward system),[16] in addition to energy processing and utilization,[16,17] endothelial dysfunction, ectopic fat deposition,[18] and inflammation. While we and others have made significant inroads to understanding the genetic architecture of the full range of obesity (BMI $\geq$30 kg/m^2) and its consequences in HL populations (i.e., identifying new loci, refining, and identifying causal effects of known loci on clinical health outcomes), this area is still largely understudied.[19–24] Furthermore, we still know little about the disrupted molecular physiology that underlies SO. While genome-wide association studies (GWASs) have identified >1,000 loci associated with BMI to date, we only understand the biologic mechanisms of a small portion of these variants.[25] Gene expression measures can illuminate links between genetic variation and diseases, pointing to pathways for therapeutic intervention.[26–34] It is also anticipated that the severe phenotype provides a greater ability to detect causal signals and pathways leading to obesity and from obesity to other metabolic derangements. Multiple gene expression studies of obesity-related traits have been conducted, both with continuous BMI[35] and with clinical definitions of class 1 obesity (30 kg/m^2 $\leq$ BMI <35 kg/m^2) as a binary trait.[36] However, these studies have been limited by small sample sizes (e.g., n < 20), lack of replication, failure to assess SO, and focus on European ancestry. Two studies have explored gene expression differences in small samples of bariatric surgery patients (<20 patients), before and after surgery; to date, no population-based human studies have examined the role of gene expression to identify molecular signatures associated with SO and downstream sequelae.[37,38] Thus, our understanding of the causal mechanisms of SO pathogenesis within and across relevant metabolic tissues, particularly in understudied populations at high risk, is limited. Such studies will be key to better understanding SO and its role in downstream diseases.

In this study, we aimed to improve the understanding of underlying molecular causes and clinical and molecular consequences of SO by identifying and characterizing whole-blood (WB) transcriptomic signals in adult Mexican Americans participating in the Cameron County Hispanic Cohort (CCHC). We replicated our study results in an independent sample from the same study population and hypothesized that some of these effects may also be measurable in other obesity-related tissues. We examined the specificity of detected effects in abdominal subcutaneous adipose tissue (ASAT) from community volunteers from New York City (20 kg/m^2 < BMI < 40 kg/m^2).[39–41] We also hypothesized that observed transcriptomic differences may translate to differential proteomic abundance in blood. Thus, we newly measured protein abundance using the Olink proteomics panel and assessed the evidence of association with SO in comparison to controls for differentially expressed genes. Because we expected some differentially expressed genes to exert causal effects on SO risk, we used information from expression quantitative trait loci (eQTLs) and SO GWASs to infer causal relationships. Finally, to characterize the clinical phenome associated with these differential omics, we generated genetically predicted gene expression measures in a large electronic health record-linked biobank to assess clinical outcomes associated with genes differentially expressed in SO. These analyses, outlined in Figure 1A, contribute significantly to our understanding of the range of health outcomes linked to the dysregulation of both known and novel genes associated with SO. Our cross-tissue and multiomics findings provide systemic information about metabolic homeostasis, hematopoietic development, and immune functions,[42] highlighting metabolic, inflammatory, adipose, and insulin-related mechanisms that may inform our understanding of therapeutic targets for obesity medicine and offering insights into the causes and consequences of SO biology.

RESULTS

Demographic characteristics

As shown in Figure 1A, across our discovery and validation samples (including replication and proteomic samples), we selected cases and controls from the CCHC, which includes 5,000 individuals randomly selected from households in Brownsville, Texas on the US-Mexico border. Table 1 and Figure 1C summarize the subjects' clinical characteristics. For the discovery stage, 49 CCHC participants with SO (BMI $\geq$40 kg/m^2) and 81 controls (BMI <25 kg/m^2) had complete phenotype data and high-quality RNA sequencing (RNA-seq) data. For the validation stage, 111 independent CCHC participants (52 cases, 59 controls) with complete phenotype and high-quality RNA-seq data were considered for replication. The proteomic validation considered 91 CCHC participants (49 cases, 42 controls), some of whom overlapped with the samples that had RNA-seq data (both discovery and replication; Figure 1B). As expected, all three SO case groups had elevated rates of T2D, hypertension, and hypercholesterolemia, compared to control study participants with BMI <25 kg/m^2 (Table 1).

We investigated the transcriptome profile in ASAT of our identified genes in bulk RNA-seq data obtained from n = 19 community volunteers in New York City who had needle aspirates of ASAT with BMI ranging from 21.0 to 36.7 kg/m^2. All females were premenopausal and in the follicular state of their menstrual cycle. Their characteristics are listed in Table 1.

To strengthen our inference for genes prioritized as causal for SO, we analyzed RNA-seq measures from ASAT from an

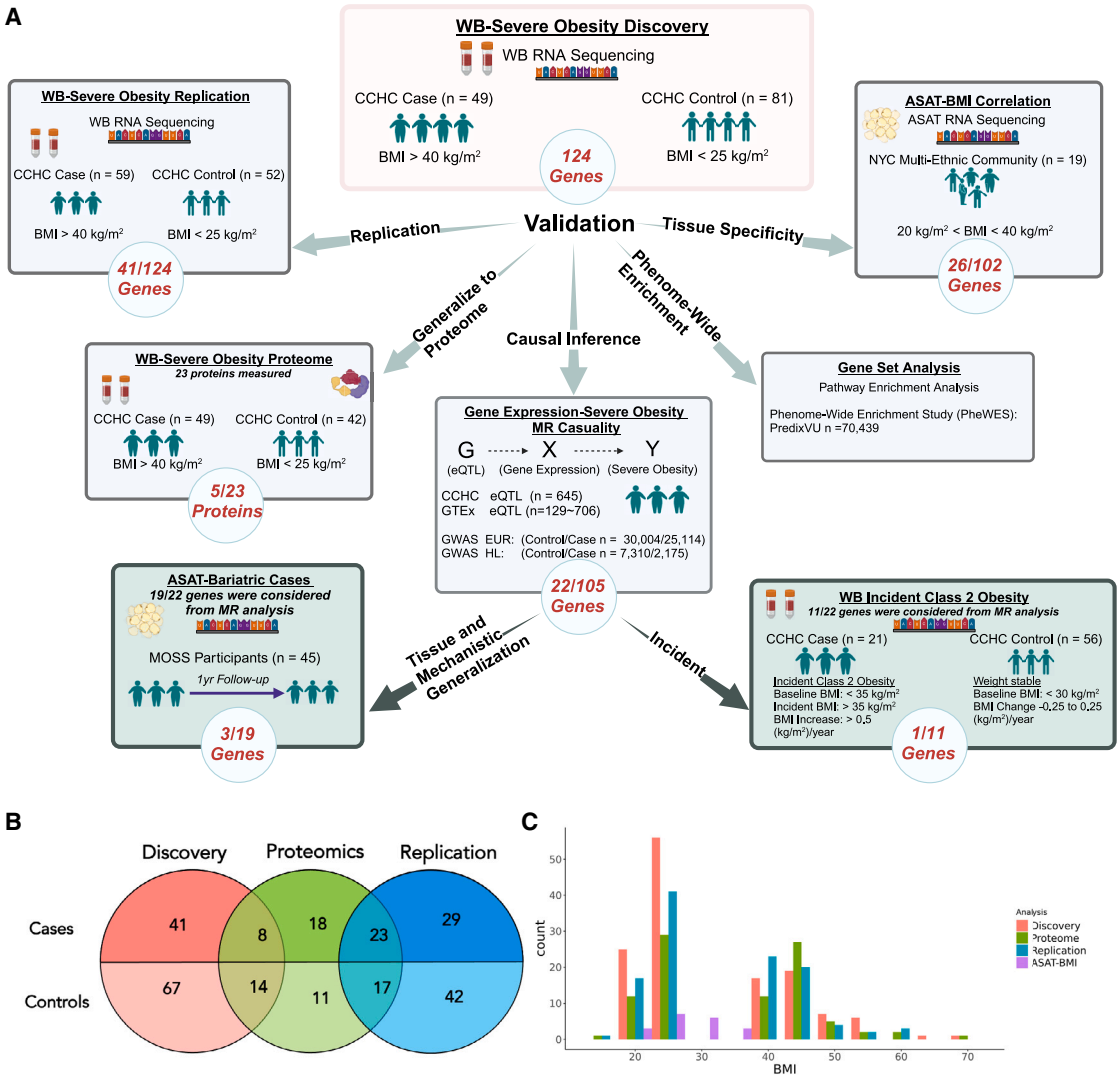


Figure 1. Study design

(A) Schematic of the study design for high-dimensional multiomics analysis. (Top, pink) Differential expression analyses in whole-blood (WB) RNA sequencing (RNA-seq) in severe obesity (SO) CCHC cohort (WB-SO discovery).

(B) Sample overlap between the discovery, replication, and proteomics analyses in the CCHC.

(C) Histogram showing the distribution of BMI in discovery, replication, proteomics, and ASAT-BMI correlation analyses.

BMI, body mass index; CCHC, Cameron County Hispanic Cohort, GReX, genetically regulated gene expression.

Generated with BioRender. See also Figure S1 and Tables S1–S4, S5, S6, S7, S8, S9, S10, and S11–S13.

additional independent ongoing longitudinal study of severely obese Mexican individuals undergoing bariatric surgery and participating in a 1-year follow-up as part of the Mexican Obesity Study (MOSS).[43,44] We analyzed a total of 45 individuals with subcutaneous adipose RNA-seq data at both time points (Table 1).

To better understand the broad clinical effects of the set of our identified genes, we conducted a phenome-wide enrichment study (PheWES) leveraging a large-scale electronic health record-linked biobank at Vanderbilt University, BioVU.

Transcriptomic discovery

We performed WB RNA-seq (49 cases; 81 controls) using standard protocols and alignment, yielding 18,565 genes after quality control (STAR Methods; Figure 2). After false discovery rate (FDR) correction, 124 genes were significantly differentially expressed (DEGs) between SO cases and controls (SO-DEGs) (Table S1), including top genes *C1RL*, *IL4R*, *OSM*, *RGS16*, and *SDC2*. Sixty-nine genes (56%) were downregulated and 55 (44%) were upregulated. Broadly, these genes regulate immune and inflammatory activation (*LILR5A*, *TLR5*, *IL4R*, *C1RL*), cytokine signaling (*SOCS2*, *SOCS3*, *CISH*), general cell growth/signaling/differentiation/proliferation (*SDC2*, *NHS*, *ZAK*), general cellular effector and energy production function (*OSM*, *PPP1R10*), and metabolism (*ENGASE*, *SLC37A3*). We conducted a sensitivity analysis, adjusting our analyses for predicted and scaled blood cell types for WB[45]; the correlation between the Wald statistics from DESeq2

Table 1. Demographic information of study participants

Trait	Discovery		Replication		Proteome		ASAT BMI (21.0–36.7)	ASAT MOSS Pre-bariatric surgery
	SO	Control	SO	Control	SO	Control		
N	49	81	52	59	49	42	19	45
Male (n, %)	12 (24)	27 (33)	14 (27)	21 (36)	14 (29)	14 (33)	7 (37)	30 (67)
Age, y (mean, SD)	42.8 (17.7)	53.0 (19.1)	50.2 (14.3)	45.7 (17.6)	46.5 (13.7)	52.4 (17.4)	31.3 (7.8)	42.5 (11.2)
BMI, kg/m^2(mean, SD)	46.2 (6.4)	23.0 (1.7)	51.3 (4.9)	22.8 (1.8)	45.6 (5.7)	22.9 (2.0)	27.5 (4.8)	46.1 (8.8)
Diabetes (%)[a]	19 (39)	20 (25)	25 (48)	11 (19)	12 (24)	6 (14)	–	20 (44)
Hypertension (%)[b]	19 (39)	17 (21)	29 (56)	16 (27)	18 (37)	11 (27)	3 (16)	–
Hypercholesterolemia (%)[c]	10 (20)	23 (28)	26 (50)	18 (31)	16 (33)	15 (36)	–	–

Severe obesity (SO) BMI >40 kg/m^2; control BMI <25 kg/m^2.
[a]Hypertension: systolic blood pressure ≥140 mm Hg or diastolic blood pressure ≥90 mm Hg or use of anti-hypertensive medication.
[b]Self-reported hypercholesterolemia.
[c]Diabetes: fasting blood glucose ≥126 mg/dL, HbA1c ≥ 6.5%, history of diabetic medication, or diagnosed with diabetes.

with and without cell-type adjustment was $r = 0.878$ and $p < 2.2 \times 10^{-16}$. All but 15 of our SO-DEGs remained significant in this analysis. To minimize confounding resulting from correcting for multiple RNA-seq-derived variables, results unadjusted for cell types are reported hereafter.

Validation

Transcriptomic replication

To replicate primary findings, we performed differential expression analysis in an additional non-overlapping set of 52 cases of SO and 59 controls from the CCHC (Table 1). Among the 124 SO-DEGs, 104 were found to have the same direction of effect, and 20 showed opposite effects (chi-squared p value for the directional consistency = 4.58×10^{-14}). Of the 43 genes with significant differential expression in the replication (FDR adjusted $p < 0.05$), 41 genes showed the same direction of effect as in the discovery (chi-squared p value for directional consistency = 2.72×10^{-9}), and two genes (*EGFL7*, *ARHGEF2*) showed an opposite direction of effect (Figure 3; Table S2). Quantification of the effect of the transcriptomic profile of SO showed that 47.6% of SO risk is explained by the 124 SO-DEGs and 38.5% by the 41 replicated genes (STAR Methods).

Although to our knowledge, no population-based human transcriptomic studies of SO have been previously published,

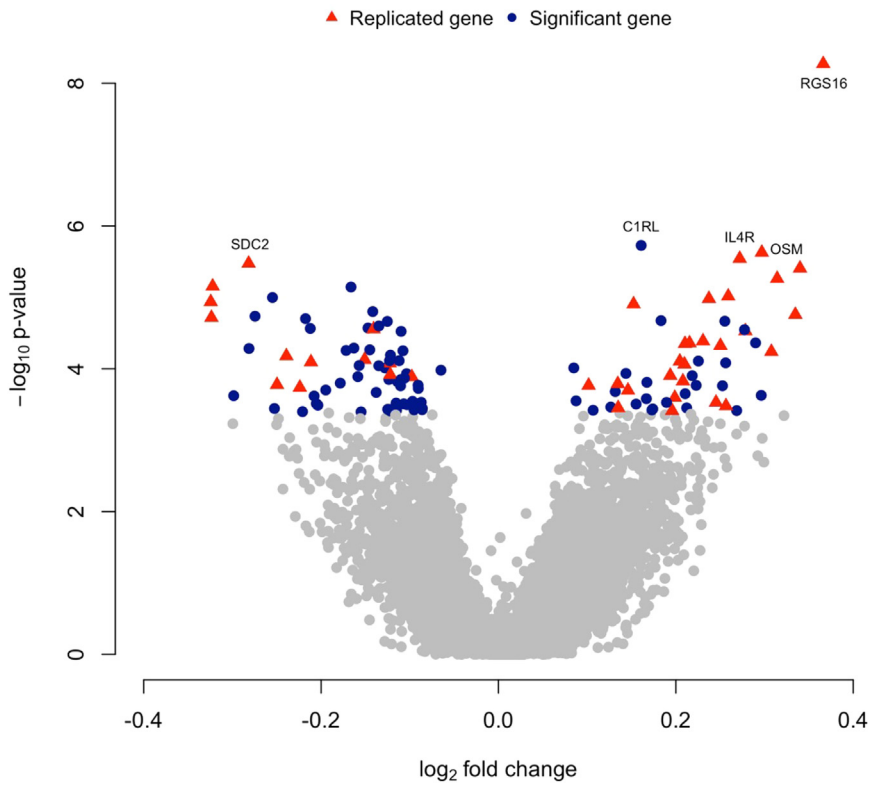


Figure 2. Discovery transcriptomics

Volcano plot for differential gene expression. Scattered points represent genes: the x axis is the log2 fold change, whereas the y axis is the –log10 of the p value. Red triangles and blue dots are genes whose expression is significantly associated with SO in discovery analysis (FDR adjusted $p < 0.05$); blue dots are also significantly associated with SO in independent replication analysis (FDR adjusted $p < 0.05$).
See also Tables S1 and S2.

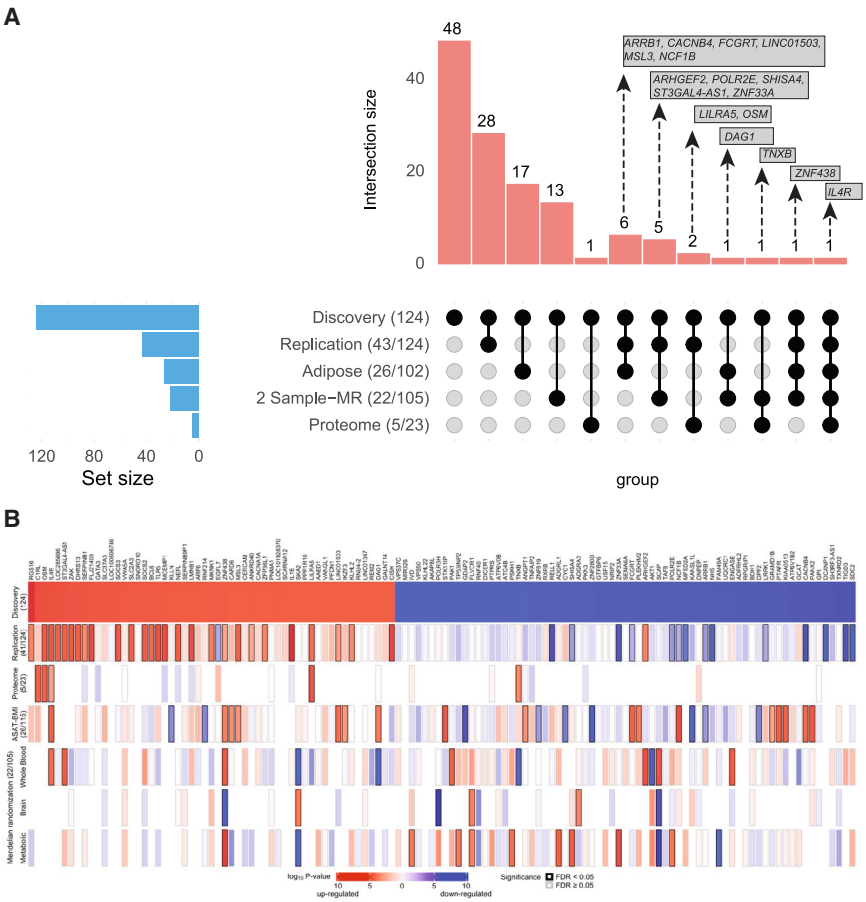


Figure 3. Replication and generalization

(A) Upset plot summarizing intersections of 124 SO-DEGs. In the discovery stage, 124 genes were significantly differentially expressed between SO cases and controls; 41 out of 124 SO-DEGs were replicated in an independent sample drawn from the same cohort. Five genes were significantly associated with SO out of 23 that overlapped in the proteomic analyses, after FDR correction. A total of 26 genes are significant in ASAT-BMI correlation analysis using ASAT RNA-seq from New York City community participants out of 102 genes that overlapped. A total of 22 genes show significant casual effects on SO in MR out of 105 genes for which strong genetic instruments were available. The histogram to the right shows the number of genes that are significant in each analysis, and the histogram to the left shows where each of the 124 genes was categorized across all analyses conducted.

(B) Heatmap plot summarizing intersections, effect estimates, and direction of effects of 124 genes differentially expressed in SO versus controls. The black box indicates FDR $p < 0.05$; the gray box means tested but not FDR significant.

See also Tables S1, S2, S5, S6, S7, and S13.

the GWAS literature, while limited by mapping to the nearest gene, provides support for many of the genes we identified. For example, 29 of the 124 SO-DEGs have been previously reported in the GWAS Catalog[46] as body shape and adipose distribution loci (Figure 4A; Table S3). Additionally, 34 of 124 SO-DEGs have been associated with cardiovascular disease and risk factors such as lipid levels (Figure 4A; Table S3). In the Common Metabolic Diseases Knowledge Portal (CMDKP),[47] the Human Genetic Evidence[48] scores for 98 genes exhibit moderate to compelling support for anthropometric traits (Figure 4B; Table S4). These genes were also implicated in hematological, hepatic, renal, and lipid traits. Notably, *FCGRT* showed compelling evidence across multiple domains, including anthropometric, hematological, hepatic, lipids, and metabolites. Similarly, *SHISA4* showed compelling evidence for anthropometric, cardiovascular, lipid, and sleep/circadian traits. *TNxB* displayed compelling evidence of genetic involvement in anthropometric, cardiovascular, and hematological traits, with extreme evidence support for glycemic traits, and very strong evidence for traits in T1D and immunological, hepatic, renal, lipid, metabolite, and ocular domains. Other genes have been implicated in the broader omics literature, for example, DNA methylation at eight genes was previously associated with BMI (*BCL6*, *GALNT14*, *PLEKHM2*, *PTPRS*, *SEMA6A*, *SOCS3*, *TP53INP2*, *VANGL1*).[49–56]

Proteome association

To further validate our observed transcriptomic effects, we tested associations in the plasma proteome for each of our SO-DEGs. Protein abundance of 5 of the 23 available were significantly associated with SO after FDR correction ($p < 0.05$: *LILRA5*, *OSM*, *CIRL*, *TNxB*, *IL4R* [Figure 3; Table S5]). Three of these five genes were significant in the replication transcriptomic study (*LILRA5*, *OSM*, and *IL4R*) (Table S5), and four show directional consistency with the regulation observed in the discovery analysis. *TNxB* shows an opposite direction of effect in the proteomic analysis when compared to the results from the discovery.

Correlation of ASAT gene expression with BMI

To explore whether the 124 SO-DEGs in WB generalized to another key obesity-related tissue, ASAT, we investigated the correlation of BMI with gene expression in ASAT in RNA-seq data from 19 community volunteers (20 kg/m^2 < BMI < 40 kg/m^2). Of the 124 SO-DEGs, 102 passed quality control, and 26 of these were significantly correlated with BMI (FDR adjusted $p < 0.05$). We observed strong associations between BMI and the ASAT expression level of several of our top genes (e.g., *IL4R*, *ZNF438*, *CACNB4* [Figure 3; Table S6]).

Mendelian randomization

We used Mendelian randomization (MR) to investigate possible causal relationships of gene dysregulation on SO using eQTLs from both CCHC WB and additional relevant tissues in the Genotype-Tissue Expression (GTEx) Portal as instrumental variables. We used a weighted median approach,[57–59] which is robust in the presence of up to 50% invalid instrumental variables.[57] We found validated instruments for 105 of the 124

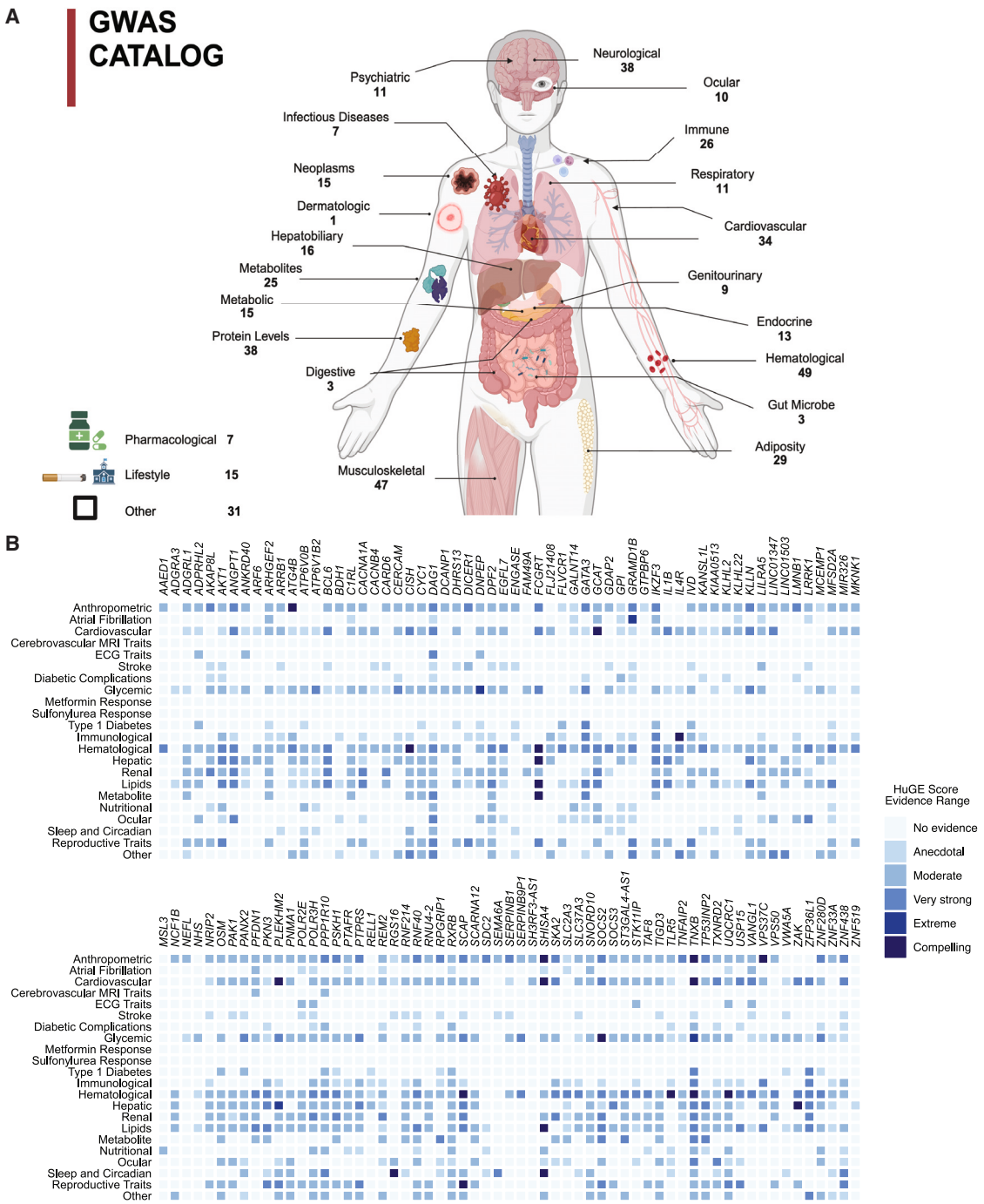


Figure 4. Literature support

(A) Diagram showing the number of SO-associated genes discovered in our analyses that were reported as mapped genes for traits contained within the GWAS Catalog. Of the 124 discovered genes, 102 were found in our GWAS Catalog search. Traits with evidence of genome-wide significant association ($p < 5 \times 10^{-8}$) identified within each category and results by association are listed in Table S3.

(B) Heatmap showing the degree of evidence of involvement of each SO-DEG in cardiometabolic trait groupings from the CMDKP. The full query is in Table S4. The image in (A) was generated with BioRender. See also Tables S3 and S4.

SO-DEGs, and after FDR correction, 22 genes demonstrated significant casual effects in SO in at least one tissue in the weighted median MR analyses. Seven of the genes with significant causal effects in SO (*ARHGEF2*, *ZNF33A*, *IL4R*, *SHISA4*, *ZNF438*, *POLR2E*, *ST3GAL4-AS1*) were among those that independently replicated after FDR correction (adjusted

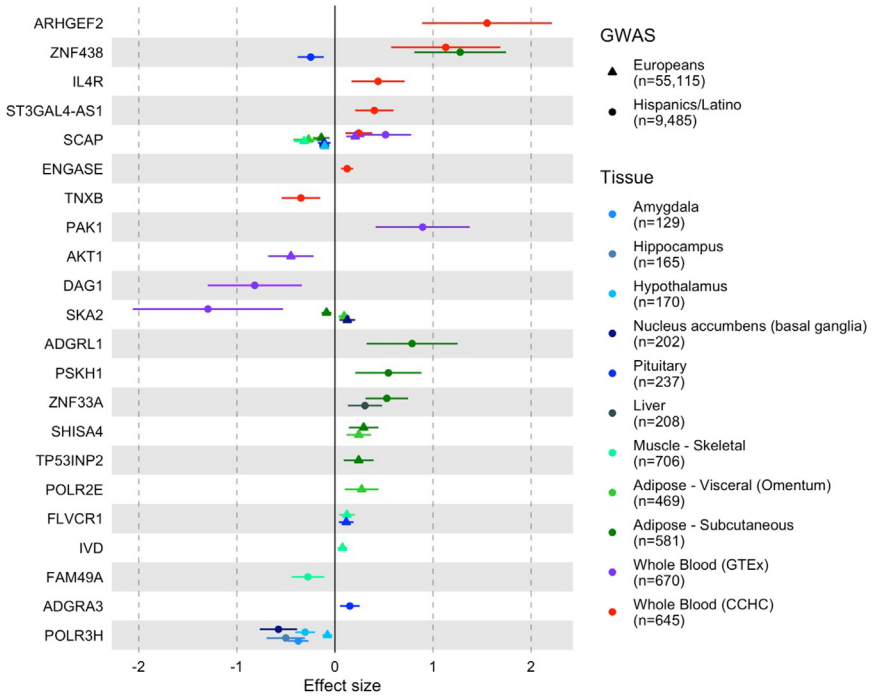


Figure 5. Causal effects

Forest plot of the causal effect size with 95% confidence interval for genes demonstrating significant casual effect for SO in two sample median weighted MR analyses (FDR adjusted $p < 0.05$). The instrumental variables (IVs) were selected from eQTLs in obesity-relevant tissues from GTEx and eQTLs developed from CCHC WB transcriptomic data ($n = 645$). The association between the IV and outcome (SO) were estimated using SO GWAS summary statistics from both a European ancestry analysis and an independent Hispanic/Latino analysis.[60]

See also Table S7.

$p < 0.05$; Figure 5). An overview of full MR results is given in Table S7.

Prospective analyses

To support our findings of causal inference, we assessed the association of the baseline expression of these 11 genes with incident class 2 obesity during follow-up (Table S8A). Among the identified causal genes identified in our MR analyses, *SKA2* expression ($\log_2$ fold change = −0.31, $p = 0.0032$) showed significant effects on obesity incidence after Bonferroni multiple test correction. Nominal significance was observed at *AKT1*. The results for all tested genes are shown in Table S8B.

Differential gene expression in ASAT of bariatric patients followed longitudinally

To additionally support our findings of causal inference, we assessed change in expression in abdominal ASAT from an ongoing longitudinal study of 45 Mexican individuals with SO undergoing bariatric surgery and participating in a 1-year follow-up as part of MOSS. Nineteen genes prioritized by MR were measured in this study, and of these, *AKT1*, *SCAP*, and *POLR2E* showed significantly altered expression across time pre- and post-bariatric surgery (FDR adjusted $p < 0.05$). The results for all tested genes are shown in Table S9.

Pathway enrichment analysis

We conducted pathway enrichment analyses to explore the biological systems underlying our 124 SO-DEGs. We performed an over-representation test with two established datasets, Gene Ontology (GO) and WikiPathways. After multiple test correction (FDR adjusted $p < 0.05$), we found 52 significant GO terms—2 from molecular function and 50 from biological process—and 3 pathways from WikiPathways (Table S10). The top five GO terms include response to peptide (GO: 1901652, raw $p = 1.58 \times 10^{-8}$), response to peptide hormone (GO: 0043434, $p = 1.7 \times 10^{-7}$), response to organonitrogen compound (GO: 0010243, $p = 3.20 \times 10^{-7}$), insulin receptor signaling pathway (GO: 0008286, $p = 1.15 \times 10^{-6}$), and response to nitrogen compound (GO: 1901698, $p = 1.29 \times 10^{-6}$). The three significant WikiPathways are the prolactin signaling pathway (WP2037, $p = 2.41 \times 10^{-4}$), the leptin signaling pathway (WP2034, $p = 2.41 \times 10^{-4}$), and the overlap of cellular insulin and leptin signaling (WP3935, $p = 2.51 \times 10^{-4}$). These pathways highlight key molecular mechanisms that may be disrupted in SO (Figure S1). Notably, the insulin and leptin signaling overlap, insulin signaling, and prolactin signaling pathways have well-established roles in obesity pathogenesis and were enriched in a recent study of gene expression and BMI.[61]

PheWESs

We then investigated the broader clinical relevance of our findings in the electronic health record-linked DNA biobank, BioVU. We tested for the enrichment of significant associations between a phecode and imputed tissue-specific genetically regulated expression (GReX) of 124 SO-DEGs, phenome-wide (1,766 phecodes). This analysis revealed significant enrichment of 25 phecodes after Bonferroni correction ($p < 2.8 \times 10^{-5}$) (Figure 6; Table S11). These included anomalies of pupillary function (phecode 379.4, 56 GRexassociations, $p < 1 \times 10^{-6}$), fracture of tibia and fibula (800.3, 55 GReX associations, $p < 1 \times 10^{-6}$), mixed hyperlipidemia (272.13, 54 GReX associations, $p = 1 \times 10^{-5}$), obstructive sleep apnea (327.32, 54 GReX associations, $p < 1 \times 10^{-6}$), and regional enteritis (555.1, 54 GReX associations, $p = 1 \times 10^{-5}$).

To prioritize gene expression profiles that are estimated to be causal for SO and clinical phenotypes identified in our PheWES, we conducted MR using tissue-specific GReX from causal genes as instrumental variables. Among the 25 identified clinical phenotypes, we observed a significant causal effect of SO on fracture of the upper limb after FDR correction (phecode 803, raw $p = 0.0019$) (Table S12).

DISCUSSION

Here, we performed a genome-wide transcriptomic analysis of individuals with SO versus controls (BMI <25 kg/m^2) and

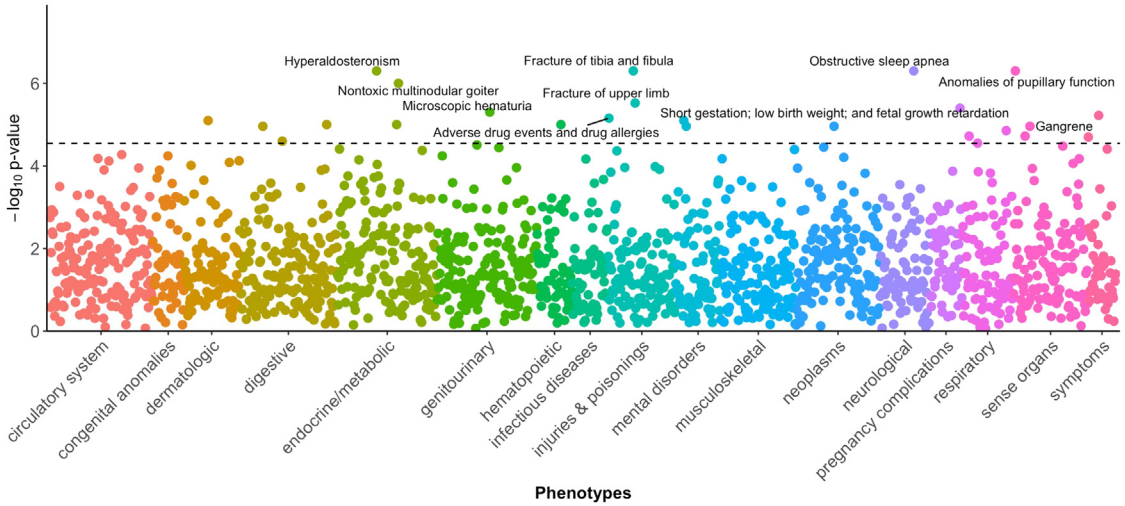


Figure 6. PheWES

PheWES analysis of electronic health record-linked DNA biobank at Vanderbilt University, BioVU permutation-based testing of the enrichment of significant association between a phecode and imputed tissue-specific genetically regulated expression of SO genes identified in our discovery analysis (124 genes) phenome-wide (1,766 phecodes).

See also Tables S11 and S12.

evaluated the replication of findings in an independent transcriptomic sample. In downstream analyses, we assessed evidence for cross-tissue dysregulation, explored whether transcriptomic differences translated to proteome abundance, established evidence for causal effects, and investigated the clinical phenome associated with molecular dysfunction in SO. Our findings undercover novel genes influencing SO in a minoritized ancestrally admixed population and yield heretofore undiscovered insights into SO biology.

There is a paucity of WB transcriptomic studies of body mass-related traits, with most of these studies restricted to European ancestry participants or lacking replication and validation. A 2019 study measured microarray gene expression data and revealed seven genes associated with BMI, although none of these overlap with our discovered genes.[36] A recent large multi-population meta-analysis identified 45 genes whose expression in cryopreserved peripheral blood mononuclear cells was significantly associated with BMI; of these, *SOCS2*, *SOCS3*, *PTPRS*, and *GRAMD1B* overlap with our discovered genes.[61] Despite the dearth of prior transcriptomic studies, many of the genes identified in our study have been implicated previously in genomic studies of SO and obesity-related traits (Figure 4; Tables S3 and S4; Table S13). Previously, over 1,100 independent loci have been identified for obesity and BMI in GWASs, candidate gene, and family-based studies; however, in many cases, the effect genes are uncharacterized or poorly characterized. We identified 124 genes whose expression is associated with SO, and 29 of these have been described in the BMI/obesity GWAS literature (Figure 4A; Tables S3 and S4). Additionally, the CMDKP indicates that 57 of our discovered genes have compelling, moderate, or very strong evidence for a role in BMI or obesity (Figure 4B). Previously identified genes that were not observed as significantly associated with SO in our analyses may have tissue-specific function, impact BMI through a mechanism other than RNA abundance, have modest effects on RNA abundance requiring larger studies, or have different effect sizes in HL populations.

Our differential expression analyses identified 66 genes associated with SO that have not previously been described in the GWAS literature for BMI or obesity phenotypes and do not already have compelling evidence for a role in obesity in the CMDKP. Twenty-eight of these genes replicated after multiple test correction in a second sample drawn from the same population. Because our approach identifies genes whose expression differences may be the result of SO or causal for SO, it is likely that some of these genes are dysregulated as a result of SO and may not harbor genetic variation associated with risk. Nonetheless, four of these replicated genes also show statistical evidence of causality and highlight the power of differential expression analyses for identification of new biology in SO.

Our integrative approach facilitated novel discovery of inflammatory genes involved in SO pathogenesis (Figure 3A; Table S13). *IL4R* displayed significant evidence of association in all lines of study inquiry. *IL4R* is a ubiquitously expressed[62,63] protein coding gene that inhibits lipid deposit and promotes insulin sensitivity and glucose tolerance. While *IL4R* has not been associated previously with obesity, it has been associated with asthma and inflammation in prior GWASs.[64] Importantly, dupilumab inhibits interleukin-4 (IL-4) intracellular signaling and is used as a treatment for asthma and eczema, and a previous study found weight gain associated with dupilumab use.[65] In addition, IL-4 reduced insulin secretion in the islets of healthy donors but not in individuals living with T2D, supporting a mechanistic role in metabolism.[66] Replicated gene *ZNF438* also has evidence from all lines of inquiry in our study with available data (ZNF438 was not assayed on the Explore 3072 panel). *ZNF438* demonstrated significant causal effects on SO in MR and has been associated previously with weight, BMI adjusted for hip and waist circumference, and

lean body mass.[67,68] *ZNF438* is also a well-known asthma susceptibility gene. Two additional genes, *LILRA5* and *OSM*, showed evidence of significant differential abundance between people with SO-associated and controls across the transcriptome and proteome. The protein encoded by *LILRA5* is a member of the leukocyte immunoglobulin-like receptor family, with activating and inhibitory functions in innate immunity. An association between *LILR5A* gene expression and BMI has been reported in Black children with obesity.[69] *OSM* regulates cytokine production, including IL-6, which plays a role in obesity.[70,71] Elevated *OSM* expression has been observed in ASAT and visceral (omental) adipose tissue in patients with SO and hyperglycemia[72]; it is thought that elevated *OSM* expression may inhibit adipogenesis, reducing *GLUT4* expression, increasing inflammation in human adipocytes, and leading to the development of insulin resistance.[73] Collectively, these findings strongly support the role of inflammatory genes in SO pathogenesis.

Our assessment of generalizability of replicated transcriptomic effects across tissues revealed seven genes (*ZNF438*, *LINC01503*, *MSL3*, *ARRB1*, *CACNB4*, *FCGRT*, and *NCF1B*) with significant effects in WB and ASAT. *ARRB1* is a key component of the superfamily of G-protein-coupled receptors, where it is responsible for agonist-mediated desensitization, dampening cellular responses to hormones, neurotransmitters, and sensory signals, all of which play critical roles in obesity pathogenesis. It is also known to be a scaffold for downstream signaling proteins and leads to a sustained cell response after stimulation.[74] Furthermore, *ARRB1* is thought to play a major role in receptor-mediated immune functions. *ARRB1* is a known triglyceride GWAS signal, has been associated with metabolic activity in cancer cells,[75] and is required for adaptive β cell expansion in obesity.[76] *CACNB4* is a β subunit of a voltage-dependent calcium channel that modulates G-protein inhibition, which may play a role in sensory cues related to obesity.[77] An indel in *CACNB4* has been associated with a heightened risk of cardiovascular disease.[78,79] *FCGRT* encodes a receptor that binds the Fc region of monomeric immunoglobulin G. It has been associated with myriad cardiovascular disease traits in GWASs, including metabolic traits such as T2D, triglyceride concentrations, low-density lipoprotein cholesterol, and sex hormone-binding globulin, among many others.[80–84] Lastly, methylation at *NCF1B* has been previously associated with maternal weight in early pregnancy in WB.[85]

SKA2 and *AKT1* provide the strongest evidence for causality in MR and prospective analyses. *SKA2* encodes spindle and kinetochore-associated complex subunit 2, and variation in this gene has been associated with height[86] and lung cancer.[87] Primarily expressed in the cerebellar hemisphere, epigenetic modification and expression of *SKA2* in the brain has been associated with suicidality and post-traumatic stress disorder.[88–90] *AKT1* has been associated with *APOL1*,[82] *HDL*,[91] and height.[92] Furthermore, candidate gene studies have implicated the phosphatidylinisitol-3-kinase (PI3K)/AKT/mTOR signaling pathway with diabetes and obesity.[93,94]

Collectively, our gene-level findings highlight key inflammatory pathways in SO pathogenesis and explain a large proportion of risk. Cumulatively, our 124 SO-DEGs explain more (47.6%) of the liability to SO than the recent Multi-Ethnic Study of Atherosclerosis (MESA) transcriptomics study, which explained 29.7% of BMI variance. Pathway analyses exploring the functional relationships of all 124 SO-DEGs highlight key molecular mechanisms that may be disrupted in SO (Table S10). One of our significant GO terms, "phosphatidylinositol 3-kinase regulator activity," was also identified by the MESA study. This GO term includes *SOCS2* and *SOCS3*, genes identified in both the MESA study and our prior methylation studies.[95] These genes, as well as *CISH*, exhibited elevated expression in SO and also play key roles in leptin receptor and prolactin signaling. Leptin acts as a multifunctional hormone and has many receptors that activate intracellular signaling pathways in the central nervous system and peripheral tissues, including the JAK/STAT pathway, PI3K pathway, mitogen-activated protein kinase pathway, and 5′-AMP-activated protein kinase pathway.[96] Expression of *SOCS2*, *SOCS3*, and *CISH* is also stimulated by leptin, and this negative feedback mechanism prevents overactivation of leptin-signaling pathways, thus highlighting potential mechanisms of risk or consequences of SO.[97,98] IL-4R also signals through the JAK/STAT pathway after linking with either IL-4 or IL-13.[99] Notably, both the insulin and leptin signaling pathways were enriched in a recent study of gene expression and BMI.[61] These mechanistic insights into inherent factors that impact SO and downstream disease highlight novel targets for early prevention and pharmaceutical intervention, and increase the representativeness of the multiomics literature. Further studies of SO will increase our knowledge about its progression, mechanistic pathways, and intervention targets.

Another gene identified in our studies, *IL1B*, joins *CISH*, *SOCS2*, and *SOCS3* in the enrichment of the prolactin signaling pathway. *IL1B* is produced by activated macrophages as a proprotein, influencing other cytokines in an autoimmune network, and may destroy β cells in the pancreas, causing insulin-dependent diabetes.[100] In addition, in humans, *SOCS2* has been associated with body weight,[91] appendicular lean mass,[101] and initiation of puberty.[67] In addition to *SOCS2*'s key roles in adipose tissue remodeling, lipid metabolism, and energy metabolism, *SOCS2* knockout mice have reduced cytokine secretion and changes in both innate and adaptive immunity, further highlighting the role of inflammation and immunoregulation in SO.[102] Of note, prolactin receptors and dopamine receptors (the targets of prolactin) are present in human adipocytes and pancreatic β cells, and excess circulating prolactin is linked to increased obesity and glucose/insulin dysregulation.[103]

The broad array of outcomes associated with these genes and the molecular systems impacted by their dysregulation highlights the importance of understanding the broad clinical consequences of their collective function. To that end, in a phenome-wide enrichment analysis, we identified phenotypes more often associated with our discovered genes than expected by chance. These analyses identified 27 outcomes significantly enriched in our gene set, many of which have been supported through clinical and epidemiological studies of obesity. Sleep apnea has been implicated both in PheWAS[104] and MR studies of obesity[105] and was one of our strongest associations. We also identified strong enrichment of associations with hyperlipidemia, which is expected given the critical role of adipose tissue in lipid biology and metabolism and consequences of T2D. For example, anomalies of pupillary function are associated with diabetic autonomic neuropathy,[106]

and risk of osteomyelitis, gangrene, and disturbance of skin sensation as a consequence of T2D is well established.[107] In line with known relationships between obesity, diabetes, and bone health, we also observed enrichments with musculoskeletal diseases, including knee fractures, and hernia, with leg fractures and fractures of the upper limb, and with infectious disease risk, hospitalization, and mortality, all of which have been extensively documented.[108,109] The distribution of risk of morbidity and mortality from infections is often J-shaped, with subjects that are underweight and those with SO harboring the highest risk. Associations of SO with preterm birth,[4] extremely low birth weight,[4] and fetal growth retardation[5] are clinically well established. Our work adds to this literature by connecting transcriptomic dysregulation of these serious pregnancy complications with SO in adulthood.

Our phenome-wide enrichment analysis also revealed new associations. Although there is compelling clinical evidence linking obesity with adverse drug events,[110] the significant enrichment of GReX associations to adverse drug events and drug allergies among our identified genes highlight that shared regulation may impact drug metabolism and drug allergies and SO.[111] Additionally, psychiatric disorders have been clinically associated with obesity,[112] and previous PheWASs revealed associations between genetically predicted BMI and depression.[105] Here, we report shared genetic regulation associated with SO and bipolar disorder. Together, these clinical associations highlight the complex consequences of genetic dysregulation of genes associated with SO.

Limitations of the study

Although our study leverages high-dimensional multiomics to discover, replicate, generalize, and translate DEGs in SO, it is important to note several limitations. BMI is a surrogate measure of adiposity, and although widely used, does not distinguish lean and fat mass or distribution. Our subsequent analyses are designed to further substantiate the role of this gene in SO; however, lack of validation in these downstream analyses could occur for many reasons—the original finding might have been spurious, the gene/protein might not be measured on the assay, "winner's curse,"[113] insufficient power in validation analyses, or poor molecular typing. Because null findings are not equivalent to negative findings for the aforementioned reasons, we have focused our discussion on genes with significant evidence across multiple analyses. For example, our proteomics assays, while state of the art, tested only a small subset of our discovered genes, limiting our ability to assess the translation of differential transcript abundance through to measured protein for some genes. Additionally, WB, although the key circulating metabolically sensitive tissue, may not capture well some tissue-specific changes that occur as a cause or consequence of SO elsewhere in the body. Future studies will be needed to test the generalizability of our findings in larger samples and in additional obesity-relevant tissues beyond ASAT. Lastly, due to limited sample size, we were not adequately powered to conduct sex-stratified analyses.

Despite these limitations, our study brings together multiple lines of evidence across omics measures and tissues to characterize the DEGs in SO versus people classified as normal weight. We provide compelling evidence for genes whose expression is associated with SO in an underrepresented and disproportionately impacted population, observing highly concordant effects in an independent replication, generalization of effects in tissue (i.e., ASAT) with BMI, and translation effects in the proteome. Furthermore, our causal inference and PheWES analyses provide insight about the clinical consequences of genetic dysregulation of transcripts associated with SO. Our findings illuminate new mechanisms of risk and consequences of SO revealing novel targets for prevention, intervention, and treatment of SO. Exploring the heterogeneity of outcomes in those impacted by SO will be an important component of future research.

RESOURCE AVAILABILITY

Lead contact

For further information and requests for resources, reagents should be directed and will be fulfilled by the lead contact, Jennifer E. Below (jennifer.e.below@vumc.org).

Materials availability

This study did not generate any new unique reagents.

Data and code availability

- The CCHC transcriptomics and proteomics datasets have been submitted to the Database of Genotypes and Phenotypes (dbGaP) dbGaP: phs003894.v1.p1. Due to the scope of consent, data access to deidentified information requires dbGaP application to verify adequate protection of human subjects.
- The New York City transcriptomics data are available on GEO: GSE287627.
- MOSS summary data from trancriptomic analyses as described in Miao et al. are results included in Table S9.[43]
- Summary statistics for the complete discovery analysis are available at Zenodo: https://zenodo.org/records/14713355.

ACKNOWLEDGMENTS

We thank all participants and community partners without whom none of this work would be possible. Funding support for the present study was generously provided by the Center for Clinical and Translational Sciences; National Institutes of Health (NIH) Clinical and Translational Award (grant no. UL1TR000371) from the National Center for Advancing Translational Sciences; predoctoral fellowship from the American Heart Association (no. 18PRE34060101 to H.-H.C.); and NIH grant nos. U01CA288325, R01HL142302, R35HG010718, R01HG011138, and R01GM140287 (to E.R.G.), R01HL142302, R01DK127084, R01AG078452, and R01HL163262 (to J.E.B., H.-H.C., L.E.P., K.E.N., A.G., J.B.M., and S.P.F.-H.). K.E.N., P.G.-L., H.M.H., P.S., J.S., M.Y.A., M.K., M.G., D.K., and K.L.Y. are further supported by R01HL151152, R01DK122503, R01HD057194, R01HG010297, R01HL143885. The dataset(s) used for the analyses described were obtained from Vanderbilt University Medical Center's BioVU, which is supported by institutional funding, private agencies, and federal grants. These sources include the NIH-funded Shared Instrumentation grant no. S10RR025141 and Clinical and Translational Science Awards grant nos. UL1TR002243, UL1TR000445, and UL1RR024975. Genomic data are also supported by investigator-led projects from NIH that include U01HG004798, R01NS032830, RC2GM092618, P50GM115305, U01HG006378, U19HL065962, and R01HD074711, and additional funding sources listed at https://victr.vumc.org/biovu-funding/. P.J.M. is supported by Barts Charity grant no. MGU0587 and the UKRI Biotechnology and Biological Sciences Research Council grant no. BB/R006946/1. S.K.F. and K.K. are supported by NIH grant nos. R01DK121547 and R01DK080448 and by Internal funds (Mount Sinai School of Medicine).

AUTHOR CONTRIBUTIONS

Formal analysis: H.-H.C., X.Z., P.S., M.G., and K.L.Y. Data curation and visualization: H.-H.C., H.M.H., X.Z., R.R., P.S., H.G.P., L.E.P., J.S., M.Y.A., D.K.,

W.Z., D.M.S., A.E.J., L.F.-R., E.G.F., A.C.S., A.K., V.L.B., and M.K. Conceptualization, supervision, and investigation: A.G., P.J.M., C.A.A.-S., M.T.T.-L., L.L.M.-H., M.H.-H., P.P., R.V.S., P.G.-L., M.L., H.M.H., E.R.G., N.J.C., S.K.F., S.P.F.-H., J.B.M., K.E.N., and J.E.B. Resources: S.K.F., S.P.F.-H., and J.B.M. Validation: K.K. Writing – original draft: H.-H.C., H.M.H., P.S., K.E.N., and J.E.B. Writing – review & editing: all authors.

DECLARATION OF INTERESTS

D.M.S. is currently employed by AstraZeneca; all contributions to this work occurred prior to employment at AstraZeneca. N.J.C. is on the advisory board of Cell Genomics.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - CCHC study subjects
  - NYC ASAT
  - Mexican Obesity Study (MOSS) cohort
  - BioVU
- METHOD DETAILS
  - CCHC genetic data
  - CCHC RNA sequencing
  - CCHC proteomics
  - NYC ASAT collection and RNA sequencing
  - MOSS RNA sequencing
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - CCHC differential expression
  - Percent variance explained
  - CCHC eQTL mapping
  - CCHC proteomics
  - Correlation of ASAT gene expression and BMI
  - Mendelian Randomization
  - CCHC prospective analyses
  - MOSS cohort longitudinal differential expression
  - Pathway enrichment analysis
  - PheWES
  - Validation from extant body of literature
- ADDITIONAL RESOURCES

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.xgen.2025.100784.

Received: April 25, 2024
Revised: October 8, 2024
Accepted: February 6, 2025
Published: March 4, 2025

REFERENCES

1. Fryar, C.D., Carroll, M., and Ogden, C.L. (2018). Prevalence of overweight, obesity, and severe obesity among adults aged 20 and over: United States, 1960–1962 through 2015–2016.
2. Petersen, R., Pan, L., and Blanck, H.M. (2019). Racial and Ethnic Disparities in Adult Obesity in the United States: CDC's Tracking to Inform State and Local Action. Prev. Chronic Dis. *16*, E46. https://doi.org/10.5888/pcd16.180579.
3. Powell-Wiley, T.M., Poirier, P., Burke, L.E., Després, J.-P., Gordon-Larsen, P., Lavie, C.J., Lear, S.A., Ndumele, C.E., Neeland, I.J., Sanders, P., et al. (2021). Obesity and Cardiovascular Disease: A Scientific Statement From the American Heart Association. Circulation *143*, e984–e1010. https://doi.org/10.1161/CIR.0000000000000973.
4. McDonald, S.D., Han, Z., Mulla, S., and Beyene, J.; Knowledge Synthesis Group (2010). Overweight and obesity in mothers and risk of preterm birth and low birth weight infants: systematic review and meta-analyses. Br. Med. J. *341*, c3428. https://doi.org/10.1136/bmj.c3428.
5. Hong, Y.H., and Chung, S. (2018). Small for gestational age and obesity related comorbidities. Ann. Pediatr. Endocrinol. Metab. *23*, 4–8. https://doi.org/10.6065/apem.2018.23.1.4.
6. Theel, W., Boxma-de Klerk, B.M., Dirksmeier-Harinck, F., van Rossum, E.F.C., Kanhai, D.A., Apers, J., van Dalen, B.M., de Knegt, R.J., Holleboom, A.G., Tushuizen, M.E., et al. (2022). Evaluation of nonalcoholic fatty liver disease (NAFLD) in severe obesity using noninvasive tests and imaging techniques. Obes. Rev. *23*, e13481. https://doi.org/10.1111/obr.13481.
7. Malik, P.R.A., Doumouras, A.G., Malhan, R.S., Lee, Y., Boudreau, V., Barlow, K., Karpinski, M., and Anvari, M. (2021). Obesity, Cancer, and Risk Reduction with Bariatric Surgery. Surg. Clin. *101*, 239–254. https://doi.org/10.1016/j.suc.2020.12.003.
8. Nehus, E. (2018). Obesity and chronic kidney disease. Curr. Opin. Pediatr. *30*, 241–246. https://doi.org/10.1097/MOP.0000000000000586.
9. Rosser, F.J., Forno, E., Cooper, P.J., and Celedón, J.C. (2014). Asthma in Hispanics. An 8-year update. Am. J. Respir. Crit. Care Med. *189*, 1316–1327. https://doi.org/10.1164/rccm.201401-0186PP.
10. Ogden, C.L., Fryar, C.D., Martin, C.B., Freedman, D.S., Carroll, M.D., Gu, Q., and Hales, C.M. (2020). Trends in Obesity Prevalence by Race and Hispanic Origin—1999-2000 to 2017-2018. JAMA *324*, 1208–1210. https://doi.org/10.1001/jama.2020.14590.
11. Kitahara, C.M., Flint, A.J., Berrington de Gonzalez, A., Bernstein, L., Brotzman, M., MacInnis, R.J., Moore, S.C., Robien, K., Rosenberg, P.S., Singh, P.N., et al. (2014). Association between class III obesity (BMI of 40-59 kg/m2) and mortality: a pooled analysis of 20 prospective studies. PLoS Med. *11*, e1001673. https://doi.org/10.1371/journal.pmed.1001673.
12. Daviglus, M.L., Talavera, G.A., Aviles-Santa, M.L., Allison, M., Cai, J., Criqui, M.H., Gellman, M., Giachello, A.L., Gouskova, N., Kaplan, R.C., et al. (2012). Prevalence of major cardiovascular risk factors and cardiovascular diseases among Hispanic/Latino individuals of diverse backgrounds in the United States. JAMA, J. Am. Med. Assoc. *308*, 1775–1784. https://doi.org/10.1001/jama.2012.14517.
13. Robertson, C., Aceves-Martins, M., Cruickshank, M., Imamura, M., and Avenell, A. (2022). Does weight management research for adults with severe obesity represent them? Analysis of systematic review data. BMJ Open *12*, e054459. https://doi.org/10.1136/bmjopen-2021-054459.
14. Cotsapas, C., Speliotes, E.K., Hatoum, I.J., Greenawalt, D.M., Dobrin, R., Lum, P.Y., Suver, C., Chudin, E., Kemp, D., Reitman, M., et al. (2009). Common body mass index-associated variants confer risk of extreme obesity. Hum. Mol. Genet. *18*, 3502–3507. https://doi.org/10.1093/hmg/ddp292.
15. Clyde, D. (2019). Making the case for more inclusive GWAS. Nat. Rev. Genet. *20*, 500–501. https://doi.org/10.1038/s41576-019-0160-0.
16. Locke, A.E., Kahali, B., Berndt, S.I., Justice, A.E., Pers, T.H., Day, F.R., Powell, C., Vedantam, S., Buchkovich, M.L., Yang, J., et al. (2015). Genetic studies of body mass indexyield new insights for obesity biology. Nature *518*, 197–206. https://doi.org/10.1038/nature14177.
17. Ghosh, S., and Bouchard, C. (2017). Convergence between biological, behavioural and genetic determinants of obesity. Nat. Rev. Genet. *18*, 731–748. https://doi.org/10.1038/nrg.2017.72.
18. Diaz-Melean, C.M., Somers, V.K., Rodriguez-Escudero, J.P., Singh, P., Sochor, O., Llano, E.M., and Lopez-Jimenez, F. (2013). Mechanisms of adverse cardiometabolic consequences of obesity. Curr. Atheroscler Rep. *15*, 364. https://doi.org/10.1007/s11883-013-0364-2.

19. Graff, M., Emery, L.S., Justice, A.E., Parra, E., Below, J.E., Palmer, N.D., Gao, C., Duan, Q., Valladares-Salgado, A., Cruz, M., et al. (2017). Genetic architecture of lipid traits in the Hispanic community health study/study of Latinos. Lipids Health Dis. *16*, 200. https://doi.org/10.1186/s12944-017-0591-6.

20. Fernández-Rhodes, L., Young, K.L., Lilly, A.G., Raffield, L.M., Highland, H.M., Wojcik, G.L., Agler, C., Love, S.M., Okello, S., Petty, L.E., et al. (2020). Importance of Genetic Studies of Cardiometabolic Disease in Diverse Populations. Circ. Res. *126*, 1816–1840. https://doi.org/10.1161/CIRCRESAHA.120.315893.

21. Downie, C.G., Dimos, S.F., Bien, S.A., Hu, Y., Darst, B.F., Polfus, L.M., Wang, Y., Wojcik, G.L., Tao, R., Raffield, L.M., et al. (2022). Multi-ethnic GWAS and fine-mapping of glycaemic traits identify novel loci in the PAGE Study. Diabetologia *65*, 477–489. https://doi.org/10.1007/s00125-021-05635-9.

22. Chen, H.H., Petty, L.E., North, K.E., McCormick, J.B., Fisher-Hoch, S.P., Gamazon, E.R., and Below, J.E. (2022). Novel diabetes gene discovery through comprehensive characterization and integrative analysis of longitudinal gene expression changes. Hum. Mol. Genet. *31*, 3191–3205. https://doi.org/10.1093/hmg/ddac039.

23. Below, J.E., Parra, E.J., Gamazon, E.R., Torres, J., Krithika, S., Candille, S., Lu, Y., Manichakul, A., Peralta-Romero, J., Duan, Q., et al. (2016). Meta-analysis of lipid-traits in Hispanics identifies novel loci, population-specific effects, and tissue-specific enrichment of eQTLs. Sci. Rep. *6*, 19429. https://doi.org/10.1038/srep19429.

24. Parra, E.J., Mazurek, A., Gignoux, C.R., Sockell, A., Agostino, M., Morris, A.P., Petty, L.E., Hanis, C.L., Cox, N.J., Valladares-Salgado, A., et al. (2017). Admixture mapping in two Mexican samples identifies significant associations of locus ancestry with triglyceride levels in the BUD13/ZNF259/APOA5 region and fine mapping points to rs964184 as the main driver of the association signal. PLoS One *12*, e0172880. https://doi.org/10.1371/journal.pone.0172880.

25. Sollis, E., Mosaku, A., Abid, A., Buniello, A., Cerezo, M., Gil, L., Groza, T., Güneş, O., Hall, P., Hayhurst, J., et al. (2023). The NHGRI-EBI GWAS Catalog: knowledgebase and deposition resource. Nucleic Acids Res. *51*, D977–D985. https://doi.org/10.1093/nar/gkac1010.

26. Amundadottir, L.T., Sulem, P., Gudmundsson, J., Helgason, A., Baker, A., Agnarsson, B.A., Sigurdsson, A., Benediktsdottir, K.R., Cazier, J.B., Sainz, J., et al. (2006). A common variant associated with prostate cancer in European and African populations. Nat. Genet. *38*, 652–658. https://doi.org/10.1038/ng1808.

27. Fritsche, L.G., Chen, W., Schu, M., Yaspan, B.L., Yu, Y., Thorleifsson, G., Zack, D.J., Arakawa, S., Cipriani, V., Ripke, S., et al. (2013). Seven new loci associated with age-related macular degeneration. Nat. Genet. *45*, 433–439e4392. https://doi.org/10.1038/ng.2578.

28. Genovese, G., Friedman, D.J., Ross, M.D., Lecordier, L., Uzureau, P., Freedman, B.I., Bowden, D.W., Langefeld, C.D., Oleksyk, T.K., Uscinski Knob, A.L., et al. (2010). Association of trypanolytic ApoL1 variants with kidney disease in African Americans. Science *329*, 841–845. https://doi.org/10.1126/science.1193032.

29. Jonsson, T., Stefansson, H., Steinberg, S., Jonsdottir, I., Jonsson, P.V., Snaedal, J., Bjornsson, S., Huttenlocher, J., Levey, A.I., Lah, J.J., et al. (2013). Variant of TREM2 Associated with the Risk of Alzheimer's Disease. N. Engl. J. Med. *368*, 107–116. https://doi.org/10.1056/NEJMoa1211103.

30. Jostins, L., Ripke, S., Weersma, R.K., Duerr, R.H., McGovern, D.P., Hui, K.Y., Lee, J.C., Schumm, L.P., Sharma, Y., Anderson, C.A., et al. (2012). Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. Nature *491*, 119–124. https://doi.org/10.1038/nature11582.

31. Morris, A.P., Voight, B.F., Teslovich, T.M., Ferreira, T., Segrè, A.V., Steinthorsdottir, V., Strawbridge, R.J., Khan, H., Grallert, H., Mahajan, A., et al. (2012). Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. Nat. Genet. *44*, 981–990. https://doi.org/10.1038/ng.2383.

32. International Multiple Sclerosis Genetics Consortium; Wellcome Trust Case Control Consortium 2; Sawcer, S., Hellenthal, G., Pirinen, M., Spencer, C.C.A., Patsopoulos, N.A., Moutsianas, L., Dilthey, A., and Su, Z. (2011). Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. Nature *476*, 214–219. https://doi.org/10.1038/nature10251.

33. Scott, R.A., Lagou, V., Welch, R.P., Wheeler, E., Montasser, M.E., Luan, J., Mägi, R., Strawbridge, R.J., Rehnberg, E., Gustafsson, S., et al. (2012). Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. Nat. Genet. *44*, 991–1005. https://doi.org/10.1038/ng.2385.

34. Smyth, D.J., Cooper, J.D., Bailey, R., Field, S., Burren, O., Smink, L.J., Guja, C., Ionescu-Tirgoviste, C., Widmer, B., Dunger, D.B., et al. (2006). A genome-wide association study of nonsynonymous SNPs identifies a type 1 diabetes locus in the interferon-induced helicase (IFIH1) region. Nat. Genet. *38*, 617–619. https://doi.org/10.1038/ng1800.

35. Demerath, E.W., Guan, W., Grove, M.L., Aslibekyan, S., Mendelson, M., Zhou, Y.H., Hedman, Å.K., Sandling, J.K., Li, L.A., Irvin, M.R., et al. (2015). Epigenome-wide association study (EWAS) of BMI, BMI change and waist circumference in African American adults identifies multiple replicated loci. Hum. Mol. Genet. *24*, 4464–4479. https://doi.org/10.1093/hmg/ddv161.

36. Joseph, P.V., Jaime-Lara, R.B., Wang, Y., xiang, L., and Henderson, W.A. (2019). Comprehensive and Systematic Analysis of Gene Expression Patterns Associated with Body Mass Index. Sci. Rep. *9*, 7447. https://doi.org/10.1038/s41598-019-43881-5.

37. Berisha, S.Z., Serre, D., Schauer, P., Kashyap, S.R., and Smith, J.D. (2011). Changes in whole blood gene expression in obese subjects with type 2 diabetes following bariatric surgery: a pilot study. PLoS One *6*, e16729. https://doi.org/10.1371/journal.pone.0016729.

38. Pinhel, M.A.S., Noronha, N.Y., Nicoletti, C.F., Pereira, V.A., de Oliveira, B.A., Cortes-Oliveira, C., Salgado, W., Jr., Barbosa, F., Jr., Marchini, J.S., Souza, D.R., and Nonino, C.B. (2020). Changes in DNA methylation and gene expression of insulin and obesity-related gene PIK3R1 after Roux-en-Y gastric bypass. Int. J. Mol. Sci. *21*, 4476. https://doi.org/10.3390/ijms21124476.

39. Frederiksen, L., Nielsen, T.L., Wraae, K., Hagen, C., Frystyk, J., Flyvbjerg, A., Brixen, K., and Andersen, M. (2009). Subcutaneous rather than visceral adipose tissue is associated with adiponectin levels and insulin resistance in young men. J. Clin. Endocrinol. Metab. *94*, 4010–4015. https://doi.org/10.1210/jc.2009-0980.

40. Wajchenberg, B.L. (2000). Subcutaneous and visceral adipose tissue: their relation to the metabolic syndrome. Endocr. Rev. *21*, 697–738. https://doi.org/10.1210/edrv.21.6.0415.

41. Fain, J.N., Madan, A.K., Hiler, M.L., Cheema, P., and Bahouth, S.W. (2004). Comparison of the release of adipokines by adipose tissue, adipose tissue matrix, and adipocytes from visceral and subcutaneous abdominal adipose tissues of obese humans. Endocrinology *145*, 2273–2282. https://doi.org/10.1210/en.2003-1336.

42. Li, S., Todor, A., and Luo, R. (2016). Blood transcriptomics and metabolomics for personalized medicine. Comput. Struct. Biotechnol. J. *14*, 1–7. https://doi.org/10.1016/j.csbj.2015.10.005.

43. Miao, Z., Alvarez, M., Ko, A., Bhagat, Y., Rahmani, E., Jew, B., Heinonen, S., Muñoz-Hernandez, L.L., Herrera-Hernandez, M., Aguilar-Salinas, C., et al. (2020). The causal effect of obesity on prediabetes and insulin resistance reveals the important role of adipose tissue in insulin resistance. PLoS Genet. *16*, e1009018. https://doi.org/10.1371/journal.pgen.1009018.

44. Pan, D.Z., Miao, Z., Comenho, C., Rajkumar, S., Koka, A., Lee, S.H.T., Alvarez, M., Kaminska, D., Ko, A., Sinsheimer, J.S., et al. (2021). Identification of TBx15 as an adipose master trans regulator of abdominal obesity genes. Genome Med. *13*, 123. https://doi.org/10.1186/s13073-021-00939-2.

45. Aguirre-Gamboa, R., de Klein, N., di Tommaso, J., Claringbould, A., van der Wijst, M.G., de Vries, D., Brugge, H., Oelen, R., Võsa, U., Zorro, M.M., et al. (2020). Deconvolution of bulk blood eQTL effects into immune cell subpopulations. BMC Bioinf. *21*, 243. https://doi.org/10.1186/s12859-020-03576-5.

46. Buniello, A., MacArthur, J.A.L., Cerezo, M., Harris, L.W., Hayhurst, J., Malangone, C., McMahon, A., Morales, J., Mountjoy, E., Sollis, E., et al. (2019). The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. Nucleic Acids Res. *47*, D1005–D1012. https://doi.org/10.1093/nar/gky1120.

47. Common Metabolic Diseases Knowledge Portal (cmdkp.org). https://hugeamp.org/.

48. Dornbos, P., Singh, P., Jang, D.K., Mahajan, A., Biddinger, S.B., Rotter, J.I., McCarthy, M.I., and Flannick, J. (2022). Evaluating human genetic support for hypothesized metabolic disease genes. Cell Metab. *34*, 661–666. https://doi.org/10.1016/j.cmet.2022.03.011.

49. Heberle, H., Meirelles, G.V., da Silva, F.R., Telles, G.P., and Minghim, R. (2015). InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. BMC Bioinf. *16*, 169. https://doi.org/10.1186/s12859-015-0611-3.

50. Ali, O., Cerjak, D., Kent, J.W., Jr., James, R., Blangero, J., Carless, M.A., and Zhang, Y. (2016). Methylation of SOCS3 is inversely associated with metabolic syndrome in an epigenome-wide association study of obesity. Epigenetics *11*, 699–707. https://doi.org/10.1080/15592294.2016.1216284.

51. Wahl, S., Drong, A., Lehne, B., Loh, M., Scott, W.R., Kunze, S., Tsai, P.C., Ried, J.S., Zhang, W., Yang, Y., et al. (2017). Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. Nature *541*, 81–86. https://doi.org/10.1038/nature20784.

52. Sun, D., Zhang, T., Su, S., Hao, G., Chen, T., Li, Q.Z., Bazzano, L., He, J., Wang, x., Li, S., and Chen, W. (2019). Body Mass IndexDrives Changes in DNA Methylation: A Longitudinal Study. Circ. Res. *125*, 824–833. https://doi.org/10.1161/CIRCRESAHA.119.315397.

53. Chen, Y., Kassam, I., Lau, S.H., Kooner, J.S., Wilson, R., Peters, A., Winkelmann, J., Chambers, J.C., Chow, V.T., Khor, C.C., et al. (2021). Impact of BMI and waist circumference on epigenome-wide DNA methylation and identification of epigenetic biomarkers in blood: an EWAS in multi-ethnic Asian individuals. Clin. Epigenet. *13*, 195. https://doi.org/10.1186/s13148-021-01162-x.

54. Mendelson, M.M., Marioni, R.E., Joehanes, R., Liu, C., Hedman, Å.K., Aslibekyan, S., Demerath, E.W., Guan, W., Zhi, D., Yao, C., et al. (2017). Association of Body Mass Indexwith DNA Methylation and Gene Expression in Blood Cells and Relations to Cardiometabolic Disease: A Mendelian Randomization Approach. PLoS Med. *14*, e1002215. https://doi.org/10.1371/journal.pmed.1002215.

55. Vehmeijer, F.O.L., Küpers, L.K., Sharp, G.C., Salas, L.A., Lent, S., Jima, D.D., Tindula, G., Reese, S., Qi, C., Gruzieva, O., et al. (2020). DNA methylation and body mass indexfrom birth to adolescence: meta-analyses of epigenome-wide association studies. Genome Med. *12*, 105. https://doi.org/10.1186/s13073-020-00810-w.

56. xu, K., Zhang, x., Wang, Z., Hu, Y., and Sinha, R. (2018). Epigenome-wide association analysis revealed that SOCS3 methylation influences the effect of cumulative stress on obesity. Biol. Psychol. *131*, 63–71. https://doi.org/10.1016/j.biopsycho.2016.11.001.

57. Bowden, J., Davey Smith, G., Haycock, P.C., and Burgess, S. (2016). Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. Genet. Epidemiol. *40*, 304–314. https://doi.org/10.1002/gepi.21965.

58. Davies, N.M., Holmes, M.V., and Davey Smith, G. (2018). Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. Br. Med. J. *362*, k601. https://doi.org/10.1136/bmj.k601.

59. Haycock, P.C., Burgess, S., Wade, K.H., Bowden, J., Relton, C., and Davey Smith, G. (2016). Best (but oft-forgotten) practices: the design, analysis, and interpretation of Mendelian randomization studies. Am. J. Clin. Nutr. *103*, 965–978. https://doi.org/10.3945/ajcn.115.118216.

60. Krishnan M,. Anwar M.Y., Justice A.E., Chittoor G., Chen H., Roshani R., Smit R.A.J., Preuss M.H., Chami N., Hadad B.S, et al. Genome-wide association study provides novel insight into the genetic architecture of severe obesity. (2025) medRXiV (MEDRXIV/2025/322870).

61. Vargas, L.B., Lange, L.A., Ferrier, K., Aguet, F., Ardlie, K., Gabriel, S., Gupta, N., Smith, J.D., Blackwell, T.W., Ding, J., et al. (2023). Gene expression associations with body mass indexin the Multi-Ethnic Study of Atherosclerosis. Int. J. Obes. *47*, 109–116. https://doi.org/10.1038/s41366-022-01240-x.

62. Tsao, C.H., Shiau, M.Y., Chuang, P.H., Chang, Y.H., and Hwang, J. (2014). Interleukin-4 regulates lipid metabolism by inhibiting adipogenesis and promoting lipolysis. J. Lipid Res. *55*, 385–397. https://doi.org/10.1194/jlr.M041392.

63. GTEx Consortium (2020). The GTExConsortium atlas of genetic regulatory effects across human tissues. Science *369*, 1318–1330. https://doi.org/10.1126/science.aaz1776.

64. Massoud, A.H., Charbonnier, L.M., Lopez, D., Pellegrini, M., Phipatanakul, W., and Chatila, T.A. (2016). An asthma-associated IL4R variant exacerbates airway inflammation by promoting conversion of regulatory T cells to TH17-like cells. Nat. Med. *22*, 1013–1022. https://doi.org/10.1038/nm.4147.

65. Johansson, E.K., Ivert, L.U., Bradley, B., Lundqvist, M., and Bradley, M. (2020). Weight gain in patients with severe atopic dermatitis treated with dupilumab: a cohort study. BMC Dermatol. *20*, 8. https://doi.org/10.1186/s12895-020-00103-0.

66. Westholm, E., Edlund, A., Karagiannopoulos, A., Wendt, A., and Eliasson, L. (2023). Interleukin-4 reduces insulin secretion in human islets from healthy but not type-2 diabetic donors. Biochem. Biophys. Res. Commun. *649*, 87–92. https://doi.org/10.1016/j.bbrc.2023.01.092.

67. Hübel, C., Gaspar, H.A., Coleman, J.R.I., Finucane, H., Purves, K.L., Hanscombe, K.B., Prokopenko, I., investigators, M., Graff, M., Ngwa, J.S., et al. (2019). Genomics of body fat percentage may contribute to sexbias in anorexia nervosa. Am J Med Genet B Neuropsychiatr Genet. *180*, 428–438. https://doi.org/10.1002/ajmg.b.32709.

68. Pei, Y.F., Liu, Y.Z., Yang, x.L., Zhang, H., Feng, G.J., Wei, x.T., and Zhang, L. (2020). The genetic architecture of appendicular lean mass characterized by association analysis in the UK Biobank study. Commun. Biol. *3*, 608. https://doi.org/10.1038/s42003-020-01334-0.

69. Rajakumar, K., Yan, Q., Khalid, A.T., Feingold, E., Vallejo, A.N., Demirci, F.Y., and Kamboh, M.I. (2019). Gene Expression and Cardiometabolic Phenotypes of Vitamin D-Deficient Overweight and Obese Black Children. Nutrients *11*, 2016. https://doi.org/10.3390/nu11092016.

70. Lantieri, F., and Bachetti, T. (2022). OSM/OSMR and Interleukin 6 Family Cytokines in Physiological and Pathological Condition. Int. J. Mol. Sci. *23*, 11096. https://doi.org/10.3390/ijms231911096.

71. Bailey, J.L., Hang, H., Boudreau, A., and Elks, C.M. (2022). Oncostatin M Induces Lipolysis and Suppresses Insulin Response in 3T3-L1 Adipocytes. Int. J. Mol. Sci. *23*, 4689. https://doi.org/10.3390/ijms23094689.

72. Sanchez-Infantes, D., and Stephens, J.M. (2020). Adipocyte Oncostatin Receptor Regulates Adipose Tissue Homeostasis and Inflammation. Front. Immunol. *11*, 612013. https://doi.org/10.3389/fimmu.2020.612013.

73. Piquer-Garcia, I., Campderros, L., Taxerås, S.D., Gavaldà-Navarro, A., Pardo, R., Vila, M., Pellitero, S., Martínez, E., Tarascó, J., Moreno, P., et al. (2020). A Role for Oncostatin M in the Impairment of Glucose Homeostasis in Obesity. J. Clin. Endocrinol. Metab. *105*, e337–e348. https://doi.org/10.1210/clinem/dgz090.

74. Ahn, S., Shenoy, S.K., Luttrell, L.M., and Lefkowitz, R.J. (2020). SnapShot: beta-Arrestin Functions. Cell *182*, 1362–1362.e1. https://doi.org/10.1016/j.cell.2020.07.034.

75. Zecchini, S., Proietti Serafini, F., Catalani, E., Giovarelli, M., Coazzoli, M., Di Renzo, I., De Palma, C., Perrotta, C., Clementi, E., Buonanno, F., et al. (2018). Dysfunctional autophagy induced by the pro-apoptotic natural compound climacostol in tumour cells. Cell Death Dis. *10*, 10. https://doi.org/10.1038/s41419-018-1254-x.

76. Barella, L.F., Rossi, M., Pydi, S.P., Meister, J., Jain, S., Cui, Y., Gavrilova, O., Fulgenzi, G., Tessarollo, L., and Wess, J. (2021). β-Arrestin-1 is required for adaptive β-cell mass expansion during obesity. Nat. Commun. *12*, 3385. https://doi.org/10.1038/s41467-021-23656-1.

77. Bueichekú, E., Aznarez-Sanado, M., Diez, I., d'Oleire Uquillas, F., Ortiz-Teran, L., Qureshi, A.Y., Sunol, M., Basaia, S., Ortiz-Teran, E., Pastor, M.A., and Sepulcre, J. (2020). Central neurogenetic signatures of the visuomotor integration system. Proc. Natl. Acad. Sci.USA *117*, 6836–6843. https://doi.org/10.1073/pnas.1912429117.

78. Cao, Y., Li, R., Li, Y., Zhang, T., Wu, N., Zhang, J., and Guo, Z. (2017). Identification of Transcription Factor-Gene Regulatory Network in Acute Myocardial Infarction. Heart Lung Circ. *26*, 343–353. https://doi.org/10.1016/j.hlc.2016.06.1209.

79. xu, H., Dorn, G.W., 2nd, Shetty, A., Parihar, A., Dave, T., Robinson, S.W., Gottlieb, S.S., Donahue, M.P., Tomaselli, G.F., Kraus, W.E., et al. (2018). A Genome-Wide Association Study of Idiopathic Dilated Cardiomyopathy in African Americans. J. Personalized Med. *8*, 11. https://doi.org/10.3390/jpm8010011.

80. Kanai, M., Akiyama, M., Takahashi, A., Matoba, N., Momozawa, Y., Ikeda, M., Iwata, N., Ikegawa, S., Hirata, M., Matsuda, K., et al. (2018). Genetic analysis of quantitative traits in the Japanese population links cell types to complexhuman diseases. Nat. Genet. *50*, 390–400. https://doi.org/10.1038/s41588-018-0047-6.

81. Hoffmann, T.J., Theusch, E., Haldar, T., Ranatunga, D.K., Jorgenson, E., Medina, M.W., Kvale, M.N., Kwok, P.Y., Schaefer, C., Krauss, R.M., et al. (2018). A large electronic-health-record-based genome-wide study of serum lipids. Nat. Genet. *50*, 401–413. https://doi.org/10.1038/s41588-018-0064-5.

82. Richardson, T.G., Sanderson, E., Palmer, T.M., Ala-Korpela, M., Ference, B.A., Davey Smith, G., and Holmes, M.V. (2020). Evaluating the relationship between circulating lipoprotein lipids and apolipoproteins with risk of coronary heart disease: A multivariable Mendelian randomisation analysis. PLoS Med. *17*, e1003062. https://doi.org/10.1371/journal.pmed.1003062.

83. Ruth, K.S., Day, F.R., Tyrrell, J., Thompson, D.J., Wood, A.R., Mahajan, A., Beaumont, R.N., Wittemans, L., Martin, S., Busch, A.S., et al. (2020). Using human genetics to understand the disease impacts of testosterone in men and women. Nat. Med. *26*, 252–258. https://doi.org/10.1038/s41591-020-0751-5.

84. Vujkovic, M., Keaton, J.M., Lynch, J.A., Miller, D.R., Zhou, J., Tcheandjieu, C., Huffman, J.E., Assimes, T.L., Lorenz, K., Zhu, x., et al. (2020). Discovery of 318 new risk loci for type 2 diabetes and related vascular outcomes among 1.4 million participants in a multi-ancestry meta-analysis. Nat. Genet. *52*, 680–691. https://doi.org/10.1038/s41588-020-0637-y.

85. Sharp, G.C., Salas, L.A., Monnereau, C., Allard, C., Yousefi, P., Everson, T.M., Bohlin, J., xu, Z., Huang, R.C., Reese, S.E., et al. (2017). Maternal BMI at the start of pregnancy and offspring epigenome-wide DNA methylation: findings from the pregnancy and childhood epigenetics (PACE) consortium. Hum. Mol. Genet. *26*, 4067–4085. https://doi.org/10.1093/hmg/ddx290.

86. Yengo, L., Vedantam, S., Marouli, E., Sidorenko, J., Bartell, E., Sakaue, S., Graff, M., Eliasen, A.U., Jiang, Y., Raghavan, S., et al. (2022). A saturated map of common genetic variants associated with human height. Nature *610*, 704–712. https://doi.org/10.1038/s41586-022-05275-y.

87. Wang, Y., Zhang, Y., Zhang, C., Weng, H., Li, Y., Cai, W., xie, M., Long, Y., Ai, Q., Liu, Z., et al. (2015). The gene pair PRR11 and SKA2 shares a NF-Y-regulated bidirectional promoter and contributes to lung cancer development. Biochim. Biophys. Acta *1849*, 1133–1144. https://doi.org/10.1016/j.bbagrm.2015.07.002.

88. Guintivano, J., Brown, T., Newcomer, A., Jones, M., Cox, O., Maher, B.S., Eaton, W.W., Payne, J.L., Wilcox, H.C., and Kaminsky, Z.A. (2014). Identification and replication of a combined epigenetic and genetic biomarker predicting suicide and suicidal behaviors. Am. J. Psychiatr. *171*, 1287–1296. https://doi.org/10.1176/appi.ajp.2014.14010008.

89. Carithers, L.J., and Moore, H.M. (2015). The Genotype-Tissue Expression (GTEx) Project. Biopreserv. Biobanking *13*, 307–308. https://doi.org/10.1089/bio.2015.29031.hmm.

90. Pandey, G.N., Rizavi, H.S., Zhang, H., Bhaumik, R., and Ren, x. (2016). The Expression of the Suicide-Associated Gene SKA2 Is Decreased in the Prefrontal Cortexof Suicide Victims but Not of Nonsuicidal Patients. Int. J. Neuropsychopharmacol. *19*, pyw015. https://doi.org/10.1093/ijnp/pyw015.

91. Sakaue, S., Kanai, M., Tanigawa, Y., Karjalainen, J., Kurki, M., Koshiba, S., Narita, A., Konuma, T., Yamamoto, K., Akiyama, M., et al. (2021). A cross-population atlas of genetic associations for 220 human phenotypes. Nat. Genet. *53*, 1415–1424. https://doi.org/10.1038/s41588-021-00931-x.

92. Kichaev, G., Bhatia, G., Loh, P.R., Gazal, S., Burch, K., Freund, M.K., Schoech, A., Pasaniuc, B., and Price, A.L. (2019). Leveraging Polygenic Functional Enrichment to Improve GWAS Power. Am. J. Hum. Genet. *104*, 65–75. https://doi.org/10.1016/j.ajhg.2018.11.008.

93. Yin, x., xu, Z., Zhang, Z., Li, L., Pan, Q., Zheng, F., and Li, H. (2017). Association of PI3K/AKT/mTOR pathway genetic variants with type 2 diabetes mellitus in Chinese. Diabetes Res. Clin. Pract. *128*, 127–135. https://doi.org/10.1016/j.diabres.2017.04.002.

94. Eshaghi, F.S., Ghazizadeh, H., Kazami-Nooreini, S., Timar, A., Esmaeily, H., Mehramiz, M., Avan, A., and Ghayour-Mobarhan, M. (2019). Association of a genetic variant in AKT1 gene with features of the metabolic syndrome. Genes Dis. *6*, 290–295. https://doi.org/10.1016/j.gendis.2019.03.002.

95. Lavens, D., Montoye, T., Piessevaux, J., Zabeau, L., Vandekerckhove, J., Gevaert, K., Becker, W., Eyckerman, S., and Tavernier, J. (2006). A complexinteraction pattern of CIS and SOCS2 with the leptin receptor. J. Cell Sci. *119*, 2214–2224. https://doi.org/10.1242/jcs.02947.

96. Hu, x., li, J., Fu, M., Zhao, x., and Wang, W. (2021). The JAK/STAT signaling pathway: from bench to clinic. Signal Transduct. Targeted Ther. *6*, 402. https://doi.org/10.1038/s41392-021-00791-1.

97. Howard, J.K., and Flier, J.S. (2006). Attenuation of leptin and insulin signaling by SOCS proteins. Trends Endocrinol. Metabol. *17*, 365–371. https://doi.org/10.1016/j.tem.2006.09.007.

98. Park, H.K., and Ahima, R.S. (2014). Leptin signaling. F1000Prime Rep. *6*, 73. https://doi.org/10.12703/P6-73.

99. Hu, x., Li, J., Fu, M., Zhao, x., and Wang, W. (2021). The JAK/STAT signaling pathway: from bench to clinic. Signal Transduct. Targeted Ther. *6*, 402. https://doi.org/10.1038/s41392-021-00791-1.

100. Pociot, F., Mølvig, J., Wogensen, L., Worsaae, H., and Nerup, J. (1992). A TaqI polymorphism in the human interleukin-1 beta (IL-1 beta) gene correlates with IL-1 beta secretion in vitro. Eur. J. Clin. Invest. *22*, 396–402. https://doi.org/10.1111/j.1365-2362.1992.tb01480.x.

101. Hernandez Cordero, A.I., Gonzales, N.M., Parker, C.C., Sokolof, G., Vandenbergh, D.J., Cheng, R., Abney, M., Sko, A., Douglas, A., Palmer, A.A., et al. (2019). Genome-wide Associations Reveal Human-Mouse Genetic Convergence and Modifiers of Myogenesis, CPNE1 and STC2. Am. J. Hum. Genet. *105*, 1222–1236. https://doi.org/10.1016/j.ajhg.2019.10.014.

102. Val, C.H., de Oliveira, M.C., Lacerda, D.R., Barroso, A., Batista, N.V., Menezes-Garcia, Z., de Assis, D.R.R., Cramer, A.T., Brant, F., Teixeira, M.M., et al. (2020). SOCS2 modulates adipose tissue inflammation and expansion in mice. J. Nutr. Biochem. *76*, 108304. https://doi.org/10.1016/j.jnutbio.2019.108304.

103. Pirchio, R., Graziadio, C., Colao, A., Pivonello, R., and Auriemma, R.S. (2022). Metabolic effects of prolactin. Front. Endocrinol. *13*, 1015520. https://doi.org/10.3389/fendo.2022.1015520.

104. Zheng, N.S., Feng, Q., Kerchberger, V.E., Zhao, J., Edwards, T.L., Cox, N.J., Stein, C.M., Roden, D.M., Denny, J.C., and Wei, W.Q. (2020). PheMap: a multi-resource knowledge base for high-throughput phenotyping within electronic health records. J. Am. Med. Inf. Assoc. *27*, 1675–1687. https://doi.org/10.1093/jamia/ocaa104.

105. Hyppönen, E., Mulugeta, A., Zhou, A., and Santhanakrishnan, V.K. (2019). A data-driven approach for studying the role of body mass in multiple diseases: a phenome-wide registry-based case-control study in the UK Biobank. Lancet Digit Health. *1*, e116–e126. https://doi.org/10.1016/S2589-7500(19)30028-7.

106. Tan, T.E., Finkelstein, M.T., Tan, G.S.W., Tan, A.C.S., Chan, C.M., Mathur, R., Wong, E.Y.M., Cheung, C.M.G., Wong, T.Y., Milea, D., and Najjar, R.P. (2022). Retinal neural dysfunction in diabetes revealed with handheld chromatic pupillometry. Clin. Exp. Ophthalmol. *50*, 745–756. https://doi.org/10.1111/ceo.14116.

107. Van, G.H., Amouyal, C., Bourron, O., Aubert, C., Carlier, A., Mosbah, H., Fourniols, E., Cluzel, P., Couture, T., and Hartemann, A. (2021). Diabetic foot ulcer management in a multidisciplinary foot centre: one-year healing, amputation and mortality rate. J. Wound Care *30*, S34–S41. https://doi.org/10.12968/jowc.2021.30.Sup6.S34.

108. Yang, W.S., Chang, Y.C., Chang, C.H., Wu, L.C., Wang, J.L., and Lin, H.H. (2021). The Association Between Body Mass Indexand the Risk of Hospitalization and Mortality due to Infection: A Prospective Cohort Study. Open Forum Infect. Dis. *8*, ofaa545. https://doi.org/10.1093/ofid/ofaa545.

109. Premaor, M.O., Comim, F.V., and Compston, J.E. (2014). Obesity and fractures. Arq. Bras. Endocrinol. Metabol. *58*, 470–477. https://doi.org/10.1590/0004-2730000003274.

110. Bettonte, S., Berton, M., and Marzolini, C. (2022). Magnitude of Drug-Drug Interactions in Special Populations. Pharmaceutics *14*, 789. https://doi.org/10.3390/pharmaceutics14040789.

111. Dias de Castro, E., Pinhão, S., Paredes, S., Cernadas, J.R., and Ribeiro, L. (2021). Obesity markers in patients with drug allergy and body fat as a predictor. Ann. Allergy Asthma Immunol. *127*, 100–108. https://doi.org/10.1016/j.anai.2021.03.014.

112. Perry, C., Guillory, T.S., and Dilks, S.S. (2021). Obesity and Psychiatric Disorders. Nurs. Clin. *56*, 553–563. https://doi.org/10.1016/j.cnur.2021.07.010.

113. xiao, R., and Boehnke, M. (2009). Quantifying and correcting for the winner's curse in genetic association studies. Genet. Epidemiol. *33*, 453–462. https://doi.org/10.1002/gepi.20398.

114. Staples, J., Qiao, D., Cho, M.H., Silverman, E.K., University of Washington Center for Mendelian Genomics; Nickerson, D.A., and Below, J.E. (2014). PRIMUS: rapid reconstruction of pedigrees from genome-wide estimates of identity by descent. Am. J. Hum. Genet. *95*, 553–564. https://doi.org/10.1016/j.ajhg.2014.10.005.

115. Staples, J., Witherspoon, D.J., Jorde, L.B., Nickerson, D.A., University of Washington Center for Mendelian Genomics; Below, J.E., and Huff, C.D. (2016). PADRE: Pedigree-Aware Distant-Relationship Estimation. Am. J. Hum. Genet. *99*, 154–162. https://doi.org/10.1016/j.ajhg.2016.05.020.

116. Conomos, M.P., Miller, M.B., and Thornton, T.A. (2015). Robust inference of population structure for ancestry prediction and correction of stratification in the presence of relatedness. Genet. Epidemiol. *39*, 276–293. https://doi.org/10.1002/gepi.21896.

117. Slenter, D.N., Kutmon, M., Hanspers, K., Riutta, A., Windsor, J., Nunes, N., Mélius, J., Cirillo, E., Coort, S.L., Digles, D., et al. (2018). WikiPathways: a multifaceted pathway database bridging metabolomics to other omics research. Nucleic Acids Res. *46*, D661–D667. https://doi.org/10.1093/nar/gkx1064.

118. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21. https://doi.org/10.1093/bioinformatics/bts635.

119. Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923–930. https://doi.org/10.1093/bioinformatics/btt656.

120. Ewels, P., Magnusson, M., Lundin, S., and Käller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics *32*, 3047–3048. https://doi.org/10.1093/bioinformatics/btw354.

121. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550. https://doi.org/10.1186/s13059-014-0550-8.

122. Zhang, D. (2017). A Coefficient of Determination for Generalized Linear Models. Am. Statistician *71*, 310–316. https://doi.org/10.1080/00031305.2016.1256839.

123. DeLuca, D.S., Levin, J.Z., Sivachenko, A., Fennell, T., Nazaire, M.D., Williams, C., Reich, M., Winckler, W., and Getz, G. (2012). RNA-SeQC: RNA-seq metrics for quality control and process optimization. Bioinformatics *28*, 1530–1532. https://doi.org/10.1093/bioinformatics/bts196.

124. Ongen, H., Buil, A., Brown, A.A., Dermitzakis, E.T., and Delaneau, O. (2016). Fast and efficient QTL mapper for thousands of molecular phenotypes. Bioinformatics *32*, 1479–1485. https://doi.org/10.1093/bioinformatics/btv722.

125. Privé, F., Aschard, H., Ziyatdinov, A., and Blum, M.G.B. (2018). Efficient analysis of large-scale genome-wide data with two R packages: bigstatsr and bigsnpr. Bioinformatics *34*, 2781–2787. https://doi.org/10.1093/bioinformatics/bty185.

126. Yavorska, O.O., and Burgess, S. (2017). MendelianRandomization: an R package for performing Mendelian randomization analyses using summarized data. Int. J. Epidemiol. *46*, 1734–1739. https://doi.org/10.1093/ije/dyx034.

127. Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. *43*, e47. https://doi.org/10.1093/nar/gkv007.

128. Law, C.W., Chen, Y., Shi, W., and Smyth, G.K. (2014). voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol. *15*, R29. https://doi.org/10.1186/gb-2014-15-2-r29.

129. Smyth, G.K., Michaud, J., and Scott, H.S. (2005). Use of within-array replicate spots for assessing differential expression in microarray experiments. Bioinformatics *21*, 2067–2075. https://doi.org/10.1093/bioinformatics/bti270.

130. Gu, Z., Eils, R., and Schlesner, M. (2016). Complexheatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics *32*, 2847–2849. https://doi.org/10.1093/bioinformatics/btw313.

131. Taliun, D., Harris, D.N., Kessler, M.D., Carlson, J., Szpiech, Z.A., Torres, R., Taliun, S.A.G., Corvelo, A., Gogarten, S.M., Kang, H.M., et al. (2021). Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program. Nature *590*, 290–299. https://doi.org/10.1038/s41586-021-03205-y.

132. Liao, Y., Wang, J., Jaehnig, E.J., Shi, Z., and Zhang, B. (2019). WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. Nucleic Acids Res. *47*, W199–W205. https://doi.org/10.1093/nar/gkz401.

133. Unlu, G., Gamazon, E.R., Qi, x., Levic, D.S., Bastarache, L., Denny, J.C., Roden, D.M., Mayzus, I., Breyer, M., Zhong, x., et al. (2019). GRIK5 Genetically Regulated Expression Associated with Eye and Vascular Phenomes: Discovery through Iteration among Biobanks, Electronic Health Records, and Zebrafish. Am. J. Hum. Genet. *104*, 503–519. https://doi.org/10.1016/j.ajhg.2019.01.017.

134. Fisher-Hoch, S.P., Rentfro, A.R., Salinas, J.J., Pérez, A., Brown, H.S., Reininger, B.M., Restrepo, B.I., Wilson, J.G., Hossain, M.M., Rahbar, M.H., et al. (2010). Socioeconomic status and prevalence of obesity and diabetes in a Mexican American community, Cameron County, Texas, 2004-2007. Prev. Chronic Dis. *7*, A53.

135. Zhou, D., Jiang, Y., Zhong, x., Cox, N.J., Liu, C., and Gamazon, E.R. (2020). A unified framework for joint-tissue transcriptome-wide association and Mendelian randomization analysis. Nat. Genet. *52*, 1239–1246. https://doi.org/10.1038/s41588-020-0706-2.

136. Gogarten, S.M., Sofer, T., Chen, H., Yu, C., Brody, J.A., Thornton, T.A., Rice, K.M., and Conomos, M.P. (2019). Genetic association testing using the GENESIS R/Bioconductor package. Bioinformatics *35*, 5346–5348. https://doi.org/10.1093/bioinformatics/btz567.

137. Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics *34*, i884–i890. https://doi.org/10.1093/bioinformatics/bty560.

138. Frankish, A., Carbonell-Sala, S., Diekhans, M., Jungreis, I., Loveland, J.E., Mudge, J.M., Sisu, C., Wright, J.C., Arnan, C., Barnes, I., et al. (2023). GENCODE: reference annotation for the human and mouse genomes in 2023. Nucleic Acids Res. *51*, D942–D949. https://doi.org/10.1093/nar/gkac1071.

139. Picard Toolkit.(No Title) (2019). https://broadinstitute.github.io/picard/.

140. Stegle, O., Parts, L., Piipari, M., Winn, J., and Durbin, R. (2012). Using probabilistic estimation of expression residuals (PEER) to obtain increased power and interpretability of gene expression analyses. Nat. Protoc. *7*, 500–507. https://doi.org/10.1038/nprot.2011.457.

141. Staples, J., Nickerson, D.A., and Below, J.E. (2013). Utilizing graph theory to select the largest set of unrelated individuals for genetic analysis. Genet. Epidemiol. *37*, 136–141. https://doi.org/10.1002/gepi.21684.

142. Robinson, M.D., and Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biol. *11*, R25. https://doi.org/10.1186/gb-2010-11-3-r25.

143. GTEx Consortium; Laboratory Data Analysis &Coordinating Center LDACC—Analysis Working Group; Statistical Methods groups—Analysis Working Group; Enhancing GTEx eGTEx groups; NIH Common Fund; NIH/NCI; NIH/NHGRI; NIH/NIMH; NIH/NIDA; Biospecimen Collection Source Site—NDRI (2017). Genetic effects on gene expression across human tissues. Nature *550*, 204–213. https://doi.org/10.1038/nature24277.

144. 1000 Genomes Project Consortium; Auton, A., Brooks, L.D., Durbin, R.M., Garrison, E.P., Kang, H.M., Korbel, J.O., Marchini, J.L., McCarthy, S., McVean, G.A., and Abecasis, G.R. (2015). A global reference for human genetic variation. Nature *526*, 68–74. https://doi.org/10.1038/nature15393.

145. Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., et al. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat. Genet. *25*, 25–29. https://doi.org/10.1038/75556.

146. Gene Ontology Consortium (2021). The Gene Ontology resource: enriching a GOld mine. Nucleic Acids Res. *49*, D325–D334. https://doi.org/10.1093/nar/gkaa1113.

147. Gamazon, E.R., Wheeler, H.E., Shah, K.P., Mozaffari, S.V., Aquino-Michaels, K., Carroll, R.J., Eyler, A.E., Denny, J.C., and GTEx Consortium; and Nicolae, D.L. (2015). A gene-based association method for mapping traits using reference transcriptome data. Nat. Genet. *47*, 1091–1098. https://doi.org/10.1038/ng.3367.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Critical commercial assays		
Infinium® Expanded Multi-Ethnic Genotyping Array (MEGA^{Ex})	Illumina	MEGA^{Ex}
TruSeq Stranded mRNA sample kits	Illumina	N/A
KAPA Library Quantification Kits	KAPA Biosystems	N/A
Olink Explore 3072 panel	Olink	N/A
Qiagen RNAseq Lipid Tissue Mini Kit	Qiagen	74804
Deposited data		
CCHC transcriptomics	This paper	dbGaP: phs003894.v1.p1
CCHC MEGA array	This paper	dbGaP: phs003894.v1.p1
CCHC proteomics	This paper	dbGaP: phs003894.v1.p1
MOSS transcriptomics	Miao et al.[43]	N/A
NYC ASAT transcriptomics	This paper	GEO: GSE287627
Software and algorithms		
PRIMUS	Staples et al.[114]	https://primus.gs.washington.edu/primusweb/
PADRE	Staples et al.[115]	https://hufflab.org/software/padre/
GENESIS R package	Conomos et al.[116]	https://bioconductor.org/packages/release/bioc/html/GENESIS.html
DeconCell R package	Aguirre-Gamboa et al.[45]	https://github.com/molgenis/systemsgenetics/tree/master/Decon2
fastp R package	Chen et al.[117]	https://github.com/OpenGene/fastp
STAR R package	Dobin et al.[118]	https://github.com/alexdobin/STAR
featureCounts R package	Liao et al.[119]	https://subread.sourceforge.net/featureCounts.html
MultiQC R package	Ewels et al.[120]	https://github.com/MultiQC/MultiQC
DESeq2 R package	Love et al.[121]	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
Picard Tools v2.13.2	The Broad Institute	https://broadinstitute.github.io/picard/
rsq R package	Zhang et al.[122]	https://cran.r-project.org/web/packages/rsq/index.html
RNA-SeQC	DeLuca et al.[123]	https://github.com/getzlab/rnaseqc
FastQTL	Ongen et al.[124]	https://github.com/francois-a/fastqtl
Bigsnpr R package	Privé et al.[125]	https://privefl.github.io/bigsnpr/
MendelianRandomization R package	Yavorska et al.[126]	https://cran.r-project.org/web/packages/MendelianRandomization/index.html
limma	Ritchie et al.[127]	https://bioconductor.org/packages/release/bioc/html/limma.html
voom	Law et al.[128]	https://rdrr.io/bioc/limma/man/voom.html
duplicateCorrelation	Smyth et al.[129]	https://rdrr.io/bioc/limma/man/dupcor.html
ComplexHeatmap R package	Gu et al.[130]	https://bioconductor.org/packages/release/bioc/html/ComplexHeatmap.html
Other		
TOPMed Imputation server with the TOPMed r2 panel	Taliun et al.[131]	https://imputation.biodatacatalyst.nhlbi.nih.gov/#!

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
NHGRI-EBI GWAS Catalog		https://www.ebi.ac.uk/gwas/
Common Metabolic Diseases Knowledge Portal	Dornbos et al.[48]	https://hugeamp.org/
GEne SeT AnaLysis Toolkit	Liao et al.[132]	https://www.webgestalt.org/
PredixVU	Unlu et al.[133]	https://my.vanderbilt.edu/mbgd/additional-resources/

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

CCHC study subjects

In this study, we employed a case–control study design. The CCHC was established on the Texas-Mexico border in 2004.[134] This randomly ascertained community cohort currently comprises over 5000 people and is approximately 60% female. Individuals are randomly selected from households in Brownsville, Tx, a city with a population that is >90% Mexican American (30% of whom live below the poverty line, 60% lack health insurance). At visits, extensive examinations included blood samples drawn following a confirmed 8-h fast. This study was approved by the Committee for the Protection of Human Subjects of the University of Texas Health Science Center at Houston. The mean age of the CCHC participants at baseline was 45.2 years and 61% were >40 years old. Blood samples were collected for biochemical risk factors, and DNA and RNA extraction. Anthropometric measures, including BMI, were collected at baseline and over the course of each five-year follow-up examination. BMI is defined as weight in kilograms divided by height in meters squared. Participants <18 years of age, pregnant women, individuals with major illnesses, and those on weight altering medications were excluded from analyses. Severe obesity was defined as BMI $\geq$40 kg/m^2 and controls BMI <25 kg/m^2 at the visit when the RNA specimen was collected (Figure 1C). All participants provided informed consent to be included in genomic studies. This study was approved by the Committee for the Protection of Human Subjects of the University of Texas Health Science Center, Houston.

NYC ASAT

RNAseq from ASAT was obtained from ancestrally diverse adult community volunteers in New York City (20 kg/m^2 < BMI <40 kg/m^2). Observations with BMI <20.0 kg/m^2 (to ensure adequate adipose sampling), weight loss or gain of ±3% over past 2 months, used statins or medications that may affect glucose or lipid metabolism including beta blockers, thiazide diuretics, hypolipidemic agents, thyroid hormone, or weight loss medications or formulas and/or have allergy to lidocaine were excluded from the adipose tissue sampling. The ASAT data were collected under IRB protocols, STUDY-16-01150: Adipose tissue depots and metabolism, STUDY-19-00122: Cellular mechanisms regulating depot differences in subcutaneous adipose tissue in women.

Mexican Obesity Study (MOSS) cohort

Individuals patients were recruited into MOSS at the Instituto Nacional de Ciencias Medicas y Nutricion (INCMN), Mexico City, as described in detail previously.[43,44] MOSS is an on-going longitudinal study of severely obese Mexican individuals undergoing bariatric surgery and participating in a 1-year follow-up. We included a total of 45 individuals with subcutaneous adipose RNA-seq data at both time points. The study was approved by the local ethics committee, and all participants provided written informed consent.

BioVU

BioVU is a hospital-based biobank at Vanderbilt University Medical Center. The majority of participants are from Nashville metropolitan area and middle Tennessee. In our analysis, we performed the enrichment test in an established dataset of results (“PredixVU”) from an analyses of the association of genetically regulated predicted gene expression with phecodes (n = 70,439 genotyped European).[133] PredixVU contains gene-phenotype associations derived from a genetics-based transcriptome-wide association analysis of BioVU.[135] BioVU data use was reviewed by the institutional review board at Vanderbilt University Medical Center and determined that the study does not qualify as ‘human subject’ research per §46.102(f)(2) under IRB# 151187.

METHOD DETAILS

CCHC genetic data

All CCHC individuals were genotyped using the MEGA^{Ex}array at VANTAGE at Vanderbilt. We measured batch effects, removed specimens with high levels of missingness, extreme heterozygosity, misspecification of sex, or duplicated specimens and pruned variants with high levels of missingness (call rate <0.95), low minor allele frequency (MAF <0.01), or those that deviate from Hardy-Weinberg

equilibrium (p-value $<10^{-6}$). Genetic relatedness was estimated in PRIMUS and PADRE and individuals related at first and second degree were removed from further analysis.[114,115] Genetic imputation for eQTL mapping was performed on all CCHC participants using the TOPMed Imputation server with the TOPMed r2 panel.[131] Verification of imputation quality by minor allele frequency has been performed and preliminary measures of alternate allele concordance show >95% agreement between directly measured and imputed genotypes across all MAF categories. Principal component analysis (PCA) was applied for capturing the population genetic structure. For a robust and unbiased estimation, we conducted PC-AiR within the R package 'GENESIS', which estimates the weights of PCs within unrelated set and predicted PCs for the related subset.[116,136]

CCHC RNA sequencing

A fasting peripheral blood specimen was collected using PaxGene tubes for each participant at both baseline and at each follow-up exam. Specimens were stored at −80°C within 20 min of the draw. RNA was extracted from all specimens and stored at −80°C. RNA sequencing of CCHC participants was conducted using stored WB with sufficient quantity and quality at VANTAGE at Vanderbilt. Total RNA quality was assessed using the 2100 Bioanalyzer (Agilent). Each sample had at least 200ng of DNase-treated total RNA with RNA integrity number greater than 6 to generate polyA (mRNA) enriched libraries using TruSeq Stranded mRNA sample kits with indexed adaptors (Illumina). Library quality was assessed using the 2100 Bioanalyzer (Agilent) and libraries were quantitated using KAPA Library Quantification Kits (KAPA Biosystems). Pooled libraries were subjected to 150 bp paired-end sequencing according to protocol (Illumina NovaSeq). Sixblood cell types were predicted and scaled using the R package DeconCell, a method that quantifies cell types using expression of marker genes.[45]

We performed fastp for quality control and filtering the unqualified sequencing reads based on the read length and uncertain sequence.[137] The QC-passed reads were aligned to human genome reference (hg38) with STAR,[118] and those aligned reads were assigned and counted for each gene by featureCounts.[119] We then performed MultiQC for quality control of overall sequencing library, and samples with aligned reads <15M, alignment rate <40%, or assigned reads <15M were excluded.[120] DESeq2 was performed for library size normalization and gene-specific dispersion estimation.[121]

CCHC proteomics

Proteomic typing of frozen plasma samples was performed by Olink Proteomics (Waltham, MA) using the Explore 3072 panel (2,921 proteins including low-abundance inflammation proteins, proteins actively secreted into blood circulation, approved and ongoing drug target proteins, organ-specific proteins that have leaked into blood circulation, and proteins representing more exploratory potential biomarkers) in 270 CCHC subjects, with each subject measured at 1 or 2 timepoints (total $n = 573$ samples). The protein abundance level was normalized in the full dataset. Olink data is presented as NPx(Normalized Protein expression) values, where NPxis Olink's relative protein quantification unit on a log2 scale. NPxvalues are derived from the number of matched read counts in sequencing. Data values for measurements below limit of detection (LOD) are reported for all samples.

NYC ASAT collection and RNA sequencing

ASAT was collected at 9:00 a.m. (fasting state) for transcriptomic analysis and banking. A blood sample was additionally collected. ASAT was collected approximately 5 cm lateral to the umbilicus with a 2.5-mm cannula. RNA was extracted from tissues quick-frozen in liquid nitrogen using the Qiagen RNAseq Lipid Tissue Mini Kit (Cat no 74804). QC included RNA quality (bioanalyzer traces and RIN). RNAseq was carried out using highly standardized procedures with strict quality controls (Genewiz.com).

MOSS RNA sequencing

Previously generated subcutaneous adipose RNA-sequencing data from the MOSS cohort were aligned to the GRCh38 genome with GENCODE v42 annotations[138] using STAR v.2.7.10a.[118] To account for novel splice junction sites, we ran STAR[118] in a two-pass mode, where splice junctions identified from the first pass were provided with the –sjdbFileChrStartEnd flag to the second pass as an additional input to further improve the mapping process. We next obtained technical metrics for the reads using the CollectRnaSeqMetrics command from Picard Tools v2.13.2,[139] and counted fragments at the gene name level using featureCounts v2.0.2.[119]

QUANTIFICATION AND STATISTICAL ANALYSIS

CCHC differential expression

PEER factor analysis was used to capture the unobserved confounders of transcriptome profiles.[140] We then implemented a negative binomial model in DESeq2 to identify severe obesity-related genes with covariate adjustment for sex, age, 10 PEER factors and/or comorbidities.[121] Covariates in the regression model included type 2 diabetes, hypertension, and hypocholesterolemia. CCHC used standard protocols and clinical definitions for hypertension (SBP $\geq$140 mmHg or DBP $\geq$90 mmHg or use of anti-hypertensive medication), and type 2 diabetes diagnosis included participants using diabetes-related medications. The status of hypercholesterolemia was queried from the questionnaire, ''have you been diagnosed with high cholesterol?'' Three seated resting systolic blood pressure readings were obtained using random zero sphygmomanometers, retaining the mean of the last two measures. Fasting glucose was measured in a CLIA-approved laboratory. After association testing, an independent filtering of lowly expressed genes using the

default thresholds in DESeq2 was performed and multiple test correction, FDR, was applied. For the 124 SO-DEGs identified in the Discovery data, the same analysis was conducted in a second independent sub-sample of the CCHC.

Percent variance explained

To further quantify the effect of our identified transcripts on severe obesity, we estimated the percentage of explained variance (adjusted R^2), contributed by our identified genes in our replication dataset. To avoid collinearity among the severe obesity genes, we conducted a PCA in our identified gene set, and determined the number of PCs that accounted for over 90% of overall variance. A logistic regression on severe obesity with selected PCs and covariates (sex, age, 10 PEER factors, and comorbidities) was used to determine the adjusted R^2. We compared this estimate with the null model (only included covariates) to obtain the explained variance of the genes of interest using the "rsq" R package.[122] We estimated the percent variance explained for both the 124 SO-DEGs, as well as the 41 genes that replicated.

CCHC eQTL mapping

We performed eQTL mapping using the GTExv8 pipeline.[63] In brief, we selected 645 unrelated CCHC individuals ($\geq$3th degree of relatedness) with both genotyping and RNA-seq data available for Hispanic/Latino-specific eQTL mapping.[141] We aligned the RNA-seq reads to the human genome reference using STAR, and then quantified each gene using RNA-SeQC.[118,123] Read counts were normalized by trimmed mean of M values (TMM), and inverse variance normalization was performed.[142] PEER factor analysis was conducted and 60 PEER factors were used in further analyses.[140,143] FastQTL was performed to identify eQTLs in cis ($\pm$1MB) for each gene with adjustment for sex, RNA-seq batch, 5 genetic PCs and 60 PEER factors. We determined the gene-specific significance threshold using 10,000 permutations.[124]

CCHC proteomics

For proteomics, we implemented logistic regression analysis in a single time point measure from 49 individuals meeting criteria for severe obesity and 42 controls with BMI <25 kg/m^2 CCHC participants with adjustment for sex, age, type 2 diabetes, hypertension, hypercholesterolemia, and 10 genetic principal components in the 23 protein products assayed from the genes in our discovery analysis. We applied FDR correction to determine the statistical significance of our results.

Correlation of ASAT gene expression and BMI

RNAseq from ASAT was obtained from ancestrally diverse adult community volunteers in New York City (20 kg/m^2 < BMI <40 kg/m^2). Due to the limited sample size, we performed a non-parametric approach and assessed Spearman's rank correlation between gene expression and BMI to explore the effect of our discovered genes in ASAT. Only 102 of the 124 SO-DEGs were available for correlation analyses. Statistical significance was assessed using FDR correction for multiple tests.

Mendelian Randomization

We conducted two sample Mendelian randomization (in which exposure and outcome measures are from independent samples) to estimate the causality of gene expression on severe obesity.[57] The eQTLs of the 106 SO-DEGs with validated instruments (see STAR methods section on CCHC eQTL mapping) were used as instrumental variables for Mendelian randomization following the protocols outlined in Zhou et al.,[135] and we used two different sources: eQTLs estimated from CCHC WB and from obesity-relevant tissues in GTEx, including WB, subcutaneous adipose, visceral omentum adipose, liver, skeletal muscle, amygdala, hippocampus, hypothalamus, pituitary, nucleus accumbens in basal ganglia, and substantia nigra. Linkage disequilibrium clumping was used to identify the instrumental variables in each tissue separately, using the LD panels from the 1000 Genome European or Admixed Americans, and the R package, "bigsnpr".[125,144] The severe obesity GWAS data for these analyses were derived from the HL and European summary statistics from a large meta-analysis of severe obesity.[60] For these analyses we used a median weighted MR analysis approach from the R package "MendelianRandomization," which offers an unbiased and reliable estimation even with half of instrumental variables violating the MR assumption or with genetic pleiotropy (e.g., if the variant is not predictive of the exposure, the variant is dependent on confounders, or the variant is not conditionally independent of the outcome given the exposure and confounders).[57,126] To test the robustness of our approach, we performed a secondary analysis using MR-Egger. We then compared MR estimates using a mixed effect model, controlling for the cross-tissue correlation of each gene and weighted by the number of validated instrumental variables. The weighted median and Egger MR estimates (standardized z-statics) were significantly concordant ($r = 0.298$, p-value = 3.5×10^{-11}).

CCHC prospective analyses

To validate the genes estimated to have causal effects on severe obesity by MR in whole blood, we examined the association of their baseline expression with weight gain and incidence of class 2 obesity during follow-up. In an independent set of data from the CCHC, we identified 21 cases without class 2 obesity at the time of RNA collection (BMI <35 kg/m^2) who later developed class 2 obesity (BMI >35 kg/m^2), with a rate of BMI increase >0.5 kg/m^2 per year. We also identified 55 weight-stable controls (rate of BMI change $-$0.25 to 0.25 kg/m^2 per year), with BMI <30 kg/m^2 at baseline who did not develop obesity over at least five years follow-up. Demographic data for the prospective analyses are shown in Table S8A. In these data, we performed differential expression analysis by DESeq2

with obesity incidence as an outcome and including age, sex, BMI at baseline, and 10 PEER factors as covariates. We used Bonferroni adjustment for multiple testing correction.

MOSS cohort longitudinal differential expression

We similarly aimed to strengthen the generalizability of the MR implicated genes in SAT tissue from an on-going longitudinal study of severely obese Mexican individuals undergoing bariatric surgery and participating in a 1-year follow-up as part of MOSS. To identify the genes DE by bariatric surgery status (at time of operation vs. 1 year postoperative), we used limma[127] with the voom[128] normalization method. We first filtered the bulk expression data to retain only expressed genes, with at least 1 count per million mapped reads in at least 10% of the samples. We then estimated the effects of bariatric surgery status on gene expression, where we included the first three genetic PCs, as described previously,[43] and the first PC of gene expression as covariates in the model to account for the admixed population structure and correct for technical factors. To account for repeated measures and thus test for differences per individual patient, we used the duplicateCorrelation method,[129] with patient ID as a blocking factor. We fit the described model for 19 of the 22 genes that were implicated as causal in MR analyses. We corrected for multiple testing using Bonferroni correction.

Pathway enrichment analysis

Pathway enrichment analysis was performed on the 124 SO-DEGs analysis using WEB-based GEne SeT AnaLysis Toolkit.[132] Genes were tested for enrichment in the pathways and gene sets from WikiPathway[117] and the Gene Ontology Resources (GO) databases: Biological Process (GO: BP), Cellular Component (GO:CC) and Molecular Function (GO:MF).[145,146]

PheWES

To estimate the enrichment of phecodes associated with genetically regulated expression of the identified severe obesity genes we imputed tissue-specific genetically regulated expression levels (GReX) for all the genotyped BioVU European participants (n = 70,439) using GTEx data-derived Predixcan models.[147] Next, we estimated disease-GReX associations using logistic regression, phenome-wide.[133] We then extracted the set of all phecodes significantly associated with the tissue-specific GReX of the 124 SO-DEGs after FDR correction and counted how many times each phecode was observed. To establish an empirical null expected distribution of phecode counts, we performed 1,000,000 permutations of random sets of 124 genes. We then derived p-values by comparing the observed count for each phecode in the severe obesity associated gene list to its null distribution with a minimum attainable p-value of 1×10^{-6}.

We applied MR to estimate the causal effect of severe obesity on the identified clinical phenotypes from the PheWES. The GReX measures that we selected as instrumental variables for these analyses were calculated for the set of genes with significant causal effects on severe obesity ($p < 0.05$, see MR above). As tissue-specific GReX may be highly correlated across tissues, for each gene, we selected the tissue with the most significant GReX as the instrument variable. The association between GReX and clinical outcome (identified clinical phenotypes) were obtained from PredixVU, a dataset contains phenome-wide and transcriptome-wide analysis results from BioVU.[133]

Validation from extant body of literature

To assess prior evidence in the literature from GWAS studies, we queried the 124 SO-DEGs for previous association with obesity-related traits in the NHGRI-EBI GWAS Catalog and the Common Metabolic Diseases Knowledge Portal (CMDKP).[48] The GWAS Catalog was searched by mapped gene (recognizing that mapped genes are often based on physical location and thus may not represent the functional element underlying the GWAS signal) and all resulting genome-wide significant trait associations ($p < 5 \times 10^{-8}$) were recorded (Table S3). We categorized phenotypes by broader trait categories (Table S3A) and display results in Figure 4. We retrieved the Human Genetic Evidence (HuGE) scores from the CMDKP across all cataloged traits. HuGE scores are categorized into seven evidence ranges: compelling ($\geq$350), extreme ($\geq$100), very strong ($\geq$30), strong ($\geq$10), moderate ($\geq$3), anecdotal (>1), and no evidence ($\leq$1). We visualized the highest range of evidence for each gene within the 22 CMDKP phenotype groups using the ComplexHeatmap R package.[130] HuGE scores for each gene-trait are in Table S4, and are,CM aggregated by category in Figure 4B.

General:

All parametric tests were two-sided.

ADDITIONAL RESOURCES

Full discovery transcriptomics summary statistics: https://zenodo.org/records/14713355.

GWAS summary statistics for severe obesity: https://doi.org/10.5281/zenodo.14927159.[60]