HHS Public Access

Author manuscript

Int J Obes (Lond). Author manuscript; available in PMC 2019 March 19.

Published in final edited form as:

Int J Obes (Lond). 2019 March ; 43(3): 457–467. doi:10.1038/s41366-018-0190-2.

Integrating Genetic, Transcriptional and Biological Information Provides Insights into Obesity

Lan Wang[#1], **Jeremiah Perez**[#1], **Nancy Heard-Costa**[2], **Audrey Y. Chu**[3,4,#], **Roby Joehanes**[5], **Peter J. Munson**[6], **Daniel Levy**[3,4], **Caroline S. Fox**[3,4,#], **L. Adrienne Cupples**[1,4], and **Ching-Ti Liu**[1,*]

[1]Department of Biostatistics, Boston University School of Public Health, Boston MA 02118

[2]Department of Medicine, Boston University, Boston MA 02118

[3]Population Sciences Branch, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD 20892

[4]The Framingham Heart Study, Framingham, Massachusetts 01702

[5]Hebrew SeniorLife, Harvard Medical School, Boston, MA 02131

[6]Mathematical and Statistical Computing Laboratory, Office of Intramural Research, Center for Information Technology, National Institutes of Health, Bethesda, MD 20892

[#] These authors contributed equally to this work.

Abstract

Objective: Indices of body fat distribution are heritable, but few genetic signals have been reported from genome-wide association studies (GWAS) of computed tomography (CT) imaging measurements of body fat distribution. We aimed to identify genes associated with adiposity traits and the key drivers that are central to adipose regulatory networks.

Subjects: We analyzed gene transcript expression data in blood from participants in the Framingham Heart Study, a large community-based cohort (n up to 4,303), as well as implemented an integrative analysis of these data and existing biological information.

Results: Our association analyses identified unique and common gene expression signatures across several adiposity traits, including body mass index, waist-hip ratio, waist circumference, and CT-measured indices, including volume and quality of visceral and subcutaneous adipose tissues. We identified six enriched KEGG pathways and two co-expression modules for further exploration of adipose regulatory networks. The integrative analysis revealed four gene sets (Apoptosis, p53 signaling pathway, Proteasome, Ubiquitin mediated proteolysis) and two co-expression modules with significant genetic variants and 94 key drivers/genes whose local

[#]This author has moved to Merck & Co., Inc. [*]Author to whom all correspondence should be addressed. Dr. Ching-Ti Liu, ctliu@bu.edu; Postal address: Department of Biostatistics, School of Public Health, Boston University, 801 Massachusetts Ave CT3, Boston, MA 02118, USA.

Conflict of Interest:
The authors declare no competing interests.

networks were enriched with adiposity-associated genes, suggesting that these enriched pathways or modules have genetic effects on adiposity. Most identified key driver genes are involved in essential biological processes such as controlling cell cycle, DNA repair and degradation of regulatory proteins and are cancer related.

Conclusion: Our integrative analysis of genetic, transcriptional and biological information provides a list of compelling candidates for further follow-up functional studies to uncover the biological mechanisms underlying obesity. These candidates highlight the value of examining CT-derived and central adiposity traits.

Introduction

Studies have shown that obesity is associated with increased risk for a variety of cardiometabolic diseases and premature mortality[1]. The prevalence of obesity among adults worldwide has nearly doubled since 1980.[2] The prevalence of obesity in the U.S. is estimated to be 36%, with 69% of U.S. adults being overweight or obese.[3] The obesity epidemic in the U.S. and worldwide contributes to a major public health burden.[4–10] Obesity is a heterogeneous condition with inter-individual variability in fat depots that confer differing metabolic risks[11–13] . Most genetic research in obesity, however, has focused on generalized obesity, measured by body mass index (BMI), and abdominal obesity, measured by waist-hip ratio (WHR) or waist circumference (WC). It is becoming clear that standard metrics of adiposity used in the clinical setting do not adequately reflect pathologic visceral or subcutaneous fat and the corresponding risk of cardiometabolic disease.[14] Subcutaneous adipose tissue (SAT) and visceral adipose tissue (VAT) are considered to be unique pathogenic fat depots that can be imaged using computed tomography (CT). VAT, in particular, has been reported to put individuals at greater risk of cardiometabolic disease than BMI.[14] To date few studies have focused on volume of adipose tissue in individual depots. Furthermore, there are no large-scale genetic studies directly linking genetic variants to visceral and subcutaneous adipose tissue quality (density) measured in Hounsfield units (VATHU and SATHU), even though fat quality may provide insight into cardiometabolic risk independent of fat volume[15] and fat density has been shown to be a unique marker of mortality risk unrelated to inflammation.[16]

Previous studies have shown that indices of body fat distribution are heritable[11,17]. For example, in one study, the heritability for SAT and VAT volumes were estimated to be 57% and 36%, respectively[11,17]. Few genetic signals have been reported from GWAS for directly measured SAT and VAT volume using CT imaging, due in part to the lack of CT imaging measurements in large enough groups of genotyped individuals[18,19]. We hypothesized that gene expression profiling would reveal transcriptomic signatures of the adiposity traits of interest, provide insights into the biology of adiposity, and highlight compelling targets for therapeutic intervention. However, prior transcriptomic studies of CT measured adiposity indices were limited to small groups of selected samples[20,21], which may not be representative of non-morbid populations. Thus, there is a need for further examination in larger samples.

We hypothesized that there is information of the underlying mechanisms of obesity beyond what is explained by BMI. We sought to investigate associations of adiposity traits, including centralized fat depots directly measured by CT-imaging, with gene expression in a large community-based cohort. Because genes in regulatory networks may affect adiposity by acting in concert, instead of acting individually, we also integrated genetic and transcriptomic data and used biological databases to identify pathways and genes that underlie key regulatory mechanisms for adiposity.

Materials and Methods

Study Samples

The Framingham Heart Study (FHS) began in 1948 through enrollment of the Original cohort, with the goal to evaluate the multi-factorial nature of risk factors for coronary heart disease.[22] In 1971 the Offspring cohort (offspring of the Original cohort and the offsprings' spouses) was recruited.[23] In 2002, the Third Generation cohort (grandchildren of the Original cohort and children of the Offspring cohort) was recruited.[24] Our study was limited to participants from the FHS Offspring cohort who attended their eighth examination cycle (Exam 8, 2005 – 2008) and participants from the FHS Third Generation cohort who attended their second examination cycle (Exam 2, 2008–2011) and had blood samples available for RNA collection and measurements for adiposity related traits including BMI, WHR, WC, and CT measures. In total 4,303 study participants were included. The sample characteristics are presented in Supplemental Table 1 (**Table S1**). All participants consented to participate in the study and the study was approved by the Institutional Review Board at Boston University Medical Center.

Adiposity Traits

We primarily focus on six CT-measured adiposity traits: volume of 1) subcutaneous adipose tissue (SAT) and 2) visceral adipose tissue (VAT) and 3) their ratio (VSRAT=VAT/SAT); and quality of 4) subcutaneous adipose tissue (SATHU) and 5) visceral adipose tissue (VATHU) measured in Hounsfield units and 6) their ratio (VSRATHU = VATHU/SATHU). We also report on BMI, WC and WHR measurements.

All CT-measurements of fat were collected with the Aquarius 3D Workstation software (TeraRecon Inc., San Mateo, CA, USA)[25] between 2008 and 2011. Specifically, the participants underwent radiographic assessment with an 8-slice multidetector computed tomography (MDCT) scanning of the abdomen in the supine position[25,26]. Twenty-five contiguous 5-mm slices were obtained. Subcutaneous and visceral adipose tissue volumes were acquired by manually outlining the visceral and subcutaneous fat depots and fat was defined as the image display window of −195 to −45 HU. This method has >0.99 inter-reader and intra-reader correlations for VAT and SAT[25]. More details can be found in Fox et al.[11]

Gene Expression Profiling

Fasting peripheral whole blood samples (2.5ml) were collected from FHS participants during the clinic examinations: Offspring eighth examination (2005–2008) and Third

generation second examination (2008–2011). In total, FHS has 5726 participants with available gene expression data. Total RNA was prepared from frozen PAXgene blood tubes (PreAnalytiX, Hombrechtikon, Switzerland) using the WT-Ovation Pico RNA Amplification System (NuGEN, San Carlos, CA). The obtained cDNA was hybridized to the Human Exon 1.0 ST Array and exon-level intensity values were collected as CEL files using Affymetrix Expression Console Software (Affymetrix, Santa Clara, CA). Gene annotations were obtained from Affymetrix NetAffx Analysis Center (Release 31). We only used the most reliable probe sets derived from RefSeq and GenBank records, including 17,873 distinct transcripts. Exons with signals lower than the background and transcript clusters that were not mapped to RefSeq transcripts were excluded. The CEL file data were quantile-normalized, log2 transformed, and summarized using Robust Multi-array Average[27] from Affymetrix Power Tools version 1.12. Samples with low RNA quality number (<3.0) and principal component outliers were excluded. The resulting expression data were then further adjusted using linear mixed-effects models for technical covariates (first principal component of the expression data, batch effect, the all probe set-mean residuals) and complete blood count (i.e. white blood cells, red blood cells, lymphocytes, neutrophils, platelets, monocytes and eosinophils). Complete blood count was measured in 2,138 Third Generation FHS participants, but not for all samples used in this study. Therefore, blood cell counts of the Offspring cohort and the remaining Third Generation cohort were imputed using a partial least-squares regression method[28,29] based on the gene expression data. Details of technical covariate selection were previously described[30]. We used the adjusted expression data for further analyses detailed below. Details of the design, sampling, RNA isolation, and mRNA measurement were previously described[30,31]. The complete expression dataset is available through dbGaP accession number phs000363 (https://www.ncbi.nlm.nih.gov/gap)

Analysis Strategy

We performed two sets of primary analyses (**Figure 1**). In the first set of analyses, we investigated the associations of adiposity related traits with gene expression (**Part I**). In the second set of analyses, we explored the underlying regulatory mechanisms of our findings obtained from Part I of the analysis (**Part II**). Below we provide a brief overview of the two sets of analyses. More details are provided in the **Supplemental Text**.

There are several studies that have identified significant genes related to BMI.[32,33] In order to further elucidate the underlying mechanisms of obesity beyond what is explained by BMI and to provide novel insights on other adiposity pathways, our primary analyses are adjusted for BMI. As results from both BMI-adjusted and BMI-unadjusted analyses could be biologically relevant (the adjusted analysis may reflect BMI-independent signals and the unadjusted analysis may represent BMI-dependent signals), we also provide results of BMI-unadjusted analyses in the supplemental materials, e.g. Supplemental Tables S2-S10, but focus our discussion on the BMI-adjusted analysis to simplify interpretation.

Our primary analyses focus on sex-combined analyses as a larger sample has greater power to detect important signals. We also provide sex-specific analyses in the Supplement as some

traits have been reported to have sexual dimorphism, but these analyses are exploratory. Hence, unless stated otherwise, all analyses are based on sex-combined data.

Part I of the Analysis: Investigate Association of Adiposity with Gene Expression

The first part of the analysis consisted of three steps (**Figure 1**, Analysis Flow Chart, I-A, I-B, I-C). (I-A) First, we performed linear mixed effects regression to identify genes whose expression levels were associated with adiposity traits with gene expression levels as the dependent variable and an adiposity trait as the independent variable. We adjusted for age, sex, BMI, and cohort of recruitment as fixed effects and familial relationships as a random effect using a kinship coefficient matrix. Our primary analyses were based on the sex-combined data. As an exploratory analysis, we also conducted these analyses with adiposity traits separately for men and women to explore whether there were sex-specific associations between gene expression and adiposity traits. Additionally, we performed a Wald Test to determine if there were sexual dimorphic effects, i.e. the effect of each adiposity trait on gene expression levels (measured by the regression coefficients from the linear mixed effects model) in men was significantly different from the effect in women. (I-B) Next, we performed gene set enrichment analyses (GSEA) using a bioinformatics web-based tool called WebGestalt[34,35] to explore whether the gene expression signatures identified for each of the adiposity-related traits in Part I-A were enriched with KEGG (Kyoto Encyclopedia of Genes and Genomes)[36] pathways. Hypergeometric tests were used to identify enriched KEGG pathways using a Benjamini-Hochberg FDR[37] adjustment to correct for multiple testing. KEGG pathways having a hypergeometric FDR corrected p-value less than 0.05 were considered to be significant gene set enriched pathways. (I-C) We also constructed a co-expression network from the gene expression data using weighted gene co-expression network analysis (WGCNA)[38] in order to identify co-expression network modules (coEMs) consisting of highly correlated genes, i.e. genes that have similar expression. Modules are clustered genes and are assigned arbitrary labels represented by colors by WGCNA. The associations of coEMs to adiposity traits were evaluated by correlating the eigengene (the first principal component representing the expression patterns of all genes in a given module) of each coEM with each adiposity trait of interest via Pearson's correlation; a p-value < 0.05 was considered significant.

Part II of the Analysis: Explore Regulatory Network

In the second part of the analysis, we explored the underlying regulatory mechanisms by integrating multiple levels of data, including pathways and modules identified from Part I, previously published GWAS results and a protein interaction database. (II-D) For significant gene sets identified from part I, we tested for genetically driven associations with adiposity, using SNP set enrichment analysis (SSEA). We first identified cis SNPs (eSNPs) that are significantly associated with expression levels of genes in each gene set. Then we investigated the contributions of genetic variants in the gene sets to adiposity associations by testing whether the set of eSNPs was enriched with low GWAS p-values of corresponding adiposity traits. (II-E) Finally, we applied Key Driver (KD) analysis[39,40] to the candidate genes pooled from gene sets identified from SSEA. The key regulatory gene was identified if its first three-degree neighbors of candidate genes in reference network were enriched. (**Figure 1**, Analysis Flow Chart, II-D, II-E)

Results

Part I: Association of Adiposity Traits with Gene Expression

Study participant adiposity characteristics are presented in Supplemental Table 1 (**Table S1**). We analyzed associations between gene expression and adiposity traits, including SAT, SATHU, VAT, VATHU, VSRAT, VSRATHU, BMI, WC and WHR adjusting for age, sex, cohort, and family relatedness, and BMI (when appropriate). Traits adjusted for BMI have the added subscript 'BMI', such as SAT_{BMI}. The number of genes in each adiposity trait's gene expression signature set (genes with FDR < 0.05) is displayed in **Table 1** for both BMI adjusted and unadjusted analysis. The full lists of significant genes associated with each adiposity trait are in **Tables S2-S15** (unadjusted for BMI **Tables S2A-S10B,** adjusted for BMI **Tables S11A-S15B**).

For all traits, the number of signature genes identified after adjusting for BMI was smaller compared with the number when not adjusting for BMI. After adjusting for BMI, no signature genes were identified for WC_{BMI}, SAT_{BMI} and $SATHU_{BMI}$. For those traits with signature genes, a larger number of signature genes were identified in the sex-combined sample than in the men-only and women-only samples, mainly due to the larger sample size in the sex-combined sample giving more statistical power to detect a significant association (**Figure 2)**. However, for WHR_{BMI}, VAT_{BMI}, $VSRAT_{BMI}$, there were some genes that were identified solely in the men-only sample or women-only sample data (**Figure 2, a-e**). When not adjusting for BMI, all traits had some signature genes identified solely in the men-only sample or women-only sample data (Figure S1, a-i).

We also compared gene expression signature sets of different adiposity measurements. Only a small number of signature genes for CT-measures (15% for VAT_{BMI}, 9% for $VATHU_{BMI}$, 13% for $VSRAT_{BMI}$ and 30% for $VSRATHU_{BMI}$) were also identified with WHR_{BMI} (**Figure 2, f-g**). Among the CT-measures, VAT_{BMI}, $VATHU_{BMI}$ and $VSRAT_{BMI}$ had unique signature genes identified. 55 (49%) of the 113 signature genes for VAT_{BMI} were not signature genes for $VATHU_{BMI}$; 31 (35%) of the 89 signature genes for $VATHU_{BMI}$ were not signature genes for VAT_{BMI}; 98 (88%) of the 111 signature genes for $VSRAT_{BMI}$ were not signature genes for $VSRATHU_{BMI}$; and 7 (35%) of the 20 signature genes for $VSRATHU_{BMI}$ were not signature genes for $VSRAT_{BMI}$ (**Figure 2, h**). Without adjusting for BMI, there were also unique signature genes for CT-measures that did not overlap with BMI (Figure S1, j-l).

The number of genes that passed the Bonferroni corrected p-value threshold for sex effect differences are also displayed in **Table 1**. The gene signature set for $VSRAT_{BMI}$ had more than half of the genes with sexual dimorphic effect sizes (defined as significant differences in sex-specific regression coefficients). Full details on differences in effects by sex are provided in **Tables S2B-S15B**.

Gene set enrichment analysis (GSEA) using the Kyoto Encyclopedia of Genes and Genomes (KEGG)[34–36] database was performed on five BMI-adjusted adiposity traits with gene signature sets identified in the sex-combined data: WHR_{BMI}, VAT_{BMI}, $VSRAT_{BMI}$, $VATHU_{BMI}$, and $VSRATHU_{BMI}$ (**Table 2**). BMI is included in this table for comparison.

The full list of pathways that were significant in at least one trait is displayed in **Table S16**. Six pathways including ABC transporters, Apoptosis, Jak-STAT signaling pathway, p53 signaling pathway, Proteasome, and Ubiquitin mediated proteolysis pathways were identified for subsequent SNP set enrichment analysis (SSEA)[30, 41] analysis due to their statistical significance and potential biological relevance based on the description of the pathway.

We identified adiposity-associated co-expression network modules (coEMs) using weighted gene co-expression network analysis (WGCNA)[38]. Network analysis was based on sex-combined BMI-adjusted traits in order to simplify the interpretation of the results. Modules are denoted by arbitrary colors. We found 24 coEMs (23 coEMs and a grey module composed of all other genes) and investigated their associations with eight adiposity traits: WHR_{BMI}, WC_{BMI}, SAT_{BMI}, VAT_{BMI}, $VSRAT_{BMI}$, $SATHU_{BMI}$, $VATHU_{BMI}$ and $VSRATHU_{BMI}$ (**Figure 3**). The black and royalblue modules correlated with the most traits at $p<0.05$. The black module consisted of 467 genes and was positively correlated with VAT_{BMI}, $VSRAT_{BMI}$, and $VSRATHU_{BMI}$, and negatively correlated with $VATHU_{BMI}$. The royalblue module consisted of 106 genes and was positively correlated with WHR_{BMI}, VAT_{BMI}, $VSRAT_{BMI}$, and $VSRATHU_{BMI}$, and negatively correlated with $VATHU_{BMI}$ and $SATHU_{BMI}$. These two modules were also selected for SSEA. For comparison, we presented these coEMs' associations with BMI in **Figure 3**.

Part II: Explore Regulatory Network

In Part I, we identified significant associations between gene sets (either enriched pathway–trait pairs or significant coEM–trait pairs) and BMI-adjusted adiposity traits. With SSEA we assessed whether the associations were affected by cis genetic variants. Specifically, we performed SSEA on each gene set – adiposity trait pair that was identified as significant in Part I-B or I-C, using available GWAS results. In total, 20 pairs were retained for testing. We identified cis SNPs that are significantly associated with expression levels of genes in each gene set (eSNPs) for 974 genes in eight sets (**Table S17**). The eSNP sets corresponding to six subsets of genes (Apoptosis, p53 signaling pathway, Proteasome, Ubiquitin mediated proteolysis, black coEM and royalblue coEM) showed significantly lower GWAS p-values compared to GWAS p-values of all SNPs for at least one trait. The Bonferroni-corrected P threshold was 0.05/20=0.0025. (**Table 3**) For comparison we also performed SSEA on BMI with all eight gene sets. The results are displayed in **Table 3.** All the gene sets significantly enriched for BMI were also found to be significant in at least one other trait; however, two gene sets, the p53 signaling pathway and the royalblue module, were identified for one or more CT-measure but not for BMI.

We combined gene sets that were significant in SSEA into one single combined set for the identification of key drivers. In total, we had 912 unique genes considered to be potential regulatory gene candidates in the combined set. Using the local networks from Human Protein Reference Database (HPRD)[42], we identified 94 genes whose local networks were significantly enriched with adiposity-associated regulatory gene candidates at Bonferroni-corrected P threshold of 0.05/912 = 5.48E-5, and thus identified these 94 as key drivers (**Table S18**).

Discussion

By integrating genetic, transcriptional, and biological information, we identified several significantly enriched pathways in the set of gene signatures and co-expression modules for a variety of adiposity traits. We further investigated six enriched KEGG pathways (ABC transporters, Apoptosis, Jak-STAT signaling pathway, p53 signaling pathway, Proteasome and Ubiquitin mediated proteolysis pathways) and two co-expression modules. Our results suggest that these enriched pathways or modules have genetic effects on adiposity. Genes in these sets may interact within a network and co-regulate adiposity.

The ABC transporters pathway contains genes that encode proteins from the superfamily of ATP-binding cassette (ABC) transporters, which couple ATP hydrolysis to active transport of a wide variety of substrates across cellular membranes. About half of the 48 human ABC transporters are thought to transport lipids or lipid-related compounds[43]. ABCA1 and ABCG1 are identified as signature genes for adiposity traits in our data. The protein encoded by ABCG1 may be involved in macrophage cholesterol and phospholipids transport and may regulate cellular lipid homeostasis[43]. The protein encoded by ABCA1 acts as a cholesterol efflux pump in the cellular lipid removal pathway[44].

Apoptosis is a process of programmed cell death that is highly regulated. In adipose tissue, it was found that adipocyte apoptosis may be associated with metabolic disorders, including insulin resistance, hepatic steatosis, and obesity associated inflammation[45].

The Janus kinase-signal transducer and activator of transcription pathway (JAK-STAT signaling pathway) was enriched only with the gene expression signature set for WHR_{BMI}. This result suggests that additional adiposity measures beyond BMI are necessary in order to understand the physiological mechanisms underlying obesity. This pathway was found to be highly related to adipose tissue function and to regulate various functions (for example adipocyte development) by transmitting extracellular polypeptide signals (such as leptin in adipose tissue) directly to target gene promoters in the nucleus.[46,47]

The p53 signaling pathway, Proteasome and Ubiquitin mediated proteolysis pathways are essential pathways. The last two form a major pathway of selective protein degradation[48]. Usually short-lived proteins, many of which are regulatory proteins, are marked by multiple ubiquitins and then degraded by the proteasome. Tumor protein p53 that activates in response to multiple stressors binds DNA and activates expression of several genes and hundreds of other down-stream genes and is thus linked to other pathways, for example apoptosis.[49] These pathways are well known to be related to cancer, and potentially to affect adiposity traits due to their universal functions.

By performing SSEA and KD tests on the signature gene sets and coEMs, we aimed to identify genes central to the regulatory networks related to adiposity traits. Most key driver genes that we identified involve essential biological processes such as controlling cell cycle, DNA repair and degradation of regulatory proteins. For example, the key driver CDC26 was found to stabilize a cell cycle regulator APC6[50]. Another key driver DET1 was found to assemble a multi-subunit ubiquitin ligase to promote ubiquitination and degradation of transcription factor c-Jun.[51] Among the 94 identified key drivers, 43 genes were included in

the black module or royalblue module with more than half involving cell cycle regulation or protein degradation. The remaining genes were from one of four pathways (Apoptosis, p53 signaling pathway, Proteasome, Ubiquitin mediated proteolysis), which connect to a broad range of regulation effects. Interestingly, many key driver genes and biological processes are cancer related. For example one KD, CCNG2, which belongs to p53 signaling pathway, has been shown to contribute to signaling networks that limit breast cancer by restricting breast cancer cell proliferation[52] and play important roles as a negative regulator to esophageal cancer cell.[53] The protein levels of CCNG2 are inversely associated with glucose and insulin resistance in adipose tissue.[54] Another KD, EMSY (C11orf30), which belongs to the royalblue module, has been shown to interact with BRCA2[55], a tumor suppressor gene. There is an abundance of evidence from observational studies suggesting that higher amounts of body fat are associated with increased risks of a number of cancers. For example, BMI is reported to be significantly associated with cancers of the colon, rectum, gastric cardia, liver, gallbladder, pancreas, and kidney.[56] The results from our BMI-adjusted analyses indicate that the underlying mechanisms of adiposity and cancer may be closely linked even when controlling for BMI. Therefore, the underlying mechanisms of how adiposity increases the risk of cancer are still unclear; further research into key genes may provide insight.

It was interesting that two pathways that seem highly relevant to adiposity, ABC transporters and the JAK-STAT signaling pathway, did not pass the SSEA test, indicating that the associations between adiposity traits and expression of these genes may not be directly linked to genetic effects. Instead, the expression levels may be affected by environmental factors such as diet or smoking.

There are many strengths of this study. We have a large sample size with mRNA expression profiles and adiposity traits, including CT-measured indices and traditional biomarkers such as BMI. Nevertheless, there are several potential limitations. In particular, there may be concerns regarding tissue specificity due to the lack of expression data measured in adipose tissue. We acknowledge this limitation as blood is usually not considered to be a target organ for obesity although prior studies have reported > 50% sharing of cis-eQTLs in blood and adipose tissue[57,58]. In addition, expression data obtained from obesity relevant tissues with reasonable sample size are lacking and thus, an unbiased and comprehensive scan for gene discovery has not been possible to date. Blood is a sentinel tissue and a system integrator of tissue and organ-level perturbations; so all major metabolic perturbations may lead to adaptive responses in blood[59]. Therefore, utilization of accessible tissues is necessary to push forward the field. Additionally, each fat depot may have unique or common underlying biological mechanisms even though we have pooled different adiposity traits together to identify key driver genes. We have illustrated this by providing the numbers of signature genes in **Figure 2** that are unique to an adiposity trait or overlap across adiposity traits. For instance, VAT_{BMI}, $VSRAT_{BMI}$, $VATHU_{BMI}$, and $VSRATHU_{BMI}$ had 27, 32, 19, and 0 unique signature genes, respectively.

With our large sample size, we integrated genomics data and adiposity traits in genome-wide analyses and provided further insight into the interplay among DNA variation, gene expression and adiposity traits. In summary, we have identified a few sets of genes

associated with adiposity related traits and also identified key drivers/genes that are potentially central to the regulatory networks related to adiposity. Some results observed for CT traits were not seen for BMI, such as the p53 signaling pathway. Thus, these findings provide a list of candidates for further follow-up in experiments to uncover the biological mechanisms underlying obesity beyond BMI.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

This research was in part support by the grant NIH R01 DK089256, NIGMS T32GM074905 and by NHLBI contracts N01-HC-25195 and HHSN268201500001I.

Reference

1. Global BMI Mortality Collaboration, Di Angelantonio E, Bhupathiraju SN, Wormser D, Gao P, Kaptoge S et al. Body-mass index and all-cause mortality: individual-participant-data meta-analysis of 239 prospective studies in four continents. Lancet 2016; 388(10046)**:** 776–86. [PubMed: 27423262] **:**
2. Farzadfar F, Finucane MM, Danaei G, Pelizzari PM, Cowan MJ, Paciorek CJ et al. National, regional, and global trends in serum total cholesterol since 1980: systematic analysis of health examination surveys and epidemiological studies with 321 country-years and 3·0 million participants. Lancet 2011; 377(9765): 578–86. [PubMed: 21295847]
3. Flegal KM, Carroll MD, Kit BK, Ogden CL. Prevalence of obesity and trends in the distribution of body mass index among US adults, 1999–2010. JAMA 2012; 307(5): 491–7. [PubMed: 22253363]
4. Ng M, Fleming T, Robinson M, Thomson B, Graetz N, Margono C et al. Global, regional, and national prevalence of overweight and obesity in children and adults during 1980–2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet 2014; 384(9945): 766–81. [PubMed: 24880830]
5. Ogden CL, Carroll MD, Kit BK, Flegal KM. Prevalence of childhood and adult obesity in the United States, 2011–2012. JAMA 2014; 311(8): 806–14. [PubMed: 24570244]
6. Rosito GA, Massaro JM, Hoffmann U, Ruberg FL, Mahabadi AA, Vasan RS et al. Pericardial fat, visceral abdominal fat, cardiovascular disease risk factors, and vascular calcification in a community-based sample: the Framingham Heart Study. Circulation 2008; 117(5): 605–13. [PubMed: 18212276]
7. Ding J, Hsu FC, Harris TB, Liu Y, Kritchevsky SB, Szklo M et al. The association of pericardial fat with incident coronary heart disease: the Multi-Ethnic Study of Atherosclerosis (MESA). Am J Clin Nutr 2009; 90(3): 499–504. [PubMed: 19571212]
8. Wormser D, Kaptoge S, Di Angelantonio E, Wood AM, Pennells L, Thompson A et al. Separate and combined associations of body-mass index and abdominal adiposity with cardiovascular disease: collaborative analysis of 58 prospective studies. Lancet 2011; 377(9771): 1085–95. [PubMed: 21397319]
9. Pischon T, Boeing H, Hoffmann K, Bergmann M, Schulze MB, Overvad K et al. General and abdominal adiposity and risk of death in Europe. N Engl J Med 2008; 359(20): 2105–20. [PubMed: 19005195]
10. Rexrode KM, Carey VJ, Hennekens CH, Walters EE, Colditz GA, Stampfer MJ et al. Abdominal adiposity and coronary heart disease in women. JAMA 1998; 280(21): 1843–8. [PubMed: 9846779]
11. Fox CS, Massaro JM, Hoffmann U, Pou KM, Maurovich-Horvat P, Liu CY et al. Abdominal visceral and subcutaneous adipose tissue compartments: association with metabolic risk factors in the Framingham Heart Study. Circulation 2007; 116(1): 39–48. [PubMed: 17576866]

12. McLaughlin T, Abbasi F, Lamendola C, Reaven G. Heterogeneity in the prevalence of risk factors for cardiovascular disease and type 2 diabetes mellitus in obese individuals: effect of differences in insulin sensitivity. Arch Intern Med 2007; 167(7): 642–8. [PubMed: 17420421]
13. McLaughlin T, Lamendola C, Liu A, Abbasi F. Preferential fat deposition in subcutaneous versus visceral depots is associated with insulin sensitivity. J Clin Endocrinol Metab 2011; 96(11): E1756–60. [PubMed: 21865361]
14. Shah RV, Murthy VL, Abbasi SA, Blankstein R, Kwong RY, Goldfine AB et al. Visceral adiposity and the risk of metabolic syndrome across body mass index: the MESA Study. JACC Cardiovasc Imaging 2014; 7(12): 1221–35. [PubMed: 25440591]
15. Rosenquist KJ, Pedley A, Massaro JM, Therkelsen KE, Murabito JM, Hoffmann U et al. Visceral and subcutaneous fat quality and cardiometabolic risk. JACC Cardiovasc Imaging 2013; 6(7): 762–71. [PubMed: 23664720]
16. Murphy RA, Register TC, Shively CA, Carr JJ, Ge Y, Heilbrun ME et al. Adipose tissue density, a novel biomarker predicting mortality risk in older adults. J Gerontol A Biol Sci Med Sci 2014; 69(1): 109–17. [PubMed: 23707956]
17. Sellers TA, Drinkard C, Rich SS, Potter JD, Jeffery RW, Hong CP et al. Familial aggregation and heritability of waist-to-hip ratio in adult women: the Iowa Women's Health Study. Int J Obes Relat Metab Disord 1994; 18(9): 607–13. [PubMed: 7812414]
18. Fox CS, Liu Y, White CC, Feitosa M, Smith AV, Heard-Costa N et al. Genome-wide association for abdominal subcutaneous and visceral adipose reveals a novel locus for visceral fat in women. PLoS Genet 2012; 8(5): e1002695. [PubMed: 22589738]
19. Chu AY, Deng X, Fisher VA, Drong A, Zhang Y, Feitosa MF et al. Multiethnic genome-wide meta-analysis of ectopic fat depots identifies loci associated with adipocyte development and differentiation. Nat Genet 2017; 49(1): 125–130. [PubMed: 27918534]
20. Aguilera CM, Gomez-Llorente C, Tofe I, Gil-Campos M, Cañete R, Gil Á. Genome-wide expression in visceral adipose tissue from obese prepubertal children. Int J Mol Sci 2015; 16(4): 7723–37. [PubMed: 25856673]
21. Linder K, Arner P, Flores-Morales A, Tollet-Egnell P, Norstedt G. Differentially expressed genes in visceral or subcutaneous adipose tissue of obese men and women. J Lipid Res 2004; 45(1): 148–54. [PubMed: 14563828]
22. Dawber TR, Meadors GF, Moore J, Felix E. Epidemiological Approaches to Heart Disease: The Framingham Study. American Journal of Public Health and the Nation's Health 1951; 41(3): 279–286.
23. Kannel WB, Feinleib M, McNamara PM, Garrison RJ, Castelli WP. An investigation of coronary heart disease in families. The Framingham offspring study. Am J Epidemiol 1979; 110(3): 281–90. [PubMed: 474565]
24. Splansky GL, Corey D, Yang Q, Atwood LD, Cupples LA, Benjamin EJ et al. The Third Generation Cohort of the National Heart, Lung, and Blood Institute's Framingham Heart Study: design, recruitment, and initial examination. Am J Epidemiol 2007; 165(11): 1328–35. [PubMed: 17372189]
25. Maurovich-Horvat P, Massaro J, Fox CS, Moselewski F, O'Donnell CJ, Hoffmann U. Comparison of anthropometric, area- and volume-based assessment of abdominal subcutaneous and visceral adipose tissue volumes using multi-detector computed tomography. Int J Obes (Lond) 2007; 31(3): 500–6. [PubMed: 16953256]
26. Fox CS, Massaro JM, Schlett CL, Lehman SJ, Meigs JB, O'Donnell CJ et al. Periaortic fat deposition is associated with peripheral arterial disease: the Framingham heart study. Circ Cardiovasc Imaging 2010; 3(5): 515–9. [PubMed: 20639302]
27. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 2003; 4(2): 249–64. [PubMed: 12925520]
28. Joehanes R, Zhang X, Huan T, Yao C, Ying SX, Nguyen QT et al. Integrated genome-wide analysis of expression quantitative trait loci aids interpretation of genomic association studies. Genome Biol 2017; 18(1): 16. [PubMed: 28122634]

29. Zhang X, Joehanes R, Chen BH, Huan T, Ying S, Munson PJ et al. Identification of common genetic variants controlling transcript isoform variation in human whole blood. Nat Genet 2015; 47(4): 345–52. [PubMed: 25685889]
30. Joehanes R, Ying S, Huan T, Johnson AD, Raghavachari N, Wang R et al. Gene expression signatures of coronary heart disease. Arterioscler Thromb Vasc Biol 2013; 33(6): 1418–26. [PubMed: 23539218]
31. Huan T, Zhang B, Wang Z, Joehanes R, Zhu J, Johnson AD et al. A systems biology framework identifies molecular underpinnings of coronary heart disease. Arterioscler Thromb Vasc Biol 2013; 33(6): 1427–34. [PubMed: 23539213]
32. Wang W, Jiang W, Hou L, Duan H, Wu Y, Xu C et al. Weighted gene co-expression network analysis of expression data of monozygotic twins identifies specific modules and hub genes related to BMI. BMC Genomics 2017; 18(1): 872. [PubMed: 29132311]
33. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR et al. Genetic studies of body mass index yield new insights for obesity biology. Nature 2015; 518(7538): 197–206. [PubMed: 25673413]
34. Wang J, Duncan D, Shi Z, Zhang B. WEB-based GEne SeT AnaLysis Toolkit (WebGestalt): update 2013. Nucleic Acids Res 2013; 41(Web Server issue): W77–83. [PubMed: 23703215]
35. Zhang B, Kirov S, Snoddy J. WebGestalt: an integrated system for exploring gene sets in various biological contexts. Nucleic Acids Res 2005; 33(Web Server issue): W741–8. [PubMed: 15980575]
36. Kanehisa M, Goto S, Furumichi M, Tanabe M, Hirakawa M. KEGG for representation and analysis of molecular networks involving diseases and drugs. Nucleic Acids Res 2010; 38(Database issue): D355–60. [PubMed: 19880382]
37. Y B, Y H Controlling the false discovery rate: a practical and powerful approach to multiple testing. In: Journal of the Royal Statistical Society Series B (methodological), 1995 pp 289–300.
38. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 2008; 9: 559. [PubMed: 19114008]
39. Huan T, Meng Q, Saleh MA, Norlander AE, Joehanes R, Zhu J et al. Integrative network analysis reveals molecular mechanisms of blood pressure regulation. Mol Syst Biol 2015; 11(4): 799. [PubMed: 25882670]
40. Mäkinen VP, Civelek M, Meng Q, Zhang B, Zhu J, Levian C et al. Integrative genomics reveals novel molecular pathways and gene networks for coronary artery disease. PLoS Genet 2014; 10(7): e1004502. [PubMed: 25033284]
41. Zhong H, Yang X, Kaplan LM, Molony C, Schadt EE. Integrating pathway analysis and genetics of gene expression for genome-wide association studies. Am J Hum Genet 2010; 86(4): 581–91. [PubMed: 20346437]
42. Keshava Prasad TS, Goel R, Kandasamy K, Keerthikumar S, Kumar S, Mathivanan S et al. Human Protein Reference Database--2009 update. Nucleic Acids Res 2009; 37(Database issue): D767–72. [PubMed: 18988627]
43. Tarling EJ. Expanding roles of ABCG1 and sterol transport. Curr Opin Lipidol 2013; 24(2): 138–46. [PubMed: 23340182]
44. Tarling EJ, de Aguiar Vallim TQ, Edwards PA. Role of ABC transporters in lipid transport and human disease. Trends Endocrinol Metab 2013; 24(7): 342–50. [PubMed: 23415156]
45. Alkhouri N, Gornicka A, Berk MP, Thapaliya S, Dixon LJ, Kashyap S et al. Adipocyte apoptosis, a link between obesity, insulin resistance, and hepatic steatosis. J Biol Chem 2010; 285(5): 3428–38. [PubMed: 19940134]
46. Gurzov EN, Stanley WJ, Pappas EG, Thomas HE, Gough DJ. The JAK/STAT pathway in obesity and diabetes. FEBS J 2016; 283(16): 3002–15. [PubMed: 26972840]
47. Richard AJ, Stephens JM. The role of JAK-STAT signaling in adipose tissue function. Biochim Biophys Acta 2014; 1842(3): 431–9. [PubMed: 23735217]
48. Reinstein E, Ciechanover A. Narrative review: protein degradation and human diseases: the ubiquitin connection. Ann Intern Med 2006; 145(9): 676–84. [PubMed: 17088581]

49. Stracquadanio G, Wang X, Wallace MD, Grawenda AM, Zhang P, Hewitt J et al. The importance of p53 pathway genetics in inherited and somatic cancer genomes. Nat Rev Cancer 2016; 16(4): 251–65. [PubMed: 27009395]
50. Wang J, Dye BT, Rajashankar KR, Kurinov I, Schulman BA. Insights into anaphase promoting complex TPR subdomain assembly from a CDC26-APC6 structure. Nat Struct Mol Biol 2009; 16(9): 987–9. [PubMed: 19668213]
51. Wertz IE, O'Rourke KM, Zhang Z, Dornan D, Arnott D, Deshaies RJ et al. Human De-etiolated-1 regulates c-Jun by assembling a CUL4A ubiquitin ligase. Science 2004; 303(5662): 1371–4. [PubMed: 14739464]
52. Zimmermann M, Arachchige-Don AP, Donaldson MS, Patriarchi T, Horne MC. Cyclin G2 promotes cell cycle arrest in breast cancer cells responding to fulvestrant and metformin and correlates with patient survival. Cell Cycle 2016; 15(23): 3278–3295. [PubMed: 27753529]
53. Chen JQ, Liu CJ, Wen HX, Shi CL, Zhang HS, Li M et al. Changes in the expression of cyclin G2 in esophageal cancer cell and its significance. Tumour Biol 2014; 35(4): 3355–62. [PubMed: 24297335]
54. Garrido-Sánchez L, Roca-Rodríguez MeM, Fernández-Veledo S, Vendrell J, Yubero-Serrano EM, Ocaña-Wilhelmi L et al. CCNG2 and CDK4 is associated with insulin resistance in adipose tissue. Surg Obes Relat Dis 2014; 10(4): 691–6. [PubMed: 24708911]
55. Hughes-Davies L, Huntsman D, Ruas M, Fuks F, Bye J, Chin SF et al. EMSY links the BRCA2 pathway to sporadic breast and ovarian cancer. Cell 2003; 115(5): 523–35. [PubMed: 14651845]
56. Lauby-Secretan B, Scoccianti C, Loomis D, Grosse Y, Bianchini F, Straif K et al. Body Fatness and Cancer--Viewpoint of the IARC Working Group. N Engl J Med 2016; 375(8): 794–8. [PubMed: 27557308]
57. Emilsson V, Thorleifsson G, Zhang B, Leonardson AS, Zink F, Zhu J et al. Genetics of gene expression and its effect on disease. Nature 2008; 452(7186): 423–8. [PubMed: 18344981]
58. Montgomery SB, Dermitzakis ET. From expression QTLs to personalized transcriptomics. Nat Rev Genet 2011; 12(4): 277–82. [PubMed: 21386863]
59. Ghosh S, Dent R, Harper ME, Gorman SA, Stuart JS, McPherson R. Gene expression profiling in whole blood identifies distinct biological pathways associated with obesity. BMC Med Genomics 2010; 3: 56. [PubMed: 21122113]

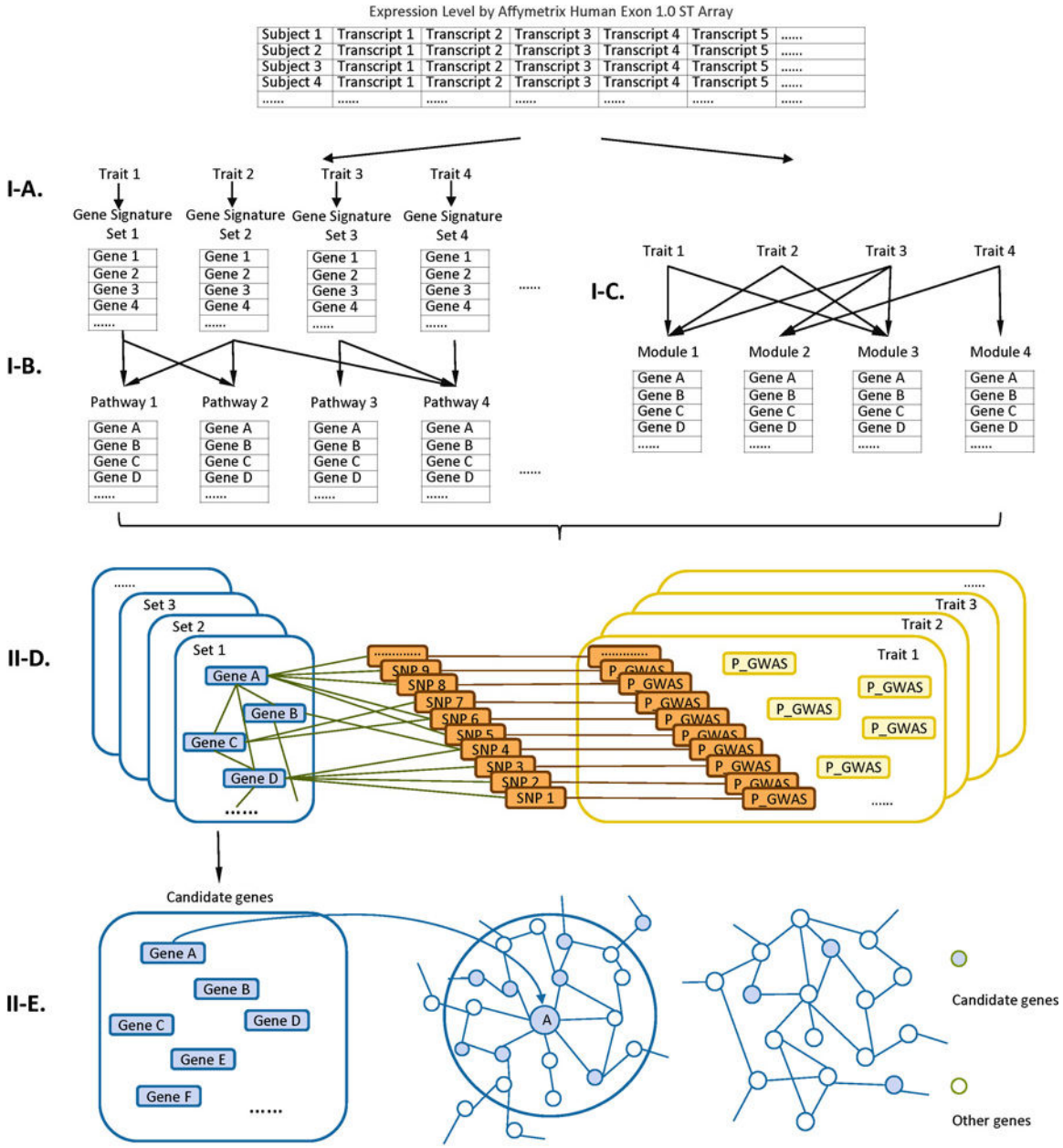

Figure 1. Flow Chart of Analysis Process.

Part I-A. Identified gene signature for each adiposity trait. **I-B. GSEA**. Identified pathways that are enriched in gene signature sets. **I-C.WGCNA**. Identified co-expression modules and tested association with each adiposity trait. **Part II-D. SSEA** on each gene set-trait pair from I-B and I-C. We first identified eSNPs for each gene set and then compared the distribution of GWAS-derived p-values for eSNPs to the distribution of p values for all variants. **II-E. Key driver analysis.** We combined all gene sets that passed II-D into a candidate gene set. For each gene in this set, we compared percentage of candidate genes in their local networks to that in entire network.

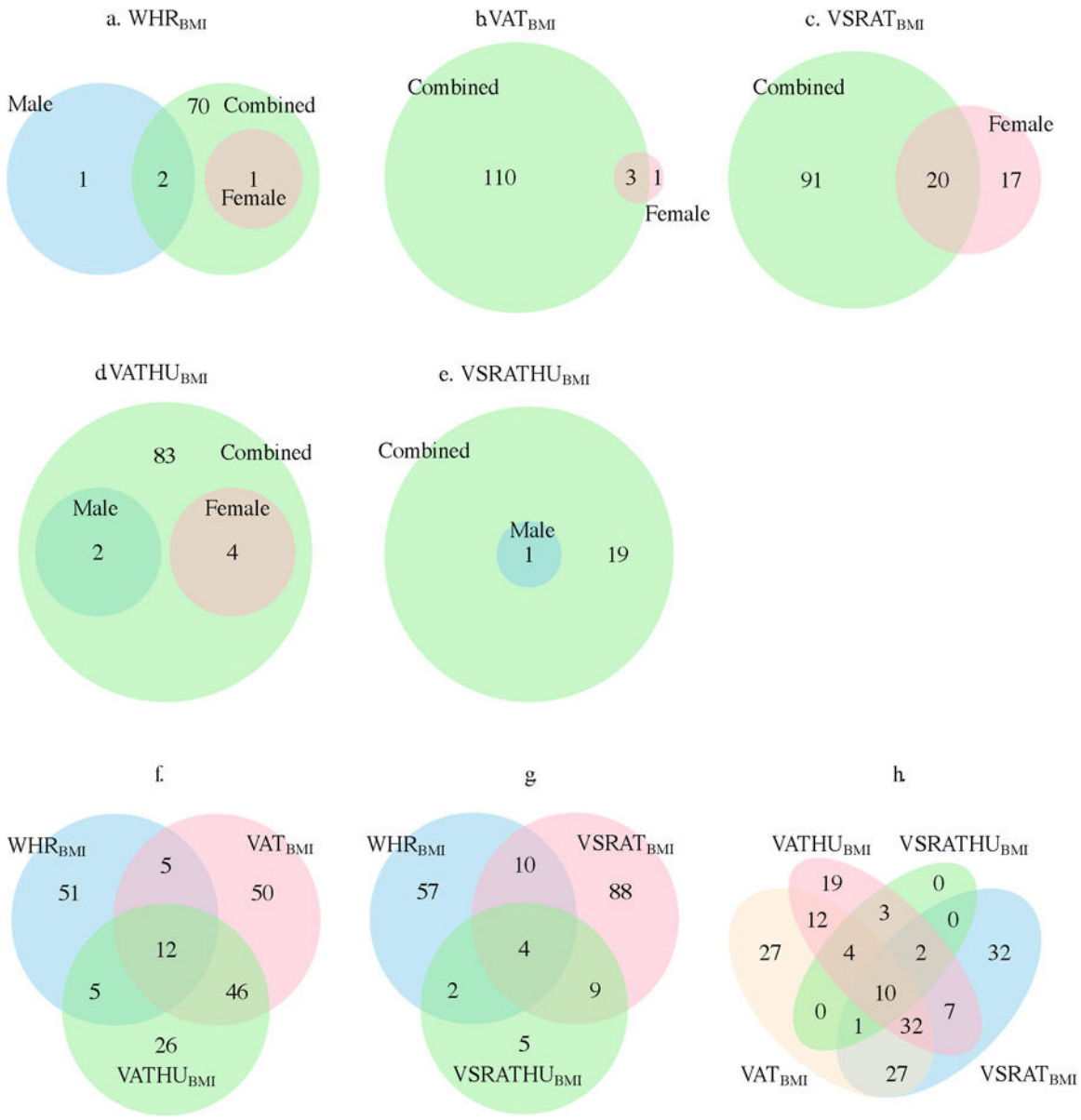

Figure 2. Venn diagram for number of signature genes.

The number of signature genes uniquely identified in men (blue), women (pink) and sex-combined (green) analyses and the number of genes overlapped for each trait.

Figure 3. Heat map of coEM module-trait relationships.

The heat map displays the strength and direction of the correlation (based on Pearson's correlation coefficient) of each coEM (via the eigengene) with each adiposity trait of interest for all coEMs and traits with adjustments of age, sex, BMI, cohort, family relatedness, and technical covariates. The rows correspond to each coEM and the columns correspond to the adiposity traits. The larger the correlation, the darker the color, with red representing a positive correlation and blue representing a negative correlation. The colors labeling each coEM was assigned arbitrarily by the software running the WGCNA and have no specific meaning.

Table 1.

Number[1] of genes in gene signature for each trait.

	Sex Combined	Women	Men
BMI	3239/342	2099/191	1534/299
WHR	2060/211	911/90	890/210
WC	3320/272	2174/203	1312/211
SAT	1244/13	817/20	228/10
VAT	2130/549	1315/593	771/261
VSRAT	367/205	406/327	60/48
SATHU	531/1	268/13	53/0
VATHU	2080/85	1034/91	613/24
VSRATHU	1507/43	630/24	500/37
WHR_{BMI}	73/4	1/0	3/3
WC_{BMI}	0	0	0
SAT_{BMI}	0	0	0
VAT_{BMI}	113/2	4/2	0
$VSRAT_{BMI}$	111/24	37/32	0
$SATHU_{BMI}$	0	0	0
$VATHU_{BMI}$	89/0	4/0	2/0
$VSRATHU_{BMI}$	20/0	0	1/0

[1]The first number in each cell is the number of genes identified as significantly associated with traits based on FDR (criteria described in Methods Section). The second number in each cell is the number of genes with significant sexual dimorphisms using Bonferroni correction (e.g. for BMI in sex-combined sample, p-threshold = 0.05/3239).

Table 2.

FDR p value for significantly enriched pathways in GSEA analysis[1].

Pathway	WHR_{BMI}	VAT_{BMI}	$VSRAT_{BMI}$	$VATHU_{BMI}$	BMI
ABC transporters	0.0428	0.0143		0.0389	0.0249
Apoptosis	0.0184				0.00005
Jak-STAT signaling pathway	0.0137				
p53 signaling pathway	0.0013	0.0242			0.0005
Proteasome			0.038	0.0389	0.0101
Ubiquitin mediated proteolysis		0.0242		0.0389	0.00005

[1]Only significant signals (FDR p<0.05) are displayed. Note that there is no significant signal observed for WC_{BMI} and $VSRATHU_{BMI}$. These enriched pathways are then selected for subsequent SSEA analyses.

Table 3.

Results (p-values) of SNP set enrichment analysis[1,2].

Gene set	WHR_{BMI}	VAT_{BMI}	$VSRAT_{BMI}$	$VATHU_{BMI}$[3]	$SATHU_{BMI}$[3]	BMI
ABC transporters	2.06E-02	6.58E-01		1.24E-01		2.95E-03
Apoptosis	**8.94E-10**					**1.59E-47**
Jak-STAT signaling pathway	3.00E-01					3.22E-01
p53 signaling pathway	**1.57E-04**	**5.25E-09**				9.53E-01
Proteasome			**1.89E-06**	1.31E-01		**4.04E-12**
Ubiquitin mediated proteolysis		**9.54E-17**		3.12E-01		**1.00E-15**
black module	**2.59E-04**	**5.58E-04**	**2.50E-14**	**1.98E-05**		**3.44E-08**
royalblue module	9.23E-01	1.65E-02	1.44E-01	**1.34E-03**	7.98E-01	9.18E-01

[1] The p-values from the Kolmogorov–Smirnov test are shown for 20 gene set-trait pairs which were significant in GSEA and WGCNA.

[2] Bolded p- values are statistically significant at $p < 0.05/20 = 0.0025$, the Bonferroni adjustment for the number of tests, excluding BMI results which are provided for comparison.

[3] GWAS results of VATHU, SATHU were used for $VATHU_{BMI}$, $SATHU_{BMI}$ respectively, since no GWAS results were available for $VATHU_{BMI}$ and $SATHU_{BMI}$