THE UNIVERSITY OF CHICAGO

REPEAT ELEMENTS AND NOVEL GENES CONTRIBUTE TO SEX-SPECIFIC MECHANISMS OF METABOLISM

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF BIOLOGICAL SCIENCES
AND THE PRITZKER SCHOOL OF MEDICINE
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

INTERDISCIPLINARY SCIENTIST TRAINING PROGRAM:
GENETICS, GENOMICS, AND SYSTEMS BIOLOGY

BY
GRACE HANSEN

CHICAGO, ILLINOIS
AUGUST 2021

"Information is physical."

-Rolf Landauer

"With the tools and the knowledge, I could turn a developing snail's egg into an elephant. It is not so much a matter of chemicals because snails and elephants do not differ that much; it is a matter of timing the action of genes."

-Barbara McClintock

Table of Contents

List of Figures

List of Tables

ACKNOWLEDGMENTS

The first thanks go to my advisor, Marcelo Nobrega, for shepherding me through the unknowns and uncertainties of getting a PhD. Thanks also to my committee, for their faith that my research would develop into something interesting. Particular thanks go to Yang Li, who I relied on for help with coding when nobody else could help, and to Marcus Clark, for reassuring me during the inevitable existential crises of graduate school. Thanks to all the members of my lab, for offering constructive advice that developed my thinking as a scientist. A heartfelt thanks goes out to Amelia Joslin, who as a senior graduate student in my lab offered support before I knew I needed it. Thank you for telling me that I "was doing a good job".

Thank you to the kind, empathetic administrators of the graduate program and the medical scientist training program (MSTP). Thanks to Susan Levison, who first welcomed me to the genetics community. Thanks to Kristin McCann, for making the MSTP a community.

I especially want to thank my earliest supporters. Thank you to my mom, who has boundless confidence in me. Enough of it rubbed off to make me believe I can do anything I want with my life. Thanks to my dad, who when I was a child answered my questions and took my curiosity seriously. I am a scientist because of you. Thanks to my sisters, especially Claire, who was always available to listen to me fret. I hope to be as kind to others as she is to me.

Thanks especially to my fellow graduate students, in particular Kate Farris, Selene Clay, Sammy Keyport, Katie Aracena, and Meike Lobb-Rabe, who understood what it was to go through graduate school and whose commitment to supporting each other kept me whole and cheerful in the first years of my PhD. Thank you to my friends Annabeth and Miriam, who are *not* graduate students in biology, but nevertheless gave sympathy and support in the later years of my PhD, when the pandemic separated me from all other friends.

Some credit for this PhD is due to my cat, Daddy. Much of this dissertation was written

with him on my lap.

The greatest thanks and much love goes to my partner, AJ Miller (AJ Miller-Hansen for publication purposes). Thank you for feeding me. Thank you for loving me for reasons that have little to do with my abilities as a researcher, and for helping me to live a happy, rich life.

ABSTRACT

Obesity-associated mortality is exacerbated by high abdominal obesity. Here, we identify genes associated with obesity and waist-to-hip ratio adjusted for body mass index (WHRadjBMI), and allele-sensitive enhancers which are predicted to regulate female WHRadjBMI genes. 76% (45/59) of obesity and waist-to-hip ratio-associated variants are located in primate-specific *Alu* elements which contain a motif associated with adipocyte differentiation. *SNX10*, the strongest female WHRadjBMI-associated gene, is required for adipocyte differentiation and is required for adipose expansion in female mice, but not males, and the lead variant at the *SNX10* locus affects intracellular lipids in human adipocytes. In conclusion, we find novel genes and a system of gene regulatory elements which regulate adipocyte differentiation and may affect human health.

CHAPTER 1

INTRODUCTION

1.1 Creating cellular diversity from a single genome

Many billions of cells comprise the human body. Cells must perform an astonishing variety of functions in order to create and maintain a complex human organism. The codex from which all these cells receive instruction is the genome. While human cells come in many types and accomplish a variety of functions, the genome of each cell is identical (or rather, should be identical; exceptions are cancer and deleterious somatic mutations).

The genome contains approximately 20,000 genes that code for proteins. Genes are not all expressed at once: rather, the are expressed in certain times, places, and conditions. A gene may produce a protein that is unique to certain cell types, or specific to certain actions that only certain cells perform. In this way, cells can generate diversity sufficient to create the human body from a fixed set of genes. The process of selectively expressing genes appropriate to the right cell type and context is accomplished through gene regulation. Increasingly, genetics is the study of gene regulation, as gene regulation is necessary to create a complex organism from a single genome.

Because gene regulation is fundamental to the functioning of cells, most human diseases involve, or indeed are caused by derangements of gene regulation. Therefore, in the process of understanding how genes are regulated, we can also learn the fundamental mechanisms that lie behind many human diseases. Conversely, by studying the changes that occur in gene regulation that occur in a single human disease, we can learn principles of gene regulation that apply to the whole genome. This, in part, has been the goal of this project: to use the gene misregulation that occurs in obesity and fat to understand how a gene network is constructed to control a human trait.

Whereas the sequence of genes, the 'coding' part of the genome, determines the structure

of proteins,gene regulation is largely controlled by the sequence of the non-coding part of the genome. Gene regulation is accomplished in a number of different ways. Below, I describe three major ways in which non-coding sequences are used to control gene expression, and thereby change the structure and function of cells.

1.1.1 Gene regulation by transcription factors

Perhaps the primary means of regulating gene expression is through transcription factors. Transcription factors are, proteins that recognize specific sequences of DNA, and preferentially bind at these locations. 10% of all genes are transcription factors, by this definition[93]. When tanscription factors recognized and bind to their cognate DNA sequences, called motifs, they recruit other factors and promote the formation of protein assemblies that can perform a variety of functions. In the case of the simplest transcription factor, RNA polymerase, the assembled protein complex transcribes the gene at that genomic locus. More complex transcription factors can encourage gene expression at more distal loci, by recruiting factors that promote the unwinding of DNA from nucleosomes, such as histone acetyltransferase, by recruiting transcription preinitiation proteins such as Mediator[4, 99] which will encourage the transcription of neighboring genes. Alternatively, transcription factors can suppress gene expression, by recruiting factors that induce heterochromatin formation[39]. In these ways, transcription factors can tune the degree of gene expression to a level appropriate to the function of a specific cell type. By controlling which transcription factors are present, a cell can determine which genes will be expressed, and to what degree. Through transcription factors, cells determine their identity, and perform many of their functions.

The motifs which transcription factors recognize are often located in the promoters of genes. This is sensible, as it allows transcription factors to recruit activating or repressive complexes directly to the promoter of a gene. However, gene promoters reside in a nucleosome-depleted region, rendering them constitutively accessible[42]. Transcription fac-

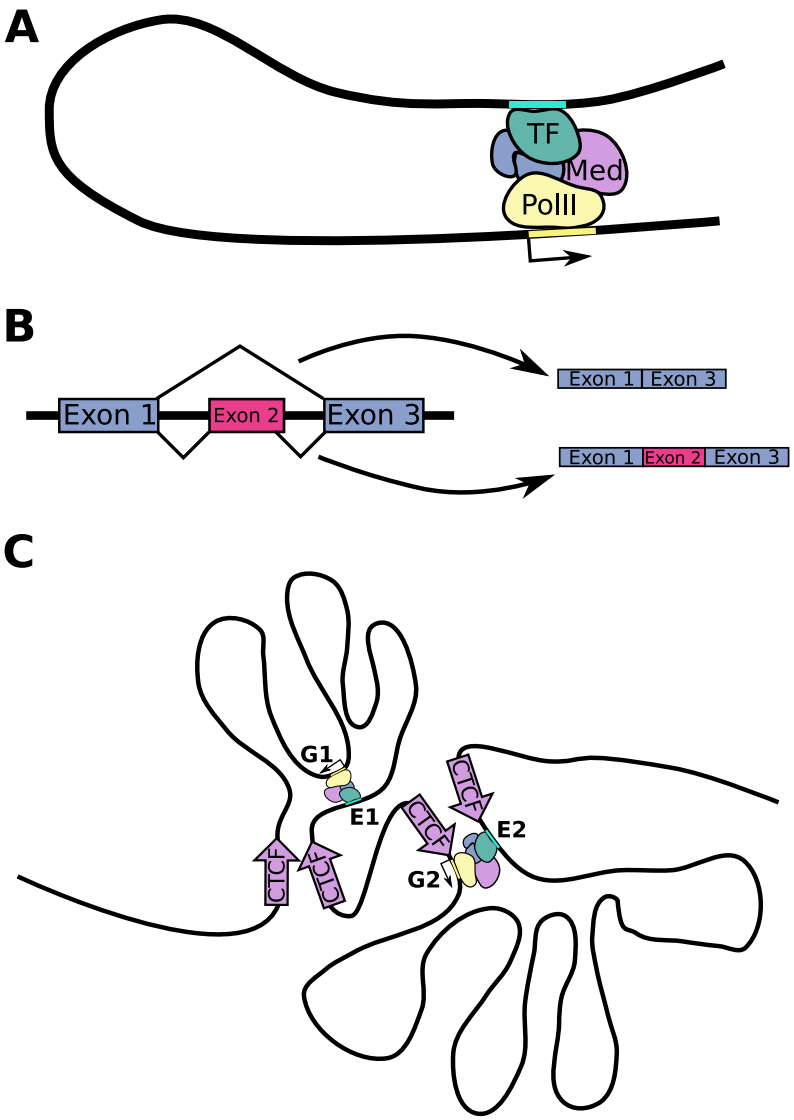

Figure 1.1: **Three major mechanisms of gene regulation.** a) Regulation of gene expression by a transcription factor. In teal is shown a transcription factor sequence, bound to a transcription factor (TF). The transcription factor recruits the transcription initiation machinery, including cofactors such as Mediator (Med), which recruits polymerase II (PolII) to the gene promoter and initiates transcription. b) Alternative gene splicing. On the left is shown the primary transcript, which contains 3 exons. Alternative splicing of exon 2 results in a longer final mRNA transcript, seen on the right. c) Expression of gene regulation via genome architecture. Shown are two topologically associating domains (TADs), each with one enhancer (E1, E2) and one gene (G1,G2). Although E1 is much closer to G2 in two-dimensional sequence space than G1, E1 regulates G1 expression and cannot regulate G2 due to genome structure.

tors binding at a gene promoter, therefore, cannot easily achieve tissue-specific expression. Therefore, many transcription factor motifs are located at non-coding sites many kilobases distant from the genes upon which they act. These distal transcription factor binding sites are often grouped together in sequences called enhancers[99], and confer exquisite restriction of gene expression to the appropriate cell type, function, or moment in time. To work appropriately, enhancers must locate the gene promoter in nuclear space. This is controlled in a number of ways. One major way is through genome architecture, which is discussed in the next section.

The "master regulator of adipogenesis", peroxisome proliferator activated receptor gamma (PPARG), is an example of many of these features of transcription factors. PPARG is sufficient to induce adipocyte(fat cell) differentiation, and therefore its behavior needs to be carefully controlled, lest the wrong cells suddenly become fat cells. This tissue specificity is achieved in a number of ways. PPARG expression is signalled by upstream transcription factors active in developing adipocytes[78]. As PPARG is expressed, it induced the expression of its partner transcription factor, CCAAT/enhancer-binding protein α (CEBPα)[51]. These two transcription factors amplify one another's expression, and co-bind to the same sites[8]. This partnership is specific to adipocytes, ensuring that the correct PPARG targets are found.

PPARG targets thousands of motifs in the genome, and activate hundreds of genes. It initiates the initiation of other transcription factors that in turn initiate the transcription of more factors, eventually producing the gene expression profile necessary for mature adipocytes to function. With relatively few but precise starting signals, developing adipocytes can achieve a specific and complex gene regulation that is unique to fat.

When a new transcription factor binding site emerges in the neighborhood of a gene, that gene's expression is now increased in the presence of that transcription factor. In this way, novel cellular functions can evolve without direct changes to the sequence of genes. Recent

literature has highlighted the important of repeat elements for the evolution of gene regulatory networks. Repeat elements are the genomic remnants of retroviruses and other elements which have replicated and inserted themselves into the genome throughout evolutionary history. Some families of repeat elements still have the capacity to insert themselves into novel locations; most, however, are inert in the genome. Because of their ubiquity and ability to amplify within a genome across evolutionary time, repeat elements have been co-opted in multiple evolutionary lineages to rapidly form gene regulatory networks regulating interferon response, mammalian pregnancy, and other traits subject to evolutionary pressures [22, 59, 49]. *Alu* elements, a class of repeat elements unique to primates, contain sequences related to a number of human binding sites, and as such are an attractive candidate for the origin of traits that have evolved along the primate and human lineages[69].

1.1.2 Alternative splicing of gene transcripts

A gene that is expressed is transcribed from genomic DNA into a primary RNA transcript, mRNA. This mRNA is not simply a chain of instructions for an amino acid sequence. Rather, some sections will be translated into amino acids, while others will be excised out, or 'spliced'. The average gene has approximately 8 exons, or segments to be transcribed, and 8 introns, to be excised[73]. Thus, most genes are composed of distinct modules, which are composed after transcription into a sequence of codons that will be translated into a protein.

Cells can use this modularity to produce complexity from a fixed set of genes through alternative splicing. In alternatively spliced gene transcripts, some exons are included that are not present in the primary gene transcript, or excluded when they are typically present. In this way, the cell can produce proteins that are sub-specialized to a given cell type, developmental stage, or function by adding or excluding domains from a protein template. As an example, in neuronal progenitor cells the protein *Nin* is localized to the centrosomes. As neurons mature, this protein becomes diffuse throughout the cytosol. Recently, Zhang

and colleagues found that exon 18 is required for *Nin* to localize to the centrioles[105]. In mature neurons, exon 18 is not included, but another exon, exon 29 is spliced in. In neurons, and *Nin* is expressed diffusely throughout the cytoplasm, where it facilitates the binding of cargo to dyenin, enabling cellular transports. Cell fate is controlled by which exons are translated from *Nin*: if exon 18 is included and exon 29 spliced out, the cell remains a progenitor. If exon 18 is removed and exon 29 is included, the cell becomes a neuron.

Which exons are included in a gene transcript is often determined by the sequence of the non-coding introns that flank each exon. These sequences can contain binding sites for RNA-binding proteins, which recruit the mRNA splicing apparatus to an exon and faciliate or prevent its inclusion. For example, in spinal motor neurons *NOVA-1* binds to a specific intronic sequence found in a number of genes, including $GABA_A$ and *GlyRα2* [48]. In these genes, *NOVA-1* facilitates the inclusion of the exon immediately downstream of its motif. If this motif is disrupted, this exon is not included. For spinal motor neurons, the correct transcript of *GlyRα2* is essential. If splicing is disrupted in this gene, these cells die, and the mice soon die themselves of motor defects.

In these ways, alternative splicing of gene transcript can change the structure and function of proteins. By controlling where and how exons are included in the processed mRNA transcript, non-coding splice site regulators refine the function of proteins in ways that are essential for specialized cellular functions and complex life.

1.1.3 The control of gene expression by genome architecture

As mentioned in the above section, enhancers consist of groups of transcription factor transcription sites at a distance from the promoters which they regulate. This distance confers the ability of enhancers to create cell type- and condition-specific gene expression, as promoters are generally always in a euchromatic state, while distal non-coding regions can be tuned between accessible and repressed states, according to the needs of a cell.

Enhancers can be as close as 5-10 kilobases from their target promoter[98]. However, in larger genomes such as the human genome, enhancers are often located far from their target promoter. For example, an enhancer located approximately 500 kilobases upstream of *SOX9* regulates its expression in the gonads in a way that is critical for sexual development [33]. Despite this distance enhancers reliably contact the correct gene promoters, achieving sophisticated control of complex patterns of gene expression.

One way that large genomes direct enhancers to their correct gene targets is by taking advantage of genome architecture. The human genome can be conceived of as 46 strings coiled into a very small space. Research has shown that this folding does not occur randomly, but is organized in a heirarchical fashion. Genomic areas which are adjacent in two-dimensional space are also clustered together in three-dimensional space[52]. Different regulatory mechanisms maintain this architecture at different scales. At a megabase scale, these territories are called topologically-associating domains (TADs), and are bounded by redundant, convergent binding sites for CCCTC-binding factor (CTCF) [25]. Within these TADs, and especially at their boundaries, are genes. The promoters of these genes preferentially make contact with enhancers within the TAD, and not without it. In this way, genes are segregated with the enhancers appropriate to their function, and protected from ectopic expression by other enhancers.

A demonstration of the control of gene expression by TAD insulation was reported by Franke et al.[33]. At the *Sox9* locus in mice, as mentioned earlier, an upstream enhancer controls *Sox9* expression in the gonads in development. Another enhancer, even further upstream, directs *Sox9* expression in developing limb buds. This gene and its enhancers are all located within the same TAD, thus ensuring appropriate expression. A gene located in the adjacent TAD, *Kcnj2*, is not expressed in limb buds. However, when the region encompassing the boundary and upstream Sox9 enhancers is duplicated, a new TAD is formed which contains *Kcnj2* and enhancers that ought to only have access to *Sox9*. This

induces *Kcnj2* expression in a *Sox9* pattern in developing limb buds. As a result of this gene misexpression, the mutated mice are absent claws and nails. This is a phenotype identical to the analogous mutation in humans, which induces an absence of fingernails and terminal digits called Cooks syndrome.

Genome architecture, while not the only determinant of gene expression, imposes a spatial logic on gene regulation. Genes are more likely to be controlled by enhancers in their neighborhoods, in terms of two-dimensional and also three-dimensional space. Researchers can take advantage of genome architecture, therefore, to identify the target genes of enhancers of unknown function. For instance, Hi-C, a sequencing strategy which identifies sequences in physical proximity, identifies TADs genome-wide. Using this, researchers can identify which genes an enhancer may regulate based on TAD structure. An evolution of HiC, called capture Hi-C, allows researchers to detect individual physical interactions between 'bait' sequences and other sections of the genome[64]. By applying this approach to gene promoters, our lab has been able to identify enhancers which directly contact genes of interest in three-dimensional nuclear space.

1.2 The effects of non-coding genetic mutations

Genome-wide association studies (GWAS) are our primary tool for identifying genetic risk for complex polygenic traits, such as obesity and waist-to-hip ratio. However, the utility of GWAS is limited by the difficulty of 1) identificating true causal variants from GWAS data, and 2) identificating how variants act to cause disease.

GWAS identifies common variants present in the population that are associated with disease. At a significant GWAS locus, not one but many variants will be statistically associated with the trait of interest[75]. This phenomenon is due to genetic linkage. Due to random rearrangement of segments of the chromosomes during meiosis, the likelihood that two variants will be co-inherited decays as a function of linear distance. Nearby variants are

thus often in close linkage: variants within 10kb of each other will have r^2 scores greater than 0.95, meaning that the same alleles are co-inherited >95% of the time. Identifying which of these variants disrupts gene regulation and leads to disease represents a significant challenge. Statistical "fine-mapping" approaches have been developed that attempt to identify the true causal variant based on the pattern of linkage and trait association[37, 96]. Experimental approaches, such as the massively parallel reporter assay used in this project, identify which of the variants disrupt active enhancers in the cell types and contexts of interest[62].

Once the causal variants have been identified, the next challenge is to identify which genes are disrupted in the context of disease. Non-coding variation may disrupt gene function via any of the processes described in the previous section. A small polymorphism, such as a single base-pair substitution, can disrupt the sequence-specific binding of a transcription factor to its binding site on DNA, or may disrupt a splice acceptor site and induce the skipping of an important exon during mRNA splicing. Larger events, such as deletions, may remove enhancers entirely, or may remove enough CTCF sites to induce the fusion of two neighboring TADs and alter the regulatory structure of an entire region of the genome. Expression- and splicing- quantitative trait loci (eQTL, sQTL) studies associate variants to changes in the expression levels and splice isoforms of specific genes, enabling the identification of gene targets associated with traits of interest. Once a variant is known to affect gene expression or splicing, investigation of the sequence surrounding it can often produce insight into how it affects gene regulation. Experimental approaches can knock down or knock out the expression of a gene of interest, and observe the effects of altered gene regulation on cellular and organismal function. In this project, I first identify obesity and WHR-associated genes from GWAS using an sQTL/eQTL-based approach. We then experimentally knock down an identified gene in fat tissue, identifying cellular phenotypes that depend on its expression. In this way, non-coding variants identified through GWAS can be associated to target genes, and their effects can be investigated.

1.3 How excess fat causes disease

The ultimate cause of obesity is caloric excess. When the body is presented with more calories than it requires, it will attempt to store this energy in the form of fat. In order to do so, new adipocytes will differentiate from preadipocytes (hyperplasia), existing adipocytes will expand (hypertrophy). At some point, adipocytes can no longer expand.

The health consequences of obesity are a result of the limits of adipose tissue expansion. Overly hypertrophic adipocytes become insulin resistant, and will no longer accept glucose [5]. As a result, glucose levels in the bloodstream increase. High glucose levels in the bloodstream cause microvascular damage, inducing damage that can cause sensory loss, coronary artery damage, and strokes in type 2 diabetes[31]. In addition to inducing high blood glucose, hypertrophic adipocytes, especially those in the visceral compartment, release free fatty acids[5]. Circulating free fatty acids, denied the appropriate storage location within adipose tissue, form ectopic fat deposits in the vasculature and in other organs. Free fatty acids within the vasculature increase the rate of myocardial infarctions and strokes; when they deposit as ectopic fat on organs they cause other diseases, such as non-alcoholic fatty liver disease[36, 11].

Genome-wide association studies have been conducted to identify genetic variation associated with a propensity to become obese. A central insight that has emerged from these GWAS is that genetic variants associated with obesity affect gene regulation in the brain[54, 30]. By affecting appetite and basal metabolic rate, these variants upset the balance between caloric intake and caloric requirement. An example is *ADCY3*, the gene that encodes adenylate cyclase. Adenylate cyclase is an important secondary messenger in many cell types. In cells in the developing brain, it localizes to primary cilia. In cilia in a subtype of neurons in the hypothalamus, it colocalizes with *MC4R*, another gene whose dysfunction leads to obesity. Disruption of either of these genes in this neuronal subtype leads to severe obesity[14].

A body mass index (BMI, measured as weight over height squared, kg/m^2) greater than 30, is the quantitative definition of obesity. However, BMI provides only limited information. It does not capture the distribution of fat between visceral and subcutaneous fat stores, or the balance between adipose tissue and ectopic fat storage. These properties vary widely between individuals, and people of the same BMI can have large differences with regard to where and how they store excess fat.

Different fat depots have different metabolic properties. Therefore, an individual's pattern of fat storage–how they become obese–largely determines their health outcomes if they gain excess weight. Adipose tissue is divided into the subcutaneous depot, the layer of fat immediately beneath our skin, and the visceral depot, which in our abdomen. Subcutaneous adipose tissue has a relatively large capacity for expansion in response to increased food intake[35, 2, 46]. Compared to visceral adipose tissue, subcutaneous adipose tissue is inert: it can accept new fat storage without large changes in cellular function, hormonal signalling, or inflammatory profile[46, 27, 85]. In comparison, visceral adipose tissue has a more limited ability to expand. In the context of obesity, overwhelmed, hypertrophic visceral adipocytes begin to lose their responsiveness to insulin, more inflammatory cytokines, and return free fatty acids to the circulation. [85, 92]. The release of free fatty acids from visceral tissue may mediate liver insulin resistance, as visceral adipose tissue drains directly into the liver via the portal vein [5]. In these ways, visceral adipose tissue, rather than subcutaneous adipose tissue, may mediate obesity-associated disease.

An individual's balance of subcutaneous versus visceral fat storage can be approximated by waist-to-hip ratio, adjusted for BMI (WHRadjBMI). High WHRadjBMI indicates that an individual is prone to storing fat in the visceral fat depot. High WHRadjBMI predicts the development of obesity-associated disease independently of obesity itself. For example, a mendelian randomization study of 228,466 women and 195,041 men from the UK Biobank showed that WHRadjBMI was causally implicated in coronary artery disease and

diabetes[18]. Another study, a meta-analysis of∼200,000 participants, causally implicated WHRadjBMI in coronoary heart disease and stroke[23].

As with obesity, WHRadjBMI is a partially heritable trait, where many genetic variants, mostly located in non-coding regions of the genome, affect WHRadjBMI risk in the general population[75]. However, where variants associated with obesity largely act in the brain, variants associated with WHRadjBMI largely act in adipose tissues[30]. For example, *DNAH10*, *L3MBTL3*, and *CCDC92*, three genes whose loci are associated with WHRadjBMI, have been shown to regulate the differentiation of adipocytes[57]. By doing this, they change the balance between proliferation of fat cells (hyperplasia) and the expansion of pre-existing fat cells (hypertrophy). Risk alleles at these loci increase the risk of familial partial lipodystrophy, a genetic disorder characterized by an inability to maintain fat tissue, resulting in type 2 diabetes, high cholesterol, and high rates or heart disease.

Given the divergent genetic regulation of obesity and WHRadjBMI, there are two main avenues to preventing human morbidity and mortality associated with disease. One is to correct the imbalance between appetite and energy expenditure, and thereby prevent the development of obesity. The other is to correct the imbalance between subcutaneous, visceral, and ectopic fat deposition, and foster metabolic health in individuals who are already obese. Both these approaches ultimately be important; however, as adipose tissue is more accessible than brain tissue, a focus on WHRadjBMI may provide more mechanisms by which to treat obesity-associated disease. In this project, we identify novel genes associated with both obesity and WHRadjBMI, which may provide therapeutic targets for both of these targets. We then characterize the genetic regulation of WHRadjBMI, identifying genetic mechanisms and gene regulatory networks associated with this trait.

1.4 Overview of dissertation

This projects seeks to synthesize data contained in GWAS of obesity and WHRadjBMI to identify new genetic regulatory mechanism that contribute to these traits. Here, we convert that which has been considered a liability of GWAS, that most significant variants are non-coding, into an asset which allows us to identify both genes and gene regulators implicated in obesity and metabolic disorder. First, we leverage non-coding genetic variants to statistically identify novel gene whose expression is associated with obesity and WHRadjBMI. For WHRadjBMI, we investigate each sex independently, as this WHRadjBMI is sexually divergent in humans, and may have gene regulatory mechanisms unique to each sex. The results of these analyses are seen in Chapter 2 and in Supplementary Tables [N] and [N]. We find many more genes whose expression is associated with WHRadjBMI in females than in males.

To better understand the sexual dimorphism of this trait, we first interrogate the locus of *SNX10* , a gene whose expression is exclusively implicated in female WHRadjBMI. We test the role of *SNX10* in human adipocyte differentiation and in organismal fat gain, and test the role of rs1534696, the lead GWAS variant at this locus, in lipid phenotypes in human adipocytes. These results are further discussed in Chapter 2, sections 3.4 and 3.5.

Next, we interrogate gene regulatory mechanisms of female WHRadjBMI on a genome-wide level. To accomplish this, we use a massively parallel reporter assay[62, 91] to test the enhancer activity of each allele 1,882 variants at the loci of genes associated with female WHRadjBMI. We identify a subset of variants that are located in enhancers and where the allele of the variant modulates enhancer activity in preadipocytes. By interrogating these loci, we determine that female WHRadjBMI-associated enhancers are associated with a motif located within *Alu* repeat elements and associated with transcription factors that coordinate adipogenesis. These results are discussed in detail in Chapter 2, section 3.3.

CHAPTER 2

REPEAT ELEMENTS AND NOVEL GENES CONTRIBUTE TO SEX-SPECIFIC MECHANISMS OF METABOLISM

2.1 Abstract

Increased obesity-associated mortality is exacerbated by abdominal obesity. Here, we identify genes associated with obesity and waist-to-hip ratio adjusted for body mass index (WHRadjBMI), and allele-sensitive enhancers which may regulate female WHRadjBMI genes. 76% (45/59) of these variants are located in primate-specific Alu elements which contain a transcription factor motif associated with adipocyte differentiation. *SNX10*, the strongest female WHRadjBMI-associated gene, is required for adipocyte differentiation, and the lead variant at the *SNX10* locus affects intracellular lipids in human adipocytes. *Snx10* is also required for adipose expansion in female mice, but not males. In conclusion, we find genes, variants, and a gene regulatory network which regulate adipocyte differentation and may affect human health.

2.2 Introduction

Several common chronic illnesses, including diabetes and cardiovascular diseases, are caused or exacerbated by obesity. However, some obese individuals are more susceptible to disease than others, as the location of excess fat deposition strongly modulates risk for obesity-related cardiometabolic disease[1, 13, 5]. Abdominal fat deposition, which represents visceral adipocyte accumulation and ectopic fat deposition on vital organs, is associated with increased insulin resistance and systemic inflammation mediated by the release of free fatty acids and inflammatory cytokines into the bloodstream[18]. Increased immune cell infiltration leads to further insulin resistance, increasing the risk of type 2 diabetes[1]. Waist-to-hip

circumference ratio, adjusted for BMI (WHRadjBMI) is a proxy for the balance of subcutaneous versus visceral and ectopic fat deposition[13]. A high WHRadjBMI indicates a high burden of abdominal fat and is an independent predictor of cardiovascular disease, type 2 diabetes, non-alcoholic fatty liver disease, renal failure, and stroke[13, 18].

Obesity and WHRadjBMI are both complex, polygenic genetic traits[102, 75]. For such traits, genetic heritability is distributed over many variants of small phenotypic effect that map primarily to non-coding regions of the genome[54]. Despite the identification of these variants[102, 75], our appreciation of the mechanisms underlying genetic heritability of metabolic health remains limited. In most loci, we have yet to determine which variants are causal, to identify which genes mediate the phenotypic effects non-coding risk variants, and to understand the role these genes play in human fat distribution and metabolism. To address these questions, we applied a series of experimental and computational approaches to identify genes that may underlie obesity and high WHRadjBMI. Using these gene loci, we identify a gene regulatory network that regulates fat deposition in females. We demonstrate the sex-specific effects in vivo of a female WHRadjBMI-associated gene, illustrating how this gene is involved in adipocyte differentiation and function.

Our work advances the understanding of sex-specific fat distribution from variants of small effect to novel genes whose roles in adipocyte biology may underlie pathologic weight gain, and identified genetic regulation of fat deposition by primate-specific elements.

2.3 Results

2.3.1 Waist-to-hip ratio and obesity show divergent patterns of tissue specificity and sexual dimorphism.

While obesity and high WHRadjBMI predispose individuals to the same chronic metabolic diseases, previous studies have demonstrated that obesity-associated genetic variation is as-

sociated primarily with gene expression in the central nervous system, while adipose tissues play a larger role in WHR heritability[54, 30, 77]. To identify specific tissues with high enrichment for each of these traits for further analyses, we performed ldsc-seg[29], identifying tissue-specific gene expression in linkage to GWAS-identified risk for obesity and WHRadjBMI, using gene expression data from the Genotype-Tissue Expression (GTEx) consortium, consisting of postmortem bulk RNA sequencing data from 53 different human tissues[56] and GWAS data from the GIANT consortium [102, 75]. We find that obesity is strongly enriched in central nervous system tissues, with the strongest enrichment in frontal cortex (Figure 2.1A)[54]. In contrast, WHRadjBMI is associated with gene expression in several tissues. Female reproductive tissues and adipose tissues, especially subcutaneous adipose tissue, show the highest enrichment for WHRadjBMI, as seen in other investigations of this trait[30] (Figure 2.1A, Supplementary Figure 2.6).

Association between tissue-specific gene expression and human traits such as obesity and WHRadjBMI is complicated by the fact that tissues are assemblies of multiple cell types. To account for this heterogeneity, we estimated cell type proportions for 63 cell types in all samples from GTEx and frontal cortex samples from the CommonMind Consortium using xCell[3, 44], which compares sample gene expression with gene set enrichments from primary human cell lines to estimate cell type prevalence. We found that samples from frontal cortex, the tissue most strongly enriched for obesity, show a high estimated proportion of neurons relative to all other cell types. In contrast, subcutaneous adipose tissue showed several cell types with proportions over 5%, including adipocytes and preadipocytes, immune cells such as dendritic cells and natural killer T cells, vascular cells, and stromal cells. (Supplementary Figure 2.7). This cell type diversity is a known attribute of adipose tissue[63]. As a result, the association of adipose bulk tissue with WHR could, in principle, be due to gene expression in any of the cell types present at a significant proportion.

To determine whether adipocytes in particular are associated with WHRadjBMI, we

quantified the estimated adipocyte/preadipocyte proportion ('fat content') for each tissue in GTEx. Across all tissues, high estimated fat content was associated with enrichment for waist-to-hip ratio heritability (Figure 2.1B). Together, these data suggest that WHRadjBMI is associated specifically with gene expression in preadipocytes and adipocytes.

Both obesity and WHRadjBMI are sexually dimorphic traits[68]. Women have lower mean WHR than men but are somewhat more likely to become obese[21, 67, 6]. Based on the sexual dimorphism of these traits at the phenotypic level, we tested whether genetic risk for obesity and WHRadjBMI is also sexually dimorphic. Genetic correlation analyses showed high correlation between male and female genetic risk for obesity, (r_g=0.923, se=0.0083), and lower correlation between male and female genetic risk for WHRadjBMI (r_g=0.6447, se=0.0262) (Figure 2.1C)[9]. Additionally, we found that total heritability is similar between women and men for obesity (female h^2=0.2029, se=0.0064, male h^2=0.2166, se=0.0075), but larger for female WHRadjBMI than male WHRadjBMI (female h^2=0.1582, se=0.0099, male h^2=0.1169, se=0.0059). These findings suggest that while genetic risk factors for obesity are likely shared between the sexes, there may be female-specific genetic determinants of fat deposition, which account for the increased heritability of WHRadjBMI seen in this and previous studies[75].

2.3.2 Identification of novel genes associated with obesity and high waist-to-hip ratio.

To identify novel genes associated with obesity and WHRadjBMI, we performed transcriptome-wide association studies (TWAS), which combine a gene expression panel from a set of individuals, their genotype data, and GWAS summary statistics[102, 75] to identify genes whose expression are associated with genetic risk for a complex trait[41]. We used frontal cortex gene expression as the panel for our obesity TWAS and subcutaneous adipose gene expression as the panel for our WHRadjBMI TWAS, using cortex samples from the CommonMind

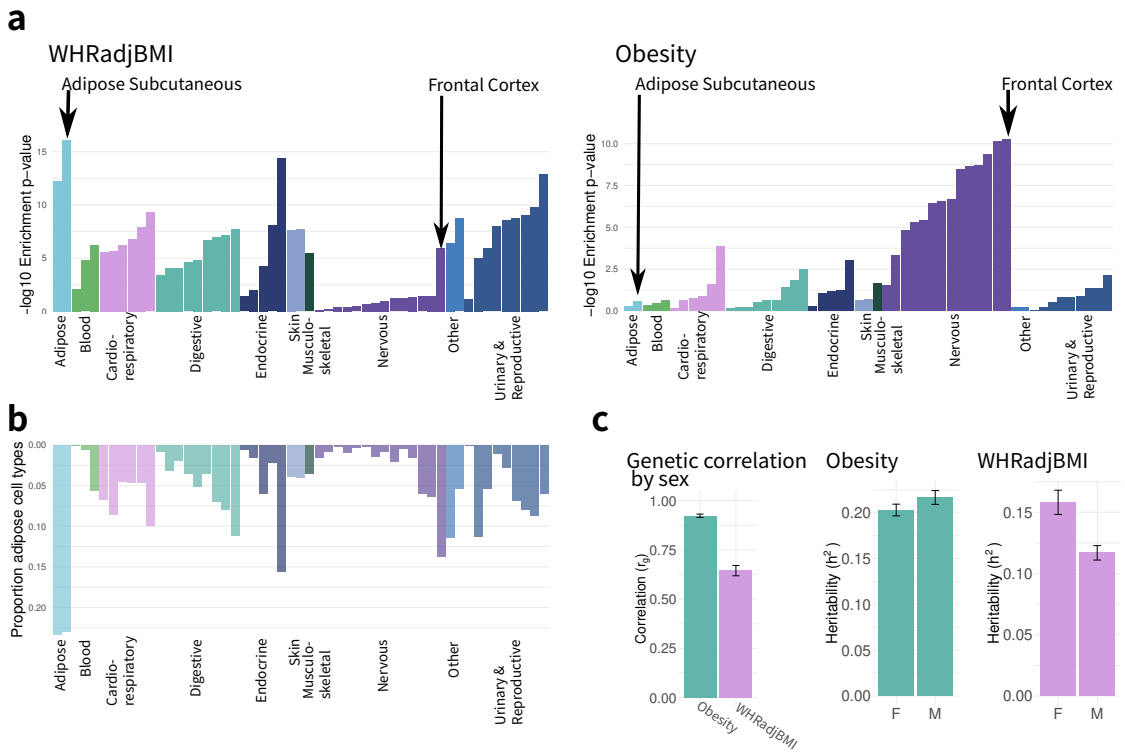

Figure 2.1: **Characteristics of obesity and WHRadjBMI genetic risk.** a) Linkage disequilibrium score regression (LDSC-seg) measuring enrichment of trait near genes with tissue-specific expression. P-values are based on standard error, and represent contribution of tissue-specific expression to trait heritability after correction for baseline annotations. b) xCell-derived cell type enrichment predictions, normalized to generate proportions and averaged across adipose cell types (preadipocyte and adipocyte. c) Genetic correlation and heritability for sex-specific genome-wide association studies (GWAS) conducted for obesity and WHRadjBMI (F=female, M=male). Genetic correlation is high between sexes in obesity and lower for WHRadjBMI. The sexual dimorphism of WHRadjBMI is partially driven by increased heritability of WHRadjBMI in women. Obesity GWAS N∼700,000 , female WHRadjBMI N∼380,000, male WHRadjBMI N∼315,000.

consortium[44] and adipose samples from GTEx[56]. To account for cell type heterogeneity in subcutaneous adipose tissue, we selected the top seven most enriched cell types, representing an average of 47.4% of the sample, and included cell type proportions for these cell types as covariates in gene expression models in the WHRadjBMI TWAS. To identify putative sex-specific mechanisms for WHRadjBMI, GWAS summary statistics from female participants were combined with gene expression samples from females, and male GWAS data were used in conjunction with male gene expression samples.

Our obesity TWAS identified 460 genes after Bonferroni correction whose expression in cortex is associated with obesity risk. The most significant genes include those whose dysfunction in the brain has been implicated in obesity, including *MC4R* and *ADCY3*[38, 26], in addition to genes with uncharacterized roles in obesity (Figure 2.2A). Our WHRadjBMI TWAS identified 91 genes whose expression is associated with high WHRadjBMI in females, and 42 genes in males. *CCDC92*, *DNAH10*, and *L3MBTL3*, which are among the most significant associations, have previously been associated with familial partial lipodystrophy, a disorder which causes subcutaneous adipose loss, visceral fat accumulation, severe insulin resistance, and diabetes[57].

Notably, our WHRadjBMI TWAS results show strong sexual dimorphism: despite equivalent sample sizes, WHRadjBMI TWAS identify 2.16-fold as many genes in women than in men. Genes that are associated with WHRadjBMI in both sexes have, on average, a 1.8-fold larger effect size in women (Figure 2.2B). To evaluate whether sexual dimorphism in WHRadjBMI is the result of sex-specific cis-regulatory effects of variants, we compared the effect size of the top eQTL of female TWAS genes in each sex with the effect size of these variants in WHRadjBMI GWAS from each sex. We observe that eQTLs have comparable effect sizes on gene expression in men and women; however, they have much large GWAS effect sizes in women (Figure 2.2D). This is in concordance with recent observations from GTEx suggesting that there are very few sex-specific eQTLs throughout the genome, despite

common sexual dimorphism in human traits[65]. Our results suggest that although genetic variation affects gene expression similarly between men and women, the consequences of genetically regulated gene expression can differ significantly based on biological sex. We illustrate this effect on *SNX10*, the strongest previously unidentified WHRadjBMI gene and the strongest female-specific WHRadjBMI TWAS gene. A variant located in an intron of *SNX10*, rs1534696, is a strong eQTL in subcutaneous adipose tissue for *SNX10* in both sexes, and is strongly associated with WHRadjBMI in women, but not in men (Figure 2.2C,2.2E, female eQTL p=$1.32*10^{-14}$, male eQTL p=$1.04*10^{-11}$). While the eQTL effect is preserved in both sexes, *SNX10* expression itself is consistently higher in women (Figure 2.22C), with a fold change of ∼1.3 across rs1534696 genotypes. These data suggest that rs1534696 affects *SNX10* expression in adipocytes in both sexes, and that this activity modulates WHRadjBMI in women, but not in men, perhaps because of higher *SNX10* expression in women. *SNX10* expression does not differ between women of typical pre-menopausal and post-menopausal age (Supplementary Figure 2.8, p=0.44), suggesting that the female-specific effect is not the result of increased systemic estrogen levels. rs1534696 is also a lead variant for triglyceride levels and HDL cholesterol in women, but not in men (Supplementary Figures 2.9 and 2.10), suggesting that rs1534696 may also have sex-specific effects on other features of metabolism.

2.3.3 Enhancers at WHRadjBMI gene loci share an adipogenesis-associated motif that is located within Alu repeat elements.

Through TWAS, we identified 91 genes whose expression is associated with high WHRadjBMI in women. Some of these genes are associated with WHRadjBMI specifically in females, but others are associated with WHRadjBMI in both sexes. By identifying enhancers which regulate the expression of these genes, we can gain insight into the gene regulatory network that controls WHRadjBMI. However, owing to extensive linkage equilibrium it is generally impossible to differentiate which variants disrupt gene expression, and which are merely in

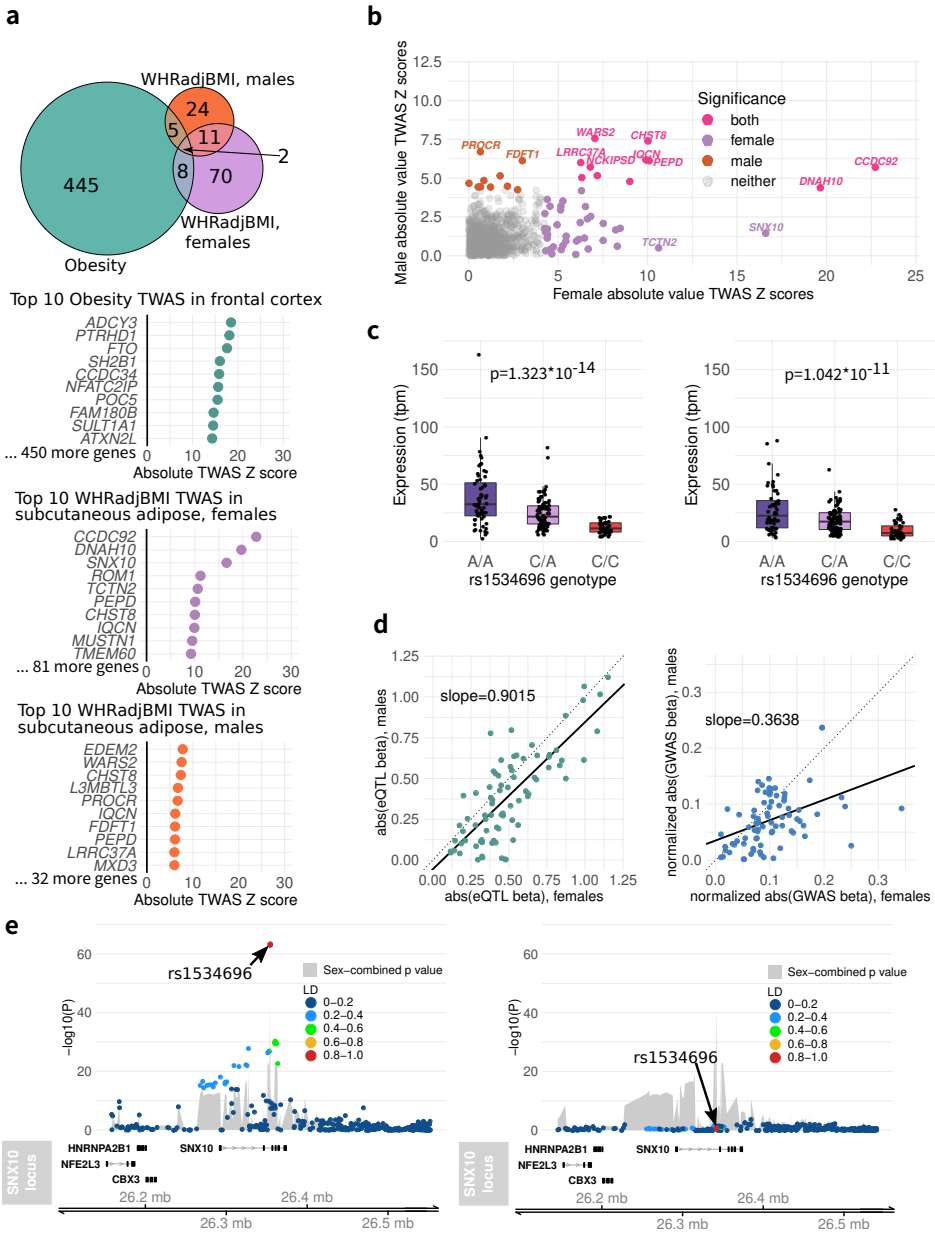

Figure 2.2: **Genes identified by TWAS.** a) Overlap between TWAS-identified gene sets, and top results for each TWAS. Significance based on permutation distrubtion with subsequent Bonferroni correction. b) Sexual dimorphism of WHRadjBMI TWAS results.c) rs1534696 eQTL effect on *SNX10* in women and men. P-values between sexes at each rs1534696 genotype obtained via pairwise t-test. Sex-specific eQTL p-values: female eQTL p=$1.32*10^{-14}$, male eQTL p=$1.04*10^{-11}$. d) Sexual dimorphism of eQTL and GWAS for female WHR. Each dot represents the top eQTL for a female WHR TWAS gene. GWAS betas are normalized by sex-specific trait heritability (h^2). On the left is the p-value of that variant as a sex-specific eQTL; on the right is the p-value of that variant in WHRadjBMI sex-specific GWAS. e) GWAS locus near *SNX10* in women and men. rs1534696 is strongly sexually dimorphic for WHRadjBMI.

linkage with true causal variants from GWAS summary statistics. Therefore, we performed a massively parallel reporter assay (MPRA) to experimentally determine the regulatory properties of all candidate variants in all loci[91, 62]. For each gene identified in our female WHRadjBMI TWAS, we derived a set of variants with a high likelihood of causality: first, we selected the lead GWAS variant, the top eQTL for the TWAS gene, and *coloc*-identified shared causal variant underlying GWAS and eQTL signal, if present. Next, we selected all variants in high LD ($r^2 > 0.95$) with the variants identified above. This resulted in a list of 1,882 biallelic single-nucleotide polymorphisms (SNPs). For each variant, we constructed 175-bp sequences from the human genome centered around each allele of that variant. We synthesized this sequence with 19 independent 10-bp DNA barcodes, allowing for independent measurements of enhancer activity per allele.

We transfected constructs representing 175bp of human genomic sequence, centered around each allele of the selected variants, ligated to a minimal promoter and barcode into 3T3-L1 preadipocytes, allowing the sequences with enhancer activity to self-transcribe and enrich their barcodes in the resulting RNA output. Using this approach, we identified 426 variants in regions that display significant enhancer activity in preadipocytes (significant enhancers). Of these, 59 harbored variants whose alternate alleles significantly modulate enhancer activity (enhancer-modulating variants, or EMVars) (Figure 2.3A, Supplementary Figure 2.11). Significant enhancer variants were more likely to lie within accessible chromatin than nonsignificant variants ($p=1.253*10^{-6}$, Supplementary Figure 2.12A) in Simpson-Golami Behmel Syndrome (SGBS) human adipocytes. Significant enhancers also participated in more physical interactions between enhancers and promoters captured by promoter capture Hi-C in SGBS adipocytes (as vs. nonsignificant variants, p=1.591*10−302, Supplementary Figure 2.12B).

To understand how regulatory active MPRA sequences can promote gene expression in adipocytes, we performed two transcription factor motif enrichment analyses. First, we

compared the 175-bp sequences centered around each of the 426 significant MPRA enhancer variants to the same-sized sequences corresponding to nonsignificant MPRA variants. This analysis yielded 13 transcription factors whose motifs are enriched in significant enhancer sequences. Of these, 7 are directly involved in adipogenesis, based on previous literature (*COUP-TFII*, *KLF14*, *LXRE*, *THRB*, *CEBP*, *RARA*, *EAR2*)[100, 83, 76, 58, 40, 15], and 3 others are involved in adipocyte-relevant metabolic processes such as cholesterol synthesis, adipocyte browning, and glucose metabolism (*SP2*, *ERRA*, *USF2*)[88, 45, 16]. In the second analysis, we looked for enrichment of transcription factor motifs in the sequences surrounding the 59 EMVars versus nonsignificant variants. Here we also found strong enrichment for factors involved in adipogenesis and adipocyte metabolism, and significant overlap with motifs enriched in enhancers. 7 of the 14 motifs are directly involved in adipogenesis (*COUP-TFII*, *LXRE*, *THRB*, *RARA*, *EAR2*, *ATF4*, *ATF1*)[32, 103], and 2 are involved in adipocyte metabolism (*ERRA*, *USF2*) (Figure 2.3B). The results of these motif enrichment analyses are available in the Supplementary Tables. Interestingly, almost all the identified adipogenic motifs are characterized by a common core sequence (Figure 3B), AGGTCA, corresponding to half of the DR1 DNA motif which binds *PPARG*, a master regulator of adipogenesis.

To further assess the functional importance of MPRA sequences, we tested whether significant MPRA enhancers were more conserved than nonsignificant sequences across mammalian species, with the expectation that conserved sequences may have greater functional importance. Surprisingly, we found that enhancer sequences are less conserved than nonsignificant sequences, and sequences containing EMVars are the least conserved sequences in the MPRA (Figure 2.3C, Student's t-test nonsignificant variants vs. enhancers p=$7.073*10^{-2}$, nonsignificant variants vs. EMVars p=$3.173*10^{-2}$). This unusual lack of conservation is due to an enrichment of primate-specific *Alu* repeat elements overlapping enhancer variants in sequences included in the MPRA (FC over nonsignifcant variants=3.17, Chi square p=$1.063*10^{-26}$). Notably, 76% of identified EMVars map within *Alu* elements (Figure 2.3C).

In our MPRA data, most *Alu* sequences tested (344 of 551 *Alu* elements containing an MPRA variant) do not drive reporter gene expression in adipocytes. However, a subset of *Alu* elements are associated with very high enhancer activity in adipocytes (Figure 2.3D). suggesting that some *Alus* have been co-opted as enhancers to stimulate the expression of nearby metabolism-related genes. *Alu* elements do not drive enhancer activity in other MPRA datasets (Supplementary Figure 2.13), suggesting that these *Alu* elements may act as enhancers specific to the context of adipocyte metabolism.

Given that some *Alu* elements stimulate gene expression in preadipoctyes and other *Alu* elements do not, we suspected that the partial DR1 motif that we found to be enriched in significant MPRA results (Figure 3B) may be part of the sequence of *Alus* with enhancer activity. To test this, we first determined that *Alu* elements and adipogenesis-associated motifs co-occur more frequently than is expected by chance (FC=1.57, chi-square p=$2.167*10^{-40}$). Next, we mapped the relative locations of *Alu* elements and motifs, and identified that partial DR1 motifs in significant MPRA sequences are consistently located within the first monomer of the *Alu* element (65-80bp from the *Alu* start site) (Figure 2.3E). The partial DR1 motif is not likely to be the product of a mutation that has emerged from a single *Alu* ancestor, as multiple sequence alignment shows that the *Alu* sequences which drive expression in the MPRA are not more closely related than nonsignificant sequences (Supplementary Figure 2.12). Our data suggest that a subset of *Alu* elements in the human genome may have served as a tool to construct a gene regulatory network that is active in adipocytes, regulates the expression of genes important for fat cell maturation, and modulates body fat distribution.

2.3.4 rs1534696 affects lipid-related morphological features in in vitro differentiated primary human adipocytes

The MPRA results suggest that many WHRadjBMI genes may be affecting body fat distribution by modulating the development and/or function of adipocytes. The strongest

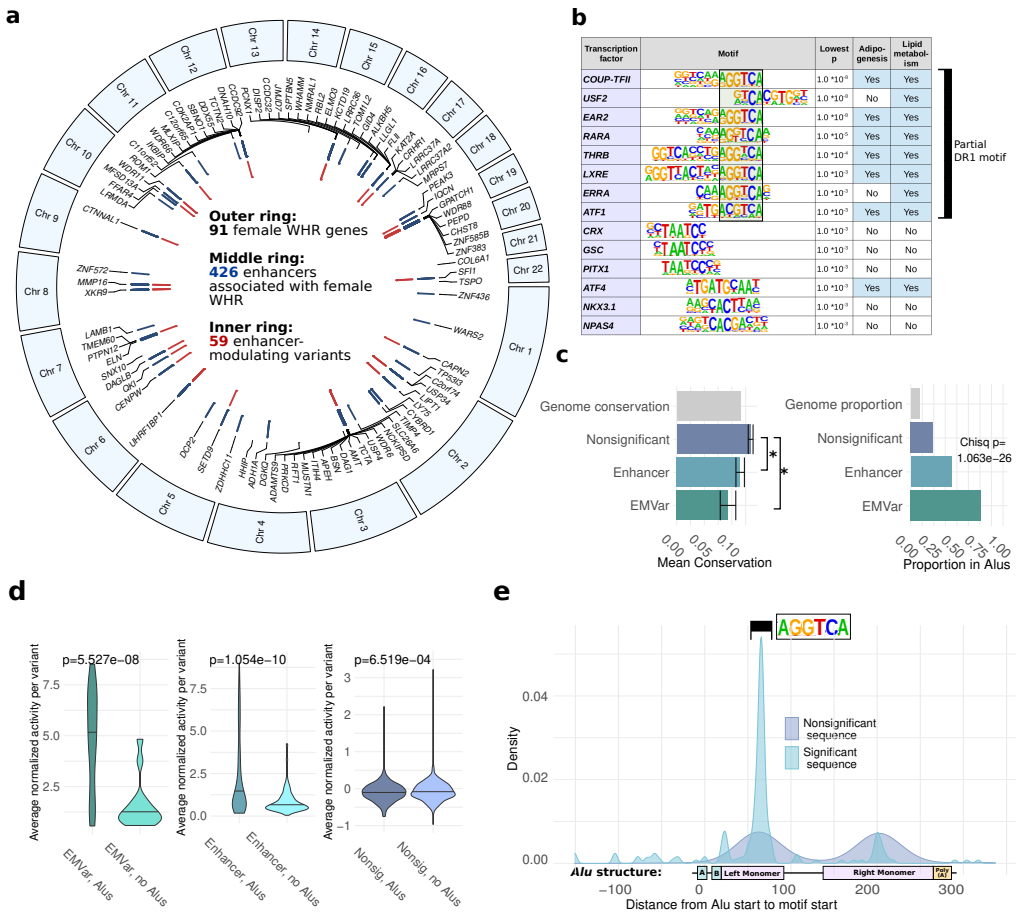

Figure 2.3: **Female WHRadjBMI MPRA.** a) MPRA results. On the outside ring are shown female WHRadjBMI TWAS genes (n=91). In the blue ring are variants that had significant enhancer activity (n=426). In the red ring are variants which modulated the activity of significant enhancers (n=59). Significance determined by Mann-Whitney U-tests and subsequent FDR correction. b) Transcription factor motifs enriched in EMVars relative to nonsignificant enhancer sequences. Transcription factors contributing to common adipose motif are highlighted in green. P-values represent probability of finding degree of enrichment by chance, based on binomial distribution. c) Conservation of MPRA sequences by significance status and presence of *Alu* repeat elements in MPRA sequences. Conservation significance represents pairwise t-tests between average conservation across MPRA sequences. *Alu* significance calculated via chi-square test comparing observed proportions to proportions expected by chance. e) Average enhancer activity of MPRA sequences with and without *Alu* elements. Significance determined by t-tests comparing average normalized activity between variants within or not within *Alu* elements per MPRA signifiance level. d) Density plot reflecting presence of enriched adipogenesis motifs within first monomer of *Alu* elements.

WHRadjBMI TWAS signal we identified was for *SNX10*, a gene and locus that has not been previously associated with fat metabolism. Although the *SNX10* locus does not have a 'AGGTCA'-associated enhancer identified through our MPRA analyses, the lead variant, rs1534696, shows strong evidence for being causal for female WHRadjBMI risk at this locus. The rs1534696 variant is the top eQTL for *SNX10*, and colocalization analyses suggest that rs1534686 is responsible for both the GWAS and eQTL signal at this locus (Supplementary Figure 2.15). rs1534696 has been previously shown to lie within an enhancer active in mature adipocytes[12], but not in preadipocytes, suggesting that this variant acts to modulate processes in mature adipocytes. We confirmed this time-point-dependent regulatory activity of this locus by applying activity-by-contact (ABC), a modeling tool that predicts interaction between enhancers and gene promoters based on chromatin accessibility and H3K27ac histone modifications[34], to pre-adipocytes at several time-points of adipogenesis. ABC analysis predicts that the rs1534696 enhancer interacts with *SNX10* in fully differentiated adipocytes, but not at earlier stages in adipocyte maturation (Figure 2.4A).

To decipher the consequences of rs1534696 on adipocyte morphological and cellular phenotypes, we next used a recently-developed high-content imaging approach, LipocyteProfiler. In brief, LipocyteProfiler generates morphological and cellular profiles describing the structure, function and relationship between cellular organelles, namely AGP (actin cytoskeleton, Golgi body, and plasma membrane), Lipid (lipid droplets and cytoplasmic RNA), Mito (mitochondria) and DNA (nucleic-acid related phenotypes). We assessed LipocyteProfiles across four time-points of adipocyte differentiation (days 0, 3, 8, and 14) in subcutaneous and visceral human adipose-derived mesenchymal stem cells (AMSCs) between different rs1534696 genotypes (Figure 2.4B). We observed significant changes in adipocyte morphological and cellular features dependent on the rs1534696 genotype only in mature subcutaneous adipocytes (Figure 2.4C). Specifically, the majority (50 of 74) of haplotype-driven differential features are Lipid associated features describing the structure of intracellular lipids (Figure 2.4D).

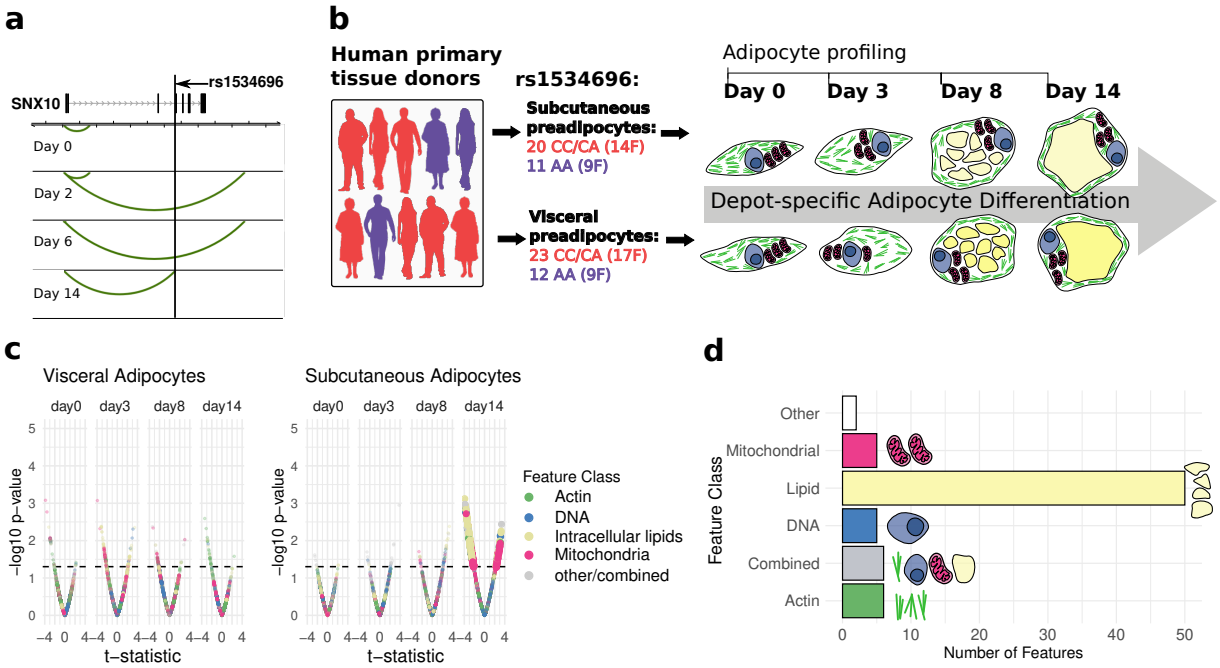

Figure 2.4: **Effects of rs1534696 allele on human primary adipocytes.**) ABC-predicted interactions between *SNX10* promoter and gene body loci in differentiating primary adipocytes. b) Adipocyte Profiler paradigm. c) Adipocyte Profiler features across primary human adipocyte differentiation. Dotted line and larger dots represent significance, defined as $p<0.05$ and $q<0.05$. Significance determined by multi-way ANOVA with subsequent FDR correction. d) Adipocyte Profiler features significant in day 14 subcutaneous adipocytes, grouped by cellular stain target and collapsed across redundant features. Subcutaneous adipocytes: CC/CA N=20, AA N=11. Visceral adipocytes: CC/CA N=23, AA N=12.

This suggests that rs1534696 affects the function of mature adipocytes by dysregulating intracellular lipids.

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

2.3.5 SNX10, a female-specific waist-to-hip ratio gene, is required for adipocyte differentiation in vitro and in vivo

To validate the function of the target gene *SNX10* in adipocyte development and function, we knocked down the expression of *SNX10* in human mesenchymal stem cells via the transfection of anti-*SNX10* short hairpin RNA (shRNA) and differentiated these cells into mature adipocytes. In comparison to wild type cells or a scrambled vector, *SNX10* knockdown resulted in impaired lipid accumulation, and reduced expression of markers of mature adipocytes (Figure 2.5A) (*ADIPOQ* ΔCt scramble shRNA=6.18, ΔCt *SNX10* shRNA=8.86, p=$1.7*10^{-4}$), confirming an important role of this gene in the formation of mature, functional adipocytes. In line with rs1534696-driven effects specific to mature adipocytes, we found that *SNX10* expression correlated with expression of markers of mature adipocytes, but not with markers of progenitor cells or stromal cells (Supplementary Figure 2.16), further suggesting that *SNX10* is the causal target gene at this locus.

To assess whether SNX10 ablation in adipose tissue has an effect on obesity-related metabolic phenotypes an organismal level, we next developed a mouse knockout (KO) of *Snx10* specific to mature adipocytes (mmΔSnx10^{Adipoq}), and then administered a 10-week high fat diet (HFD, D12492, 60kcal% fat)41 Control littermates of knockout animals became obese upon HFD administration. Male mmΔSnx10^{Adipoq} mice also became obese. However, female mmΔSnx10^{Adipoq} mice maintained did not become obese (Figures 2.5B-D, Supplementary Figures 17-18), (KO mean mass F=28.1g, M=48.5g, p=$6.5*10^{-3}$; KO mean body fat F=24.4%, M=42.1%, p=0.029). This suggests that *SNX10* is required for expansion of adipose tissue in a female cellular context, but not in a male one, which may explain its robust association with female WHRadjBMI, but not male WHRadjBMI.

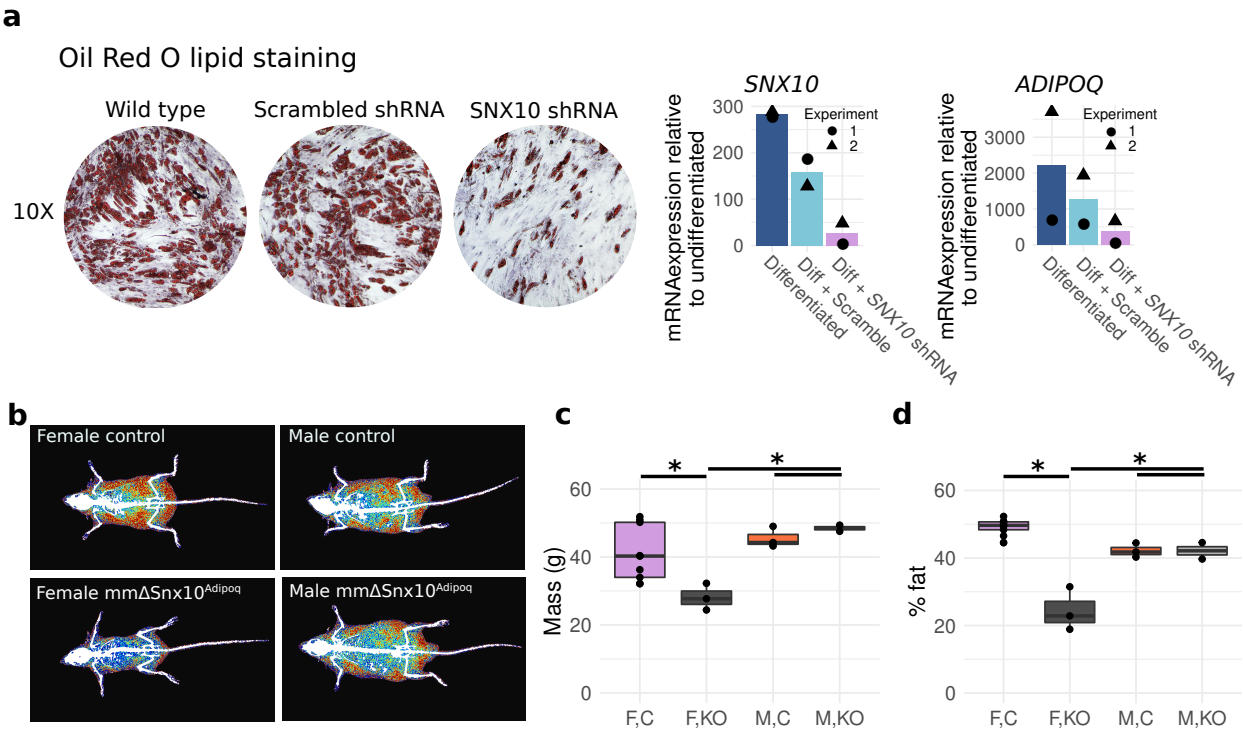

Figure 2.5: **SNX10-mediated inhibition of adipogenesis.** a) Oil Red O staining and qPCR quantification representing lipid accumulation after adipocyte differentiation in human mesenchymal stem cells upon *SNX10* shRNA introduction. Oil Red O images are representative of samples from that condition. b) DEXA scans from wild-type mice (top) and female and male mmΔSnx10^{Adipoq} mice. c) Body weight of WT and mmΔSnx10^{Adipoq} mice after high-fat diet administration. Asterisk indicate significance between groups (F-WT N=8, F-KO N=3, M-WT N=3, M-KO N=2, pairwise t-test, $p<0.05$). d) Body fat percentage of WT and mmΔSnx10^{Adipoq} mice after high-fat diet administration. Asterisks represent results of pairwise t-testing.

2.4 Discussion

This study then illustrates how sexual dimorphism on cellular and organismal traits can emerge from a locus despite variant-mediated gene regulation being shared between both sexes. By closely investigating a genetic locus that is associated with fat metabolism in women, we found gene regulation of *SNX10* to be depot-specific, developmental time point-specific, and sexually dimorphic, potentially explaining the female-specific association of this gene with changes in fat distribution in humans.

Our data also suggest that human patterns of fat deposition may be regulated through co-opted, primate-specific *Alu* transposable elements. Recent reports have demonstrated examples of repeat elements being used to form novel transcriptional networks to create new functions, including mammalian pregnancy and species-specific interferon-mediated immune responses[87, 61]. Our data indicate that *Alu* elements in the vicinity of WHRadjBMI genes often display enhancer-like regulatory properties, as well as enrichment for accessible chromatin. Future work will determine the extent to which *Alu* repeats regulate the development of mature adipocytes in humans.

To explore the role of defects of mature adipocyte function in WHRadjBMI, we chose the strongest novel WHRadjBMI association, *SNX10*. *SNX10* is a sorting nexin that regulates early endosome trafficking and is essential for the function of osteoclasts and gastric zymogenic cells[71, 101]. Here, we demonstrate that *SNX10* is required for adipocyte differentiation, for lipid accumulation, and for expansion of adipose tissue in response to a high-fat diet.

Interestingly, the role of *Snx10* in adipocyte function was female-specific: when exposed to a high-fat diet, male mmΔSnx10^{Adipoq} mice gained fat mass, but females did not. The likely causal variant for *SNX10* expression in human mature adipocytes, rs1534696, affects gene expression in both sexes. However, the GWAS signal at this locus shows a strong relationship with WHRadjBMI in women, but not in men. These findings illuminate a

vexing problem in genetics: how sexually dimorphic traits emerge from gene regulation that is common to both sexes. Sexual dimorphism is present in many human traits, from height and weight to lifespan and disease burden[87, 61]. Recently, the GTEx consortium[65] showed that 37% of genes have sex-biased expression, most only in a few tissues. However, this sex-biased expression is not driven by sex-biased effects of allelic variation. Of the ∼500,000 cis-eQTLs identified by GTEx, only 369 had effects that were biased by sex[65].

In this work, we describe how *SNX10* expression affects adipocyte maturation and fat accumulation in a female-specific manner, explaining a GWAS locus for WHRadjBMI that is also specific to females. In this respect, *SNX10* bears remarkable resemblance to *KLF14*, a gene that is causal at a female-specific locus for type 2 diabetes, and has been described by Small et al[83]. *KLF14* is a master transcription factor that likely acts through disrupting later stages of adipocyte maturation. Small et al. attribute many of the effects of *KLF14* disruption to reduced expression of *SLC2A4*, which encodes the glucose transporter GLUT4. Without sufficient GLUT4, *KLF14*-deficient adipocytes cannot perform lipogenesis, resulting in limited adipose capacity, ectopic fat deposition, and predisposition to diabetes.

SNX10 differs from *KLF14* in a number of important ways. *SNX10* is not imprinted; while *KLF14* is expressed only from the maternal chromosome[90]. *SNX10* is not regulated by *KLF14*, not is it a transcription factor itself. *SNX10* is a sorting nexin, a class of proteins that regulate intracellular trafficking. *SNX10* in particular localizes to early endosomes and is critical for their function[71, 101]. In adipocytes, early endosomes are responsible for trafficking GLUT4 to and from the membrane, regulating glucose intake in response to diet. Therefore, whereas *KLF14* disrupts GLUT4 expression, *SNX10* may disrupt GLUT4 on a post-translational level. Disruption of glucose intake and lipogenesis, process critical to mature adipocyte function, may disrupt the cellular function of adipocytes globally, resulting in reduced expression of mature adipocyte markers seen in this study.

Importantly, *SNX10* shows sexual dimorphic effects on adipose tissue both in human

and in mice. This is another way in which *SNX10* differs from *KLF14*: mice with a conditional *Klf14* knockout in adipose tissue knockout (driven by *Adipoq*, as in our study) showed metabolic phenotypes in both sexes. More research needs to be done to understand how *Snx10* depletion affects adipocyte metabolism in female cells, but not male cells. Our data show modestly increased *SNX10* expression in females relative to males (Figure 2.2C): therefore, the sexual dimorphism could be a property that emerges at a threshold level of *SNX10* expression, as Small et al. theorize. *SNX10* expression is not significantly different between younger and older women, suggesting that sexual dimorphism is not a property of systemic estrogen signalling. As X chromosome dosage has been shown to affect adiposity independent of gonadal hormones[19], *SNX10* could also be acting through interaction with X chromosome genes. In any case, because *SNX10* affects fat metabolism in a sex-specific manner in both human women and female mice, this gene is a valuable resource for understanding sex differences in adipocyte physiology and metabolism.

With this work, we how recently evolved gene regulatory networks formed from pre-existing repeat elements regulate adipocyte differentiation, an important determinant of human waist-to-hip ratio. We also describe how adipocyte differentiation is disrupted in females by depletion of *SNX10*, a gene that may regulate intracellular GLUT4 trafficking. In these ways, we expand our knowledge about how the genome regulates adipocytes to control fat deposition, determining the impact of caloric excess on human health.

2.5 Methods

2.5.1 LDSC-seg:

LDSC-seg analyses were conducted as described by Finucane et al[30]. For each tissue, enrichment for a set of specifically-expressed genes was calculated in comparison to a baseline model measuring non-specific enrichment near genes, histone annotations, and open chromatin, and enhancers. The set of specifically expressed genes were used as a functional annotation, and the enrichment of this annotation in linkage to GWAS variants of low p-value, conditioning on the baseline model and the expression of all genes. Tissues included all 53 tissues available from the GTEx v8[56].

2.5.2 xCell:

xCell analyses were performed as described in Aran et al.[3]. All samples from GTEx v8 (N=17382), representing 53 tissues, were combined with frontal cortex data from the CommonMind Consortium (N=613)[44]. Enrichment analyses for 63 cell types and cellular landscapes were calculated, transformed to proportion estimates, and corrected for spillover between similar cell types. Estimations for similar cell types (B cells, CD4+ T-cells, CD8+ T-cells, dendritic cells, macrophages, endothelial cells) were averaged, and outliers (smooth muscle) and stem-type cell types removed prior to visualization.

2.5.3 Genetic correlation:

Heritability estimation and genetic correlation was performed using cross-trait LD score regression as described by Bulik-Sullivan et al[9, 10]. Summary statistics were gathered from sex-specific GWAS meta-analyses of obesity and WHRadjBMI[102, 75]. Heritability was obtained by solving the LD score regression equation for h^2, given chi-square statistics from GWAS variants and linkage disquilibrium data. Genetic covariance was obtained by

obtaining the product of the Z-scores for each variant, accounting for linkage disequilibrium and sample size. Genetic correlation was obtained by normalizing genetic covariance by SNP heritability.

2.5.4 Transcriptome-wide association studies:

For the TWAS investigating waist-to-hip ratio adjusted for BMI (WHRadjBMI), we used the following data: 1) subcutaneous adipose tissue samples from GTEx v8 segregated by sex, 2) Sex-specific WHRadjBMI GWAS summary statistics, and 3) xCell estimates of cell type proportion derived from GTEx v8 and CommonMind Consortium data as described above. Estimates for the 6 most prevalent cell types in subcutaneous adipose tissue (preadipocytes, adipocytes, dendritic cells, chondrocytes, epithelial cells, and fibroblasts) were included as covariates, as were participant age and race.

We used FUSION scripts to perform TWAS on both gene expression and intron splicing data. First, we estimated the heritability of each feature (gene expression or intron boundary usage) as a function of the genotypes of the variants in cis, using GCTA-GREML to compute h^2 and a p-value representing the result of a likelihood ratio test. Genes and intron boundaries where $p>0.05$ after Bonferroni correction for multiple testing were included in further analyses. Predictive models were calculated estimating the effect of cis-variation on gene expression and intron boundary usage. Four different models were used: lasso, elastic net, top single eQTL, and best linear unbiased predictor. For the TWAS investigating WHRadjBMI, cell type proportions from the seven most prevalent cell types (adipocytes, chondrocytes, epithelial cells, fibroblasts, preadipocytes, dendritic cells)in adipose tissue from xCell predictions were included in the expression model. The best-performing model was selected via cross-validation, using shuffled gene identifiers to generate a null distribution. The imputed gene expression or boundary usage under the model was then correlated with WHRadjBMI summary statistics for the relevant sex. Bonferroni correction was applied to resulting set of

genes to yield a set of genes that are statistically associated with sex-specific WHRadjBMI under strict criteria.

For the TWAS investigating obesity, we used the following data: 1) frontal cortex samples derived from the CommonMind Consortium, and 2) Sex-combined GWAS summary statistics derived from the same samples as the WHRadjBMI GWAS. Cell type estimation was performed but cell types proportions were not included in obesity cortex TWAS due to the homogeneity of frontal cortex samples (see Figure 2.1). We performed a sex-combined TWAS for obesity based on the high genetic correlation of obesity risk between the sexes (Figure 2.1). Obesity TWAS was otherwise performed as described for WHRadjBMI.

Colocalization analyses were performed using *coloc*[37], version 3.1 to detect shared causal variants between eQTL and GWAS signal. For each significant gene in each TWAS, *coloc* was performed between the GWAS summary statistics that went into that TWAS (e.g. female, male, or combined) and the sex-combined eQTL data for that gene. eQTL data used was provided by the GTEx Consortium. The locus used was +/-0.5 MB of the canonical transcription start site(TSS) of the gene. A variant was considered a single shared causal variant if 1) it had the maximum posterior probability of being the single shared causal variant, and 2) the posterior probability of there being a single shared causal variant (PP.H4) was > 0.5. At the *SNX10* locus, a secondary colocalization analysis was performed to account for multiple female WHRadjBMI GWAS signals located within +/-0.5 MB of the *SNX10* TSS. The locus boundaries of the new locus were chr7:26200000-26400000(hg19).

Sex-specific eQTL analyses (Figure 2.2C) were performed on rank-normalized expression data from GTEx and matching genotypes using fastQTL[66], including as the first three principal components from the genotype data as covariates. P-value represents nominal p-value derived from linear modeling of relationship of *SNX10* expression to genotype.

2.5.5 UKBB GWAS locus zoom:

UKBB GWAS locus zooms were generated from sex-specific GWAS data generated by the Neale lab (http://www.nealelab.is/uk-biobank). Summary statistics were combined with linkage disequilibrium information generated from GTEx v8 genotype data. Points shown represent variants available in both UKBB[86] and GTEx v8.

2.5.6 Massively parallel reporter assay:

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

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

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

3T3-L1 cell culture and transfection: 3T3-L1 preadipocyte cells were grown at 37°C in 5% CO2 in DMEM (Gibco # 11995-065) media supplemented with 10% FBS, 1% penicillin-streptomycin solution (10,000 U/ml; Gibco #15140122), 0.8mg/ml Biotin (Sigma; #B4639)

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

MPRA library preparation: MPRA library design and preparation was performed as previously described in Ulirsch et al[91]. Briefly, oligonucleotides representing each allele of selected variants (see 'Variant selection' above) were synthesized on an Agilent 100k oligonucleotide array. Homology arms were added onto the fragments using emulsion PCR, which reduces PCR amplification bias. Oligos were then cloned into a linearized pMPRA1 vector (addgene #49349) using Gibson Assembly. This pool of pMPRA1 vector plus inserts was then linearized and a truncated eGFP reporter gene containing a minimal promoter and spacer sequence was then cloned into the vector using T4 DNA ligase (NEB; # B0202S) to create the input DNA library. Once the final constructs were produced, they were transfected into 3T3-L1 preadipocytes (3 technical replicates) as described above. At least 10 million cells were transfected per replicate. Cells were collected 48 hours after transfection and flash frozen in liquid nitrogen. After all replicates were transfected, RNA was extracted using a Qiagen RNeasy mini kit. mRNA was isolated from total RNA using Invitrogen Dynabeads (ThermoFisher #61006). All mRNA was then converted to cDNA using Superscript III reverse transcriptase and then PCR amplified using Illumina multiplexing primers. Two reactions of input DNA (DNA used as transfection material) were also PCR amplified with Illumina multiplexing primers at this point. Libraries were amplified with 10-11 PCR cycles and pooled before being cleaned. Library quality was assessed using the Agilent DNA 1000 bioanalyzer chip, where a single sharp peak of around 250bp is expected. Samples were then sent for paired end NGS sequencing. A 25% PhiX genome spike in must be added to each sequencing run due to low complexity.

Significance testing: For each technical replicate of the MPRA, counts per barcode were log2cpm format. Barcodes where log2cpm $<=$ -3 were removed, as were variants where either

allele was left with fewer than 7 barcodes after removal. Enhancer activity per barcode was defined as log2cpm(replicate) – log2cpm(input). A nonparametric (Mann-Whitney U) test was performed to identify variants with significant enhancer activity, comparing each allele of a variant with all other sequences included in the analysis. Subsequent FDR correction was performed on the set of significant enhancers to correct for multiple comparisons. Enhancers are considered significant if FDR q<=0.05 in two or more replicates. Within the set of significant enhancers, subsequent Mann-Whitney U tests were performed to identify significant enhancer-modulating variants (EMVars), comparing one allele of the enhancer versus the other. Subsequent FDR correction was performed and EMVars with FDR q<=0.05 in two or more replicates were considered to be significant.

2.5.7 Enrichment of MPRA enhancers in functional chromatin and physical interactions:

ATAC-seq was performed on Simpson Golabi Behmel Syndrome (SGBS) adipocytes at four stages of differentiation (preadipocytes, day 2, day 8, day 16). Overlap was defined as an MPRA variant lying within a chromatin accessibility peak in any of the four time points. Promoter capture Hi-C was performed on the same time points across the differentiation time course. An MPRA variant was defined as participating in a physical interaction if the variant lay within the restriction enzyme fragment on either end of the physical interaction. The number of physical interactions an MPRA variant participated in was averaged across time points. For both ATAC-seq and promoter capture Hi-C, significance was determined via chi-square testing.

2.5.8 MPRA motif analysis:

Significantly enriched motifs were identified using HOMER, using the set of known motifs included in the HOMER motif database[43]. Motif detection was done using the 175-bp

sequence surrounding each variant in hg38, using the reference allele of each variant. To correct for strand-specific biases introduced by autonormalization in Homer v3, the analysis was performed on both the sequences and their reverse complements. Motifs were considered enriched if FDR q<0.05. Motif enrichment was identified in a) the set of significant enhancers compared to nonsignificant tested enhancers, and b) the set of EMVars compared to nonsignificant tested enhancers. The location of the common adipose motif was defined as the span from the earliest start of any of the transcription factor motifs which contributed to this motif (*COUP-TFII*, *EAR2*, *ERRA*, *LXRE*, *RARA*, *THRB*, *ATF1*, *USF2*) to the latest end of any of these motifs.

2.5.9 Repeat elements:

Repeat element location and class were obtained from the RepeatMasker data available for assembly GrCh38 from UCSC (http://hgdownload.cse.ucsc.edu/goldenpath/hg38/database/). The *Alu* canonical sequence was obtained from RepeatMasker: http://www.repeatmasker.org/AluSubfamilies/humanAluSubfamilies.html. The sequences of all *Alu* elements overlapping an MPRA variant were aligned using Clustal Omega[79], version 1.2.4, using full distance matrix calculation. The resulting guidetree was visualized as a cladogram (no branch length information) using ggtree, with leaves labeled according to MPRA significance.

2.5.10 Repeat Element Conservation:

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

categories of variants.

2.5.11 Lipocyte Profiler in human adipose-derived mesenchymal stem cells:

Lipocyte profiles of morphological and cellular features were generated from a genotyped cohort of subcutaneous and visceral adipose-derived mesenchymal stem cells (AMSCs) from patients undergoing a range of abdominal laparoscopic surgeries (sleeve gastrectomy, fundoplication or appendectomy). The visceral adipose tissue was derived from the proximity of the angle of His and subcutaneous adipose tissue was obtained from beneath the skin at the site of surgical incision. Each participant gave written informed consent before inclusion and the study protocol was approved by the ethics committee of the Technical University of Munich (Study # 5716/13). Isolation of AMSCs was performed as previously described[82]. Pre-adipocytes were differentiated for 14 days. LipocytePainting was performed at four time-points of differentiation (days 0, 3, 8 14) and features were extracted. We compared the lipocyte profiles in females between rs1534696 CC/CA and AA allele carriers using a multi-way analysis of variance (ANOVA) adjusted for BMI, age and batch. p-values were corrected for multiple testing using false positive rate (FDR) described in R package "qvalue" (qvalue). Features with FDR $< 5\%$ were classified to be significantly different between risk and non-risk haplotype carriers.

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

Activity-by-contact (ABC) was performed as previously described[34, 80]. In brief, ABC predictions were modelled based on H3K27ac ChIPseq and ATACseq55 from primary human adipose-derived mesenchymal stem cells from one female donor at four time-points of differentiation (days 0, 2, 7, 14).

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

Adipogenesis differentiation in human mesenchymal stem cells (hMSC) was followed as per the manufacturer's instruction (Lonza, Walkersville, MD, USA, Document #AA-2501-16 07/11). hMSC 2.1 x 104 cells/well were plated on a 24 well plate in MSCGMTM at 370C, 5% CO2 until confluency. At 100% confluence, the MSCGMTM were replaced with adipogenic induction media supplemented with SingleQuots^{TM} of h-insulin (recombinant), L-glutamine, MCGS, dexamethasone, indomethacin, IBMX (3-isobutyl-l-methyl-xanthine) and GA-1000 (Lonza, Walkersville, MD, USA). hMSC cells were transfected with shRNA control (Cat #SHC202V) and *SNX10* MISSION shRNA Lentiviral Transduction Particles (Cat #SHCLNV) 5 hours prior to adding the adipogenic induction media. hMSC adipogenic cultures were incubated with shRNA for 3 days. After 3 days, the medium was replaced with Adipogenic Maintenance Medium supplemented with SingleQuots^{TM} of h-insulin (recombinant), L-glutamine, MCGS and GA-1000 (Lonza, Walkersville, MD, USA). The cells again were transfected with *SNX10* MISSION shRNA Lentiviral Transduction Particles 5 hours prior to adding the Adipogenic Maintenance Media and cultured for 3 days. The hMSC cells were maintained in induction/maintenance media for two more cycles without shRNA for optimal adipogenic differentiation. The cells were then maintained in Adipogenic Mainteinance Medium for 7 more days (differentiation was carried for a period of 21 days). hMSC adipogenic cultures were harvested for RNA as well as stained for Oil Red O.

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

We obtained the *Snx10* targeting vector, PG00216_Z_2_C06, from the European Conditional Mouse Mutagenesis Program (EUCOMM). This vector is a "knockout first" gene trap which inserts a flippase site-flanked Neo selection cassette with an IRES and LacZ reporter into intron 3 and inserts LoxP sites flanking exons 4 and 5. Exons 4 and 5 contain the

functional domain of *Snx10*, the PX domain required for phospholipid interactions. This allele is designated Snx10tm1a(EUCOMM)Hmgu. Hereafter in this paper, we refer to the resulting targeted allele as Snx10Neo-f. The vector was electroporated into V6.5 ES cells. Infection was carried out at a multiplicity of infection (MOI) =1. Control cells were infected with scrambled Snx10 shRNA virus. Both SNX10 MISSION shRNA Lentiviral Transduction Particles and shRNA Control Transduction Scrambled Particles were purchased from Sigma-Aldrich. Neomycin resistant clones were picked, expanded and screened for correct insertion by Long Range Genomic PCR using the following vector-specific primers and gene-specific primers:

5' Integration

Gene Specific Forward (GF3). 5'-GCTTATGGTCGACTCATCGGAGAATC-3'

LacZ Reverse (LAR7), 5'-GGTGTGGGAAAGGGTTCGAAGTTCCTAT-3'

Amplicon size=5,170 bp

3' Integration

LacZ Forward (LAF), 5'-GAGATGGCGCAACGCAATTAATG-3'

Gene Specific Reverse (GR4),5'-CACAGAAGTAATGTACGCTAATGGCAACG-3'

Amplicon size=5,680 bp

The resulting clones were injected into host blastocysts to generate mouse chimeras. Two male chimeras were bred with C57BL/6J females. Germ line transmission was confirmed by PCR, using primers flanking the third loxP site:

LoxP3 Forward: 5'-ATAACTAACCCAGGCAAACA-3'

LoxP3 Reverse: 5'- TTGTCAAGTGCGT GTGTCGT-3'

Snx10Neo-f/+ offspring were bred to homozygosity to generate animals for experiments at the expected Mendelian ratio of approximately 25%. PCR genotyping using the preceding primer pair produced bands of 213 bp for the WT (+/+), 273 bp for the Snx10Neo-f/Neo-f, and both bands for the heterozygotes Snx10Neof/+. Adipose tissue-specific *Snx10*-deficient

mice were generated in two steps. First, the Neo cassette was removed by crossing Snx10Neo-f/+ males with females homozygous for *Rosa26*-driven FLP recombinase (strain B6.129S4-Gt(ROSA)26 Sortm1(FLP1)Dym/RainJ), resulting in an allele with exons 4 and 5 flanked by loxP sites (Snx10fl/+). Snx10fl/fl animals were viable and fertile. Then, homozygous males were crossed with females mice hemizygous for the *Adipoq*-Cre BAC transgene (B6.FVB-Tg(Adipoq-cre)1Evdr/J Stock # 028020), in order to breed Adipoq;Snx10fl/+ males which were backcrossed with Snx10fl/+ females. The resulting mice carry an adipose tissue-specific *Snx10* allele with exons 4 and 5 deleted, for adipose tissue-specific knockout.

2.5.15 In vivo measurement of fat and lean body mass:

Dual Energy X-ray Absorptiometry (DEXA) analysis was performed using the InAlyzer DEXA (InAlyzer, Serial # INZ-201021-S002, MEDIKORS). The values for mass, body fat percentage, and femur length were obtained. Magnetic resonance imaging of body composition estimation was acquired.

2.6 Appendix A: Supplementary Figures

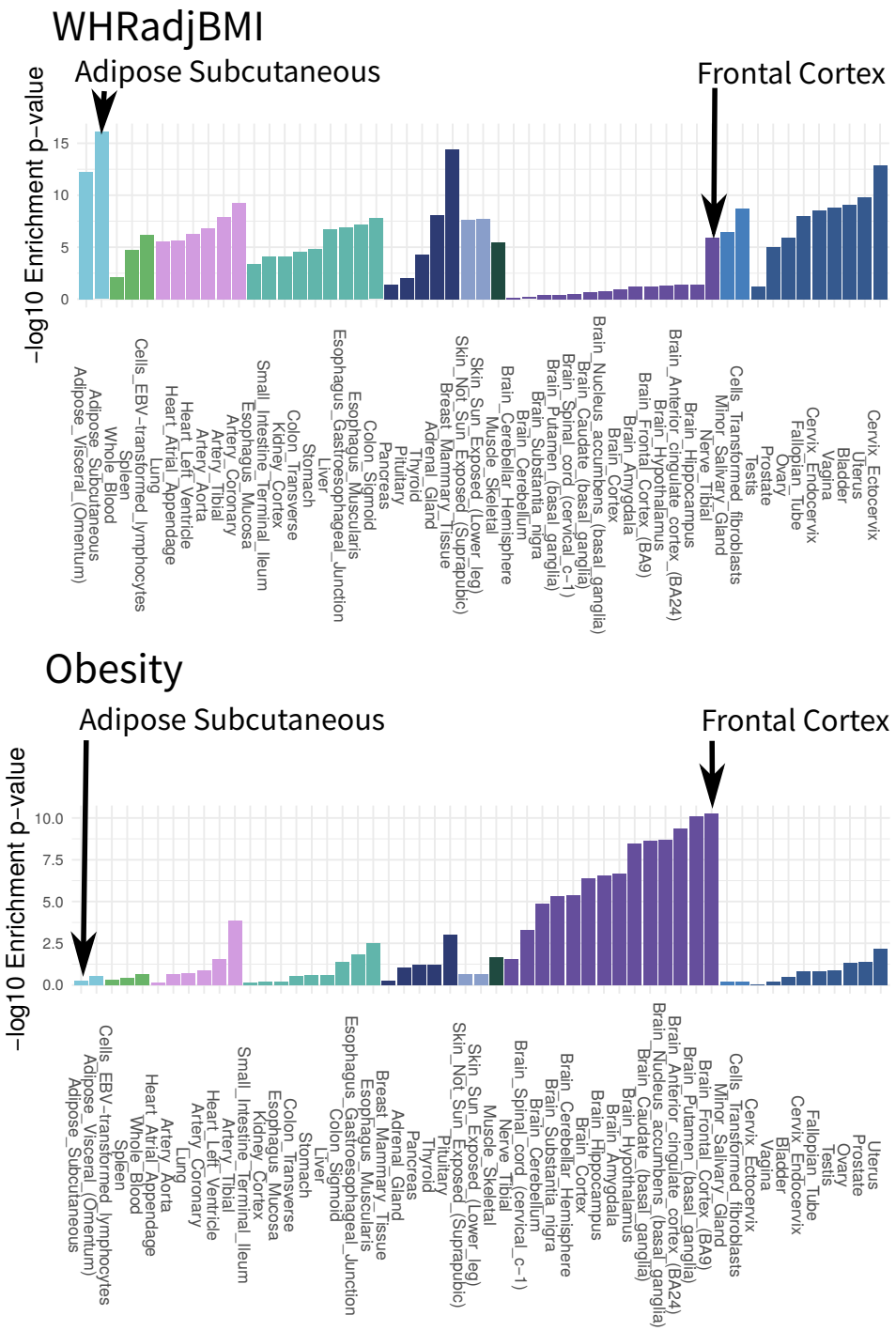

Figure 2.6: **Enrichment for trait by GTEx tissue.** ldsc-seg results for each tissue tested in ldsc-seg enrichment analysis. y-axis is -log10(p) for the enrichment of trait in linkage with genes specifically expressed in that tissue. P-values are based on standard error, and represent contribution of tissue-specific expression to trait heritability after correction for baseline annotations.

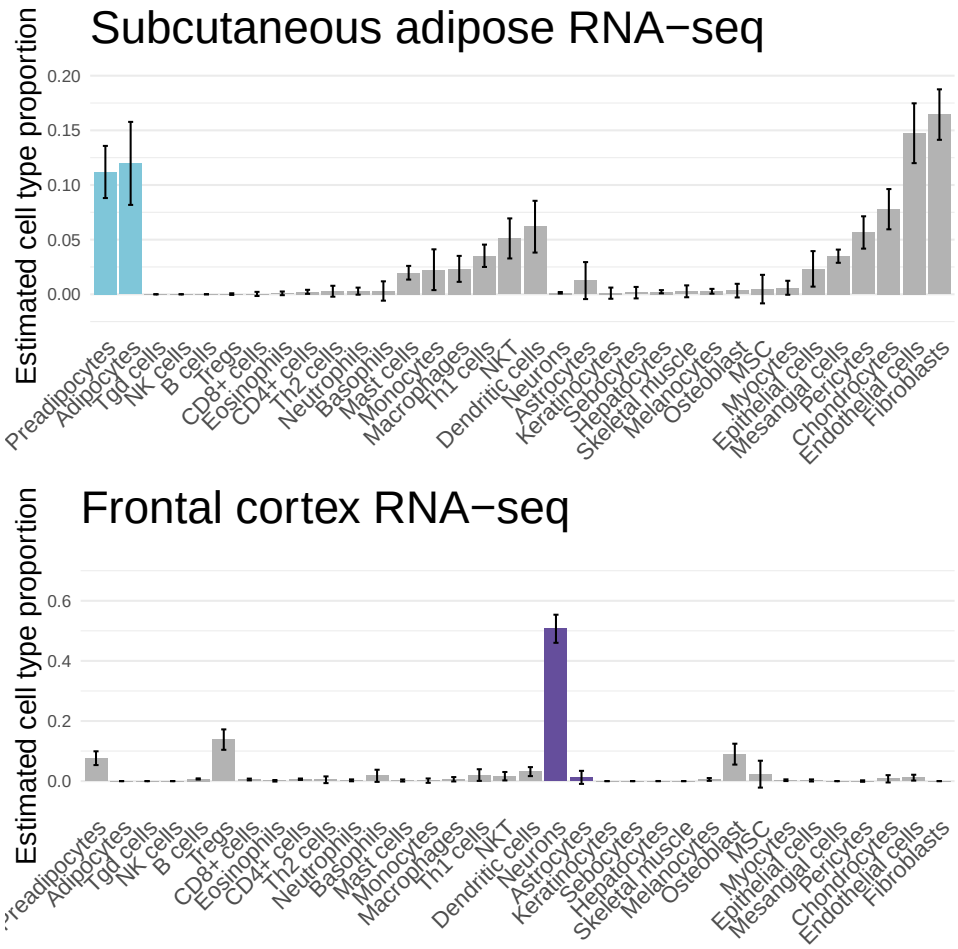

Figure 2.7: **Estimated cell type proportions in subcutaneous adipose.** Y axis is estimated and normalized cell type proportions from xCell for subcutaneous adipose(GTEx) and frontal cortex(CMC) samples.

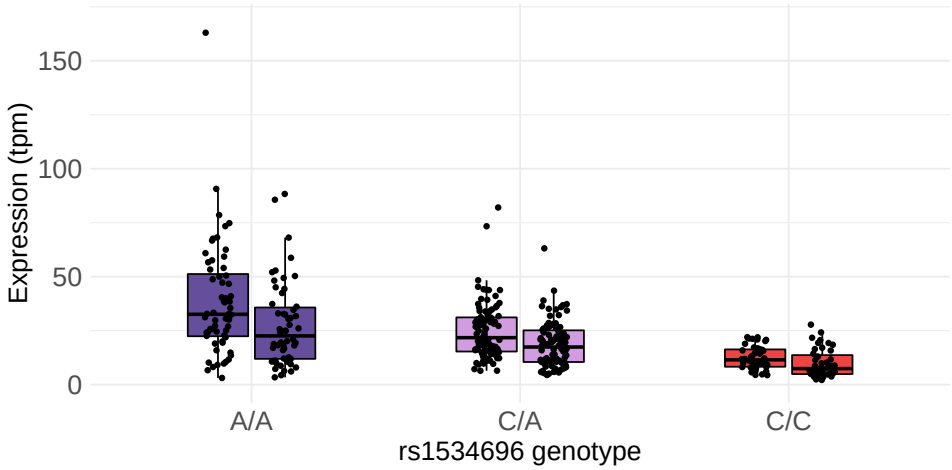

Figure 2.8: **SNX10 expression in women of premenopausal and postmenopausal age ranges.** Shown is TPM values representing SNX10 expression in GTEx subcutaneous adipose tissue in women ages 20-39 (n=32) and 60-79 (n=61). p-value represents pairwise t-test between the two groups.

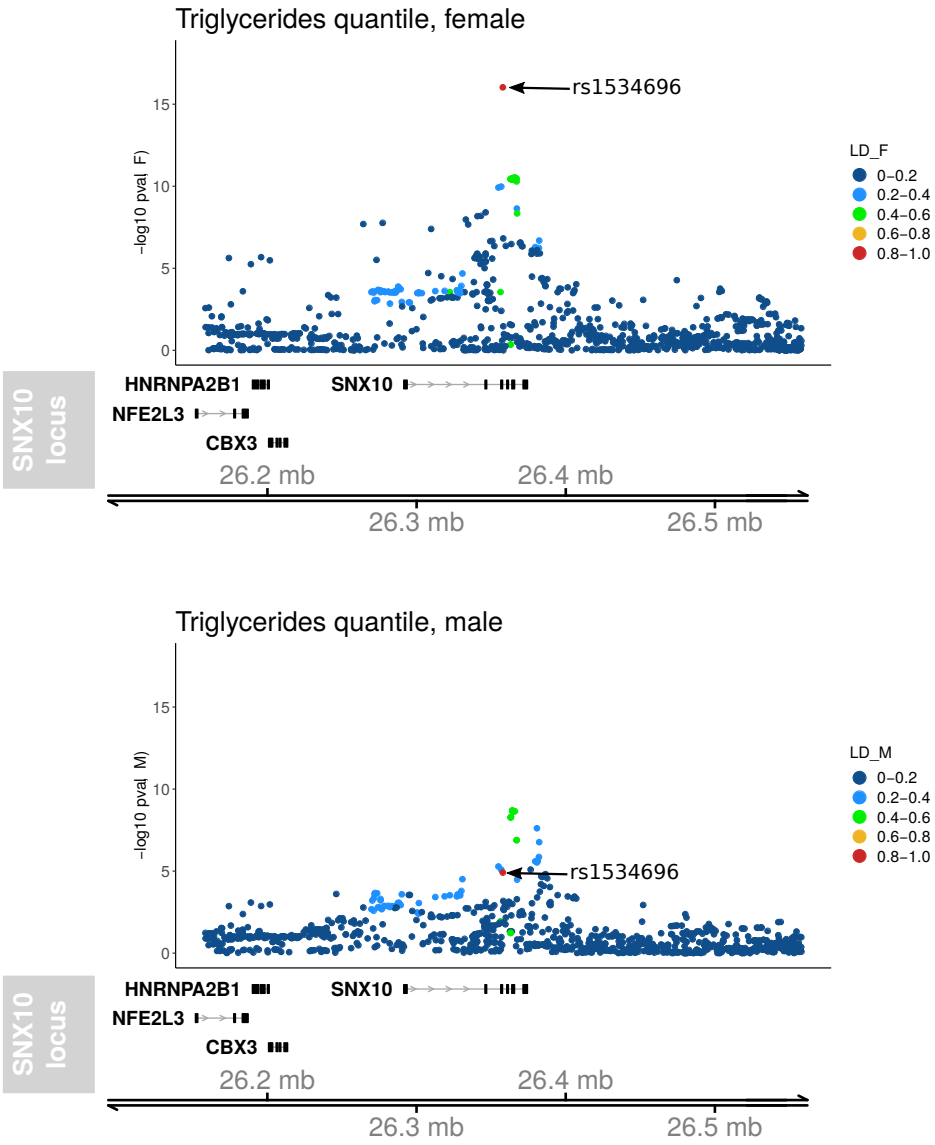

Figure 2.9: **UKBB GWAS summary statistics at SNX10 locus, triglycerides.** Plots represent UK biobank GWAS summary statistics for quantile-binned blood triglyceride levels, separated by sex. x- axis represents genomic position; y axis represents -log10(p) for association with trait. Color of points represents linkage with rs1534696.

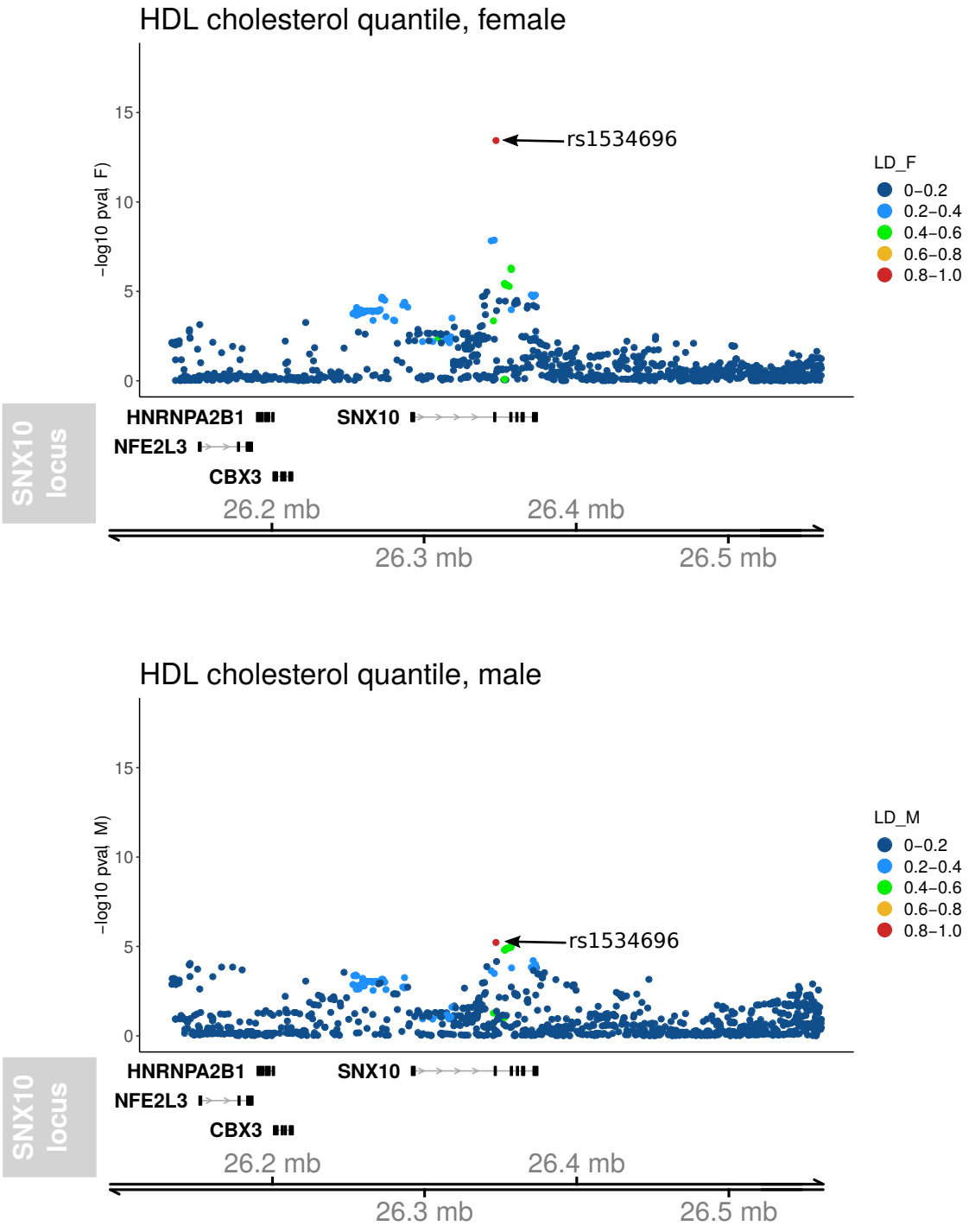

Figure 2.10: **UKBB GWAS summary statistics at SNX10 locus, HDL cholesterol.** Plots represent UK biobank GWAS summary statistics for quantile-binned HDL cholesterol levels, separated by sex. x- axis represents genomic position; y axis represents -log10(p) for association with trait. Color of points represents linkage with rs1534696.

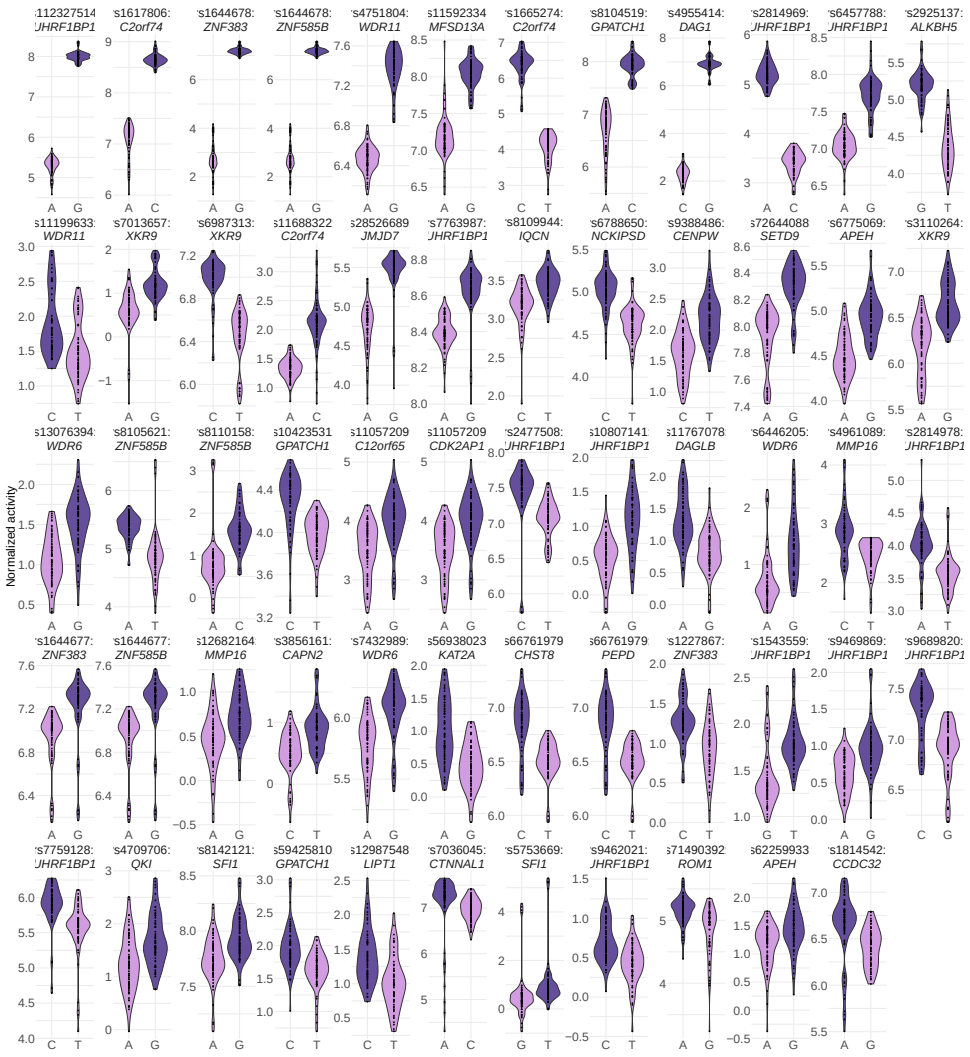

Figure 2.11: **EMVars by allele.** Normalized activity level of all 59 EMVars. y axis represents normalized activity level; x axis represents allele. EMVars are labeled by rsid and by female WHRadjBMI TWAS gene from which they were identified (see MPRA methods).

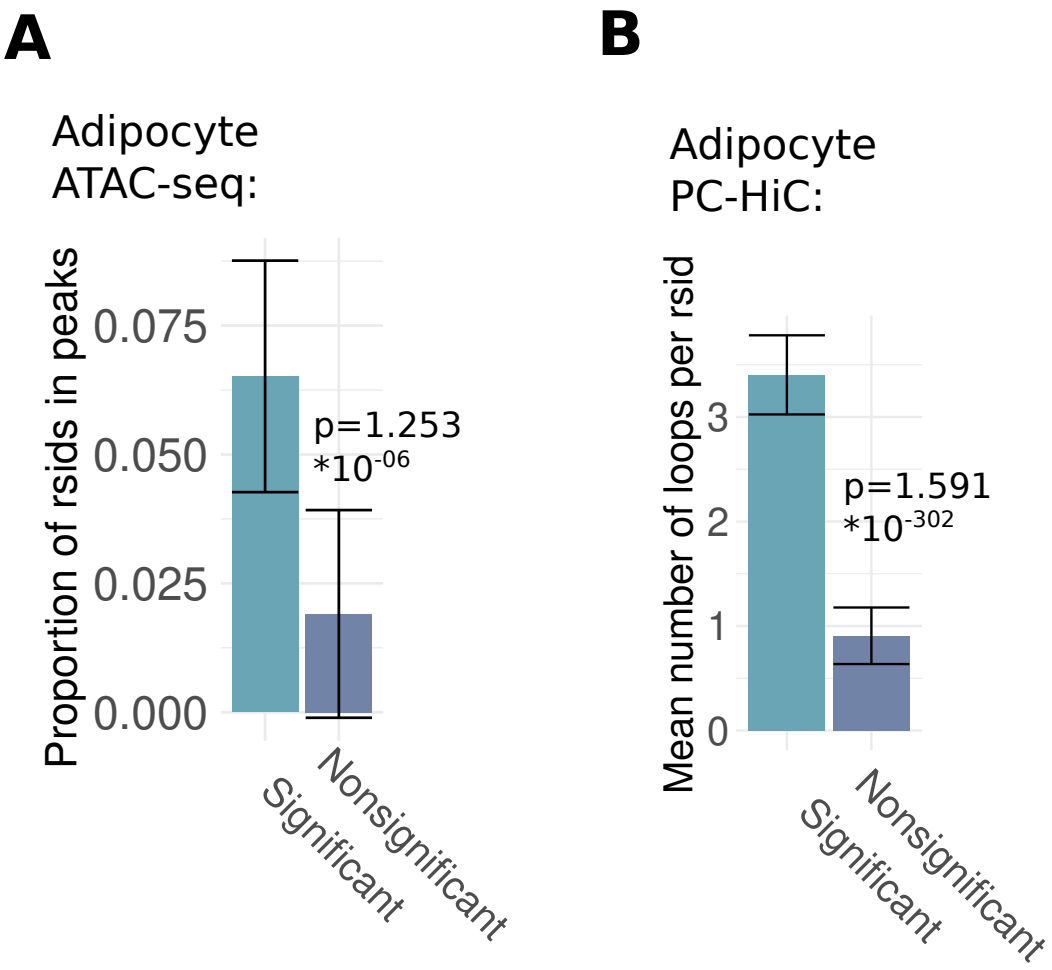

Figure 2.12: **Enrichment of significant MPRA enhancers in open chromatin and enhancer-promoter interactions.** a) y axis represents proportion of variants in ATAC-seq peaks across adipocyte differentiation (preadipocytes, day 2, day 8, day 14); x axis represents variant category (significant/nonsignificant). b) y axis represents number of loop ends overlapping variant across adipocyte differentiation; x axis represents variant category. P-values represent chi-square testing of observed values to values expected by chance.

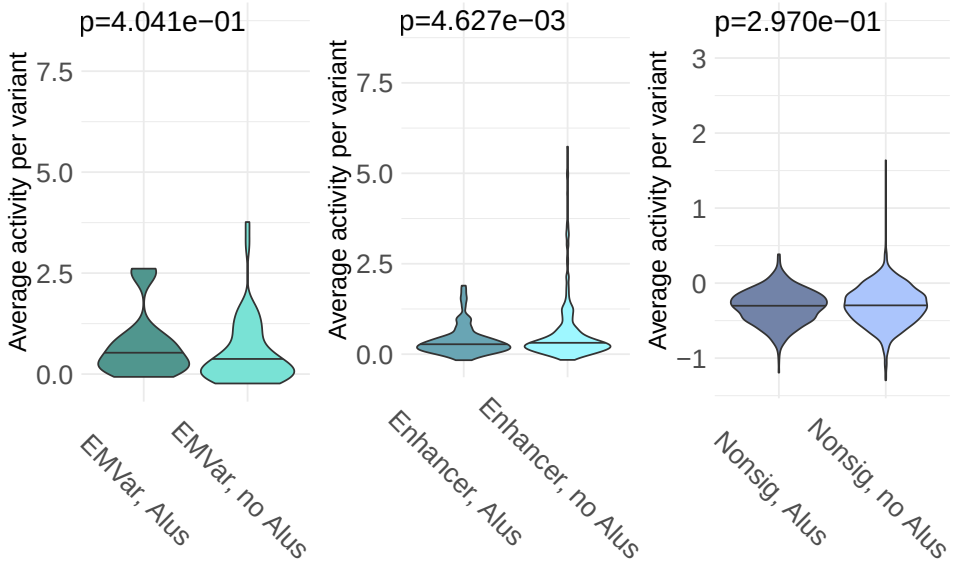

Figure 2.13: **Expression of MPRA Alus, Ulirsch et al.** Variant set from an MPRA conducted in erythroid cells to identify enhancers regulating blood cell traits. Variants separated by significance and by whether variant lies in an *Alu* element. y-axis represents normalized expression. P-values represent pairwise t-tests between normalized activity, averaged across all barcodes corresponding to each variant in condition and significance group.

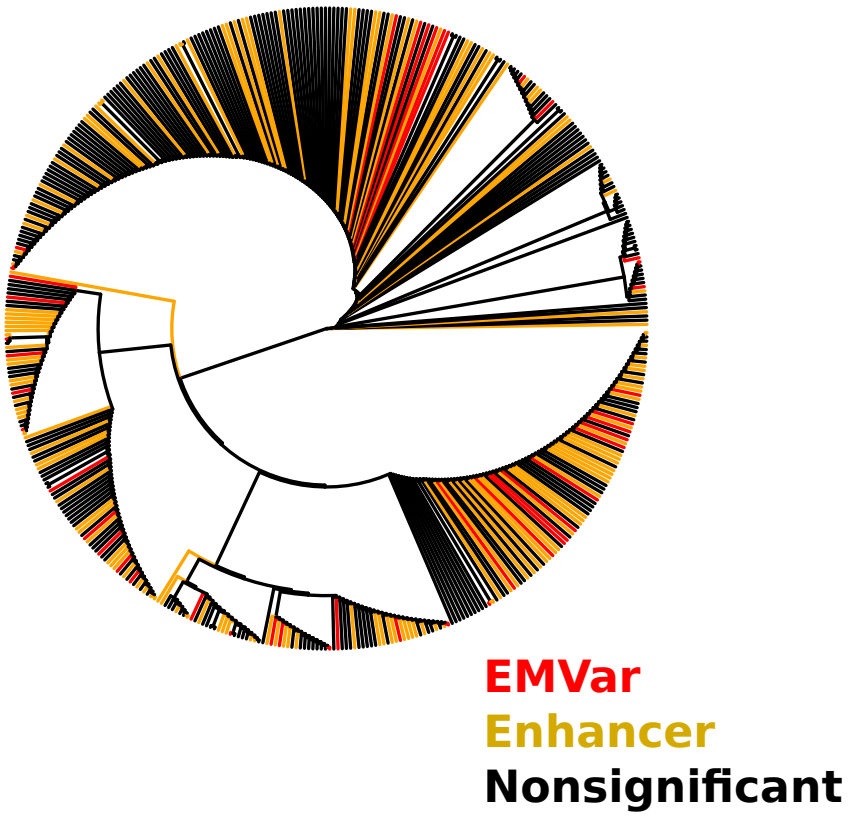

Figure 2.14: **Multiple sequence alignment of *Alu* elements overlapping MPRA variants.** Cladogram represents guide tree generated from multiple sequence alignment of all *Alu* elements overlapping an MPRA variant. Leaves are colored according to MPRA variant significance. Sequences included are all unique *Alu* elements containing MPRA variants (n=478).

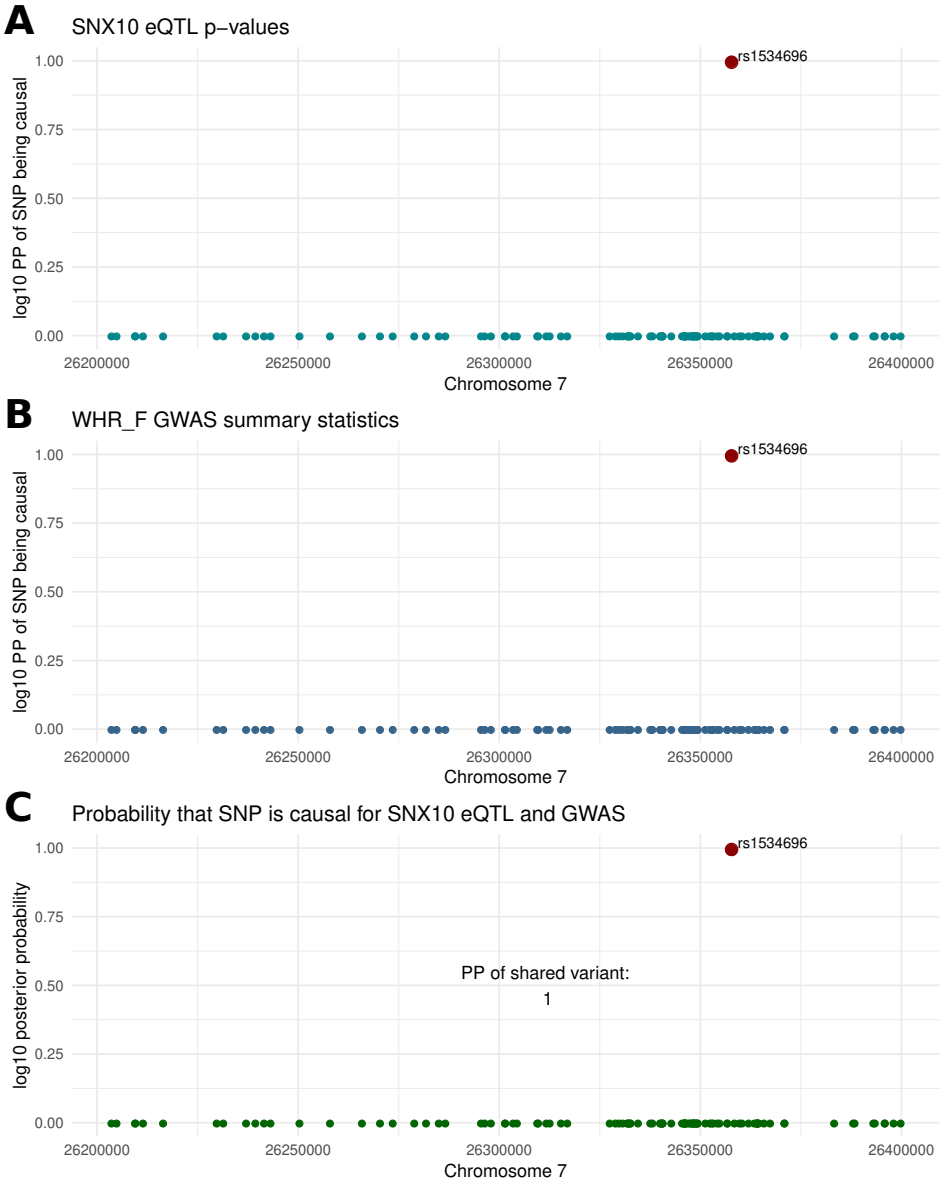

Figure 2.15: **Colocalization of female WHRadjBMI GWAS and *SNX10* eQTL signal at rs1534696.** a) posterior probability from fine-mapping of causality for eQTL effect for *SNX10* expression in GTEx subcutaneous adipose samples. b) posterior probability of GWAS effect for female WHRadjBMI. c) prior probability of shared causal variant for *SNX10* eQTL and female WHRadjBMI GWAS signal.

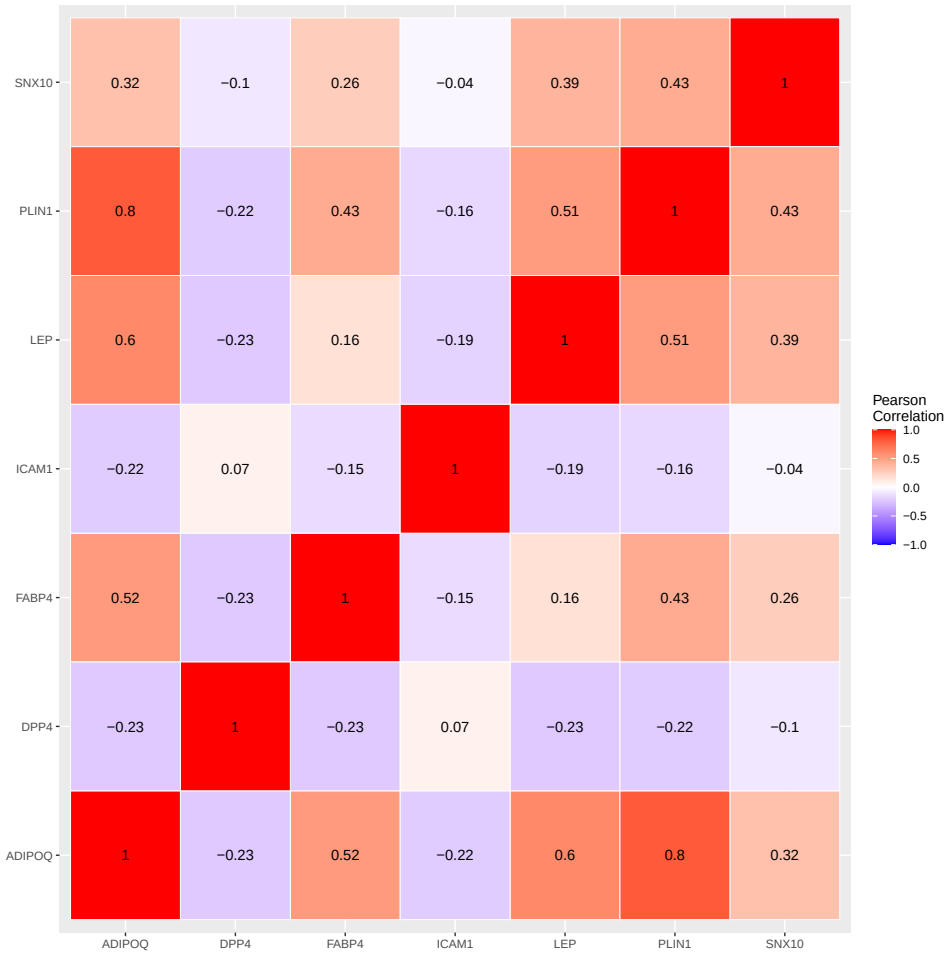

Figure 2.16: **Correlation of *SNX10* expression with adipocyte markers.** Numbers represent Pearson's r values. Genes included are mature adipocyte markers *PLIN1*, *LEP*, *FABP4*, and *ADIPOQ*, adipocyte progenitor marker *ICAM1*, and multipotent progenitor marker *DPP4*. Samples are all those included in GTEx v8 (n=17382).

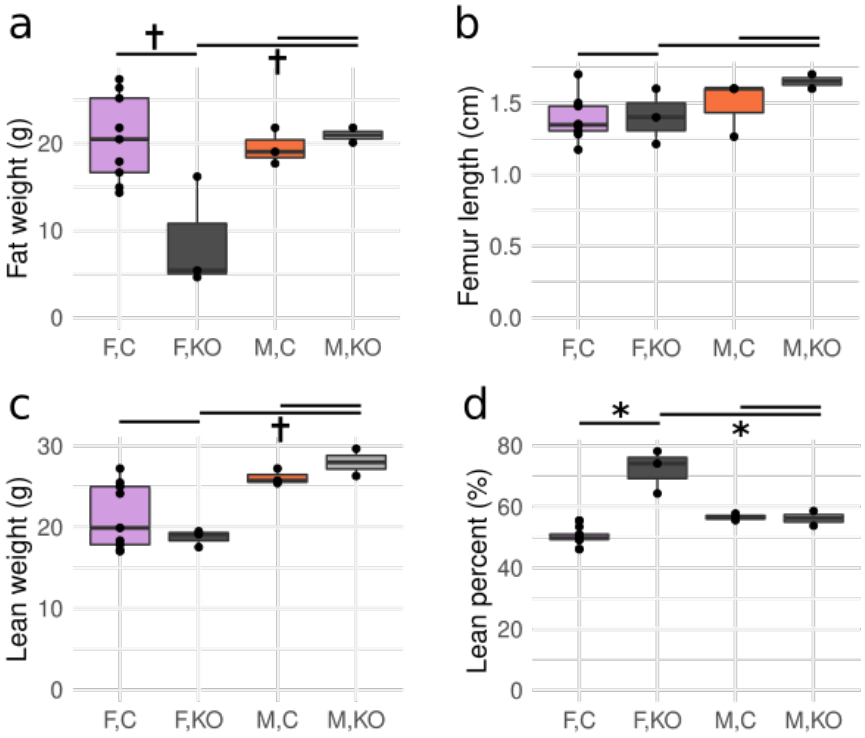

Figure 2.17: **Metabolic and anthropometric attributes of mmΔSnx10^{Adipoq} mice and control littermates.** Asterisk indicates p<0.05, cross indicates p < 0.10. All statistical tests are pairwise t-tests between sex-genotype groups. F-WT N=8, F-KO N=3, M-WT N=3, M-KO N=2. a) Mass of body fat of mmΔSnx10^{Adipoq} mice and littermate controls by sex. b) Femur length of mmΔSnx10^{Adipoq} mice and littermate controls by sex. c) Lean mass of mmΔSnx10^{Adipoq} mice and littermate controls by sex. d) Lean percent of mmΔSnx10^{Adipoq} mice and littermate controls by sex.

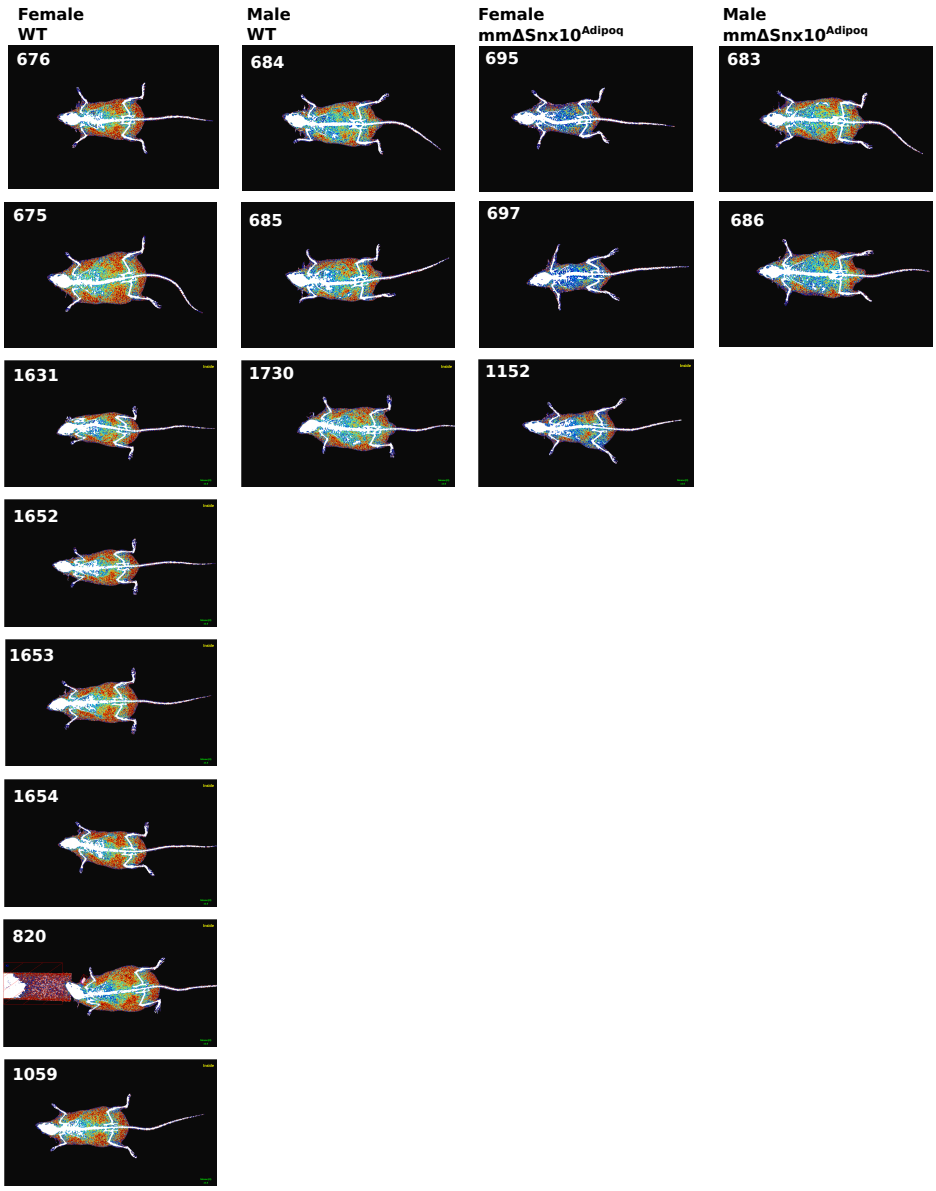

Figure 2.18: **DEXA scans of mmΔSnx10^{Adipoq} mice and control littermates.** DEXA composition scans of all mice after high-fat diet. Scans are organized by genotype status and sex. F-WT N=8, F-KO N=3, M-WT N=3, M-KO N=2.

2.7 Appendix B: Supplementary Tables

Transcription factor	**Consensus**	**p-value**	**q-value (Benjamini)**	**% Target with Motif**	**% Background with Motif**	**Known role in adipose**
Usf2	GTCACGTGGT	1.00E-12	0	13.40%	2.58%	Lipid metabolism
ZNF711	AGGCCTAG	1.00E-06	0.0002	28.09%	16.03%	None
COUP-TFII	AGRGGTCA	1.00E-05	0.0004	34.02%	21.51%	Adipogenesis
Sp5	RGKGGGCGGAGC	1.00E-05	0.0019	14.95%	7.12%	None
EAR2	NRBCARRGGTCA	1.00E-05	0.0005	31.19%	19.29%	Adipogenesis
Npas4	NHRTCACGACDN	1.00E-05	0.0014	18.56%	9.74%	None
KLF14	RGKGGGCGKGGC	1.00E-04	0.0036	27.58%	17.43%	Adipogenesis
LXRE	RGGTTACTANA-GGTCA	1.00E-04	0.0036	9.79%	3.86%	Adipogenesis
THRb	TRAGGTCA	1.00E-04	0.0043	47.94%	36.17%	Adipogenesis
Sp2	YGGCCCCGCCCC	1.00E-04	0.0044	17.78%	9.83%	Lipid metabolism
RARa	TTGAMCTTTG	1.00E-04	0.0056	52.06%	40.62%	Adipogenesis
CEBP	TNATGCAAYMN-NHTGMAAY	1.00E-03	0.0366	3.61%	0.82%	Adipogenesis
Erra	CAAAGGTCAG	1.00E-03	0.0392	33.51%	24.35%	Lipid metabolism

Table 2.1: Motif enrichment results, enhancer sequences versus nonsignificant sequences.

Transcription factor	**Consensus**	**p-value**	**q-value (Benjamini)**	**% Target with Motif**	**% Background with Motif**	**Known role in adipose**
COUP-TFII	AGRGGTCA	1.00E-08	0	62.07%	24.64%	Adipogenesis
Usf2	GTCACGTGGT	1.00E-08	0	25.86%	3.47%	Lipid Metabolism
EAR2	NRBCARRGGTCA	1.00E-08	0	60.34%	22.70%	Adipogenesis
RARa	TTGAMCTTTG	1.00E-05	0.0011	75.86%	45.71%	Adipogenesis
THRb	TRAGGTCA	1.00E-04	0.0077	67.24%	39.19%	Adipogenesis
LXRE	RGGTTACTANA-GGTCA	1.00E-03	0.0301	20.69%	5.52%	Adipogenesis
Erra	CAAAGGTCAG	1.00E-03	0.0433	51.72%	28.37%	Lipid Metabolism
Atf4	MTGATGCAAT	1.00E-03	0.0371	12.07%	2.01%	Adipogenesis
CRX	GCTAATCC	1.00E-03	0.0371	68.97%	45.32%	None
GSC	RGGATTAR	1.00E-03	0.0371	65.52%	42.03%	None
Atf1	GATGACGTCA	1.00E-03	0.0371	31.03%	12.85%	Adipogenesis
Nkx3.1	AAGCACTTAA	1.00E-03	0.0371	60.34%	37.12%	None
Pitx1	TAATCCCN	1.00E-03	0.0371	81.03%	59.17%	None
Npas4	NHRTCACGACDN	1.00E-03	0.0429	27.59%	11.06%	None

Table 2.2: Motif enrichment results, EMVar sequences versus nonsignificant sequences.

CHAPTER 3

DISCUSSION

3.1 Principal findings

In this work, we identified genes whose functions may contribute to the maintenance of healthy weight and the prevention of obesity-associated disease. We identified a set of enhancers that regulates waist-to-hip ratio across the genome, which is defined by a common motif which regulates adipogenesis. Many of these enhancers lie within *Alu* repeat elements. Our data indicate that impaired differentiation of fat cells from adipocyte precursors is central to high waist-to-hip ratio. *SNX10*, the strongest waist-to-hip ratio-associated novel gene that we identified, prevents the maturation of adipocytes, and also prevents expansion of fat in response to diet. The lead variant in the *SNX10* locus, rs1534696, regulates *SNX10* expression and affects lipid traits in adipocytes in humans. Interestingly, *SNX10* appears to only regulate fat accumulation in females, which is consistent with the female-specific significance of this locus in waist-to-hip GWAS.

3.1.1 Regulation of metabolism through adipocyte differentiation

Although obesity is a disorder defined by caloric excess, obesity-associated disease is most likely to emerge when there is a defect in fat accumulation. Lotta et al. find that polygenic score based on the 53 insulin resistanace risk loci predicts the development of familial partial lipodystrophy type 1, a disorder characterized by the inability to store subcutaneous fat [57], and further show that individual genes at these loci, in particular *CCDC92*, *DNAH10* and *L3MBTL3*, inhibit the differentiation of adipocytes from mesenchymal progenitor cells.

Upon caloric excess, fat tissue can either expand by increasing the number of fat cells (hypertrophy) or increasing cell size (hyperplasia). When adipogenesis is impaired, hypertrophic adipocytes predominate. Hypertrophic adipocytes promote the development of

obesity-associated disease in a number of ways. Hypertrophic adipocytes secrete more inflammatory markers and fewer anti-inflammatory markers, encouraging inflammation-associated systemic insulin resistance[81]. Hypertrophic adipocytes also secrete fatty acids, resulting in ectopic fat accumulation on visceral organs[50]. This results in disease directly, as in fatty liver disease, and indirectly, by promoting further insulin resistance and inflammation. Hypertrophic adipocytes also move GLUT4 receptors from the cell surface to the cytoplasm[20]. This limits their ability to respond to accept glucose in response to insulin, thus limiting lipogenesis and accelerating insulin resistance.

For many years, researchers in metabolism have suggested that fat depots in different parts of the body may have different adipogenic capacities. In 2013, Wang et al. administered a high-fat diet(HFD) to male mice, allowed their adipose tissue to expand, then examined adipose tissue from the perigonadal visceral adipose tissue to subcutaneous adipose tissue[97]. In subcutaneous adipose tissue, almost no new adipocytes were formed, and the depot expanded through hypertrophy. In contrast, in visceral adipose tissue HFD triggered adipogenesis, and new fat cells were formed. Wang et al., also showed that while subcutaneous fat appears to form in prenatally, visceral fat adipogenesis occurs after birth, further suggesting that the adipogenesis is regulated differently based on the location of adipocytes.

Later research shows that there may also be sex differences in adipogenesis between different fat depots: in female mice, both visceral and subcutaneous fat depots experience adipogenesis upon HFD administration[47]. Thus, differences in fat distribution may be explained by differences in the adipogenic capacity of different fat stores in the body. Similarly, sex differences in fat distribution may be explained by differences in adipogenic capacity of fat stores in men and in women.

Here, we find that a previously undescribed gene, *SNX10*, is required for adipocyte differentiation. *SNX10* does not act globally in adipocytes: rather, its effects are restricted

by depot, differentiation time point, and sex. *SNX10* expression correlates with markers of mature adipocytes, but not with adipocyte precursors or progenitors. Allele dosage of rs1534696 affects lipids in mature adipocytes, but not at earlier stages of differentiation. rs1534696 allele dosage changes lipid features in subcutaneous human adipocytes, but has no significant effects in visceral human adipocytes. Finally, *SNX10* knockout prevents fat tissue expansion after HFD in females, but not in males.

If adipogenesis is controlled differently between subcutaneous and visceral depots, and occurs at different times during development in these depots *in vivo*, then there are likely to be distinct gene regulators that differentiate adipogenesis in these two depots. rs1534696 may participate in controlling adipogenesis in subcutaneous adipocytes, but not in visceral adipocytes. If this is true, then this may explain sex-specific effects of *SNX10* on adipose tissue expansion: if *SNX10* only participates in adipogenesis in subcutaneous adipocytes, and subcutaneous adipocytes only differentiate in response to HFD in female mice, then the effect of *SNX10* in adipose tissue is likely to be specific to females.

More experiments are necessary to determine whether *SNX10*'s effect on adipogenesis is depot-specific, whether rs1534696 mediates depot specificity, and whether depot-specific effects on adipogenesis mediate the sex-specific effect of *SNX10* on adipose tissue expansion. To address whether *SNX10* expression is required for adipogenesis in either depot, we could perform an experiment using primary human mesenchymal stem cells from subcutaneous and visceral depots. We would perform shRNA-mediated knockdown of *SNX10* in these two stem cell types, then differentiate them into adipocytes. By evaluating whether *SNX10* knockdown impaired adipogenesis in both cell types, we could determine whether its effects are depot-specific, or shared between both kinds of adipocyte.

An *in vivo* investigation of the effects of *SNX10* on adipogenesis could help evaluate whether *SNX10* has an effect on adipogenesis in different fat depots, as well as allow us to evaluate whether any depot-specific effects of *SNX10* knockout relate to sex. By perform-

ing a bromodeoxyuridine (BrdU) pulse-chase experiment as described in Jeffery et al., we could determine which fat depots experience adipogenesis in mmΔSnx10^{Adipoq} mice after HFD. If both male and female mmΔSnx10^{Adipoq} mice experience adipogenesis in the visceral adipose tissue upon HFD, but only female mice experience subcutaneous adipogenesis after HFD, then this suggests that the effects of *SNX10* may have to do with sex-specific adipogenesis. An *in vivo* analysis would also allow us to examine the complexities of fat location, adipogenesis, and the effects of *SNX10*. While fat storage is often simplified into two depots (subcutaneous and visceral), in reality subcutaneous and visceral fat both consist of subtypes which differ based on bodily location and which exhibit different properties[94]. By performing experiments in a whole animal, we will be able to examine the behavior of fat tissue from more parts of the body.

3.1.2 Repeat elements as a substrate for evolution

The transcriptome-wide association studies performed here used statistical relationships between genotype at single-nucleotide polymorphisms and the expression of nearby genes to identify genes associated with waist-to-hip ratio. For each of these loci, genetic linkage limits our knowledge of which variants actually disrupt regulators of the target gene. Nearby variants which are co-inherited with the causal variant will share the same statistical relationship with target gene expression.

To identify which variants lie within enhancers at female waist-to-hip ratio loci, and may directly regulate target genes in adipose tissue, we chose to perform a massively parallel reporter assay. By evaluating all variants in linkage with candidate regulatory variants at female WHRadjBMI gene loci, we identified 59 enhancers with allele-sensitivity active in preadipocytes. Given that these variants 1) lie within active enhancers in fat cells, and 2) are strongly associated with female WHRadjBMI in GWAS, a close investigation of how these variants may modulate gene expression in fat cells is warranted.

To identify features that these variants may have in common, we performed transcription factor motif analysis, comparing the 175 base pairs surrounding each allele-sensitive enhancer with the 175 base pairs surrounding variants that were not enhancers. We identified a specific sequence, 'AGGTCA', that is enriched near allele-sensitive enhancer variant. This sequence is central to the binding sites of many transcription factors that regulate adipogenesis. It is also not brought to these regions by chance: rather, this sequence is most commonly located within the first monomer of *Alu* repeat elements.

Repeat elements make up about 45% of the genome. They are defined by their massive copy number: some repeat element families have hundreds of thousands of members, scattered throughout the genome by a long evolutionary history of transposition and reinsertion. *Alu* elements are one such large family, and themselves represent 9% of the genome. Repeat elements are the remnants of ancient viral infections, and potentially of pathogenic self-replicating processes more ancient than viruses. Paradoxically, the product of this massive infection burden, a tremendous copy number of near-identical elements, has given genomes the raw substrate to form complex and dynamic gene regulatory systems. In this way, repeat elements are crucial to the creation of cellular diversity, and the evolution of new cellular functions.

Systems of gene regulation formed by repeat elements control many human traits. A compelling recent example, published by Chuong and Elde in 2016, demonstrated how interferon-responsive elements have been dispersed throughout the human genome by mouse endogeneous retrovirus 41B (MER41B)[22]. MER41B contains a transcription factor binding site that is responsive to interferon. At sites where MER41B has been transposed, interferon drives gene expression, creating a gene regulatory network unique to the human lineage. This system is an example of 'copy-and-paste' gene regulatory networks formed by repeat elements: along the evolutionary lineage, MER41B was transposed to new gene loci, adding them to a pre-existing gene regulatory network (Figure 3.1A).

The 'amplify-then-mutate' approach is another way that repeat elements form gene regulatory networks during evolution[49]. In contrast to the 'copy-and-paste' model, where identical copies of a repeat element are transposed to new loci, the 'amplify-then-mutate' model takes advantage of the massive copy number of repeat elements left over from previous transpositions (Figure 3.1B). In this model, latent, imperfect binding sites are mutated, creating new transcription factor binding sites. In this model, repeat elements form the substrate for evolution of the binding sites. The gene regulatory network described here is an example of this form of evolution. *Alu* elements near genes which regulate adipogenesis have 'AGGTCA' located within the first monomer, two base pairs from the ancestral 'AGCCCA' sequence. Along the human lineage, mutations which form an enhancer active in differentiating adipocytes have occured at these loci, and because these mutations are advantageous, they have been kept.

Alu sequences contain short sequences that are a few base pairs away from canonical transcription factor binding motifs[69], and *Alu* elements have been shown to bind to some of these transcription factors, including *GATA3* and *PITX2*. Thus, *Alu* elements may be involved in the regulation of other human traits, such as hematopoeitic cell differentiation and heart development. In the case of adipogenesis, the involvement of *Alu* elements is suggested based on the enrichment of *Alus* within our identified MPRA enhancers, but the role of *Alu* elements in adipogenesis is not yet understood genome-wide. To evaluate this, the set of all *Alu* elements with the adipogenesis-associated motif in the vicinity of genes differentially expressed in adipogenesis should first be identified. We could then evaluate whether these Alus are enriched in the neighborhood of differentially expressed genes relative to all genes. Finally, to experimentally confirm the involvement of selected *Alu* elements in adipogenesis-related gene expression, CRISPR knockout studies could be performed, evaluating the expression of genes prior to and after removal of a candidate regulatory *Alu* element.

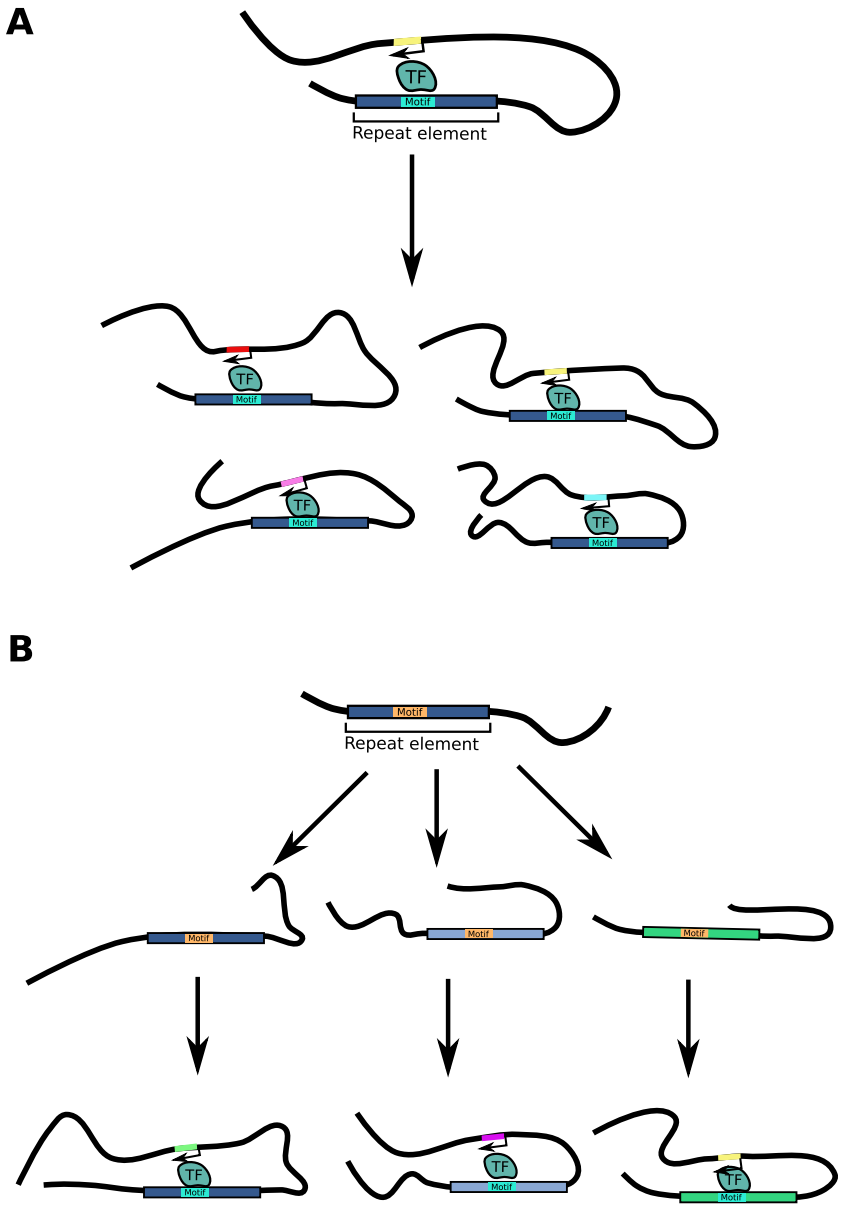

Figure 3.1: **Gene regulatory network formation by repeat elements: copy-and-paste versus amplify-then-mutate.** a) Copy-and-paste gene regulatory network formation: a repeat element which contains a functional transcription factor binding site is transposed to new loci, where neighboring genes will now be controlled by the same transcription factor. b) Amplify-then-mutate gene regulatory network formation: a repeat element with a motif similar to a transcription factor binding site is transposed to many loci throughout the genome. Mutations within the proto-motif convert it to an transcription factor binding site, adding neighboring genes to the regulatory network.

The origins of gene regulatory networks in ancient viral infections are interesting, but it is not immediately apparent what this avenue has to offer for human health. However, if we understand the evolutionary principles driving gene network formation, we may someday understand the fundamental logic of these systems, and anticipate certain features common to many genetic systems. These features will have implications for any gene therapy designed to treat human disease. For example, repeat-element constructed gene regulatory networks exhibit widespread redundancy: many repeat element copies throughout the genome contain the necessary sequence bind a transcription factor. A disadvantage of this, from a therapeutic perspective, is that precise gene editing therapies will be difficult to design. In contrast, an advantage is that a gene therapy targeted towards a specific short sequence may be able to efficiently target a whole gene regulatory network active in a cell type of interest.

3.1.3 Emergence of sexually dimorphic traits from gene regulation common to both sexes

Our findings in this project illuminate a vexing problem in genetics: namely, how sexually dimorphic traits emerge from gene regulation that is common to both sexes. Sexual dimorphism is present in many human traits, from height and weight to lifespan and disease burden[102, 87]. Discovering the gene regulatory underpinnings of these large phenotypic differences has been a challenge. A recent paper published by the Genotype-Tissue-Expression consortium [65] showed that 37% of genes have sex-biased expression, most only in a few tissues. However, this sex-biased expression is not driven by sex-biased effect of allelic variation. Of the ∼500,000 cis-eQTLs identified by GTEx, only 369 had effects that were biased by sex. Sexually dimorphic gene expression, therefore, may be driven by other factors than *cis*-regulated gene expression, such as the extracellular hormonal signalling.

Our study supports this finding: in our data the eQTLs associated with female WHRadjBMI are strong eQTLs in both sexes (Figure 2.2D). However, their association with WHRad-

jBMI is much stronger in females than in males. The variant we chose to experimentally interrogate, rs1534696, is an example of this: it is a strong eQTL for its target gene, *SNX10*, in both sexes (Figur 2.2C). However, the GWAS signal at this locus shows a strong relationship with WHRadjBMI in women, and no relationship in men (Figure 2.2E). This suggests that genetically-mediated *SNX10* expression plays a role in WHRadjBMI in women, but it not involved in WHRadjBMI regulation in men.

As biologists attempting to understand sexually dimorphic phenotypes, our goals are to 1) understand how the phenotype emerges from cellular processes, and 2) understand how sexual dimorphism can be introduced to those cellular processes. With regard to the first goal, our and other research indicates that adipocyte differentiation is central to the control of waist-to-hip ratio [57, 55]. We know that *SNX10* is required for adipocyte differentiation in females, but not in males. Therefore, the task that lies before us is to understand why the gene regulatory network that controls fat cell maturation includes *SNX10* in female cells, but not in male cells.

Sexual differences in cellular functions could emerge on one of two ways. The first is indirect effects of chromosomal differences through circulating gonadal hormones. The second is direct effects of chromosomal differences, i.e. different gene expression of X and Y chromosome genes. It is well established that gonadal hormones affect fat metabolism: women's WHR changes to favor abdominal obesity after menopause, and men who lose testosterone production gain fat in a more 'gynoid' pattern in gluteofemoral subcutaneous depots[104, 74]. Estrogen receptors α and β are both present on adipocytes[24], suggesting that these effects may be mediated by estrogen signalling in adipose tissue. Overall; estrogen appears to lower inflammatory signalling and promote adipogenesis[28, 72]. Adipocyte-specific knockout of ERα increases adipose fibrosis and inflammation in mice, whereas ERβ seems to have the opposite effect [24]. Therefore, gonadal hormones play a role in adipocyte metabolism and metabolic health. However, elegant studies differentiating the effect of XX/XY karyotype

from gonadal hormones during development in mice have shown that chromosome dosage also affects metabolic phenotypes: gonadectimized mice with XX chromosomes, regardless of whether they had testes or ovaries in development, had up to 2-fold increased adiposity in comparison to XY mice[19]. These mice also have differences in miRNA abundance in adipose tissue dependent on chromosomal sex, suggesting that at least some of the differences in adiposity could be mediated locally in adipose tissue[53].

How does biological sex affect the outcome of *SNX10* depletion in fat cells? We know that it could be through sex chromosome effects, hormonal effects, or both. The studies by Jeffery et al. examining depot-specific adipogenesis in male and female mice suggest that estrogen may be responsible for determining many of the sex differences in adipogenesis: when female mice were given ovariectomy, adipogenesis only occured in visceral adipose tissue, rather than occuring in both visceral and subcutaneous depots[47]. Conversely, when administered estrogen, male mice experienced adipogenesis in both visceral and subcutaneous adipose depot, consistent with the pattern seen in female mice.

If estrogen determines the female pattern of adipogenesis, how does it mediate its effects? Estrogen is known to directly affect adipogenesis: studies have shown that estrogen increases expression of the adipgenesis-driving transcription factor $CEBP\alpha$[28]. However, our study illustrates an example of a different model of female-specific gene. *SNX10* has gene expression that does not appear dependent on estrogen, as the expression of *SNX10* is consistent between women of pre- and post-menopausal ages. In the case of these genes, estrogen may create a cellular environment that facilitates adipogenesis, setting a stage upon which *SNX10* may act. In this way, these genes may promote adipogenesis only in females, even though their expression is not determined by chromosomal sex or hormonal status. More research is needed to determine whether and how estrogen creates a pro-adipogenic environment in adipose tissue and genes that regulate metabolism in females participate in this process.

SNX10 is not the only gene that has been found to play a female-specific role in adipo-

genesis and human metabolic disorders. In 2018, Small et al. described the female-specific actions of *KLF14* [83]. *KLF14* is a transcription factor with more than 300 downstream targets. The risk allele at *KLF14* is predicted to reduce *KLF14* expression, which shifts fat storage from gluteofemoral to abdominal deposits. This is the same direction of the effect of *SNX10*, where decreased expression in mature subcutaneous adipocytes due to C allele dosage at rs1534696 is predicted to impair subcutaneous adipogenesis and favor visceral adipose expansion. However, *KLF14* and *SNX10* differ in important ways. *KLF14* is a master transcription factor. Therefore, *KLF14* is expected to have large effects on cell differentiation through a cascade of downstream targets. In contrast, *SNX10* is a sorting nexin, *SNX10* likely does not directly disrupt the expression of other genes. Rather, the role of Rather, *SNX10* may disrupt the ability of adipocytes to transport GLUT4 into the cell. Further research is needed to understand why disrupting this critical function results in a broad phenotype of reduced expression of mature adipocyte marker genes. Broader characterization of *SNX10* shRNA knockdown adipocytes could help us to understand why these cells do not express mature adipocyte markers and cannot store intracellular lipids appropriately. Specifically, RNA-seq with subsequent pathway analyses may help us to understand the genome-wide consequences of *SNX10* knockdown, and further functional studies, such as insulin stimulation followed by Western blot investigation of downstream signalling (e.g. P-Akt/total Akt) could allow us to understand the impact of *SNX10* knockdown on adipocyte-relevant cellular processes.

Although the effect of *KLF14* in humans is female-specific, mice with a *KLF14* knockout specific to fat have metabolic abnormalities that are present in both sexes. In *SNX10* mice with the same conditional knockout, reduced adipose expansion is seen in female mice, but not in males. Small et al. do not mention whether the non-coding enhancer that regulates *KLF14* is conserved in mice; perhaps female specificity at the *KLF14* locus emerged more recently in human evolution than female specific effects of *SNX10*. In humans, fat

distribution is highly sex-specific relative to other animals[70]. However, most mammals have some sexual dimorphism in body fat distribution. Therefore, some female-specific human loci may be present in animal models, while others may not. The fact that the sexual dimorphism of the *SNX10* locus persists into mouse models means that *SNX10* could be a valuable model for experimental exploration of sexual dimorphism.

The biological mechanisms of sexual dimorphism, especially female-specific biology, have historically been underfunded and understudied. One justification has been that by studying female-specific effects you are studying only a portion of the population, and resulting insights will therefore not be widely applicable. Recently, biomedical funding agencies have appreciated the opposite position: that rather than limiting science, studying differences between humans is another route to potential generally applicable therapies. Investigating female-specific fat distribution offers important, unique avenues to understanding healthy fat storage. Understanding how some bodies store excess fat with fewer metabolic complications may help us develop therapies to protect overweight people from diabetes and cardiovascular disease. Additionally, investigating sex differences will help us to understand why existing therapies don't work effectively for all bodies. Transitions from childhood through puberty to adulthood, medical transitions between male and female genders, aging, and menopause are all characterized by dramatic shifts in hormonal complement. Understanding the effect these hormones and chromosomal sex have on health will help us to develop therapies that can be effective for all people, at all stages of life.

3.1.4 Human-specific patterns of fat distribution

As mentioned earlier, humans have more body fat and more sexually dimorphic body fat than other great apes. In a study conducted by Pontzer et al., the body fat percentage of human women and men was 41.1% and 22.9%, respectively. In comparison, the body fat percentage of female and male chimpanzees was 9.0% and and 8.4%. Orangutans and gorillas have

similarly low body fat percentage to chimpanzees. This difference was measured in sedentary populations of both humans ahd chimpanzees but the difference persists in physically fit humans: for example, the body fat percentage of men and women measured in a human subsistence gathering community was 20.9% and 13.5%, respectively, still substantially larger than sedentary chimpanzees.

Although the increased adiposity of humans is a concern for modern public health, our increased fat mass is intrinsic to the evolution of humans. In comparison to chimpanzees, the most closely related species to us, humans represent an 'energetic paradox': we have a lifespan that is double that of chimpanzees, we reproduce more frequently (in the absence of contraception), and we have metabolically costly brains[70]. To achieve this, humans have evolved a basal metabolic rate that is $\sim$20% higher than that of chimpanzees[70]. A high basal metabolic rate allows us to simultaneously allocate energy to our brains, our fertility, and our lifespans; however, it also requires increased caloric input. This suggests that as humans have developed more sophisticated ways to find food, our bodies have taken advantage of these increased resources to allow us live longer, have more children, and further expand our brain size. Human-specific and female-specific increases in fat mass are most likely a part of this evolutionary history: as humans have increased our metabolic rate, our need for a 'buffer' against food scarcity has increased: especially for women, who require reliable access to energy during pregnancy.

The genetic mechanisms of human-specific expansion of fat storage have not been found. *A priori*, we would expect to see human-specific sequences or human-specific genetic variants which are associated with changes in gene expression in fat. The enrichment of primate-specific *Alu* elements in our study suggests a possible mechanism that could create these human fat phenotypes. However, the adipogenesis-associated *Alu* elements that I identify in this study are not human-specific: 75% of *Alu* elements containing an adipogenesis-associated motif are also present in chimps.

Despite the fact that the adipogenesis-associated signal located in *Alu* elements found in this study is largely not human-specific, these *Alu* elements could nevertheless contribute to human-specific fat distribution. One explanation could be recruitment and subsequent optimization: *Alus* could have been recruited along the primate lineage to regulate fat, and further modified along the human lineage to regulate human-specific fat distribution. Alternatively, the 25% of *Alu* 'AGGTCA' motifs that are human-specific in our data could represent important changes to fat biology. To globally test this, we could perform comparative RNA-seq between human and chimp adipocytes to identify genes that are differentially expressed between the two species. In the neighborhood of these genes, we could identify all *Alu* sequences with an 'AGGTCA' element. By looking for sequence variation between the species within these *Alus*, we could evaluate the human specificity of the 'AGGTCA' motif globally, as well as determine whether there are potential human-specific regulatory motifs within the *Alus* that we missed with the smaller MPRA sample. Later studies where we edit or remove human-specific *Alu* features and evaluate the effect on adipocyte differentiation and function could further support the role of any identified elements in the transition to human patterns of fat distribution.

3.1.5 Pleiotropy and complex gene regulation at SNX10

Our study suggests that the C allele of the non-coding variant rs1534696 increases risk for metabolic disorders by disrupting an enhancer that drives *SNX10* expression. This enhancer appears to regulate *SNX10* in a manner that is cell-type and time-point specific: our ABC results suggest that rs1534696 only physically interacts with the *SNX10* promoter in mature adipocytes, and our Lipocyte Profiler results suggest that rs1534696 allele only affects adipocyte features in adipocytes that are a) mature, and b) from subcutaneous adipose tissue.

SNX10 is not a gene whose expression and function is unique to fat. Indeed, our col-

laborators have described how *SNX10* expression is required for normal function of both osteoclasts and gastric zymogenic cells. In osteoclasts, *Snx10* knockout interferes with bone resorption; in gastric zymogenic cells, *Snx10* knockout interferes with potassium absorption. Both of these functions require acidification of the extracellular environment, which is accomplished by vesicular trafficking to the cell surface. Therefore, *SNX10*'s role in intracellular signalling may be affecting the function of a number of cell types, including adipocytes, osteoclasts, and gastric cells, with very different phenotypes dependent on the affected cell type.

By knocking out *SNX10* expression specifically in mature fat cells, we replicate the *SNX10* phenotype associated with the enhancer at rs1534696. This suggests that this enhancer controls *SNX10* expression in mature adipocytes, with implications for fat distribution. Other enhancers that are active in other cell types may control the expression of *SNX10* in osteoclasts, gastric zymogenic cells, and other cell types. In this way, one gene can serve to accomplish a variety of functions and be involved in many human traits. Trait- and cell-type-specific enhancers can help us develop therapies that may improve symptoms of one disorder without undesirable side-effects.

Although our research has focused on rs1534696, this may not be the only enhancer that regulates *SNX10* expression in adipocytes. In 2019, Castillejo-Lopez et al. used CAGE sequencing to identify enhancer RNAs transcribed during adipocyte differentiation. In this way, they described an enhancer 600kb upstream from *SNX10*, in a distinct haplotype block from rs1534696[17]. This enhancer is located in a GWAS locus that is associated with both WHRadjBMI and with body mass index (BMI). This is in contrast to rs1534696, which is only associated with female WHRadjBMI. The deletion of this enhancer reduces adipogenesis of SGBS preadipocytes. The authors associate this enhancer to a *trans*-gene, *CBI3L1*. However, based on the proximity of this enhancer to *SNX10* and the role of *SNX10* in adipogenesis described in this paper, we believe that it is likely that this enhancer may

primarily exerting its effect through action on *SNX10*. If so, it is likely that this enhancer regulates the expression of *SNX10* in other contexts than does rs1534696. Based on the fact that this enhancer lies in a locus associated with both WHRadjBMI and with obesity, perhaps this enhancer has a more broad regulatory profile. Regulation of *SNX10* in adipocytes in all fat depots to varying degrees, or regulation of *SNX10* in both brain and adipose tissue, could explain the association of this enhancer with both obesity and WHRadjBMI.

3.1.6 Limitations of current study

Our approach to identifying genes and regulatory elements associated with metabolic disorder is necessarily limited. In this section, I will describe the major limitations associated with this project. Some limitations can be resolved through further statistical analyses and further experiments; others cannot be resolved, but are described so that the results can be appropriately interpreted.

Transcriptome-wide association studies take advantage of the statistical association between the allele of non-coding genetic variants and gene expression to identify genes whose expression is associated with traits of interest. These studies are vulnerable to false positives in situations where a genetic variant regulates the expression of two genes, but only one of these genes is associated with the trait of interest. Indeed, even if a non-causal gene shares no regulatory variants with a causal gene, but has *cis*-regulatory variants in linkage to variants that regulate a causal gene, this can cause false positive identification[95]. This vulnerability is present in our data: for example, at the *FTO* locus *FTO* is identified as a gene whose expression is associated with obesity. Previous research from our lab indicates that the causal variant at the *FTO* locus acts through long-range regulation of *IRX3*, not *FTO*[84]. However, weak eQTL association with *FTO* regulation combined with a strong association of the locus with obesity results in identification of *FTO* expression as associated with obesity. Fine-mapping approaches have been developed which reduce the number of

false-positive gene identifications in TWAS[60], but in cases of close linkage and co-regulation of genes this remains a fundamental limitation that must be overcome with experimental follow-up analyses.

Our massively parallel reporter assay was designed with the intent to identify regulatory elements which drive expression of female WHRadjBMI genes in adipocytes and potentially explain their association with WHRadjBMI. As adipocytes mature and accumulate lipids, transfection of material into the cell becomes progressively difficult. Therefore, we chose to perform the MPRA in pre-adipocytes. Our results, which highlight the role of adipogenesis in gene regulation at WHRadjBMI-associated loci, are consistent with previous research describing the role of adipogenesis in WHRadjBMI[57]. However, an MPRA performed in mature adipocytes may reveal gene regulation that regulates later stages of adipogenesis or that regulates other processes relevant to WHRadjBMI. For example, the causal variant at *SNX10* was not identified in our preadipocyte MPRA. This variant appears to regulate adipogenesis through disrupting the activity of *SNX10* late in adipocyte differentiation.

Our transcription factor motif enrichment analyses identified a number of transcription factors whose binding sites are enriched in MPRA enhancer-modifying variant sequences, compared to MPRA nonsignificant sequences. Interestingly, nine of these transcription factors are associated with adipogenesis and/or lipid metabolism. Eight of these nine transcription factors share the same core 'AGGTCA' element in their binding motif. This suggest that 'AGGTCA' is the sequence that is enriched in our data. Rather than all eight transcription factors being active, one, a few, or any combination of the identified transcription factors could actually be binding to these sites *in vivo*. Identifying which transcription factors are responsible for driving expression at female WHRadjBMI-associated loci could greatly help us to understand the biology of adipogenesis and fat distribution. However, as with the MPRA, we are limited by the difficulty of working with adipocytes. Subsequent attempts at ChIP-seq type assays in earlier stages of adipocyte differentiation, or with adipocyte nuclei,

may allow us to identify which transcription factors are important for fat distribution.

Lipocyte Profiler, an imaging and analysis pipeline, was used here to detect depot- and time course-specific changes in adipocyes dependent on rs1534696 allele. This approach offers great power to detect cellular changes of small scale, as would be expected from non-coding variants such as rs1534696. However, the information that Lipocyte Profiler provides is limited. The features identified by Lipocyte Profiler represent broad descriptions of the qualities of flourescent stain images: e.g. one significant feature of our data is a difference in the 'granularity' of BODIPY(intracellular) stain images between CC/CA and AA allele rs1534696 carriers. What granularity represents here–wether this represents a difference in the size of intracellular lipid vacuoles, the identity of synthesized fatty acids, or other properties–is unclear, and experimental follow-up is required to identify the cellular processes which underlie morphological changes identified with this approach.

In this study, we experimentally interrogated the function of a gene identified in the female WHRadjBMI TWAS, *SNX10*, in adipocytes in cellular culture and in adipose tissue *in vivo*. We determined that *SNX10* is required for adipocyte differentiation, and that *Snx10* is required for expansion of adipose tissue in response to high-fat diet, but only in female mice. We also determined that rs1534696, the lead causal variant at the *SNX10* locus, is an eQTL for *SNX10*, and is associated with intracellular lipid phenotypes, but only in mature subcutaneous adipocytes. Together, this information suggests that an enhancer at rs1534696 regulates *SNX10* in adipocytes in a fat depot- and developmental time point-specific manner that may modulate a woman's fat distribution.However, many portions of this story are incomplete. We know that the female mmΔSnx10^{Adipoq} mice do not gain adipose mass when fed a high-fat diet, but we do not know what the metabolic consequences are for these animals. In humans, limited adipogenesis is associated with metabolic dysregulation, including elevated blood glucose, insulin resistance, and elevated blood triglycerides. In the next phase of this project, we plan to evaluate these features in mmΔSnx10^{Adipoq} mice, and

evaluate whether *Snx10* knockout in mice leads to female-specific liabilities for metabolic disease.

Another open question is how the sex specific effects of *Snx10* knockout are mediated. As mentioned earlier in this discussion, one hypothesis is that estrogen may change the microenvironment of adipose tissue, allowing female-specific adipogenesis-associated genes such as *Snx10* and *Klf14* to promote adipocyte differentiation in subcutaneous adipose tissue. Future experiments introducing estrogen to male mmΔSnx10^{Adipoq} mice, and introducing estrogen antagonists to female mmΔSnx10^{Adipoq} mice may help us to understand the role *Snx10* and other genes have to play in sexually dimorphic body fat distribution.

3.2 Concluding remarks

In this work, we investigated the genetic risk associated with obesity and waist-to-hip ratio, with the goal of identifying novel genes and gene regulators that contribute to metabolic disease. In the process of identifying these elements, we learned critical new information about the regulation of fat distribution. Specifically, we idenfied a motif, 'AGGTCA', that is located within *Alu* repeat elements, is associated with adipogenic transcription factors, drives gene expression, and is located near female WHRadjBMI genes. We also identified a gene, *SNX10*, at a locus is strongly associated with female WHRadjBMI. Experimental interrogation of this gene identified that it is critical for adipocyte maturation and that adipocyte-specific knockout prevents fat accumulation, but only in females.

Examples of gene regulatory networks formed by repeat elements are accumulating in the genetic literature[59, 22, 49, 89]. As early as 1969, Britten and Davidson predicted that the transposable, highly repetitive, highly prevalent elements that characterized complex genomes could be used to form gene regulatory networks[7]. Regulatory network construction by repeat element networks is proving to be a feature of many human traits. Here, the outcome is that many genes that control waist-to-hip ratio respond to a common signal.

Going forward, we need research on how to exploit broad signals that define gene regulatory networks, as well as increased research on genetic features that provide locus-level specificity, such as histone modifications and transcription factor motif syntax. In this way, we may be able to create gene therapies in the future with high effectiveness and few off-target effects.

Our finding that *SNX10* is required for fat accumulation in females, but not in males, is the first experimental confirmation of a sex-specific GWAS locus. However, most of the work to understand *why* this gene plays a sex-specific role in adipocyte maturation remains to be done. Repeating shRNA-mediated knockout of *SNX10* in mesenchymal progenitor cells of both sexes in the presence of estrogen, estrogen receptor α antagonists, and control would allow us to determine whether estrogens or sex chromosome dosage determine *SNX10*'s role in adipocyte maturation.

The control of when, where, and how genes are expressed is essential for the formation of complex organisms. It is also central to every human disease state. We are just beginning to understand how genetic sequence, organismal signalling, and the environment controls gene expression. This dissertation contributes toward the knowledge of how sexually dimorphic states emerge, and do not emerge, as a product of differential gene expression, and how gene expression control can emerge across evolutionary time.

References

[1] Nikolaos Alexopoulos, Demosthenes Katritsis, and Paolo Raggi. Visceral adipose tissue as a source of inflammation and promoter of atherosclerosis. *Atherosclerosis*, 233(1):104–112.

[2] Maud Alligier, Laure Gabert, Emmanuelle Meugnier, Stéphanie Lambert-Porcheron, Emilie Chanseaume, Frank Pilleul, Cyrille Debard, Valérie Sauvinet, Béatrice Morio, Antonio Vidal-Puig, Hubert Vidal, and Martine Laville. Visceral fat accumulation during lipid overfeeding is related to subcutaneous adipose tissue characteristics in healthy men. *The Journal of Clinical Endocrinology & Metabolism*, 98(2):802–810.

[3] Dvir Aran, Zicheng Hu, and Atul J. Butte. xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biology*, 18(1):220–220. Publisher: BioMed Central.

[4] Longchuan Bai and Juanita L. Merchant. Transcription factor ZBP-89 cooperates with histone acetyltransferase p300 during butyrate activation of p21 waf1 transcription in human cells*. *Journal of Biological Chemistry*, 275(39):30725–30733.

[5] Richard N. Bergman, Stella P. Kim, Karyn J. Catalano, Isabel R. Hsu, Jenny D. Chiu, Morvarid Kabir, Katrin Hucking, and Marilyn Ader. Why visceral fat is bad: Mechanisms of the metabolic syndrome. *Obesity*, 14:16S–19S. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1038/oby.2006.277.

[6] H. Borgeraas. Association of time of obesity onset with comorbidities in treatment-seeking men and women with severe obesity. *Obesity Science & Practice*, 4(5):427–436.

[7] R. J. Britten and E. H. Davidson. Gene regulation for higher cells: a theory. *Science (New York, N.Y.)*, 165(3891):349–357.

[8] R. P. Brun, P. Tontonoz, B. M. Forman, R. Ellis, J. Chen, R. M. Evans, and B. M. Spiegelman. Differential activation of adipogenesis by multiple PPAR isoforms. *Genes & Development*, 10(8):974–984. Company: Cold Spring Harbor Laboratory Press Distributor: Cold Spring Harbor Laboratory Press Institution: Cold Spring Harbor Laboratory Press Label: Cold Spring Harbor Laboratory Press Publisher: Cold Spring Harbor Lab.

[9] Brendan Bulik-Sullivan, Hilary K Finucane, Verneri Anttila, Alexander Gusev, Felix R Day, Po-Ru Loh, ReproGen ReproGen Consortium, Psychiatric Genomics Psychiatric Genomics Consortium, Genetic Consortium for Anorexia Nervosa of the Wellcome Trust Case Control Consortium Genetic Consortium for Anorexia Nervosa of the Wellcome Trust Case Control Consortium 3, Laramie Duncan, John R B Perry, Nick Patterson, Elise B Robinson, Mark J Daly, Alkes L Price, and Benjamin M Neale. An atlas of genetic correlations across human diseases and traits. *Nature genetics*, 47(11):1236–41. Publisher: NIH Public Access.

[10] Brendan K. Bulik-Sullivan, Po-Ru Loh, Hilary K. Finucane, Stephan Ripke, Jian Yang, Nick Patterson, Mark J. Daly, Alkes L. Price, and Benjamin M. Neale. LD score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature Genetics*, 47(3):291–295. Number: 3 Publisher: Nature Publishing Group.

[11] Byrne Christopher D. and Targher Giovanni. Ectopic fat, insulin resistance, and nonalcoholic fatty liver disease. *Arteriosclerosis, Thrombosis, and Vascular Biology*, 34(6):1155–1161. Publisher: American Heart Association.

[12] Maren E. Cannon, Kevin W. Currin, Kristin L. Young, Hannah J. Perrin, Swarooparani Vadlamudi, Alexias Safi, Lingyun Song, Ying Wu, Martin Wabitsch, Markku Laakso, Gregory E. Crawford, and Karen L. Mohlke. Open chromatin profiling in adipose tissue marks genomic regions with functional roles in cardiometabolic traits. *G3: Genes/Genomes/Genetics*, 9(8):2521–2533.

[13] Dexter Canoy. Distribution of body fat and risk of coronary heart disease in men and women. *Current opinion in cardiology*, 23(6):591–8.

[14] Hong Cao, Xuanmao Chen, Yimei Yang, and Daniel R Storm. Disruption of type 3 adenylyl cyclase expression in the hypothalamus leads to obesity. *Integrative obesity and diabetes*, 2(2):225–228.

[15] Jun Cao, Yuhong Ma, Weiqi Yao, Xiaoye Zhang, and Dongcheng Wu. Retinoids regulate adipogenesis involving the TGFβ/SMAD and wnt/β-catenin pathways in human bone marrow mesenchymal stem cells. *International Journal of Molecular Sciences*, 18(4).

[16] Marta Casado, Virginie S. Vallet, Axel Kahn, and Sophie Vaulont. Essential role in vivo of upstream stimulatory factors for a normal dietary response of the fatty acid synthase gene in the liver. *Journal of Biological Chemistry*, 274(4):2009–2013. Publisher: American Society for Biochemistry and Molecular Biology.

[17] Casimiro Castillejo-Lopez, Milos Pjanic, Anna Chiara Pirona, Susanne Hetty, Martin Wabitsch, Claes Wadelius, Thomas Quertermous, Erik Arner, and Erik Ingelsson. Detailed functional characterization of a waist-hip ratio locus in 7p15.2 defines an enhancer controlling adipocyte differentiation. *iScience*, 20:42–59.

[18] Jenny C. Censin, Sanne A. E. Peters, Jonas Bovijn, Teresa Ferreira, Sara L. Pulit, Reedik Mägi, Anubha Mahajan, Michael V. Holmes, and Cecilia M. Lindgren. Causal relationships between obesity and the leading causes of death in women and men. *PLOS Genetics*, 15(10):e1008405. Publisher: Public Library of Science.

[19] Xuqi Chen, Rebecca McClusky, Jenny Chen, Simon W. Beaven, Peter Tontonoz, Arthur P. Arnold, and Karen Reue. The number of x chromosomes causes sex differences in adiposity in mice. *PLoS Genetics*, 8(5).

[20] Sung Sik Choe, Jin Young Huh, In Jae Hwang, Jong In Kim, and Jae Bum Kim. Adipose tissue remodeling: Its role in energy metabolism and metabolic disorders. *Frontiers in Endocrinology*, 7. Publisher: Frontiers.

[21] Yu Chung Chooi, Cherlyn Ding, and Faidon Magkos. The epidemiology of obesity. *Metabolism: Clinical and Experimental*, 92:6–10.

[22] Edward B. Chuong, Nels C. Elde, and Cédric Feschotte. Regulatory evolution of innate immunity through co-option of endogenous retroviruses. *Science*, 351(6277):1083–1087. Publisher: American Association for the Advancement of Science Section: Report.

[23] Dale Caroline E., Fatemifar Ghazaleh, Palmer Tom M., White Jon, Prieto-Merino David, Zabaneh Delilah, Engmann Jorgen E.L., Shah Tina, Wong Andrew, Warren Helen R., McLachlan Stela, Trompet Stella, Moldovan Max, Morris Richard W., Sofat Reecha, Kumari Meena, Hyppönen Elina, Jefferis Barbara J., Gaunt Tom R., Ben-Shlomo Yoav, Zhou Ang, Gentry-Maharaj Aleksandra, Ryan Andy, Mutsert Renée de, Noordam Raymond, Caulfield Mark J., Jukema J. Wouter, Worrall Bradford B., Munroe Patricia B., Menon Usha, Power Chris, Kuh Diana, Lawlor Debbie A., Humphries Steve E., Mook-Kanamori Dennis O., Sattar Naveed, Kivimaki Mika, Price Jacqueline F., Davey Smith George, Dudbridge Frank, Hingorani Aroon D., Holmes Michael V., and Casas Juan P. Causal associations of adiposity and body fat distribution with coronary heart disease, stroke subtypes, and type 2 diabetes mellitus. *Circulation*, 135(24):2373–2388. Publisher: American Heart Association.

[24] Kathryn E. Davis, Michael D. Neinast, Kai Sun, William M. Skiles, Jessica D. Bills, Jordan A. Zehr, Daniel Zeve, Lisa D. Hahner, Derek W. Cox, Lana M. Gent, Yong Xu, Zhao V. Wang, Sohaib A. Khan, and Deborah J. Clegg. The sexually dimorphic role of adipose and adipocyte estrogen receptors in modulating adipose tissue expansion, inflammation, and fibrosis. *Molecular Metabolism*, 2(3):227–242.

[25] Jesse R. Dixon, Siddarth Selvaraj, Feng Yue, Audrey Kim, Yan Li, Yin Shen, Ming Hu, Jun S. Liu, and Bing Ren. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*, 485(7398):376–380. Number: 7398 Publisher: Nature Publishing Group.

[26] B. Dubern, K. Clément, V. Pelloux, P. Froguel, J. P. Girardet, B. Guy-Grand, and P. Tounian. Mutational analysis of melanocortin-4 receptor, agouti-related protein, and alpha-melanocyte-stimulating hormone genes in severely obese children. *The Journal of Pediatrics*, 139(2):204–209.

[27] John N. Fain, Atul K. Madan, M. Lloyd Hiler, Paramjeet Cheema, and Suleiman W. Bahouth. 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(5):2273–2282.

[28] Luciana A. Fatima, Raquel S. Campello, João N. Barreto-Andrade, Marisa Passarelli, Roberta S. Santos, Deborah J. Clegg, and Ubiratan F. Machado. Estradiol stimulates adipogenesis and slc2a4/GLUT4 expression via ESR1-mediated activation of CEBPA. *Molecular and Cellular Endocrinology*, 498:110447.

[29] Hilary K. Finucane, Brendan Bulik-Sullivan, Alexander Gusev, Gosia Trynka, Yakir Reshef, Po-Ru Loh, Verneri Anttila, Han Xu, Chongzhi Zang, Kyle Farh, Stephan Ripke, Felix R. Day, Shaun Purcell, Eli Stahl, Sara Lindstrom, John R. B. Perry, Yukinori Okada, Soumya Raychaudhuri, Mark J. Daly, Nick Patterson, Benjamin M. Neale, and Alkes L. Price. Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nature Genetics*, 47(11):1228–1235. Number: 11 Publisher: Nature Publishing Group.

[30] Hilary K. Finucane, Yakir A. Reshef, Verneri Anttila, Kamil Slowikowski, Alexander Gusev, Andrea Byrnes, Steven Gazal, Po-Ru Loh, Caleb Lareau, Noam Shoresh, Giulio Genovese, Arpiar Saunders, Evan Macosko, Samuela Pollack, John R. B. Perry, Jason D. Buenrostro, Bradley E. Bernstein, Soumya Raychaudhuri, Steven McCarroll, Benjamin M. Neale, and Alkes L. Price. Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. *Nature Genetics*, 50(4):621–629. Publisher: Nature Publishing Group.

[31] Michael J. Fowler. Microvascular and macrovascular complications of diabetes. *Clinical Diabetes*, 26(2):77–82. Publisher: American Diabetes Association Section: Departments.

[32] Keith E. Fox, Dana M. Fankell, Paul F. Erickson, Susan M. Majka, Joseph T. Crossno, and Dwight J. Klemm. Depletion of cAMP-response element-binding protein/ATF1 inhibits adipogenic conversion of 3t3-l1 cells ectopically expressing CCAAT/enhancer-binding protein (c/EBP) alpha, c/EBP beta, or PPAR gamma 2. *The Journal of Biological Chemistry*, 281(52):40341–40353.

[33] Martin Franke, Daniel M. Ibrahim, Guillaume Andrey, Wibke Schwarzer, Verena Heinrich, Robert Schöpflin, Katerina Kraft, Rieke Kempfer, Ivana Jerković, Wing-Lee Chan, Malte Spielmann, Bernd Timmermann, Lars Wittler, Ingo Kurth, Paola Cambiaso, Orsetta Zuffardi, Gunnar Houge, Lindsay Lambie, Francesco Brancati, Ana Pombo, Martin Vingron, Francois Spitz, and Stefan Mundlos. Formation of new chromatin domains determines pathogenicity of genomic duplications. *Nature*, 538(7624):265–269. Number: 7624 Publisher: Nature Publishing Group.

[34] Charles P. Fulco, Joseph Nasser, Thouis R. Jones, Glen Munson, Drew T. Bergman, Vidya Subramanian, Sharon R. Grossman, Rockwell Anyoha, Benjamin R. Doughty, Tejal A. Patwardhan, Tung H. Nguyen, Michael Kane, Elizabeth M. Perez, Neva C. Durand, Caleb A. Lareau, Elena K. Stamenova, Erez Lieberman Aiden, Eric S. Lander, and Jesse M. Engreitz. Activity-by-contact model of enhancer–promoter regulation from thousands of CRISPR perturbations. *Nature Genetics*, 51(12):1664–1669. Number: 12 Publisher: Nature Publishing Group.

[35] Gealekman Olga, Guseva Nina, Hartigan Celia, Apotheker Sarah, Gorgoglione Matthew, Gurav Kunal, Tran Khanh-Van, Straubhaar Juerg, Nicoloro Sarah, Czech Michael P., Thompson Michael, Perugini Richard A., and Corvera Silvia. Depot-specific differences and insufficient subcutaneous adipose tissue angiogenesis in human obesity. *Circulation*, 123(2):186–194. Publisher: American Heart Association.

[36] Arijit Ghosh, Lei Gao, Abhimanyu Thakur, Parco M. Siu, and Christopher W. K. Lai. Role of free fatty acids in endothelial dysfunction. *Journal of Biomedical Science*, 24(1):50.

[37] Claudia Giambartolomei, Damjan Vukcevic, Eric E. Schadt, Lude Franke, Aroon D. Hingorani, Chris Wallace, and Vincent Plagnol. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genetics*, 10(5):e1004383–e1004383. Publisher: Public Library of Science.

[38] Niels Grarup, Ida Moltke, Mette K. Andersen, Maria Dalby, Kristoffer Vitting-Seerup, Timo Kern, Yuvaraj Mahendran, Emil Jørsboe, Christina V. L. Larsen, Inger K. Dahl-Petersen, Arthur Gilly, Daniel Suveges, George Dedoussis, Eleftheria Zeggini, Oluf Pedersen, Robin Andersson, Peter Bjerregaard, Marit E. Jørgensen, Anders Albrechtsen, and Torben Hansen. Loss-of-function variants in ADCY3 increase risk of obesity and type 2 diabetes. *Nature Genetics*, 50(2):172–174. Publisher: Nature Publishing Group.

[39] Deborah J. Greenway, Miyoko Street, Aaron Jeffries, and Noel J. Buckley. RE1 silencing transcription factor maintains a repressive chromatin environment in embryonic hippocampal neural stem cells. *STEM CELLS*, 25(2):354–363. _eprint: https://stemcellsjournals.onlinelibrary.wiley.com/doi/pdf/10.1634/stemcells.2006-0207.

[40] Liang Guo, Xi Li, and Qi-Qun Tang. Transcriptional regulation of adipocyte differentiation: A central role for CCAAT/enhancer-binding protein (c/EBP) β. *The Journal of Biological Chemistry*, 290(2):755–761.

[41] Alexander Gusev, Arthur Ko, Huwenbo Shi, Gaurav Bhatia, Wonil Chung, Brenda W J H Penninx, Rick Jansen, Eco J C de Geus, Dorret I Boomsma, Fred A Wright, Patrick F Sullivan, Elina Nikkola, Marcus Alvarez, Mete Civelek, Aldons J Lusis, Terho Lehtimäki, Emma Raitoharju, Mika Kähönen, Ilkka Seppälä, Olli T Raitakari, Johanna Kuusisto, Markku Laakso, Alkes L Price, Päivi Pajukanta, and Bogdan Pasaniuc. Integrative approaches for large-scale transcriptome-wide association studies. *Nature Genetics*, 48(3):245–252. Publisher: Nature Publishing Group.

[42] Paul D. Hartley and Hiten D. Madhani. Mechanisms that specify promoter nucleosome location and identity. *Cell*, 137(3):445–458.

[43] Sven Heinz, Christopher Benner, Nathanael Spann, Eric Bertolino, Yin C. Lin, Peter Laslo, Jason X. Cheng, Cornelis Murre, Harinder Singh, and Christopher K. Glass.

Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and b cell identities. *Molecular cell*, 38(4):576–589.

[44] Gabriel E. Hoffman, Jaroslav Bendl, Georgios Voloudakis, Kelsey S. Montgomery, Laura Sloofman, Ying-Chih Wang, Hardik R. Shah, Mads E. Hauberg, Jessica S. Johnson, Kiran Girdhar, Lingyun Song, John F. Fullard, Robin Kramer, Chang-Gyu Hahn, Raquel Gur, Stefano Marenco, Barbara K. Lipska, David A. Lewis, Vahram Haroutunian, Scott Hemby, Patrick Sullivan, Schahram Akbarian, Andrew Chess, Joseph D. Buxbaum, Greg E. Crawford, Enrico Domenici, Bernie Devlin, Solveig K. Sieberts, Mette A. Peters, and Panos Roussos. CommonMind consortium provides transcriptomic and epigenomic data for schizophrenia and bipolar disorder. *Scientific Data*, 6(1):180. Number: 1 Publisher: Nature Publishing Group.

[45] Janice M. Huss, Inés Pineda Torra, Bart Staels, Vincent Giguère, and Daniel P. Kelly. Estrogen-related receptor α directs peroxisome proliferator-activated receptor α signaling in the transcriptional control of energy metabolism in cardiac and skeletal muscle. *Molecular and Cellular Biology*, 24(20):9079–9091.

[46] M. Mohsen Ibrahim. Subcutaneous and visceral adipose tissue: structural and functional differences. *Obesity Reviews*, 11(1):11–18. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-789X.2009.00623.x.

[47] Elise Jeffery, Allison Wing, Brandon Holtrup, Zachary Sebo, Jennifer L. Kaplan, Rocio Saavedra-Peña, Christopher D. Church, Laura Colman, Ryan Berry, and Matthew S. Rodeheffer. The adipose tissue microenvironment regulates depot-specific adipogenesis in obesity. *Cell Metabolism*, 24(1):142–150.

[48] Kirk B. Jensen, B. Kate Dredge, Giovanni Stefani, Ru Zhong, Ronald J. Buckanovich, Hirotaka J. Okano, Yolanda Y. L. Yang, and Robert B. Darnell. Nova-1 regulates neuron-specific alternative splicing and is essential for neuronal viability. *Neuron*, 25(2):359–371. Publisher: Elsevier.

[49] Julius Judd, Hayley Sanderson, and Cédric Feschotte. Evolution of mouse circadian enhancers from transposable elements. *bioRxiv*, page 2020.11.09.375469. Publisher: Cold Spring Harbor Laboratory Section: New Results.

[50] Jurga Laurencikiene, Thomas Skurk, Agné Kulyté, Per Hedén, Gaby Åström, Eva Sjölin, Mikael Rydén, Hans Hauner, and Peter Arner. Regulation of lipolysis in small and large fat cells of the same subject. *The Journal of Clinical Endocrinology & Metabolism*, 96(12):E2045–E2049.

[51] Martina I. Lefterova, Anders K. Haakonsson, Mitchell A. Lazar, and Susanne Mandrup. PPARγ and the global map of adipogenesis and beyond. *Trends in Endocrinology & Metabolism*, 25(6):293–302.

[52] Erez Lieberman-Aiden, Nynke L. van Berkum, Louise Williams, Maxim Imakaev, Tobias Ragoczy, Agnes Telling, Ido Amit, Bryan R. Lajoie, Peter J. Sabo, Michael O.

Dorschner, Richard Sandstrom, Bradley Bernstein, M. A. Bender, Mark Groudine, Andreas Gnirke, John Stamatoyannopoulos, Leonid A. Mirny, Eric S. Lander, and Job Dekker. Comprehensive mapping of long range interactions reveals folding principles of the human genome. *Science (New York, N.Y.)*, 326(5950):289–293.

[53] Jenny C. Link, Yehudit Hasin-Brumshtein, Rita M. Cantor, Xuqi Chen, Arthur P. Arnold, Aldons J. Lusis, and Karen Reue. Diet, gonadal sex, and sex chromosome complement influence white adipose tissue miRNA expression. *BMC Genomics*, 18(1):89.

[54] A. E. Locke, B. Kahali, S. I. Berndt, A. E. Justice, T. H. Pers, F. R. Day, C. Powell, S. Vedantam, M. L. Buchkovich, J. Yang, D. C. Croteau-Chonka, T. Esko, T. Fall, T. Ferreira, S. Gustafsson, Z. Kutalik, J. Luan, R. Mägi, J. C. Randall, T. W. Winkler, A. R. Wood, T. Workalemahu, J. D. Faul, J. A. Smith, J. H. Zhao, W. Zhao, J. Chen, R. Fehrmann, K. Hedman, J. Karjalainen, E. M. Schmidt, D. Absher, N. Amin, D. Anderson, M. Beekman, J. L. Bolton, J. L. Bragg-Gresham, S. Buyske, A. Demirkan, G. Deng, G. Ehret, B. Feenstra, M. Feitosa, K. Fischer, A. Goel, J. Gong, A. U. Jackson, S. Kanoni, M. E. Kleber, K. Kristiansson, U. Lim, V. Lotay, M. Mangino, I. M. Leach, C. Medina-Gomez, S. E. Medland, M. A. Nalls, C. D. Palmer, D. Pasko, S. Pechlivanis, M. J. Peters, I. Prokopenko, D. Shungin, A. Stančáková, R. J. Strawbridge, Y. J. Sung, T. Tanaka, A. Teumer, S. Trompet, S. W. van der Laan, J. van Setten, J. V. Van Vliet-Ostaptchouk, Z. Wang, L. Yengo, W. Zhang, A. Isaacs, E. Albrecht, J. Ärnlöv, G. M. Arscott, A. P. Attwood, S. Bandinelli, A. Barrett, I. N. Bas, C. Bellis, A. J. Bennett, C. Berne, R. Blagieva, M. Blüher, S. Böhringer, L. L. Bonnycastle, Y. Böttcher, H. A. Boyd, M. Bruinenberg, I. H. Caspersen, Y. I. Chen, R. Clarke, E. W. Daw, A. J.M. de Craen, G. Delgado, M. Dimitriou, A. S.F. Doney, N. Eklund, K. Estrada, E. Eury, L. Folkersen, R. M. Fraser, M. Garcia, F. Geller, V. Giedraitis, B. Gigante, A. S. Go, A. Golay, A. Goodall, S. D. Gordon, M. Gorski, H. J. Grabe, H. Grallert, T. B. Grammer, J. Gräßler, H. Grönberg, C. J. Groves, G. Gusto, J. Haessler, P. Hall, T. Haller, G. Hallmans, C. A. Hartman, M. Hassinen, C. Hayward, N. L. Heard-Costa, Q. Helmer, C. Hengstenberg, O. Holmen, J. J. Hottenga, A. L. James, J. Jeff, Johansson, J. Jolley, T. Juliusdottir, L. Kinnunen, W. Koenig, M. Koskenvuo, W. Kratzer, J. Laitinen, C. Lamina, K. Leander, N. R. Lee, P. Lichtner, L. Lind, J. Lindström, K. S. Lo, S. Lobbens, R. Lorbeer, Y. Lu, F. Mach, P. K. Magnusson, A. Mahajan, W. L. McArdle, S. McLachlan, C. Menni, S. Merger, E. Mihailov, L. Milani, A. Moayyeri, K. L. Monda, M. A. Morken, A. Mulas, G. Müller, M. Müller-Nurasyid, A. W. Musk, R. Nagaraja, M. M. Nöthen, I. M. Nolte, S. Pilz, N. W. Rayner, F. Renstrom, R. Rettig, J. S. Ried, S. Ripke, N. Robertson, L. M. Rose, S. Sanna, H. Scharnagl, S. Scholtens, F. Schumacher, W. R. Scott, T. Seufferlein, J. Shi, A. V. Smith, J. Smolonska, A. V. Stanton, V. Steinthorsdottir, K. Stirrups, H. M. Stringham, J. Sundström, M. A. Swertz, A. J. Swift, A. C. Syvänen, S. T. Tan, B. Tayo, B. Thorand, G. Thorleifsson, J. Tyrer, H. W. Uh, L. Vandenput, F. C. Verhulst, S. H. Vermeulen, N. Verweij, J. M. Vonk, L. L. Waite, H. R. Warren, D. M. Waterworth, M. N. Weedon, L. Wilkens, C. Willenborg, T. Wilsgaard, M. K. Wojczynski, A. Wong, A. F. Wright, Q. Zhang, E. P. Brennan, M. Choi, Z. Dastani,

A. W. Drong, P. Eriksson, A. Franco-Cereceda, J. R. Gådin, A. G. Gharavi, M. E. Goddard, R. E. Handsaker, J. Huang, F. Karpe, S. Kathiresan, S. Keildson, K. Kiryluk, M. Kubo, J. Y. Lee, L. Liang, R. P. Lifton, B. Ma, S. A. McCarroll, A. J. McKnight, J. L. Min, M. F. Moffatt, G. W. Montgomery, J. M. Murabito, G. Nicholson, D. R. Nyholt, Y. Okada, J. R. Perry, R. Dorajoo, E. Reinmaa, R. M. Salem, N. Sandholm, R. A. Scott, L. Stolk, A. Takahashi, F. M. Van't Hooft, A. A.E. Vinkhuyzen, H. J. Westra, W. Zheng, K. T. Zondervan, A. C. Heath, D. Arveiler, S. J. Bakker, J. P. Beilby, R. N. Bergman, J. Blangero, P. Bovet, H. Campbell, M. Caulfield, G. Cesana, A. Chakravarti, D. Chasman, P. S. Chines, F. S. Collins, D. Crawford, L. Cupples, D. Cusi, J. Danesh, U. de Faire, H. M. Den Ruijter, A. F. Dominiczak, R. Erbel, J. Erdmann, J. G. Eriksson, M. Farrall, S. B. Felix, E. Ferrannini, J. Ferrières, I. Ford, N. G. Forouhi, T. Forrester, O. H. Franco, R. T. Gansevoort, P. V. Gejman, C. Gieger, O. Gottesman, V. Gudnason, U. B. Gyllensten, A. S. Hall, T. B. Harris, A. T. Hattersley, A. A. Hicks, L. Hindorff, A. Hingorani, A. Hofman, G. Homuth, G. Hovingh, S. E. Humphries, S. C. Hunt, E. Hyppönen, T. Illig, K. B. Jacobs, M. R. Jarvelin, K. H. Jöckel, B. Johansen, P. Jousilahti, J. Jukema, A. Jula, J. Kaprio, J. J. Kastelein, S. M. Keinanen-Kiukaanniemi, L. A. Kiemeney, P. Knekt, J. S. Kooner, C. Kooperberg, P. Kovacs, A. T. Kraja, M. Kumari, J. Kuusisto, T. Lakka, C. Langenberg, L. L. Marchand, T. Lehtimäki, V. Lyssenko, S. Männistö, A. Marette, T. Matise, C. A. McKenzie, B. McKnight, F. L. Moll, A. D. Morris, A. P. Morris, J. C. Murray, M. Nelis, C. Ohlsson, A. J. Oldehinkel, K. K. Ong, P. A.F. Madden, G. Pasterkamp, J. F. Peden, A. Peters, D. S. Postma, P. P. Pramstaller, J. F. Price, L. Qi, O. Raitakari, T. Rankinen, D. C. Rao, T. K. Rice, P. Ridker, J. D. Rioux, M. Ritchie, I. Rudan, V. Salomaa, N. Samani, J. Saramies, M. A. Sarzynski, H. Schunkert, P. E. Schwarz, P. Sever, A. R. Shuldiner, J. Sinisalo, R. P. Stolk, K. Strauch, A. Tönjes, D. A. Trégouët, A. Tremblay, E. Tremoli, J. Virtamo, M. C. Vohl, U. Völker, G. Waeber, G. Willemsen, J. C. Witteman, M. C. Zillikens, L. S. Adair, P. Amouyel, F. W. Asselbergs, T. L. Assimes, M. Bochud, B. O. Boehm, E. Boerwinkle, S. R. Bornstein, E. P. Bottinger, C. Bouchard, S. Cauchi, J. C. Chambers, S. J. Chanock, R. S. Cooper, P. I.W. de Bakker, G. V. Dedoussis, L. Ferrucci, P. W. Franks, P. Froguel, L. Groop, C. Haiman, A. Hamsten, J. Hui, D. J. Hunter, K. Hveem, R. C. Kaplan, M. Kivimaki, D. Kuh, M. Laakso, Y. Liu, N. G. Martin, W. März, M. Melbye, A. Metspalu, S. Moebus, P. Munroe, I. Njølstad, B. A. Oostra, C. N. Palmer, N. L. Pedersen, M. Perola, L. Pérusse, U. Peters, C. Power, T. Quertermous, R. Rauramaa, F. Rivadeneira, T. E. Saaristo, D. Saleheen, N. Sattar, E. Schadt, D. Schlessinger, P. E. Slagboom, H. Snieder, T. D. Spector, U. Thorsteinsdottir, M. Stumvoll, J. Tuomilehto, A. G. Uitterlinden, M. Uusitupa, P. van der Harst, M. C. Walker, H. Wallaschofski, N. Wareham, H. Watkins, D. R. Weir, H. Wichmann, J. F. Wilson, P. Zanen, I. Borecki, P. Deloukas, C. S. Fox, I. M. Heid, J. R. O'connell, D. P. Strachan, K. Stefansson, C. Van Duijn, G. Abecasis, L. Franke, T. M. Frayling, M. I. McCarthy, P. M. Visscher, A. Scherag, C. J. Willer, M. Boehnke, K. L. Mohlke, C. M. Lindgren, J. S. Beckmann, I. Barroso, K. E. North, E. Ingelsson, J. N. Hirschhorn, R. J. Loos, E. K. Speliotes, J. R. Thompson, B. A. Goldstein, I. R. König, J. B. Cazier, E. Grundberg, A. S. Havulinna, W. K.

Ho, J. C. Hopewell, N. Eriksson, P. Lundmark, L. P. Lyytikäinen, S. Rafelt, E. Tikkanen, N. Van Zuydam, B. F. Voight, A. Ziegler, D. Altshuler, A. J. Balmforth, P. S. Braund, C. Burgdorf, S. Claudi-Boehm, D. Cox, R. Do, A. S. Doney, N. El Mokhtari, P. Fontanillas, J. Hager, B. G. Han, S. E. Hunt, H. M. Kang, T. Kessler, J. W. Knowles, G. Kolovou, C. Langford, M. L. Lokki, A. Lundmark, C. Meisinger, O. Melander, S. Maouche, K. Nikus, A. Rasheed, S. Rosinger, D. Rubin, M. P. Rumpf, A. Schäfer, M. Sivananthan, C. Song, A. F. Stewart, G. Thorgeirsson, C. E. van der Schoot, P. J. Wagner, G. A. Wells, P. S. Wild, Y. Tsun-Po, H. Basart, P. Brambilla, F. Cambien, A. L. Cupples, A. Dehghan, P. Diemert, S. E. Epstein, A. Evans, M. M. Ferrario, D. Gauguier, S. L. Hazen, H. Holm, C. Iribarren, Y. Jang, M. Kähönen, F. Kee, H. S. Kim, N. Klopp, K. Kuulasmaa, R. Laaksonen, W. Ouwehand, S. Parish, J. E. Park, D. J. Rader, S. Shah, K. Stark, L. Wallentin, M. E. Zimmermann, M. S. Nieminen, M. S. Sandhu, T. Pastinen, P. A. Zalloua, A. Siegbahn, S. Schreiber, S. Ripatti, S. S. Blankenberg, C. O'donnell, M. P. Reilly, R. Collins, R. Roberts, C. Pattaro, A. Köttgen, M. Garnaas, C. A. Böger, C. Fuchsberger, M. Olden, M. H. Chen, A. Tin, D. Taliun, M. Li, X. Gao, Q. Yang, C. Hundertmark, M. C. Foster, C. M. O'seaghdha, N. Glazer, C. T. Liu, M. Struchalin, G. Li, A. D. Johnson, H. J. Gierman, S. J. Hwang, E. J. Atkinson, K. Lohman, M. C. Cornelis, V. Chouraki, E. G. Holliday, R. Sorice, H. Deshmukh, S. Ulivi, A. Y. Chu, F. Murgia, M. Imboden, B. Kollerits, G. Pistis, L. Launer, T. Aspelund, G. Eiriksdottir, B. D. Mitchell, H. Schmidt, M. Cavalieri, M. Rao, F. B. Hu, M. de Andrade, S. T. Turner, J. Ding, J. S. Andrews, B. I. Freedman, A. Döring, I. Kolcic, T. Zemunik, M. Boban, C. Minelli, H. E. Wheeler, W. Igl, G. Zaboli, S. H. Wild, D. Ellinghaus, U. Nöthlings, G. Jacobs, R. Biffar, K. Endlich, F. Ernst, H. K. Kroemer, M. Nauck, S. Stracke, H. Völzke, Y. S. Aulchenko, O. Polasek, N. Hastie, V. Vitart, C. Helmer, J. J. Wang, D. Ruggiero, S. Bergmann, J. Viikari, T. Nikopensius, M. Province, S. Ketkar, H. Colhoun, A. Doney, A. Robino, F. Giulianini, B. K. Krämer, L. Portas, B. M. Buckley, M. Adam, G. A. Thun, B. Paulweber, M. Haun, C. Sala, M. Metzger, P. Mitchell, M. Ciullo, S. K. Kim, P. Vollenweider, C. Palmer, P. Gasparini, M. Pirastu, N. M. Probst-Hensch, F. Kronenberg, D. Toniolo, J. Coresh, R. Schmidt, D. Siscovick, S. L. Kardia, G. C. Curhan, A. Franke, A. Parsa, W. Goessling, W. Kao, I. H. de Boer, C. A. Peralta, E. Akylbekova, H. Kramer, D. E. Arking, N. Franceschini, J. Egan, D. Hernandez, R. R. Townsend, T. Lumley, B. Psaty, B. Kestenbaum, T. Haritunians, V. Mooser, J. C. Florez, J. B. Meigs, X. Lu, T. S. Leak, K. Aasarød, F. Skorpen, J. Baumert, O. Devuyst, J. C. Mychaleckyj, L. Kedenko, S. Coassin, S. Hallan, G. Navis, M. G. Shlipak, S. B. Bull, A. D. Paterson, J. I. Rotter, A. W. Dreisbach, C. A. Anderson, Q. Guo, A. Henders, A. Lambert, S. H. Lee, P. Kraft, S. H. Kennedy, S. Macgregor, S. A. Missmer, J. N. Painter, F. Roseman, S. A. Treloar, L. Wallace, C. Forsblom, T. Isakova, G. J. McKay, W. W. Williams, D. M. Sadlier, V. P. Mäkinen, E. J. Swan, A. P. Boright, E. Ahlqvist, B. J. Keller, H. Huang, A. Ahola, E. Fagerholm, D. Gordin, V. Harjutsalo, B. He, O. Heikkilä, K. Hietala, J. Kytö, P. Lahermo, M. Lehto, A. M. Österholm, M. Parkkonen, J. Pitkäniemi, M. Rosengård-Bärlund, M. Saraheimo, C. Sarti, J. Söderlund, A. Soro-Paavonen, A. Syreeni, L. M. Thorn, H. Tikkanen, N. Tolonen, K. Tryggva-

son, J. Wadén, G. V. Gill, S. Prior, C. Guiducci, D. Mirel, A. Taylor, M. Hosseini, H. H. Parving, P. Rossing, L. Tarnow, C. Ladenvall, F. Alhenc-Gelas, P. Lefebvre, V. Rigalleau, R. Roussel, D. A. Tregouet, A. Maestroni, S. Maestroni, H. Falhammar, T. Gu, A. Möllsten, D. Cimponeriu, I. Mihai, M. Mota, E. Mota, C. Serafinceanu, M. Stavarachi, R. L. Hanson, R. G. Nelson, M. Kretzler, N. M. Panduru, H. F. Gu, K. Brismar, G. Zerbini, S. Hadjadj, M. Marre, M. Lajer, D. Waggott, D. A. Savage, S. C. Bain, F. Martin, C. Godson, P. H. Groop, A. P. Maxwell, S. Sengupta, G. M. Peloso, A. Ganna, S. Mora, H. Y. Chang, H. M. Den Hertog, L. A. Donnelly, D. F. Freitag, D. Gurdasani, K. Heikkilä, T. Johnson, M. Kaakinen, J. Kettunen, X. Li, M. E. Montasser, A. K. Petersen, R. Saxena, S. K. Service, C. Sidore, I. Surakka, T. M. Teslovich, E. G. Van den Herik, K. A. Volcik, Y. Wu, G. Asiki, L. F. Been, M. S. Burnett, A. Doring, P. Elliott, G. I. Eyjolfsson, M. O. Goodarzi, M. L. Gravito, A. L. Hartikainen, Y. J. Hung, M. R. Jones, P. Kaleebu, K. T. Khaw, E. Kim, P. Komulainen, T. Lehtimaki, S. Y. Lin, J. Lindstrom, G. Muller, N. Narisu, T. V. Nieminen, R. N. Nsubuga, I. Olafsson, A. Palotie, T. Papamarkou, C. Pomilla, A. Pouta, A. Ruokonen, J. Seeley, K. Silander, L. Tiret, L. van Pelt, N. Wainwright, C. Wijmenga, E. H. Young, F. Bennett, D. I. Boomsma, M. Burnier, Y. D. Chen, A. B. Feranil, J. Ferrieres, N. B. Freimer, C. A. Hsiung, A. Kesäniemi, P. J. Koudstaal, R. M. Krauss, K. O. Kyvik, P. Meneton, L. Moilanen, D. K. Sanghera, W. H. Sheu, J. B. Whitfield, B. H. Wolffenbuttel, J. M. Ordovas, S. S. Rich, A. Johnson, L. Johnson, M. G. Larson, D. Levy, C. Newton-Cheh, P. F. O'reilly, W. Palmas, K. Rice, A. Smith, H. Snider, M. Tobin, G. Verwoert, K. M. Rice, G. C. Verwoert, V. Pihur, S. Heath, S. Sõber, P. Arora, F. Zhang, G. Lucas, Y. Milaneschi, A. N. Parker, C. Fava, E. R. Fox, M. J. Go, M. Sjögren, D. Vinay, M. Alexander, Y. Tabara, S. Shaw-Hawkins, P. H. Whincup, G. Shi, M. Seielstad, X. Sim, K. D. Nguyen, G. Matullo, T. R. Gaunt, N. C. Onland-Moret, M. N. Cooper, C. G. Platou, E. Org, R. Hardy, S. Dahgam, J. Palmen, T. Kuznetsova, C. S. Uiterwaal, A. Adeyemo, B. Ludwig, M. Tomaszewski, I. Tzoulaki, N. D. Palmer, Y. P. Chang, N. I. Steinle, D. E. Grobbee, A. C. Morrison, S. Najjar, D. Hadley, M. J. Brown, J. M. Connell, I. N. Day, D. A. Lawlor, R. W. Lawrence, H. Ongen, Y. Li, J. H. Young, J. C. Bis, J. A. Bolton, N. Chaturvedi, M. Islam, T. H. Jafar, S. R. Kulkarni, P. Howard, S. Guarrera, F. Ricceri, V. Emilsson, A. S. Plump, A. B. Weder, Y. V. Sun, L. J. Scott, L. Peltonen, E. Vartiainen, S. M. Brand, J. A. Staessen, T. J. Wang, P. R. Burton, M. S. Artigas, Y. Dong, X. Wang, H. Zhu, M. E. Rudock, S. R. Heckbert, N. L. Smith, K. L. Wiggins, A. Doumatey, D. Shriner, G. Veldre, M. Viigimaa, S. Kinra, D. Prabhakaran, V. Tripathy, C. D. Langefeld, A. Rosengren, D. S. Thelle, A. M. Corsi, A. Singleton, G. Hilton, T. Salako, N. Iwai, Y. Kita, T. Ogihara, T. Ohkubo, T. Okamura, H. Ueshima, S. Umemura, S. Eyheramendy, T. Meitinger, Y. S. Cho, H. L. Kim, J. Scott, J. S. Sehmi, B. Hedblad, P. Nilsson, G. D. Smith, L. J. Raffel, J. Yao, S. M. Schwartz, M. Ikram, Longstreth W, T. H. Mosley, S. Seshadri, N. R. Shrine, L. V. Wain, P. Zitting, J. A. Cooper, W. H. van Gilst, C. S. Janipalli, K. Mani, C. S. Yajnik, F. U. Mattace-Raso, E. G. Lakatta, M. Orru, A. Scuteri, M. Ala-Korpela, A. J. Kangas, P. Soininen, T. Tukiainen, P. Würtz, R. T. Ong, M. Dörr, P. Galan, S. Hercberg, M. Lathrop, D. Zelenika, G. Zhai, J. F. Meschia, P. Sharma, J. Terzic,

M. Kumar, M. Denniff, E. Zukowska-Szczechowska, L. E. Wagenknecht, F. Fowkes, F. J. Charchar, X. Guo, C. Rotimi, M. L. Bots, E. Brand, P. J. Talmud, F. Nyberg, M. Laan, L. J. Palmer, Y. T. van der Schouw, J. P. Casas, P. Vineis, S. K. Ganesh, T. Y. Wong, E. S. Tai, R. W. Morris, M. G. Marmot, T. Miki, G. R. Chandak, X. Zhu, R. Elosua, N. Soranzo, E. J. Sijbrands, M. Uda, R. S. Vasan, M. R. Järvelin, B. Z. Alizadeh, R. A. de Boer, H. M. Boezen, H. L. Hillege, M. M. van der Klauw, J. Ormel, J. G. Rosmalen, J. P. Slaets, V. Lagou, R. P. Welch, E. Wheeler, E. Rehnberg, L. J. Rasmussen-Torvik, C. Lecoeur, P. C. Johnson, B. Sennblad, P. Salo, N. J. Timpson, D. M. Evans, B. St Pourcain, L. F. Bielak, M. Horikoshi, P. Navarro, S. Raychaudhuri, H. Chen, D. Rybin, S. M. Willems, K. Song, P. An, L. Marullo, H. Jansen, J. Pankow, S. Edkins, T. V. Varga, H. Oksa, M. Antonella, A. Kong, C. Herder, J. Antti, K. S. Small, I. Miljkovic, M. Atalay, W. Kiess, J. H. Smit, S. Campbell, G. R. Fowkes, W. Rathmann, W. Maerz, R. M. Watanabe, E. J. de Geus, B. W. Penninx, A. Toenjes, P. A. Peyser, A. Körner, J. Dupuis, F. Cucca, B. Balkau, N. Bouatia-Naji, S. Purcell, K. Musunuru, D. Ardissino, P. M. Mannucci, S. Anand, J. C. Engert, T. Morgan, J. A. Spertus, M. Stoll, D. Girelli, P. P. McKeown, C. C. Patterson, P. A. Merlini, C. Berzuini, L. Bernardinelli, F. Peyvandi, M. Tubaro, P. Celli, R. Fetiveau, N. Marziliano, G. Casari, M. Galli, F. Ribichini, M. Rossi, F. Bernardi, P. Zonzin, A. Piazza, J. Yee, Y. Friedlander, J. Marrugat, I. Subirana, J. Sala, R. Ramos, G. Williams, D. M. Nathan, C. A. Macrae, G. Berglund, R. Asselta, S. Duga, M. Spreafico, M. J. Daly, J. Nemesh, J. M. Korn, A. Surti, L. Gianniny, M. Parkin, N. Burtt, S. B. Gabriel, B. J. Wright, S. G. Ball, I. Schunkert, P. Linsel-Nitschke, W. Lieb, M. Fischer, A. Grosshennig, M. Preuss, M. Scholz, Z. Chen, R. Wilensky, W. Matthai, A. Qasim, H. H. Hakonarson, J. Devaney, A. D. Pichard, K. M. Kent, L. Satler, J. M. Lindsay, R. Waksman, C. W. Knouff, T. Scheffold, K. Berger, A. Huge, N. Martinelli, O. Olivieri, R. Corrocher, H. Hólm, C. Xie, K. R. Ahmadi, C. Ainali, V. Bataille, J. T. Bell, A. Buil, E. T. Dermitzakis, A. S. Dimas, R. Durbin, D. Glass, N. Hassanali, C. Ingle, D. Knowles, M. Krestyaninova, C. E. Lowe, E. Meduri, P. Di Meglio, S. B. Montgomery, F. O. Nestle, A. C. Nica, J. Nisbet, S. O'rahilly, L. Parts, S. Potter, M. Sekowska, S. Y. Shin, G. Surdulescu, M. E. Travers, L. Tsaprouni, S. Tsoka, A. Wilk, T. P. Yang, J. Higashio, R. Williams, A. Nato, J. L. Ambite, E. Deelman, T. Manolio, G. Heiss, K. Taylor, C. Avery, M. Graff, D. Lin, M. Quibrera, B. Cochran, L. Kao, J. Umans, S. Cole, J. Maccluer, S. Person, M. Gross, M. Fornage, P. Durda, N. Jenny, B. Patsy, A. M. Arnold, P. Buzkova, J. Haines, D. Murdock, K. Glenn, K. Brown-Gentry, T. Thornton-Wells, L. Dumitrescu, W. S. Bush, S. L. Mitchell, R. Goodloe, S. Wilson, J. Boston, J. Malinowski, N. Restrepo, M. Oetjens, J. Fowke, K. Spencer, S. Pendergrass, L. Le Marchand, L. Park, M. Tiirikainen, L. Kolonel, I. Cheng, H. Wang, R. Shohet, D. Stram, B. Henderson, K. Monroe, G. Anderson, C. Carlson, R. Prentice, A. Lacroix, C. Wu, C. Carty, S. Rosse, A. Young, J. Kocarnik, Y. Lin, R. Jackson, D. Duggan, L. Kuller, C. He, P. Sulem, M. Barbalic, L. Broer, E. M. Byrne, D. F. Gudbjartsson, P. F. McArdle, E. Porcu, S. van Wingerden, W. V. Zhuang, L. B. Lauc, F. J. Broekmans, A. Burri, C. Chen, T. Corre, A. D. Coviello, P. D'adamo, G. Davies, I. J. Deary, S. Ebrahim, B. C. Fauser, L. Ferreli, A. R. Folsom,

S. E. Hankinson, M. Hass, A. C. Janssens, D. Karasik, J. Keyzer, D. P. Kiel, J. Lahti, S. Lai, T. Laisk, J. S. Laven, J. Liu, L. M. Lopez, Y. V. Louwers, M. Marongiu, I. M. Klaric, C. Masciullo, D. Melzer, A. B. Newman, G. Paré, P. H. Peeters, V. J. Pop, K. Räikkönen, A. Salumets, S. N. Stacey, J. M. Starr, M. G. Stathopoulou, U. Styrkarsdottir, A. Tenesa, L. Tryggvadottir, K. Tsui, R. M. van Dam, C. H. van Gils, P. van Nierop, J. M. Vink, M. Voorhuis, E. Widen, C. J. Wijnands-Van Gent, L. M. Yerges-Armstrong, L. Zgaga, M. Zygmunt, J. E. Buring, L. Crisponi, E. W. Demerath, E. A. Streeten, A. Murray, J. A. Visser, K. L. Lunetta, C. E. Elks, D. L. Cousminer, D. L. Koller, P. Lin, E. N. Smith, N. M. Warrington, H. Alavere, G. S. Berenson, H. Blackburn, F. Busonero, W. Chen, D. Couper, D. F. Easton, J. Eriksson, T. Foroud, T. O. Kilpeläinen, S. Li, S. S. Murray, A. R. Ness, K. Northstone, M. Peacock, C. E. Pennell, P. Pharoah, T. Rafnar, J. P. Rice, S. M. Ring, N. J. Schork, A. V. Segrè, U. Sovio, S. R. Srinivasan, M. L. Tammesoo, J. B. van Meurs, L. Young, L. J. Bierut, and M. J. Econs. Genetic studies of body mass index yield new insights for obesity biology. *Nature*, 518(7538):197–206. Publisher: Nature Publishing Group.

[55] Nellie Y. Loh, Matt J. Neville, Kyriakoula Marinou, Sarah A. Hardcastle, Barbara A. Fielding, Emma L. Duncan, Mark I. McCarthy, Jonathan H. Tobias, Celia L. Gregson, Fredrik Karpe, and Constantinos Christodoulides. LRP5 regulates human body fat distribution by modulating adipose progenitor biology in a dose- and depot-specific fashion. *Cell Metabolism*, 21(2):262–273.

[56] John Lonsdale, Jeffrey Thomas, Mike Salvatore, Rebecca Phillips, Edmund Lo, Saboor Shad, Richard Hasz, Gary Walters, Fernando Garcia, Nancy Young, Barbara Foster, Mike Moser, Ellen Karasik, Bryan Gillard, Kimberley Ramsey, Susan Sullivan, Jason Bridge, Harold Magazine, John Syron, Johnelle Fleming, Laura Siminoff, Heather Traino, Maghboeba Mosavel, Laura Barker, Scott Jewell, Dan Rohrer, Dan Maxim, Dana Filkins, Philip Harbach, Eddie Cortadillo, Bree Berghuis, Lisa Turner, Eric Hudson, Kristin Feenstra, Leslie Sobin, James Robb, Phillip Branton, Greg Korzeniewski, Charles Shive, David Tabor, Liqun Qi, Kevin Groch, Sreenath Nampally, Steve Buia, Angela Zimmerman, Anna Smith, Robin Burges, Karna Robinson, Kim Valentino, Deborah Bradbury, Mark Cosentino, Norma Diaz-Mayoral, Mary Kennedy, Theresa Engel, Penelope Williams, Kenyon Erickson, Kristin Ardlie, Wendy Winckler, Gad Getz, David DeLuca, Daniel MacArthur, Manolis Kellis, Alexander Thomson, Taylor Young, Ellen Gelfand, Molly Donovan, Yan Meng, George Grant, Deborah Mash, Yvonne Marcus, Margaret Basile, Jun Liu, Jun Zhu, Zhidong Tu, Nancy J. Cox, Dan L. Nicolae, Eric R. Gamazon, Hae Kyung Im, Anuar Konkashbaev, Jonathan Pritchard, Matthew Stevens, Timothèe Flutre, Xiaoquan Wen, Emmanouil T. Dermitzakis, Tuuli Lappalainen, Roderic Guigo, Jean Monlong, Michael Sammeth, Daphne Koller, Alexis Battle, Sara Mostafavi, Mark McCarthy, Manual Rivas, Julian Maller, Ivan Rusyn, Andrew Nobel, Fred Wright, Andrey Shabalin, Mike Feolo, Nataliya Sharopova, Anne Sturcke, Justin Paschal, James M. Anderson, Elizabeth L. Wilder, Leslie K. Derr, Eric D. Green, Jeffery P. Struewing, Gary Temple, Simona Volpi, Joy T. Boyer, Elizabeth J. Thomson, Mark S. Guyer, Cathy Ng, As-

sya Abdallah, Deborah Colantuoni, Thomas R. Insel, Susan E. Koester, A. Roger Little, Patrick K. Bender, Thomas Lehner, Yin Yao, Carolyn C. Compton, Jimmie B. Vaught, Sherilyn Sawyer, Nicole C. Lockhart, Joanne Demchok, and Helen F. Moore. The genotype-tissue expression (GTEx) project. *Nature Genetics*, 45(6):580–585. Bandiera_abtest: a Cc_license_type: cc_y Cg_type: Nature Research Journals Number: 6 Primary_atype: Comments & Opinion Publisher: Nature Publishing Group Subject_term: Databases;Genetics Subject_term_id: databases;genetics.

[57] Luca A Lotta, Pawan Gulati, Felix R Day, Felicity Payne, Halit Ongen, Martijn van de Bunt, Kyle J Gaulton, John D Eicher, Stephen J Sharp, Jian'an Luan, Emanuella De Lucia Rolfe, Isobel D Stewart, Eleanor Wheeler, Sara M Willems, Claire Adams, Hanieh Yaghootkar, Nita G Forouhi, Kay-Tee Khaw, Andrew D Johnson, Robert K Semple, Timothy Frayling, John R B Perry, Emmanouil Dermitzakis, Mark I McCarthy, Inês Barroso, Nicholas J Wareham, David B Savage, Claudia Langenberg, Stephen O'Rahilly, Robert A Scott, Stephen O'Rahilly, and Robert A Scott. Integrative genomic analysis implicates limited peripheral adipose storage capacity in the pathogenesis of human insulin resistance. *Nature Genetics*, 49(1):17–26.

[58] Changxue Lu and Sheue-Yann Cheng. Thyroid hormone receptors regulate adipogenesis and carcinogenesis via crosstalk signaling with peroxisome proliferator-activated receptors. *Journal of Molecular Endocrinology*, 44(3):143–154. Publisher: Society for Endocrinology.

[59] Vincent J. Lynch, Mauris C. Nnamani, Aurélie Kapusta, Kathryn Brayer, Silvia L. Plaza, Erik C. Mazur, Deena Emera, Shehzad Z. Sheikh, Frank Grützner, Stefan Bauersachs, Alexander Graf, Steven L. Young, Jason D. Lieb, Francesco J. DeMayo, Cédric Feschotte, and Günter P. Wagner. Ancient transposable elements transformed the uterine regulatory landscape and transcriptome during the evolution of mammalian pregnancy. *Cell reports*, 10(4):551–561.

[60] Nicholas Mancuso, Malika K. Freund, Ruth Johnson, Huwenbo Shi, Gleb Kichaev, Alexander Gusev, and Bogdan Pasaniuc. Probabilistic fine-mapping of transcriptome-wide association studies. *Nature Genetics*, 51(4):675–682. Bandiera_abtest: a Cg_type: Nature Research Journals Number: 4 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Gene expression;Genome-wide association studies;Software Subject_term_id: gene-expression;genome-wide-association-studies;software.

[61] Franck Mauvais-Jarvis, Noel Bairey Merz, Peter J. Barnes, Roberta D. Brinton, Juan-Jesus Carrero, Dawn L. DeMeo, Geert J. De Vries, C. Neill Epperson, Ramaswamy Govindan, Sabra L. Klein, Amedeo Lonardo, Pauline M. Maki, Louise D. McCullough, Vera Regitz-Zagrosek, Judith G. Regensteiner, Joshua B. Rubin, Kathryn Sandberg, and Ayako Suzuki. Sex and gender: modifiers of health, disease, and medicine. *The Lancet*, 396(10250):565–582. Publisher: Elsevier.

[62] Alexandre Melnikov, Xiaolan Zhang, Peter Rogov, Li Wang, and Tarjei S Mikkelsen. Massively parallel reporter assays in cultured mammalian cells. *Journal of visualized experiments : JoVE*, (90). Publisher: MyJoVE Corporation.

[63] David Merrick, Alexander Sakers, Zhazira Irgebay, Chihiro Okada, Catherine Calvert, Michael P. Morley, Ivona Percec, and Patrick Seale. Identification of a mesenchymal progenitor cell hierarchy in adipose tissue. *Science*, 364(6438). Publisher: American Association for the Advancement of Science Section: Research Article.

[64] Borbala Mifsud, Filipe Tavares-Cadete, Alice N. Young, Robert Sugar, Stefan Schoenfelder, Lauren Ferreira, Steven W. Wingett, Simon Andrews, William Grey, Philip A. Ewels, Bram Herman, Scott Happe, Andy Higgs, Emily LeProust, George A. Follows, Peter Fraser, Nicholas M. Luscombe, and Cameron S. Osborne. Mapping long-range promoter contacts in human cells with high-resolution capture hi-c. *Nature Genetics*, 47(6):598–606. Number: 6 Publisher: Nature Publishing Group.

[65] Meritxell Oliva, Manuel Muñoz-Aguirre, Sarah Kim-Hellmuth, Valentin Wucher, Ariel D. H. Gewirtz, Daniel J. Cotter, Princy Parsana, Silva Kasela, Brunilda Balliu, Ana Viñuela, Stephane E. Castel, Pejman Mohammadi, François Aguet, Yuxin Zou, Ekaterina A. Khramtsova, Andrew D. Skol, Diego Garrido-Martín, Ferran Reverter, Andrew Brown, Patrick Evans, Eric R. Gamazon, Anthony Payne, Rodrigo Bonazzola, Alvaro N. Barbeira, Andrew R. Hamel, Angel Martinez-Perez, José Manuel Soria, GTEx Consortium§, Brandon L. Pierce, Matthew Stephens, Eleazar Eskin, Emmanouil T. Dermitzakis, Ayellet V. Segrè, Hae Kyung Im, Barbara E. Engelhardt, Kristin G. Ardlie, Stephen B. Montgomery, Alexis J. Battle, Tuuli Lappalainen, Roderic Guigó, and Barbara E. Stranger. The impact of sex on gene expression across human tissues. *Science*, 369(6509). Publisher: American Association for the Advancement of Science Section: Research Article.

[66] Halit Ongen, Alfonso Buil, Andrew Anand Brown, Emmanouil T. Dermitzakis, and Olivier Delaneau. Fast and efficient QTL mapper for thousands of molecular phenotypes. *Bioinformatics*, 32(10):1479–1485.

[67] S. Paeratakul, J. C. Lovejoy, D. H. Ryan, and G. A. Bray. The relation of gender, race and socioeconomic status to obesity and obesity comorbidities in a sample of US adults. *International Journal of Obesity*, 26(9):1205–1210. Number: 9 Publisher: Nature Publishing Group.

[68] Biff F Palmer and Deborah J Clegg. The sexual dimorphism of obesity. *Molecular and cellular endocrinology*, 402:113–9. Publisher: NIH Public Access.

[69] Paz Polak and Eytan Domany. Alu elements contain many binding sites for transcription factors and may play a role in regulation of developmental processes. *BMC Genomics*, 7:133.

[70] Herman Pontzer, Mary H. Brown, David A. Raichlen, Holly Dunsworth, Brian Hare, Kara Walker, Amy Luke, Lara R. Dugas, Ramon Durazo-Arvizu, Dale Schoeller, Jacob Plange-Rhule, Pascal Bovet, Terrence E. Forrester, Estelle V. Lambert, Melissa Emery Thompson, Robert W. Shumaker, and Stephen R. Ross. Metabolic acceleration and the evolution of human brain size and life history. *Nature*, 533(7603):390–392. Number: 7603 Publisher: Nature Publishing Group.

[71] Baoming Qin, Miao He, Xiao Chen, and Duanqing Pei. Sorting nexin 10 induces giant vacuoles in mammalian cells. *Journal of Biological Chemistry*, 281(48):36891–36896.

[72] Nicole H. Rogers, James W. Perfield, II, Katherine J. Strissel, Martin S. Obin, and Andrew S. Greenberg. Reduced energy expenditure and increased inflammation are early events in the development of ovariectomy-induced obesity. *Endocrinology*, 150(5):2161–2168.

[73] Meena Kishore Sakharkar, Vincent T. K. Chow, and Pandjassarame Kangueane. Distributions of exons and introns in the human genome. *In Silico Biology*, 4(4):387–393.

[74] Sylvia Santosa, Nikki C. Bush, and Michael D. Jensen. Acute testosterone deficiency alters adipose tissue fatty acid storage. *The Journal of Clinical Endocrinology and Metabolism*, 102(8):3056–3064.

[75] † Sara and L L Pulit. Meta-analysis of genome-wide association studies for body fat distribution in 694 649 individuals of european ancestry. *Human Molecular Genetics*, 28(1):166–174.

[76] Jong Bae Seo, Hyang Mi Moon, Woo Sik Kim, Yun Sok Lee, Hyun Woo Jeong, Eung Jae Yoo, Jungyeob Ham, Heonjoong Kang, Myoung-Gyu Park, Knut R. Steffensen, Thomas M. Stulnig, Jan-Åke Gustafsson, Sang Dai Park, and Jae Bum Kim. Activated liver x receptors stimulate adipocyte differentiation through induction of peroxisome proliferator-activated receptor γ expression. *Molecular and Cellular Biology*, 24(8):3430–3444.

[77] Dmitry Shungin, Thomas W. Winkler, Damien C. Croteau-Chonka, Teresa Ferreira, Adam E. Locke, Reedik Mägi, Rona J. Strawbridge, Tune H. Pers, Krista Fischer, Anne E. Justice, Tsegaselassie Workalemahu, Joseph M. W. Wu, Martin L. Buchkovich, Nancy L. Heard-Costa, Tamara S. Roman, Alexander W. Drong, Ci Song, Stefan Gustafsson, Felix R. Day, Tonu Esko, Tove Fall, Zoltán Kutalik, Jian'an Luan, Joshua C. Randall, André Scherag, Sailaja Vedantam, Andrew R. Wood, Jin Chen, Rudolf Fehrmann, Juha Karjalainen, Bratati Kahali, Ching-Ti Liu, Ellen M. Schmidt, Devin Absher, Najaf Amin, Denise Anderson, Marian Beekman, Jennifer L. Bragg-Gresham, Steven Buyske, Ayse Demirkan, Georg B. Ehret, Mary F. Feitosa, Anuj Goel, Anne U. Jackson, Toby Johnson, Marcus E. Kleber, Kati Kristiansson, Massimo Mangino, Irene Mateo Leach, Carolina Medina-Gomez, Cameron D. Palmer, Dorota Pasko, Sonali Pechlivanis, Marjolein J. Peters, Inga Prokopenko, Alena Stančáková, Yun Ju Sung, Toshiko Tanaka, Alexander Teumer, Jana V. Van Vliet-Ostaptchouk,

Loïc Yengo, Weihua Zhang, Eva Albrecht, Johan Ärnlöv, Gillian M. Arscott, Stefania Bandinelli, Amy Barrett, Claire Bellis, Amanda J. Bennett, Christian Berne, Matthias Blüher, Stefan Böhringer, Fabrice Bonnet, Yvonne Böttcher, Marcel Bruinenberg, Delia B. Carba, Ida H. Caspersen, Robert Clarke, E. Warwick Daw, Joris Deelen, Ewa Deelman, Graciela Delgado, Alex S. F. Doney, Niina Eklund, Michael R. Erdos, Karol Estrada, Elodie Eury, Nele Friedrich, Melissa E. Garcia, Vilmantas Giedraitis, Bruna Gigante, Alan S. Go, Alain Golay, Harald Grallert, Tanja B. Grammer, Jürgen Gräßler, Jagvir Grewal, Christopher J. Groves, Toomas Haller, Goran Hallmans, Catharina A. Hartman, Maija Hassinen, Caroline Hayward, Kauko Heikkilä, Karl-Heinz Herzig, Quinta Helmer, Hans L. Hillege, Oddgeir Holmen, Steven C. Hunt, Aaron Isaacs, Till Ittermann, Alan L. James, Ingegerd Johansson, Thorhildur Juliusdottir, Ioanna-Panagiota Kalafati, Leena Kinnunen, Wolfgang Koenig, Ishminder K. Kooner, Wolfgang Kratzer, Claudia Lamina, Karin Leander, Nanette R. Lee, Peter Lichtner, Lars Lind, Jaana Lindström, Stéphane Lobbens, Mattias Lorentzon, François Mach, Patrik K. E. Magnusson, Anubha Mahajan, Wendy L. McArdle, Cristina Menni, Sigrun Merger, Evelin Mihailov, Lili Milani, Rebecca Mills, Alireza Moayyeri, Keri L. Monda, Simon P. Mooijaart, Thomas W. Mühleisen, Antonella Mulas, Gabriele Müller, Martina Müller-Nurasyid, Ramaiah Nagaraja, Michael A. Nalls, Narisu Narisu, Nicola Glorioso, Ilja M. Nolte, Matthias Olden, Nigel W. Rayner, Frida Renstrom, Janina S. Ried, Neil R. Robertson, Lynda M. Rose, Serena Sanna, Hubert Scharnagl, Salome Scholtens, Bengt Sennblad, Thomas Seufferlein, Colleen M. Sitlani, Albert Vernon Smith, Kathleen Stirrups, Heather M. Stringham, Johan Sundström, Morris A. Swertz, Amy J. Swift, Ann-Christine Syvänen, Bamidele O. Tayo, Barbara Thorand, Gudmar Thorleifsson, Andreas Tomaschitz, Chiara Troffa, Floor V. A. van Oort, Niek Verweij, Judith M. Vonk, Lindsay L. Waite, Roman Wennauer, Tom Wilsgaard, Mary K. Wojczynski, Andrew Wong, Qunyuan Zhang, Jing Hua Zhao, Eoin P. Brennan, Murim Choi, Per Eriksson, Lasse Folkersen, Anders Franco-Cereceda, Ali G. Gharavi, Åsa K. Hedman, Marie-France Hivert, Jinyan Huang, Stavroula Kanoni, Fredrik Karpe, Sarah Keildson, Krzysztof Kiryluk, Liming Liang, Richard P. Lifton, Baoshan Ma, Amy J. McKnight, Ruth McPherson, Andres Metspalu, Josine L. Min, Miriam F. Moffatt, Grant W. Montgomery, Joanne M. Murabito, George Nicholson, Dale R. Nyholt, Christian Olsson, John R. B. Perry, Eva Reinmaa, Rany M. Salem, Niina Sandholm, Eric E. Schadt, Robert A. Scott, Lisette Stolk, Edgar E. Vallejo, Harm-Jan Westra, Krina T. Zondervan, Philippe Amouyel, Dominique Arveiler, Stephan J. L. Bakker, John Beilby, Richard N. Bergman, John Blangero, Morris J. Brown, Michel Burnier, Harry Campbell, Aravinda Chakravarti, Peter S. Chines, Simone Claudi-Boehm, Francis S. Collins, Dana C. Crawford, John Danesh, Ulf de Faire, Eco J. C. de Geus, Marcus Dörr, Raimund Erbel, Johan G. Eriksson, Martin Farrall, Ele Ferrannini, Jean Ferrières, Nita G. Forouhi, Terrence Forrester, Oscar H. Franco, Ron T. Gansevoort, Christian Gieger, Vilmundur Gudnason, Christopher A. Haiman, Tamara B. Harris, Andrew T. Hattersley, Markku Heliövaara, Andrew A. Hicks, Aroon D. Hingorani, Wolfgang Hoffmann, Albert Hofman, Georg Homuth, Steve E. Humphries, Elina Hyppönen, Thomas Illig, Marjo-Riitta Jarvelin, Berit Johansen, Pekka Jousi-

lahti, Antti M. Jula, Jaakko Kaprio, Frank Kee, Sirkka M. Keinanen-Kiukaanniemi, Jaspal S. Kooner, Charles Kooperberg, Peter Kovacs, Aldi T. Kraja, Meena Kumari, Kari Kuulasmaa, Johanna Kuusisto, Timo A. Lakka, Claudia Langenberg, Loic Le Marchand, Terho Lehtimäki, Valeriya Lyssenko, Satu Männistö, André Marette, Tara C. Matise, Colin A. McKenzie, Barbara McKnight, Arthur W. Musk, Stefan Möhlenkamp, Andrew D. Morris, Mari Nelis, Claes Ohlsson, Albertine J. Oldehinkel, Ken K. Ong, Lyle J. Palmer, Brenda W. Penninx, Annette Peters, Peter P. Pramstaller, Olli T. Raitakari, Tuomo Rankinen, D. C. Rao, Treva K. Rice, Paul M. Ridker, Marylyn D. Ritchie, Igor Rudan, Veikko Salomaa, Nilesh J. Samani, Jouko Saramies, Mark A. Sarzynski, Peter E. H. Schwarz, Alan R. Shuldiner, Jan A. Staessen, Valgerdur Steinthorsdottir, Ronald P. Stolk, Konstantin Strauch, Anke Tönjes, Angelo Tremblay, Elena Tremoli, Marie-Claude Vohl, Uwe Völker, Peter Vollenweider, James F. Wilson, Jacqueline C. Witteman, Linda S. Adair, Murielle Bochud, Bernhard O. Boehm, Stefan R. Bornstein, Claude Bouchard, Stéphane Cauchi, Mark J. Caulfield, John C. Chambers, Daniel I. Chasman, Richard S. Cooper, George Dedoussis, Luigi Ferrucci, Philippe Froguel, Hans-Jörgen Grabe, Anders Hamsten, Jennie Hui, Kristian Hveem, Karl-Heinz Jöckel, Mika Kivimaki, Diana Kuh, Markku Laakso, Yongmei Liu, Winfried März, Patricia B. Munroe, Inger Njølstad, Ben A. Oostra, Colin N. A. Palmer, Nancy L. Pedersen, Markus Perola, Louis Pérusse, Ulrike Peters, Chris Power, Thomas Quertermous, Rainer Rauramaa, Fernando Rivadeneira, Timo E. Saaristo, Danish Saleheen, Juha Sinisalo, P. Eline Slagboom, Harold Snieder, Tim D. Spector, Unnur Thorsteinsdottir, Michael Stumvoll, Jaakko Tuomilehto, André G. Uitterlinden, Matti Uusitupa, Pim van der Harst, Giovanni Veronesi, Mark Walker, Nicholas J. Wareham, Hugh Watkins, H-Erich Wichmann, Goncalo R. Abecasis, Themistocles L. Assimes, Sonja I. Berndt, Michael Boehnke, Ingrid B. Borecki, Panos Deloukas, Lude Franke, Timothy M. Frayling, Leif C. Groop, David J. Hunter, Robert C. Kaplan, Jeffrey R. O'Connell, Lu Qi, David Schlessinger, David P. Strachan, Kari Stefansson, Cornelia M. van Duijn, Cristen J. Willer, Peter M. Visscher, Jian Yang, Joel N. Hirschhorn, M. Carola Zillikens, Mark I. McCarthy, Elizabeth K. Speliotes, Kari E. North, Caroline S. Fox, Inês Barroso, Paul W. Franks, Erik Ingelsson, Iris M. Heid, Ruth J. F. Loos, L. Adrienne Cupples, Andrew P. Morris, Cecilia M. Lindgren, Karen L Mohlke, Mark I. McCarthy, Elizabeth K. Speliotes, Kari E. North, Caroline S. Fox, Inês Barroso, Paul W. Franks, Erik Ingelsson, Iris M. Heid, Ruth J. F. Loos, L. Adrienne Cupples, Andrew P. Morris, Cecilia M. Lindgren, and Karen L Mohlke. New genetic loci link adipose and insulin biology to body fat distribution. *Nature*, 518(7538):187–196. Publisher: Nature Publishing Group.

[78] Rasmus Siersbæk, Ronni Nielsen, and Susanne Mandrup. PPARγ in adipocyte differentiation and metabolism – novel insights from genome-wide studies. *FEBS Letters*, 584(15):3242–3249.

[79] Fabian Sievers, Andreas Wilm, David Dineen, Toby J Gibson, Kevin Karplus, Weizhong Li, Rodrigo Lopez, Hamish McWilliam, Michael Remmert, Johannes Söding, Julie D Thompson, and Desmond G Higgins. Fast, scalable generation of high-quality

protein multiple sequence alignments using clustal omega. *Molecular Systems Biology*, 7:539.

[80] Nasa Sinnott-Armstrong, Isabel S. Sousa, Samantha Laber, Elizabeth Rendina-Ruedy, Simon E. Nitter Dankel, Teresa Ferreira, Gunnar Mellgren, David Karasik, Manuel Rivas, Jonathan Pritchard, Anyonya R. Guntur, Roger D. Cox, Cecilia M. Lindgren, Hans Hauner, Richard Sallari, Clifford J. Rosen, Yi-Hsiang Hsu, Eric S. Lander, Douglas P. Kiel, and Melina Claussnitzer. A regulatory variant at 3q21.1 confers an increased pleiotropic risk for hyperglycemia and altered bone mineral density. *Cell Metabolism*, 33(3):615–628.e13.

[81] Thomas Skurk, Catherine Alberti-Huber, Christian Herder, and Hans Hauner. Relationship between adipocyte size and adipokine expression and secretion. *The Journal of Clinical Endocrinology and Metabolism*, 92(3):1023–1033.

[82] Thomas Skurk and Hans Hauner. Primary culture of human adipocyte precursor cells: expansion and differentiation. *Methods in Molecular Biology (Clifton, N.J.)*, 806:215–226.

[83] Kerrin S. Small, Marijana Todorčević, Mete Civelek, Julia S. El-Sayed Moustafa, Xiao Wang, Michelle M. Simon, Juan Fernandez-Tajes, Anubha Mahajan, Momoko Horikoshi, Alison Hugill, Craig A. Glastonbury, Lydia Quaye, Matt J. Neville, Siddharth Sethi, Marianne Yon, Calvin Pan, Nam Che, Ana Viñuela, Pei-Chien Tsai, Abhishek Nag, Alfonso Buil, Gudmar Thorleifsson, Avanthi Raghavan, Qiurong Ding, Andrew P. Morris, Jordana T. Bell, Unnur Thorsteinsdottir, Kari Stefansson, Markku Laakso, Ingrid Dahlman, Peter Arner, Anna L. Gloyn, Kiran Musunuru, Aldons J. Lusis, Roger D. Cox, Fredrik Karpe, and Mark I. McCarthy. Regulatory variants at KLF14 influence type 2 diabetes risk via a female-specific effect on adipocyte size and body composition. *Nature Genetics*, 50(4):572–580. Number: 4 Publisher: Nature Publishing Group.

[84] Scott Smemo, Juan J. Tena, Kyoung-Han Kim, Eric R. Gamazon, Noboru J. Sakabe, Carlos Gómez-Marín, Ivy Aneas, Flavia L. Credidio, Débora R. Sobreira, Nora F. Wasserman, Ju Hee Lee, Vijitha Puviindran, Davis Tam, Michael Shen, Joe Eun Son, Niki Alizadeh Vakili, Hoon-Ki Sung, Silvia Naranjo, Rafael D. Acemel, Miguel Manzanares, Andras Nagy, Nancy J. Cox, Chi-Chung Hui, Jose Luis Gomez-Skarmeta, and Marcelo A. Nóbrega. Obesity-associated variants within FTO form long-range functional connections with IRX3. *Nature*, 507(7492):371–375. Publisher: Nature Publishing Group.

[85] Katherine J. Strissel, Zlatina Stancheva, Hideaki Miyoshi, James W. Perfield, Jason DeFuria, Zoe Jick, Andrew S. Greenberg, and Martin S. Obin. Adipocyte death, adipose tissue remodeling, and obesity complications. *Diabetes*, 56(12):2910–2918. Publisher: American Diabetes Association Section: Obesity Studies.

[86] Cathie Sudlow, John Gallacher, Naomi Allen, Valerie Beral, Paul Burton, John Danesh, Paul Downey, Paul Elliott, Jane Green, Martin Landray, Bette Liu, Paul Matthews, Giok Ong, Jill Pell, Alan Silman, Alan Young, Tim Sprosen, Tim Peakman, and Rory Collins. UK biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLOS Medicine*, 12(3):e1001779. Publisher: Public Library of Science.

[87] Anatoly T. Teriokhin, Elena V. Budilova, Frederic Thomas, and Jean-Francois Guegan. Worldwide variation in life-span sexual dimorphism and sex-specific environmental mortality rates. *Human Biology*, 76(4):623–641.

[88] Gloria Terrados, Florian Finkernagel, Bastian Stielow, Dennis Sadic, Juliane Neubert, Olga Herdt, Michael Krause, Maren Scharfe, Michael Jarek, and Guntram Suske. Genome-wide localization and expression profiling establish sp2 as a sequence-specific transcription factor regulating vitally important genes. *Nucleic Acids Research*, 40(16):7844–7857. Publisher: Oxford Academic.

[89] Marco Trizzino, YoSon Park, Marcia Holsbach-Beltrame, Katherine Aracena, Katelyn Mika, Minal Caliskan, George H. Perry, Vincent J. Lynch, and Christopher D. Brown. Transposable elements are the primary source of novelty in primate gene regulation. *Genome Research*, 27(10):1623–1633.

[90] Valter Tucci, Anthony R. Isles, Gavin Kelsey, Anne C. Ferguson-Smith, Valter Tucci, Marisa S. Bartolomei, Nissim Benvenisty, Déborah Bourc'his, Marika Charalambous, Catherine Dulac, Robert Feil, Juliane Glaser, Lisa Huelsmann, Rosalind M. John, Gráinne I. McNamara, Kim Moorwood, Francoise Muscatelli, Hiroyuki Sasaki, Beverly I. Strassmann, Claudius Vincenz, Jon Wilkins, Anthony R. Isles, Gavin Kelsey, and Anne C. Ferguson-Smith. Genomic imprinting and physiological processes in mammals. *Cell*, 176(5):952–965. Publisher: Elsevier.

[91] Jacob C. Ulirsch, Satish K. Nandakumar, Li Wang, Felix C. Giani, Xiaolan Zhang, Peter Rogov, Alexandre Melnikov, Patrick McDonel, Ron Do, Tarjei S. Mikkelsen, and Vijay G. Sankaran. Systematic functional dissection of common genetic variation affecting red blood cell traits. *Cell*, 165(6):1530–1545.

[92] Lianne van Beek, Jan B. van Klinken, Amanda C. M. Pronk, Andrea D. van Dam, Eline Dirven, Patrick C. N. Rensen, Frits Koning, Ko Willems van Dijk, and Vanessa van Harmelen. The limited storage capacity of gonadal adipose tissue directs the development of metabolic disorders in male c57bl/6j mice. *Diabetologia*, 58(7):1601–1609.

[93] Juan M. Vaquerizas, Sarah K. Kummerfeld, Sarah A. Teichmann, and Nicholas M. Luscombe. A census of human transcription factors: function, expression and evolution. *Nature Reviews. Genetics*, 10(4):252–263.

[94] M. a. A. Vogel, P. Wang, F. G. Bouwman, N. Hoebers, E. E. Blaak, J. Renes, E. C. Mariman, and G. H. Goossens. A comparison between the abdominal and femoral adipose tissue proteome of overweight and obese women. *Scientific Reports*, 9(1):4202. Bandiera_abtest: a Cc_license_type: cc_by Cg_type: Nature Research Journals Number: 1 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Cardiovascular biology;Fat metabolism;Proteomics Subject_term_id: cardiovascular-biology;fat-metabolism;proteomics.

[95] Michael Wainberg, Nasa Sinnott-Armstrong, Nicholas Mancuso, Alvaro N. Barbeira, David A. Knowles, David Golan, Raili Ermel, Arno Ruusalepp, Thomas Quertermous, Ke Hao, Johan L. M. Björkegren, Hae Kyung Im, Bogdan Pasaniuc, Manuel A. Rivas, and Anshul Kundaje. Opportunities and challenges for transcriptome-wide association studies. *Nature Genetics*, 51(4):592–599. Bandiera_abtest: a Cg_type: Nature Research Journals Number: 4 Primary_atype: Reviews Publisher: Nature Publishing Group Subject_term: Genetics;Genomics Subject_term_id: genetics;genomics.

[96] Gao Wang, Abhishek Sarkar, Peter Carbonetto, and Matthew Stephens. A simple new approach to variable selection in regression, with application to genetic fine mapping. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 82(5):1273–1300. _eprint: https://rss.onlinelibrary.wiley.com/doi/pdf/10.1111/rssb.12388.

[97] Qiong A Wang, Caroline Tao, Rana K Gupta, and Philipp E Scherer. Tracking adipogenesis during white adipose tissue development, expansion and regeneration. *Nature medicine*, 19(10):1338–1344.

[98] Xing-Cheng Wei, Jun-ichi Dohkan, Hiroyuki Kishi, Chun-Xiao Wu, Sachiko Kondo, and Atsushi Muraguchi. Characterization of the proximal enhancer element and transcriptional regulatory factors for murine recombination activating gene-2. *European Journal of Immunology*, 35(2):612–621.

[99] Warren A. Whyte, David A. Orlando, Denes Hnisz, Brian J. Abraham, Charles Y. Lin, Michael H. Kagey, Peter B. Rahl, Tong Ihn Lee, and Richard A. Young. Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell*, 153(2):307–319.

[100] Zhao Xu, Songtao Yu, Chung-Hsin Hsu, Jun Eguchi, and Evan D. Rosen. The orphan nuclear receptor chicken ovalbumin upstream promoter-transcription factor II is a critical regulator of adipogenesis. *Proceedings of the National Academy of Sciences*, 105(7):2421–2426. Publisher: National Academy of Sciences Section: Biological Sciences.

[101] Liang Ye, Leslie R. Morse, Li Zhang, Hajime Sasaki, Jason C. Mills, Paul R. Odgren, Greg Sibbel, James R. L. Stanley, Gee Wong, Ariane Zamarioli, and Ricardo A. Battaglino. Osteopetrorickets due to snx10 deficiency in mice results from both failed osteoclast activity and loss of gastric acid-dependent calcium absorption. *PLOS Genetics*, 11(3):e1005057. Publisher: Public Library of Science.

[102] Loic Yengo, Julia Sidorenko, Kathryn E. Kemper, Zhili Zheng, Andrew R. Wood, Michael N. Weedon, Timothy M. Frayling, Joel Hirschhorn, Jian Yang, Peter M. Visscher, and GIANT Consortium. Meta-analysis of genome-wide association studies for height and body mass index in $\sim$700000 individuals of european ancestry. *Human Molecular Genetics*, 27(20):3641–3649.

[103] Kaifan Yu, Delin Mo, Ming Wu, Hu Chen, Luxi Chen, Mingsen Li, and Yaosheng Chen. Activating transcription factor 4 regulates adipocyte differentiation via altering the coordinate expression of CCATT/enhancer binding protein β and peroxisome proliferator-activated receptor γ. *The FEBS Journal*, 281(10):2399–2409. _eprint: https://febs.onlinelibrary.wiley.com/doi/pdf/10.1111/febs.12792.

[104] M Zamboni, F Armellini, MP Milani, M De Marchi, T Todesco, R Robbi, IA Bergamo-Andreis, and O Bosello. Body fat distribution in pre- and post-menopausal women: metabolic and anthropometric variables and their inter-relationships. *International journal of obesity and related metabolic disorders : journal of the International Association for the Study of Obesity*, 16(7):495–504.

[105] Xiaochang Zhang, Ming Hui Chen, Xuebing Wu, Andrew Kodani, Jean Fan, Ryan Doan, Manabu Ozawa, Jacqueline Ma, Nobuaki Yoshida, Jeremy F. Reiter, Douglas L. Black, Peter V. Kharchenko, Phillip A. Sharp, and Christopher A. Walsh. Cell-type-specific alternative splicing governs cell fate in the developing cerebral cortex. *Cell*, 166(5):1147–1162.e15.