CHARACTERIZING CARDIOMETABOLIC GWAS LOCI WITH REGULATORY ANNOTATION, REGULATORY ASSAYS, TRANS-ANCESTRY FINE-MAPPING, AND OPEN CHROMATIN PROFILING

Maren Ettinger Cannon

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Curriculum of Genetics and Molecular Biology in the School of Medicine.

Chapel Hill
2017

Approved by:

Karen Mohlke

Ian Davis

Praveen Sethupathy

Gregory Crawford

Christopher Mack

ABSTRACT

Maren Ettinger Cannon: Characterizing cardiometabolic GWAS loci with regulatory annotation, regulatory assays, trans-ancestry fine-mapping, and open chromatin profiling
(Under the direction of Karen Mohlke)

Cardiometabolic phenotypes, including diseases such as cardiovascular disease and type 2 diabetes (T2D) and related traits such as cholesterol levels, obesity, and lipid levels cause significant public health burden in many countries. Genetic and environmental factors contribute to the etiology of these phenotypes. While genome-wide association studies (GWAS) have identified hundreds of loci associated with cardiometabolic phenotypes, progress towards elucidating causal variants and genes has been slow. Many associated variants are located in noncoding regions, suggesting a regulatory mechanism. Identifying causal GWAS variants is a time-consuming and challenging process. Challenges arise from strong linkage disequilibrium (LD) at loci requiring multiple variants to be prioritized investigated. Combining statistical fine-mapping, overlap with genome-wide regulatory datasets, expression quantitative trait loci associations (eQTLs) and functional assays elucidates the variant(s) and gene(s) contributing to genetic mechanisms at cardiometabolic GWAS loci. I used these approaches to identify functional variants at two cardiometabolic GWAS loci and generated open chromatin profiles in adipose tissue. At the *CDC123/CAMK1D* T2D locus, one variant altered binding to FOXA1 and FOXA2 and increased transcriptional activity, suggesting it contributes to the mechanism at this locus. At the *ANGPTL8* high-density lipoprotein cholesterol (HDL-C) GWAS locus, seven variants showed allelic differences in functional assays, suggesting a more complex regulatory

mechanism that may include multiple variants. Finally, open chromatin profiles were generated from frozen human subcutaneous adipose samples and a preadipocyte strain to characterize the regulatory landscape of adipose tissue and a cellular model system used to test GWAS variants in functional assays. Thorough investigations of GWAS loci are necessary for the development of new therapies, full elucidation of direction of effect, and better understanding of the genetic contributions to cardiometabolic phenotypes.

ACKNOWLEDGMENTS

Words cannot adequately express my thanks to the many people who have supported me throughout my Ph.D. First, I must thank and recognize my advisor, Karen Mohlke. You are an incredible scientist, a brilliant writer, and a thorough and thoughtful mentor. You have contributed significantly to my scientific development and helped me improve my writing and communication skills. I am grateful for the experience I have gained in your lab through experimental endeavors, attempts at computational endeavors, and writing grants and papers. Thank you, Karen, for your unending support and incredible mentorship. I would also like to thank my committee: Ian Davis, Praveen Sethupathy, Chris Mack, and Greg Crawford for their helpful conversations, direction, mentorship, and advice.

Thank you to the fantastic members of the Mohlke lab, past and present, for being such amazing colleagues and friends! Thank you Marie Fogarty, Jen Kulzer, Tamara Roman, Rani Vadlamudi, Jim Davis, and Hannah Perrin for your support and advice, especially for helping to troubleshoot experiments. Thank you Martin Buchkovich, Ying Wu, Cassie Spracklen, Kevin Currin, and Raulie Raulerson for your help with countless computational analyses and questions. Thank you especially to Kevin Currin for your extensive work on the analyses in Chapter 4; this chapter would not have been possible without your expertise and work. Thank you to all of the Mohlke lab members for helping me in all aspects of graduate school.

Many thanks to our collaborators around the world who I have been privileged to work

with. Thank you for collecting and processing thousands of tissue and blood samples; my work would not be possible without your previous work. I have learned much from interacting with you on conference calls, at collaborator meetings, and at national conferences; I appreciate the collaborative environment you foster, and your genuine friendship and interest in my work. Thank you especially to those who have collected data, and performed experiments and statistical analyses in this dissertation.

Thank you to the many teachers, professors, and mentors throughout my life who cultivated my love of science, especially Mark Dickson, Keith Crandall, Heather Bracken-Grissom, and Maegan Leary. Your support, knowledge, and enthusiasm pushed me to try harder than I thought possible and sparked my interest in pursuing a Ph.D.

For their unwavering support throughout this process and my life, thank you Mom, Dad, Ames, and Kyle. I could not ask for a better family and words cannot express my sincere thanks for your support and love. Thank you to the more recent additions to my family- Carlie, Kimberli, John, Brooke, Jordan, Derek, Jenn, Kelsie, Seth, Lisa- for your support and love. Thanks to my daughter Quinn for being a light in my life and my very best genetics experiment.

Finally, and most importantly, thank you to my husband and best friend, Jake. Thank you for listening to practice talks and reviewing papers you didn't understand, hearing seemingly endless complaints about experiments, being a constant source of encouragement, helping to keep my spirits up, and for simply being there everyday. I love you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AA-ALIGNER	Allele aware aligner
ADCY5	adenylate cyclase 5
ADIPOQ	adiponectin
AFR	African ancestry
AGEN	Asian Genetic Epidemiology Network
AMR	Admixed American ancestry
ANGPTL3	angiopoiten-like protein 3
ANGPTL8	angiopoiten-like protein 8
ATAC-seq	assay for transposase-accessible chromatin
ATP2A1	ATPase sarcoplasmic/endoplasmic reticulum Ca2+ transporting protein 1
B2M	beta-2-microglobulin
BCA	bicinchoninic acid
BMI	body mass index
bp	base pair
CAMK1D	calcium/calmodulin-dependent protein kinase 1D
CAVIAR	Causal Variants Identication in Associated Regions
CDC123	cell division cycle protein 123
CDC42	cell division cycle protein 42
cDNA	complementary DNA
CEBPA	CCAAT/enhancer binding protein alpha
CEBPB	CCAAT/enhancer binding protein beta
CEBPD	CCAAT/enhancer binding protein delta

ChIP-seq	chromatin immunoprecipitation
CREB	cAMP response element binding protein
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CTCF	CCCTC-binding factor
CVD	cardiovascular disease
DMEM	Dulbecco's modified eagle medium
DNA	deoxyribonucleic acid
DNase-seq	DNase I hypersensitivity
DOCK6	dedicator of cytokinesis 6
EDTA	ethylenediaminetetraacetic acid
EGR1	early growth response factor 1
ELISA	enzyme-linked immunosorbent assay
EMMAX	Efficient Mixed-Model Association eXpedited
EMSA	electrophoretic mobility shift assay
ENCODE	Encyclopedia of DNA elements
EPACTS	Efficient and Parallelizable Association Container Toolbox
eQTL	expression quantitative trait locus
EUR	European ancestry
FAIRE-seq	formaldehyde-assisted isolation of regulatory elements
FBS	fetal bovine serum
FBN2	fibrillin 2
FDR	false discovery rate
FIMO	Find Individual Motif Occurances

FOXA1	forkhead box protein A1
FOXA2	forkhead box protein A2
FTO	FTO, alpha-ketoglutarate dependent dioxygenase
GALNT2	polypeptide N-acetylgalactosaminyltransferase 2
GLGC	Global Lipids Genetics Consortium
GREGOR	Genomic Regulatory Elements and GWAS Overlap algoRithm
GSIS	glucose stimulated insulin secretion
GSNAP	Genomic Short-read Nucleotide Alignment Program
GWAS	genome-wide association study
HDL-C	high-density lipoprotein cholesterol
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HOMA	Homeostasis Model Assessment
JAZF1	JAZF zinc finger 1
kb	kilobase
LD	linkage disequilibrium
LDL-C	low-density lipoprotein cholesterol
LiCl	lithium chloride
LPL	lipoprotein lipase
MACS	Model-based Analysis for ChIP-Seq
MAF	minor allele frequency
MALDI	matrix-assisted laser desorption/ionization
MANTRA	Meta-Analysis of Transethnic Association studies
MAPQ	mapping quality

Mb	megabase
MEM-a	Minimum Essential Medium Eagle alpha
METSIM	Metabolic Syndrome in Men study
$MgCl_2$	magnesium chloride
mM	millimolar
MMP16	matrix metallopeptidase 16
mRNA	messenger RNA
MTNR1B	melatonin receptor 1B
NaCl	sodium chloride
$NaHCO_3$	sodium bicarbonate
NCBI	National Center for Biotechnology Information
NE	nuclear extract
NIH	National Institutes of Health
NMR	nuclear magnetic resonance spectroscopy
PAGE	polyacrylamide gel electrophoresis
PAINTOR	Probabilistic Annotation INtegraTOR
PBS	Phosphate-buffered saline
PCR	polymerase chain reaction
PDX1	pancreatic and duodenal homeobox 1
PEER	probabilistic estimation of expresion residuals
PLA2G7	phospholipase A2 group VII
PNPLA2	patatin-like phopholipase domain containing 2
qPCR	quantitative PCR

RAC1	Rac family small GTPase 1
RIFL	refeeding-induced fat and liver protein
RMA	Robust multi-array average
RNA	ribonucleic acid
RPMI	Roswell Park Memorial Institute
RT-PCR	reverse transcription PCR
RXRα	retinoid x receptor alpha
SD	standard deviation
SE	standard error
SGBS	Simpson-Golabi-Behmel syndrome
SH2B1	SH2B adaptor protein 1
SNP	single nucleotide polymorphism
SNX10	sorting nexin 10
SPIB	Spi-B transcription factor
T2D	type 2 diabetes
TBE	Tris/Borate/EDTA
TRIB1	tribbles pseudokinase 1
TOF	time-of-flight
TPR	transposition probability ratio
TSS	transcription start site
TUFM	Tu translation elongation factor, mitochondrial protein
UCSC	University of California Santa Cruz
UTR	untranslated region

UV	ultraviolet
VGAM	Vector Generalized Linear and Additive Models
VWF	Von Willebrand factor
WHI	Women's Health Initiative cohort
WHI-SHARe	Women's Health Initiative SNP Health Association Resource
WHR	waist-hip ratio

CHAPTER 1: INTRODUCTION

Genome-wide association studies (GWAS) have identified hundreds of loci associated with cardiometabolic phenotypes (www.ebi.ac.uk/gwas/). Most of these loci are located in noncoding regions and the underlying functional mechanisms remain unknown. Statistical analyses and functional experiments are needed to determine the causal variant(s), which gene(s) they act on, and how the gene functions to alter cardiometabolic phenotypes. Robust genomic regulatory datasets exist to help inform and prioritize candidate variants, but some tissues, such as adipose, are under-represented. The aims of my research were to identify the functional variant(s) at the *CDC123/CAMK1D* type 2 diabetes (T2D) GWAS locus, identify the functional variant(s) and gene(s) at the *ANGPTL8* high-density lipoprotein cholesterol (HDL-C) GWAS locus, and to create a more robust regulatory map of adipose tissue by generating open chromatin profiles of human adipose tissue using the assay for transposase-accessible chromatin (ATAC-seq). Generating ATAC-seq open chromatin profiles from adipose tissue will enhance identification of functional variants. A thorough understanding of cardiometabolic GWAS loci is necessary to the development of therapies for the appropriate genes (1) and direction of effect (2), especially given the increasing recognition of allelic heterogeneity at GWAS loci (3-8).

Overview of cardiometabolic phenotypes

Cardiometabolic phenotypes encompass a wide range of diseases and related traits that cause significant public health burden in many countries (9-11). These phenotypes include T2D,

cardiovascular disease (CVD), cholesterol and triglyceride levels, and obesity. T2D is characterized by impaired glucose tolerance and/or impaired fasting glucose (12,13). These impairments can be caused by decreased insulin secretion from the pancreatic beta cells and insulin resistance. People affected with T2D can have multiple complications including vascular and circulatory abnormalities that can lead to coronary heart disease, congestive heart failure, or stroke (14,15).

CVD is a leading cause of death worldwide (16). CVD includes multiple conditions including heart failure, myocardial infarction, stroke, and high blood pressure (12). Risk factors for CVD include high levels of low-density lipoprotein cholesterol (LDL-C), high triglyceride levels, and low levels of high-density lipoprotein cholesterol (HDL-C) (17-19). LDL-C and triglycerides can contribute to the development of plaques in the bloodstream; the buildup of plaque in the bloodstream is called atherosclerosis and is a leading cause of CVD (20). HDL-C helps clear plaques by collecting LDL-C to take it away from the arteries and back to the liver for disposal, thus decreasing plaque buildup and risk for CVD.

Obesity is a significant public health burden and is extremely prevalent, especially in the United States. Approximately 35% of all adults in the U.S. are obese (21,22). Central adiposity, or fat deposited around the abdomen, is associated with CVD and general obesity increases risk of other comorbid conditions, including T2D, hypertension, dyslipidemia, osteoarthritis, several cancers, and many other detrimental phenotypes (21,23,24). CVD, T2D, and their related risk factors are regulated by genetic and environmental factors (25).

Genetics of cardiometabolic diseases and traits

Cardiometabolic phenotypes are multifactorial and are influenced by environmental and genetic factors. Exercise and diet are known to affect lipid and cholesterol levels and increased exercise and a healthy diet can decrease the risk of many cardiometabolic diseases. Bowes *et al.* (26) demonstrated that lifestyle intervention including dietician-led education sessions and weekly exercise programs led to weight loss and significant improvement in cardiovascular risk factors including lipid and triglyceride levels. Strong evidence exists for a genetic component to these phenotypes. Heritability estimates range from 30-77% for T2D (27,28), 40-90% for cholesterol and lipid levels (29-34), and 31-70% for waist-to-hip ratio, a measure of obesity (31,35-37). GWAS have identified hundreds of loci associated with cardiometabolic phenotypes (6,7,38). To date, at least 80 loci have been associated with CVD (39-41) and hundreds of loci (>250) have been identified for cholesterol and lipid levels and other CVD risk factors (38,42-46). At least 184 loci have been identified for T2D (47,48), and at least 200 loci have been identified for obesity (6,7). While mechanisms are known at a few loci, most mechanisms remain unknown. GWAS have been successful in identifying associated loci, but the next step is to move past association on to understanding the function of variant(s) and gene(s) at these loci (49,50).

Much more work is needed to elucidate the mechanisms at cardiometabolic GWAS loci. Of the hundreds of loci associated with cardiometabolic traits, only a few have been targeted for functional validation. One example, near the *TRIB1* gene, identified variants associated with HDL-C, LDL-C, triglyceride levels, and myocardial infarction (51). Further studies showed a connection of *TRIB1* gene function and CVD risk factors; *Trib1* overexpression in mouse liver decreased cholesterol and lipid levels (52). A second example, near *JAZF1,* identified variants

associated with T2D (53). A common variant, rs1635852 increased transcriptional activity in pancreatic beta cells by altering binding of the PDX1 transcription factor, suggesting that rs1635852 may contribute to T2D susceptibility. A third example described multiple regulatory variants at the *GALNT2* HDL-C GWAS locus (54). At least two variants, rs4846913 and rs2281721 influence *GALNT2* expression by altering transcription factor binding. Several other mechanisms at cardiometabolic GWAS loci have been elucidated, but hundreds remain unsolved. Characterizing functional mechanisms at GWAS loci is challenging and time-consuming, but is a vital step to understanding the genetics of cardiometabolic diseases and traits.

This dissertation focuses on elucidating novel genetic mechanisms of cardiometabolic diseases and related traits. As previously mentioned, GWAS are strong studies for identifying loci associated with cardiometabolic phenotypes, but they do not delineate which gene(s) and variant(s) are functioning to alter the phenotype. Many loci are located in noncoding regions, suggesting a regulatory mechanism. In most cases, multiple genes located near the association signals have plausible biological function. Similarly, any of the variants inherited together in linkage disequilibrium (LD) could be contributing to the underlying mechanism. Additional statistical analyses and functional studies are needed to determine which gene(s) and variant(s) are functional at these loci.

Prioritizing candidate variants at GWAS loci

Multiple approaches exist to prioritize candidate variants at a GWAS locus (55). As a first step, variants in LD with the most significantly associated GWAS variant, or lead variant, are considered. Variants in pairwise LD are inherited together more frequently than is expected

by chance; a statistical measure of LD is r^2, where r^2=1.0 means that two variant alleles are always inherited together. Traditionally, a LD r^2 threshold of 0.8 is used to determine candidate variants at a GWAS locus; however, various studies use different, less stringent thresholds (*i.e.*: r^2>0.5). All variants in strong LD with the lead variant are considered equally likely to be functional.

After determining the number of candidate variants in LD with the lead variant at a locus, statistical fine-mapping analyses and overlap with genomic datasets can be used to select variants for follow-up in functional experiments (55). Fine-mapping can have multiple definitions. For the purposes of this dissertation, I use two types of fine-mapping: to determine the number of signals present at an association locus and to statistically predict functional variants using association statistics and/or genomic annotations. For fine-mapping to be successful, the variants in LD with the lead variant need to be densely genotyped or imputed with high confidence in large enough sample sizes to differentiate between variants in strong pairwise LD. The first fine-mapping approach is to determine the number of signals present at a GWAS locus by performing conditional analysis. If other variants remain significant after conditioning on the lead variant, then there are additional signals present at the GWAS locus. Conditional analysis is important to ensure all candidate variants are considered. If multiple signals exist, there may be multiple mechanisms of action to alter the associated trait.

The second fine-mapping approach, statistical fine-mapping, predicts which variants are more likely to be functional. Statistical fine-mapping analyses fit in two broad categories: prioritizing variants based on association statistics and/or LD and Bayesian methods that assign posterior probabilities of functionality to each variant. In this dissertation, I employ three statistical fine-mapping analyses. CAVIAR (56) is a method that leverages association statistics

including p-values and effect sizes to determine which variants are most likely to be functional. PAINTOR (57) also leverages association strength to prioritize variants, but also includes functional genomic annotation data. One Bayesian method is MANTRA (58), which compares association signals across diverse populations to prioritize shared variants. CAVIAR and MANTRA generate a 'credible set' of variants most likely to contain the functional variant. PAINTOR predicts a set number of functional variants given by the user before analyzing the data. Statistical fine-mapping approaches can reach different predictions; it can be beneficial to compare variants predicted by multiple approaches to prioritize variants for functional followup.

Although some GWAS variants alter the protein coding sequence of genes, the majority of GWAS loci are located in noncoding regions. The variants underlying these association signals likely have regulatory function, by which the variant alters transcription of a gene through differential transcription factor binding or other mechanisms. Another approach for prioritizing variants for functional assays is to consider overlap with genomic regulatory regions. Large consortium efforts including the Encyclopedia of DNA Elements (ENCODE) (59) and the Roadmap Epigenomics Project (60) have created robust datasets for many cell and tissue types to describe regions characteristic of regulatory activity. These datasets include chromatin immunoprecipitation (ChIP-seq) of histone marks often observed at enhancers, promoters, and insulators, and transcription factors. Additionally, multiple datasets of open chromatin profiling using DNase hypersensitivity (DNase-seq), formaldehyde-assisted isolation of regulatory elements (FAIRE-seq), and ATAC-seq provide maps of 'open' regions of the genome. Regulatory datasets are useful for identifying regions of regulatory activity and the overlap of candidate variants with regulatory regions in the cell types of interest.

Regulatory datasets in cell and tissue types relevant to cardiometabolic diseases are needed to make the best prioritization of candidate variants because gene expression and regulatory regions can be cell-type specific (59). Many tissues, including liver, blood, adipose, pancreas, and others are involved in CVD, T2D, and obesity progression. Multiple assays have been performed in liver and blood cell types from the ENCODE (59) and the Roadmap Epigenomics Project (60), and recent studies have increased the amount of data for pancreatic islet (61). The Roadmap Epigenomics Project (60) has generated ChIP-seq data for histone marks, but open chromatin profiling in adipose tissue and adipocyte cell models is lacking. ATAC-seq is a particularly useful method to determine open chromatin profiles because it requires less tissue/cells and less time than DNase-seq and FAIRE-seq (62). In this dissertation, I create open chromatin profiles in three human adipose tissue samples and the SGBS preadipocyte cell strain. Characterization of the adipose open chromatin profile will help prioritize candidate variants at GWAS loci and contribute to a more complete understanding of gene regulation in adipose.

The gene(s) being acted on at regulatory GWAS loci also need to be identified. One method is by expression quantitative trait loci (eQTL) associations. eQTL analysis identifies variants associated with the RNA levels of nearby (*cis*-eQTL) or distant (*trans*-eQTL) genes. Further analysis can determine coincidence of a GWAS association signal with an eQTL signal. A coincident eQTL signal suggests that the same variants underlying the GWAS trait association are acting on the RNA levels of the eQTL-associated gene. In this dissertation, I use eQTL associations identified in the METabolic Syndrome in Men (METSIM) study (63). The METSIM study consists of ~10,000 Finnish men and includes dense genotyping, ~200 metabolic phenotypes, and eQTL associations in 770 adipose samples. The METSIM study is unique

because it includes genotypes, expression, and clinical trait data in the same samples. Combining these datasets allows us to identify gene expression associations with clinical traits; for example, increased *PLA2G7* expression level is associated with increased triglyceride levels (63). In the METSIM study, we can identify variant-trait, variant-gene, and gene-trait associations in the same set of individuals. If a variant is identified as an eQTL with a specific gene, functional experiments can be performed to confirm the role of the gene and determine the mechanism of action.

Functional assays of regulatory activity

Many approaches exist to test regulatory variants in functional experiments (64,65). The goal of all functional validation is to determine if variants show allelic differences in function, usually resulting in allele-specific effects on gene expression levels. Allelic differences can be identified in transcriptional reporter assays, protein binding assays, allelic imbalance in sequencing reads, genome editing experiments, physical chromatin interaction assays, and other experiments. Transcriptional reporter assays test variant alleles located in regulatory regions for differences in transcriptional activity. The regulatory region surrounding an associated variant is cloned into a vector containing a reporter gene, usually luciferase or GFP, and transfected in a cell line or transiently expressed in a model organism. Luciferase activity is measured and the variant alleles are compared to determine any allelic differences in transcriptional activity. In this dissertation, reporter assays are performed in cell lines, but they can also be performed in mouse, zebrafish, and other model organisms. For cardiometabolic disease loci and in this dissertation, I use cell lines derived from adipose, liver, and pancreatic islet. For adipose, mouse 3T3L1 preadipocytes, human SW872 adipocytes, human SGBS preadipocytes are used. I use two

human liver carcinoma cell lines: HepG2 and HUH7. Finally, mouse MIN6 and rat 832/13 insulinoma cells are used to model human pancreatic beta cells.

Protein binding assays, including electrophoretic mobility shift assays (EMSA), DNA-affinity pull downs, and ChIP experiments are used to identify variant alleles that bind transcription factors differentially. EMSAs are an *in vitro* approach that visualizes nuclear protein complexes that bind to ~20-bp DNA probes surrounding the candidate variant. The DNA-protein complexes are visualized on a gel and identity of the transcription factor can be determined using antibodies to transcription factors predicted by conserved transcription factor binding motifs or identified by ChIP-seq datasets. DNA-affinity pull downs are similar to EMSA; all DNA-protein complexes are captured by a ~20-bp probe surrounding the candidate variant and visualized on a gel. Proteins in allele-specific bands are identified using mass spectrometry.

Sequencing data generated from ChIP-seq, DNase-seq, FAIRE-seq, or ATAC-seq can be used to identify sites of allelic imbalance. A site of allelic imbalance occurs when a sample is heterozygous for a candidate variant and has disproportionate sequencing reads for each allele. For example, DNase-seq and CEBPB ChIP-seq reads containing rs4846913-A at the *GALNT2* HDL-C GWAS locus showed more reads than rs4846913-C, suggesting that the A allele increases CEBPB binding and open chromatin (54). Although not presented in this dissertation, genome editing and physical interaction experiments can be used to test candidate variants. Functional assays provide supporting evidence for candidate variants; however, it is necessary to use multiple assays and approaches to fully delineate the contribution of all variant(s), transcription factor(s), and gene(s) at a GWAS locus.

Aims and overview

The aims of this dissertation are to identify the functional variants at the T2D-associated *CDC123/CAMK1D* locus, the HDL-C-associated *ANGPTL8* locus, and to create ATAC-seq open chromatin profiles to further characterize the regulatory map of human adipose tissue. The variants at both *CDC123/CAMK1D* and *ANGPTL8* are located in regulatory regions. The ATAC-seq open chromatin profiles generated from human adipose tissue and preadipocytes are used to better characterize regulatory elements in these cell types and to prioritize variants at cardiometabolic GWAS loci.

In Chapter 2, I demonstrate the allelic effects on regulatory function of a single variant at the *CDC123/CAMK1D* T2D GWAS locus. To gain insight into the molecular mechanisms underlying the association signal, we used genomic regulatory overlap to identify SNPs overlapping predicted regulatory regions. Two regions containing T2D-associated variants were tested for enhancer activity using luciferase reporter assays. One SNP, rs11257655, displayed allelic differences in transcriptional enhancer activity in 832/13, MIN6, and HepG2 cells. The rs11257655 risk allele T showed greater transcriptional activity then the non-risk allele C in all cell types tested. Using EMSAs, the rs11257655 risk allele showed allele-specific binding to FOXA1 and FOXA2. FOXA1 and FOXA2 enrichment at the rs11257655 risk allele was validated using allele-specific ChIP in human islets. These results suggest that rs11257655 affects transcriptional activity through altered binding of a protein complex that includes FOXA1 and FOXA2, providing a potential molecular mechanism at this GWAS locus.

In Chapter 3, I investigate candidate regulatory variants at the HDL-C-associated *ANGPTL8* locus. The *ANGPTL8* association with HDL-C levels has been identified in multiple populations (38,43,66). Given the extensive sharing of GWAS loci across populations (50), I

hypothesized that at least one shared variant at this locus affects HDL-C. The HDL-C-associated variants are coincident with eQTLs for *ANGPTL8* and *DOCK6* in subcutaneous adipose tissue; however, only *ANGPTL8* expression levels are associated with HDL-C. A 400-bp promoter region upstream of *ANGPTL8* and enhancer regions within 5 kb contribute to regulating expression in liver and adipose. To identify variants functionally responsible for the HDL-C association, we performed fine-mapping analyses and selected 13 candidate variants that overlap putative regulatory regions to test for allelic differences in regulatory function. Of these variants, rs12463177-G increased transcriptional activity (1.5-fold, P=0.004) and showed differential protein binding. Six additional variants (rs17699089, rs200788077, rs56322906, rs3760782, rs737337, and rs3745683) showed evidence of allelic differences in transcriptional activity and/or protein binding. Taken together, these data suggest a regulatory mechanism at the *ANGPTL8* HDL-C GWAS locus involving tissue-selective expression and multiple potentially functional variants.

In Chapter 4, I describe the open chromatin profile of human adipose tissue using ATAC-seq. ATAC-seq open chromatin profiles were generated using frozen human subcutaneous adipose tissue needle biopsies from three individuals and a preadipocyte cell strain. I compared heterogeneous adipose tissue ATAC-seq to homogenous preadipocytes and observed 83 gene promoters with ATAC-seq peaks specific to adipose tissue and 437 gene promoters with peaks specific to SGBS preadipocytes. 17 transcription factors generated footprints using the ATAC-seq data. Finally, I identified enrichment of cardiometabolic GWAS loci in ATAC peaks, and 147 variants at 59 cardiometabolic GWAS loci that also harbor colocalized adipose expression quantitative trait loci overlapped ATAC peaks. Of the 147 variants, I further investigated

rs1534696 at the *SNX10* waist-to-hip ratio GWAS locus and rs7187776 at the *ATP2A1-SH2B1* BMI GWAS locus and identified allelic differences in functional assays.

Finally, in Chapter 5 I summarize the major findings of Chapters 2, 3, and 4 and discuss the scope and future directions of functionally characterizing GWAS signals for complex cardiometabolic traits.

CHAPTER 2: IDENTIFICATION OF A REGULATORY VARIANT THAT BINDS FOXA1 AND FOXA2 AT THE CDC123/CAMK1D TYPE 2 DIABETES GWAS LOCUS[1,2]

Introduction

Type 2 diabetes is a complex metabolic disease with a substantial heritable component (67). Over the past seven years, genome-wide association studies (GWAS) have successfully identified over 70 common risk variants associated with type 2 diabetes (68-71). Association signals at many of these loci localize to non-protein-coding intronic and intergenic regions and likely harbor regulatory variants altering gene transcription. In recent years great advances have facilitated identification of regulatory elements genome-wide using techniques including DNase-seq and FAIRE-seq (formaldehyde-assisted isolation of regulatory elements), which identify regions of nucleosome depleted open chromatin, and ChIP-seq (chromatin immunoprecipitation), which identify histone modifications to nucleosomes and transcription factor binding sites. Several studies have successfully integrated trait-associated variants at GWAS loci with publicly available regulatory element datasets in disease-relevant cell types to guide identification of regulatory variants underlying disease susceptibility (53,72-75).

The *CDC123* (cell division cycle protein 123) / *CAMK1D* (calcium/calmodulin-

---

[1] This chapter previously appeared as an article in *PLoS Genetics*. The citation is: Fogarty MP, Cannon ME, Vadlamudi S, Gaulton KJ, Mohlke KL. 2014. Identification of a regulatory variant that binds FOXA1 and FOXA2 at the CDC123/CAMK1D Type 2 Diabetes GWAS locus. *PLoS Genetics* Sep 11;1-(9):e1004633.

[2] Maren Cannon performed and analyzed the transcriptional reporter and electrophoretic mobility shift assays. Co-authors designed, performed, and analyzed the DNA affinity, chromatin immunoprecipitation, and RNA experiments. Marie Fogarty wrote the manuscript. Maren Cannon edited the manuscript and contributed to response and editing in review.

dependent protein kinase ID) locus on chromosome 10 contains common variants (MAF > .05) strongly associated with type 2 diabetes in Europeans (rs12779790, $P = 1.2 \times 10^{-10}$) (69), East Asians (rs10906115, $P = 1.5 \times 10^{-8}$) (70), and South Asians (rs11257622, $P = 5.8 \times 10^{-6}$) (76). Fine-mapping using the Metabochip identified rs11257655 as the lead SNP (68). The index variant and proxies ($r^2 > .7$) span an intergenic region of at least 45 kb between *CDC123* and *CAMK1D* and overlap the 3' end of *CDC123* (69). None of the type 2 diabetes-associated variants at this locus are located in exons. Analysis of the beta cell function measurements HOMA-B and insulinogenic index, derived from paired glucose and insulin measures at fasting or 30 minutes after a glucose challenge, demonstrated association of the risk allele at the *CDC123/CAMK1D* locus with reduced beta cell function, suggesting the beta cell as a candidate affected tissue (68,77). Another intronic variant (rs7068966, r^2 =0.18 EUR, 1000G Phase 1) located 50 kb away from rs12779790 is associated with lung function (78).

The transcript(s) targeted by risk variant activity at this locus remain unknown. *CDC123* is regulated by nutrient availability in yeast and is essential to the onset of mRNA translation and protein synthesis through assembly of the eukaryotic initiation factor 2 complex (79,80). Evidence from previous GWA studies suggest cell cycle dysregulation as a common mechanism in type 2 diabetes; for example, type 2 diabetes association signals are found close to the cell cycle regulator genes, CDKN2A / CDKN2B and *CDKAL1* (50). *CAMK1D* is a member of the Ca^{2+}/calmodulin-dependent protein kinase family which transduces intracellular calcium signals to affect diverse cellular processes. Upon calcium influx in granulocyte cells and hippocampal neurons, CAMK1D activates CREB-dependent gene transcription (81,82). Given the roles of cytosolic calcium in regulation of beta cell exocytotic machinery and of *CREB* in beta cell survival, *CAMK1D* may have a role in beta cell insulin secretion. In *cis*-eQTL analyses, the

rs11257655 type 2 diabetes risk allele was more strongly and directly associated with increased expression of *CAMK1D* than *CDC123* in both blood and lung (83,84).

In this study we aimed to identify the variant(s) underlying the association signal at the *CDC123*/*CAMK1D* locus using genome-wide maps of open chromatin, chromatin state and transcription factor binding in pancreatic islets, hepatocytes, adipocytes and skeletal muscle myotubes. We measured transcriptional activity of variants in putative regulatory elements using luciferase reporter assays, and identified a candidate *cis*-acting SNP driving allele-specific enhancer activity in two mammalian beta cell-lines as well as hepatocellular carcinoma cells. We then evaluated DNA-protein binding in sequence surrounding this variant and identified allele-specific binding to key islet and hepatic transcription factors. Thus, our study provides strong evidence of a functional variant underlying the type 2 diabetes association signal at the *CDC123/CAMK1D* locus acting through altered regulation in type 2 diabetes-relevant cell types.

Materials and Methods

Selection of SNPs for functional study

Variants were prioritized for functional study based on linkage disequilibrium (LD) and evidence of being in an islet or liver regulatory element based on data from the ENCODE consortium (85). Of 11 variants meeting the LD threshold ($r^2 \geq .7$, EUR, with the GWAS index SNP rs12779790, 1000G Phase 1 release), two SNPs showed evidence of open chromatin (72,74,86,87), histone modifications (88-90) or transcription factor binding and were tested for evidence of differential transcriptional activity.

Cell culture

Two insulinoma cell lines, rat-derived 832/13 (91) (C.B. Newgard, Duke University) and mouse-derived MIN6 (92) were maintained at 37°C with 5% CO_2. 832/13 cells were cultured in RPMI 1640 (Cellgro/Corning) supplemented with 10% FBS, 1 mM sodium pyruvate, 2 mM L-glutamine, 10 mM HEPES and 0.05 mM β-mercaptoethanol. MIN6 cells were cultured in DMEM (Sigma), supplemented with 10% FBS, 1 mM sodium pyruvate, 0.1 mM β-mercaptoethanol. HepG2 hepatocellular carcinoma cells were cultured in MEM-alpha (Gibco) supplemented with 10% FBS, 1 mM sodium pyruvate and 2 mM L-glutamine.

Generation of luciferase reporter constructs, transient DNA transfection and luciferase reporter assays

Fragments surrounding each of rs11257655 (151 bp) and rs34428576 (179 bp) were PCR-amplified (Table 1) from DNA of individuals homozygous for risk and non-risk alleles. Restriction sites for KpnI and XhoI were added to primers during amplification, and the resulting PCR products were digested with KpnI and XhoI and cloned in both orientations into the multiple cloning site of the minimal promoter-containing firefly luciferase reporter vector pGL4.23 (Promega, Madison, WI). Fragments are designated as 'forward' or 'reverse' based on their orientation with respect to the genome. Two to five independent clones for each allele for each orientation were isolated, verified by sequencing, and transfected in duplicate into 832/13, MIN6 and HepG2 cell lines. Missing haplotypes of rs36062557-rs11257655 constructs were created using the QuikChange site directed mutagenesis kit (Stratagene).

Approximately 1 x 10^{-5} cells per well were seeded in 24-well plates. At 80% confluency, cells were co-transfected with luciferase constructs and *Renilla* control reporter vector (phRL-

TK, Promega) at a ratio of 10:1 using Lipofectamine 2000 (Invitrogen) for 832/13, and using FUGENE-6 for MIN6 and HepG2 cells (Roche Diagnostics, Indianapolis, IN). 48 h after transfection, cells were lysed with passive lysis buffer (Promega), and luciferase activity was measured using the Dual-luciferase assay system (Promega). To control for transfection efficiency, raw values for firefly luciferase activity were divided by raw *Renilla* luciferase activity values, and fold change was calculated as normalized luciferase values divided by pGL4.23 minimal promoter empty vector control values. Data are reported as the fold change in mean (± SD) relative luciferase activity per allele. A two-sided *t*-test was used to compare luciferase activity between alleles. All experiments were carried out on a second independent day and yielded comparable allele-specific results.

Electrophoretic mobility shift assay (EMSA)

Nuclear cell extracts were prepared from 832/13, MIN6, and HepG2 cells using the NE-PER nuclear and cytoplasmic extraction kit (Thermo Scientific) according to the manufacturer's instructions. Protein concentration was measured with a BCA protein assay (Thermo Scientific), and lysates were stored at -80°C until use. 21 bp oligonucleotides were designed to the sequence surrounding rs11257655 risk or non-risk alleles: Sense 5' biotin-GGGCAAGTGT[**C/T**]TACTGGGCAT 3', antisense 5' biotin-ATGCCCAGTA[**G/A**]ACACTTGCCC 3' (SNP allele in bold). Double-stranded oligonucleotides for the risk and non risk alleles were generated by incubating 50 pmol complementary oligonucleotides at 95^{o}C for 5 minutes followed by gradual cooling to room temperature. EMSA's were carried out using the LightShift Chemiluminescent EMSA Kit (Thermo Scientific). Binding reactions were set up as follows: 1X binding buffer, 50 ng/μL poly

(dI•dC), 3 μg nuclear extract, 200 fmol of labeled probe in a final volume of 20 μL. For competition reactions, 67-fold excess of unlabeled double-stranded oligonucleotides for either the risk or non-risk allele were included. Reactions were incubated at room temperature for 25 minutes. For supershift assays, 4 μg of polyclonal antibodies against FOXA1 (ab23738; Abcam) or FOXA2 (SC6554X; Santa Cruz Biotechnology) was added to the binding reaction and incubation proceeded for a further 25 minutes. Binding reactions were subjected to non-denaturing PAGE on DNA retardation gels in 0.5X TBE (Lonza), transferred to Biodyne nylon membranes (Thermo Scientific) and cross-linked on a UV-light cross linker (Stratagene). Biotin labeled DNA-protein complexes were detected by chemiluminescence. EMSAs were carried out on a second independent day and yielded comparable.

DNA affinity capture assay

DNA affinity capture was carried out as previously described (53). Briefly, dialyzed nuclear extracts (300 μg) were pre-cleared with 100 μl of streptavidin-agarose dynabeads (Invitrogen) coupled to biotin-labeled scrambled control oligonucleotides. For DNA-protein binding reactions, 40 pmol of biotin labeled probe for either rs11257655 allele (same probe as for EMSA) or for a scrambled control were incubated with 300 μg nuclear extract, binding buffer (10 mM Tris, 50 mM KCL, 1 mM DTT), 0.055 μg/μL poly (dI•dC) and H_20 to total 450 μL at room temperature for 30 minutes with rotation. 100 μL (1mg) of streptavidin-agarose dynabeads were added and the reaction incubated for a further 20 minutes. Beads were washed and DNA-bound proteins were eluted in 1X reducing sample buffer (Invitrogen). Proteins were separated on NuPAGE denaturing gels and protein bands stained with SYPRO-Ruby. Protein bands displaying differential binding between rs11257655 alleles were excised from the gel and

subjected to matrix assisted laser desorption time-of-flight/time-of-flight tandem mass spectrometry (MS) and analysis at the University of North Carolina proteomics core facility. For peptide identification, all MS/MS spectra were searched against all entries, NCBI non-redundant (NR) database, using GPS Explorer TM Software Version 3.6 (ABI) and the Mascot (MatrixScience) search algorithm. Mass tolerances of 80 ppm for precursor ions and 0.6 Da for fragment ions were used. In addition, two missed cleavages were allowed and oxidation of methionine was a variable modification.

Chromatin Immunoprecipitation (ChIP) assays

Human islets from non-diabetic organ donors were provided by the National Disease Research Interchange (NDRI). Use of human tissues was approved by the University of North Carolina Institutional Review Board. Islet viability and purity were assessed by the NDRI. Islets were warmed to 37^{o}C and washed with calcium- and magnesium-free Dulbecco's phosphate-buffered saline (Life Technologies) prior to crosslinking. For chromatin immunoprecipitation (ChIP) studies, approximately 2000 islet equivalents (IEQs) were crosslinked for 10 min in 1% formaldehyde (Sigma-Aldrich) at room temperature. Islets were lysed and chromatin was sheared on ice using a standard bioruptor (Diagenode; 20–22 cycles of 30 s sonication with 1 min rest between cycles) to a size of 200–1000 bp. IP dilution buffer (0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA, 16.7 mM Tris at pH 8.1, 167 mM NaCl, protease inhibitors) was added, 5% of the volume was removed and used as input, and the remainder was incubated overnight at 4°C on a nutating platform with FOXA1 or FOXA2 antibody or a species-matched IgG as control. Antibodies used for ChIP were the same as for EMSA; FOXA1 (Abcam) and FOXA2 (Santa Cruz). Protein A agarose beads (Santa Cruz) were added and incubated for 3 h at 4°C.

Beads were then washed for 5 minutes at 4°C with gentle mixing, using the following solutions: Low Salt Buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris, 150 mM NaCl); High Salt Buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris, 500 mM NaCl); LiCl buffer (1 mM EDTA, 10 mM Tris, 250 mM LiCl, 1% NP-40, 1% Na-Deoxycholate), twice; and TE buffer (Sigma-Aldrich), twice. Chromatin was eluted from beads with two 15-minute washes at 65°C using freshly prepared Elution Buffer (1% SDS/0.1 M $NaHCO_3$). To reverse crosslinks, 5 M NaCl was added to each sample to a final concentration of 0.2 M, and incubated overnight at 65°C; to remove protein, samples were incubated with 10 uL 0.5 M EDTA, 20 uL 1M Tris (pH 6.5) and 3 uL of Proteinase K (10 mg/mL) at 45°C for 3 hours. DNA was extracted with 25:24:1 phenol:choloform:isoamyl alcohol, precipitated with 100% ethanol with 1 µl glycogen as a carrier, and resuspended in TE (Sigma). qPCR was performed in triplicate using SYBR Green Master Mix. Primers were designed to amplify a 99-bp region surrounding rs112576555; 5′-CTACTGCTTCTCCGGACTCG ’3′ and 5′- TGGCCTCAAGAGG GAGATAA -3′. Primers for a 133-bp control region not overlapping open chromatin and located 27 kb away were 5′-GCACCCATGGTACTGAAACC -3′ and 5′- CTTTTCCCG AGGAAGGAACT -3′. Dissociation curves demonstrated a single PCR product in each case without primer dimers. Fold enrichment was calculated as FOXA1/FOXA2 enrichment divided by IgG control. A one-sided t-test was performed to compare enrichment based on the direction of binding observed using EMSA.

Effect of glucose on Cdc123 and Camk1d transcript level

To measure effects of glucose on expression of *Cdc123* and *Camk1d*, 832/13 cells and MIN6 cells were washed with PBS and preincubated for 2.0 h in secretion buffer (114 mm NaCl, 4.7 mm KCl, 1.2 mm KH_2PO_4, 1.16 mm $MgSO_4$, 20 mm HEPES, 2.5 mm $CaCl_2$, 0.2% BSA, pH

7.2. For GSIS, cells were incubated in secretion buffer for an additional 2 hours or 16 hours in the presence of 3 mM or 20 mM glucose and then harvested for RNA.

RNA isolation and quantitative real-time reverse-transcription PCR

Total cytosolic RNA was isolated using the RNeasy Mini Kit (Qiagen). RNA concentrations were determined using a Nanodrop 1000 (Thermo Scientific, Wilmington, DE, USA). For real-time reverse transcription (RT)–PCR, first-strand cDNA was synthesized using 8 ul of total RNA in a 20 µl reverse transcriptase reaction mixture (Superscript III First strand synthesis kit; Life Technologies). cDNA was diluted to contain equivalent to 20-55 ng/µl input RNA. To measure total human mRNA levels of *CDC123*, *CAMK1D and B2M*, gene-specific primers and fast SYBR Green Master Mix (Life Technologies) were used (Table 2). TaqMan designed gene expression assays (Life Technologies) were used to measure Cdc123, Camk1D and Rsp9 (housekeeping gene) mRNA levels of mouse and rat cells. All PCR reactions were performed in triplicate in a 10-µl volume using a STEPOne Plus real-time PCR system (Life Technologies). Serial 3-fold dilutions of cDNA from pooled human tissues, 832/13 or MIN6 cells as appropriate were used as a reference for a standard curve. Statistical significance was determined by two-tailed *t*-tests.

Results

Prioritization of type 2 diabetes-associated SNPs with regulatory potential at the CDC123/CAMK1D locus.

To identify potentially functional SNPs at the *CDC123/CAMK1D* locus, we considered variants in high LD ($r^2 \geq .7$, EUR, 1000G Phase 1 release) with GWAS index SNP rs12779790.

To further prioritize variants for functional follow up, we used genome wide maps of chromatin state (Figure 1) in available type 2 diabetes-relevant cell types including pancreatic islets, liver hepatocytes, skeletal muscle myotubes and adipose nuclei. Variant position was evaluated with respect to DNase- and FAIRE-seq peaks and several histone modifications, including H3K4me1 and H3K9ac. DNase and FAIRE are established methods of identification of nucleosome depleted regulatory regions (86), while H3K4me1 and H3K9ac are post-translational chromatin marks often associated with enhancer regions (88,89). We also assessed chromatin occupancy by transcription factors using available genome wide ChIP-seq data sets. Of 11 variants meeting the LD threshold, two SNPs were found to overlap chromatin signals. One SNP, rs11257655 (r^2 = .74 with GWAS index SNP rs12779790), located 15 kb from the 3' end of *CDC123* and 84 kb from the 5' end of *CAMK1D,* was a particularly plausible candidate overlapping islet, liver and HepG2 cell line DNase peaks, islet and liver FAIRE peaks, H3K4me1 and H3K9ac (data not shown) chromatin marks, and FOXA1 and FOXA2 ChIP-seq peaks in HepG2 cells (Figure 6). A second SNP, rs34428576 (r^2 = .71 with rs12779790), overlapped a HepG2 DNase peak and displayed occupancy by FOXA1 and FOXA2 binding in HepG2 cells (Figure 1). No SNPs overlapped with DNase peaks in skeletal muscle myotubes (data not shown).

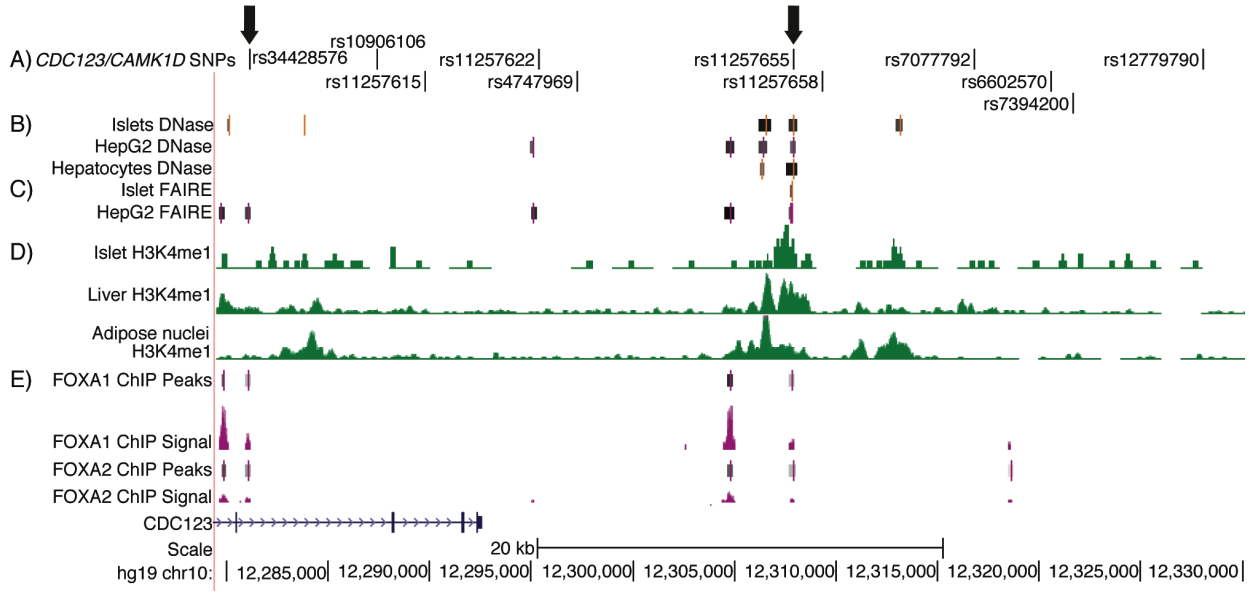

Figure 1: Regulatory potential at type 2 diabetes-associated SNPs at the *CDC123/CAMK1D* locus

A) The 11 SNPs in high LD ($r^2 \geq .7$, EUR) with GWAS index SNP rs12779790. Arrows indicate the two SNPs that overlap islet, liver, and HepG2 open chromatin and epigenomic marks and that are located near to HepG2 ChIP-seq peaks; these two SNPs were tested for allele-specific transcriptional activity. B) DNase hypersensitivity peaks identified in two pooled islet samples from the ENCODE Consortium. C) FAIRE peaks identified in one representative islet sample from the ENCODE Consortium. D) H3K4me1 histone modifications from the Roadmap Epigenomics Consortium. E) FOXA1 and FOXA2 ChIP-seq peaks and signal from ENCODE. Image is taken from the UCSC genome browser, February 2009 (GRCh37/hg19) assembly (http://genome.ucsc.edu) (93). The 5' end of *CAMK1D* begins after position 12,390,000.

Allele-specific enhancer activity of rs11257655 in islet and liver cells.

To evaluate transcriptional activity of the SNPs in predicted regulatory regions, 150 – 200 bp surrounding each SNP allele was cloned into a minimal promoter vector and luciferase activity was measured in two beta cell lines, 832/13 rat insulinoma and MIN6 mouse insulinoma cells, and in HepG2 liver hepatocellular carcinoma cells. Four to five independent clones for each allele were generated and enhancer activity was measured in duplicate for each clone. A 151-bp region including rs11257655 (and rs36062557 due to proximity, $r^2 = .38$ with rs11257655) showed differential allelic enhancer activity in both orientations in all three cell

lines (Figure 2). The risk allele rs11257655-T showed significantly increased luciferase activity compared to the non-risk allele rs11257655-C (forward: 832/13 $P = 6.3 \times 10^{-3}$, MIN6 $P = 1.7 \times 10^{-5}$; HepG2 $P = 8.0 \times 10^{-5}$; reverse: 832/13 $P = 2.2 \times 10^{-3}$, MIN6 $P = 9.9 \times 10^{-5}$; HepG2 $P = 2.0 \times 10^{-3}$). Enhancer activity represents greater than a 1.4-fold (HepG2, MIN6) to 2.1-fold (832/13) increase in transcriptional activity relative to the non-risk allele in both the forward and reverse orientations. Compared to an empty vector control, enhancer activity was greatest in the islet cell lines (risk allele: 832.13, 4-fold; MIN6, 10-fold; HepG2, 1.6-fold).

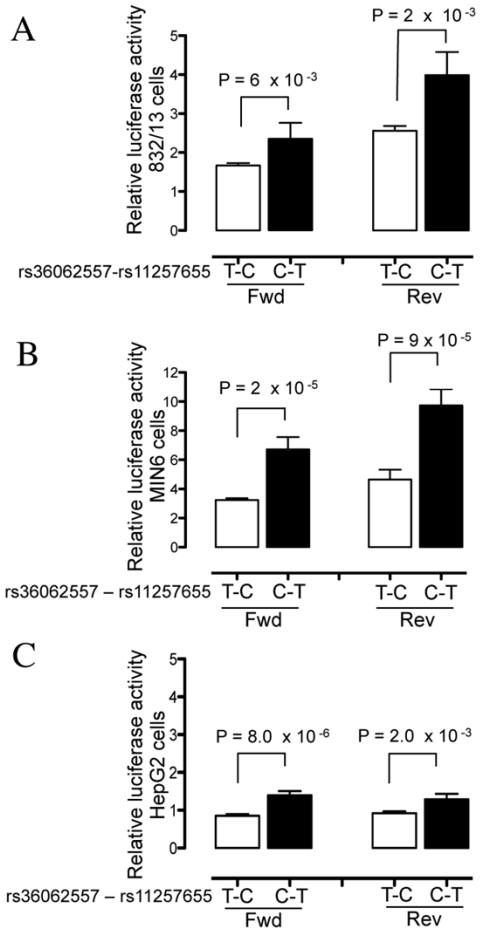

Figure 2: Haplotype containing type 2 diabetes-associated SNPs displays differential transcriptional activity

Enhancer activity was tested in 832/13, MIN6 and HepG2 cells for the type 2 diabetes non-risk (white bars) and risk (black bars) haplotypes in the forward and reverse orientations with respect to the genome. Risk refers to the rs11257655 variant; rs36062557 is included in the haplotype due to proximity. The haplotype containing risk allele rs11257655-T shows greater transcriptional activity than the non-risk allele rs11257655-C in both orientations with respect to a minimal promoter vector in 832/13 cells (A), MIN6 cells (B) and HepG2 cells (C). Error bars represent standard deviation of 4-5 independent clones for each allele. Firefly luciferase activity was normalized to *Renilla* luciferase activity, and normalized results are expressed as fold change compared to empty vector control. *P* values were calculated by a two-sided *t*-test.

A 179-bp region surrounding the second candidate SNP rs34428576 showed only moderate allele-specific activity, and only in the reverse orientation, in HepG2 cells ($P = .02$) and no allele-specific activity in islet cells (Figure 7).

To verify that rs11257655 and not rs36062557 accounted for allele-specific effects, we used site-directed mutagenesis to construct the remaining haplotype combinations. The T risk allele of rs11257655 exhibited > 1.8 fold increased transcriptional activity compared to the non-risk allele C independent of rs36062557 genotype (Figure 3A, B). In contrast, altering alleles of rs36062557 on a consistent rs11257655 background showed no significant effect on transcriptional activity. Taken together, these data confirm that rs11257655 exhibits allelic differences in transcriptional enhancer activity and suggest it functions within a *cis*-regulatory element at the *CDC123/CAMK1D* type 2 diabetes-associated locus.

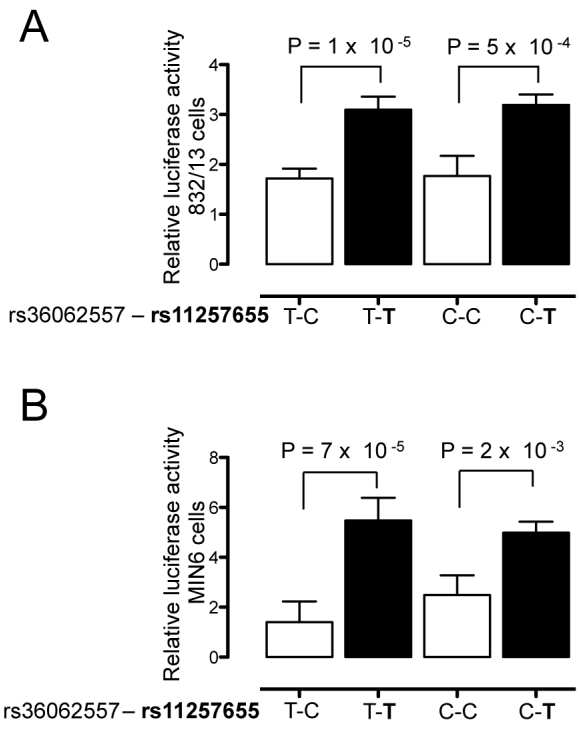

Figure 3: rs11257655 drives differential transcriptional activity
Site-directed mutagenesis was carried out to separate the effects of rs36062557 from rs11257655. Enhancer activity was tested in 832/13 and MIN6 and cells for the type 2 diabetes non-risk (white bars) and risk (black bars) haplotypes in the forward orientation. The risk allele rs11257655-T shows greater transcriptional activity than the non-risk allele rs11257655-C independent of rs36062557 genotype in 832/13 cells (A) and MIN6 cells (B). Error bars represent standard deviation of 2-4 independent clones for each allele. Results are expressed as fold change compared to empty vector control. *P* values were calculated by a two-sided *t*-test.

Alleles of rs11257655 differentially bind FOX transcription factors.

To assess whether alleles of rs11257655 differentially affect protein-DNA binding *in vitro*, biotin-labeled probes surrounding the T (risk) or C (non-risk) allele were incubated with 832/13, MIN6 or HepG2 nuclear lysate and subjected to electrophoretic mobility shift assays (EMSA). Band shifts indicative of multiple DNA-protein complexes were observed for both rs11257655 alleles (Figure 4A, 4B, 4C). In EMSAs from all three cell nuclear extracts, protein complexes were observed for the probe containing the T allele that were not present for the probe containing the C allele (832/13, arrow a; MIN6, arrows b, c, d; HepG2, arrows e, f) suggesting differential protein binding dependent on the rs11257655 allele. Competition of labeled T-allele probe with excess unlabeled T-allele probe more efficiently competed away allele-specific bands than excess unlabeled C-allele probe, demonstrating allele-specificity of the protein-DNA complexes (Figure 4A, 4B, 4C). rs11257655 did not show a differential protein binding pattern in EMSA using 3T3-L1 mouse adipocytes (data not shown). To examine transcription factor binding to rs11257655, we used a DNA-affinity capture assay. We observed one protein band showing allele-specific binding to the T allele (Figure 4D) that was identified as transcription factor FOXA2 using MALDI TOF/TOF mass spectrometry.

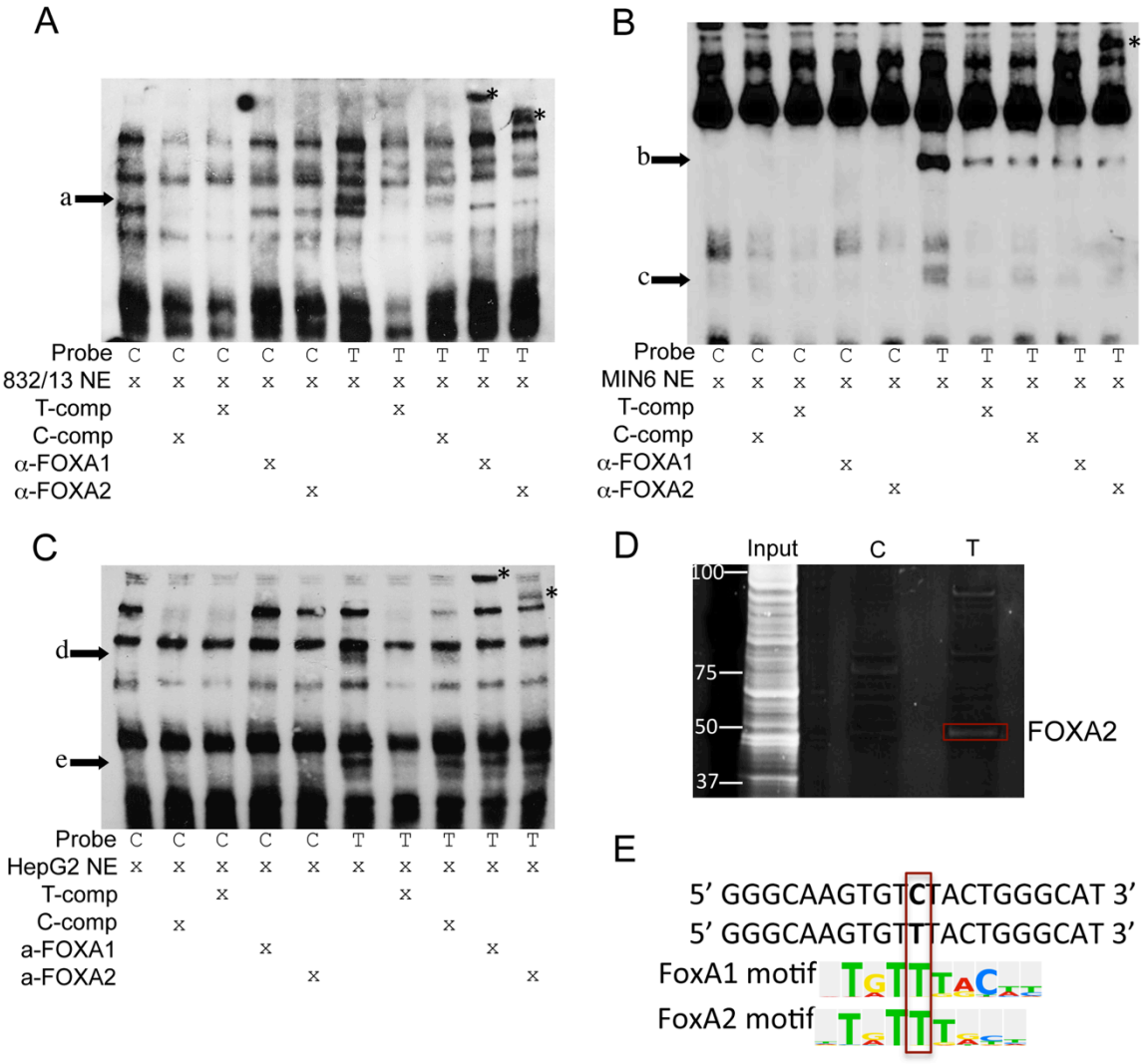

Figure 4: Alleles of rs11257655 differentially bind FOXA proteins in rat 832/13 insulinoma cells, mouse MIN6 insulinoma cells and human HepG2 hepatoma cells

EMSA using 832/13 (A), MIN6 (B) and HepG2 (C) nuclear extract shows differential protein-DNA binding of rs11257655 alleles. The probe containing risk allele rs11257655 -T shows allele-specific protein binding (arrows a - e) compared to the probe containing non-risk allele C. Excess unlabeled probe containing the T allele (T-comp) more efficiently competed away allele-specific bands than unlabeled probe for the C allele (C-comp). Incubation of 832/13 and HepG2, nuclear extract with FOXA1 / FOXA2 antibodies disrupt the DNA-protein complex formed with T allele-containing DNA probe (arrow a, d, e) and result in band supershifts (asterisks). Incubation of MIN6 nuclear extract with FOXA2 antibody decreases the DNA-protein complex formed with T allele-containing DNA probe (arrow b) and results in a band supershift. To enhance visualization of protein complexes, free biotin-labeled probe is not shown. (D) DNA affinity-capture identified differential binding of FOXA2 at rs11257655 alleles in 832/13 cells. (E) The T allele of rs11257655 is predicted as a FOXA1 and FOXA2 consensus core-binding motif.

A search in the JASPAR CORE database provided further evidence that the rs11257655 SNP is located within predicted binding sites for FOXA1 and FOXA2, with only the T risk-allele predicted to contain a FOXA1 and FOXA2 consensus core-binding motif (Figure 4E) (94). To assess binding to FOXA1 and FOXA2, we performed supershift experiments incubating DNA-protein complexes with antibodies for these factors. Incubation of the T allele-protein complex with FOXA1 antibody resulted in a band supershift in 832/13 and HepG2 cells (asterisk, Figure 4A, 4C) A FOXA2-mediated supershift was observed in 832/13, MIN6 and HepG2 cells (asterisk, Figure 4A, 4B, 4C). Differences in antibody species reactivity may account for the lack of a visible FOXA1-mediated supershift in MIN6 cells. Collectively, these results suggest that rs11257655 is located in binding sites for a transcriptional regulator complex including FOXA1 and/or FOXA2, which bind preferably to the rs11257655-T allele in beta cell and liver cell lines.

FOXA1 and FOXA2 occupancy at rs11257655 in human islets

To evaluate whether FOXA1 and FOXA2 bind differentially to rs11257655 in a native chromatin context, we performed allele-specific ChIP in human islets with different rs11257655 genotypes. FOXA1 was enriched 7.2-fold compared to IgG control in islets carrying a T allele while FOXA1 was not enriched in islets homozygous for C allele (Figure 5A). Although less robust, FOXA2 was enriched 4.2-fold in islets carrying a T allele compared to IgG control (Figure 5B). This direction of enrichment is consistent with the EMSA data (Figure 4). A region 28 kb downstream of rs11257655 with no evidence of open chromatin (chr10 control) was used as a negative control (Figure 8). These findings strengthen the conclusion that rs11257655 is part of a bona fide *cis*-regulatory complex binding FOXA1 and/or FOXA2 in human islets.

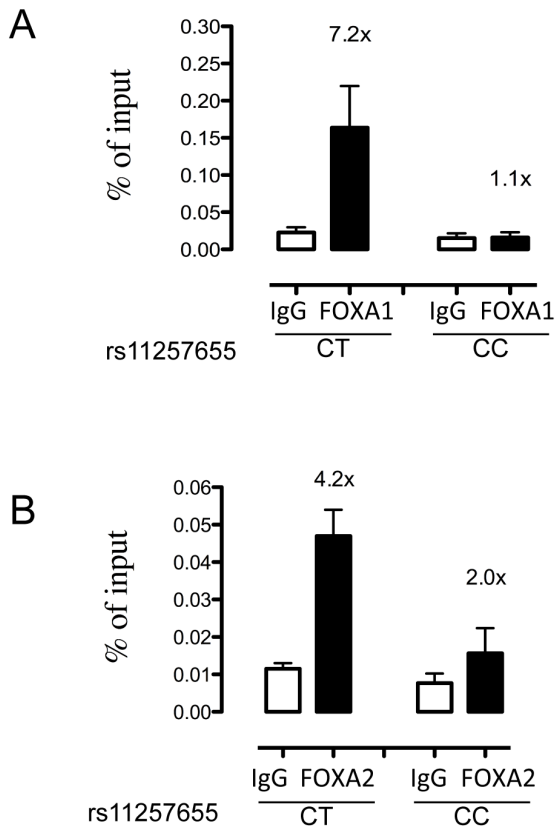

Figure 5: rs11257655-T allele shows increased binding to FOXA1 and FOXA2 in human islets
FOXA1 (A) and FOXA2 (B) ChIP in human islets shows enrichment at rs11257655 compared to IgG control. Islets containing one copy of the rs11257655-T allele show 7.2-fold greater FOXA1 enrichment and 4.2-fold greater FOXA2 enrichment. rs11257655 CT heterozygotes are more significantly enriched than rs11257655 CC homozygotes for FOXA1 (one-sided *t*-test, P = .06) and FOXA2 (one-sided *t*-test, P = .026). A negative control region 28 kb downstream of rs11257655 was not enriched in FOXA1- and FOXA2- bound chromatin (Figure 8A and 8B). Error bars represent standard error of two to three islets for each represented genotype.

CDC123 and CAMK1D transcript levels

To determine whether *CDC123* or *CAMK1D* are expressed in type 2 diabetes-relevant tissues, we measured and confirmed expression of both transcripts in human islets and hepatocytes (Figure 9A, 9B). These data are supported by RNA-seq evidence that both genes are expressed in islets (95). Based on our results showing islet beta cells as a target tissue of risk variant regulatory activity, we assessed whether glucose treatment regulated *CDC123* and *CAMK1D* transcript level. Glucose-mediated transcriptional changes in one of these genes might point to the more plausible candidate important in beta cell biology. In MIN6 cells treated with low (3 mM) and high (20 mM) concentrations of glucose for 16 hours, *CAMK1D* expression increased (P = .004; Figure 9C) while *CDC123* expression remained unchanged (P = .22; Figure 9D). In 832/13 cells, *CDC123* levels were significantly higher in cells stimulated with high

glucose ($P = 1.6 \times 10^{-5}$; Figure 9E). We could not assess the effect of glucose on *CAMK1D* levels in 832/13 cells because this transcript level was below detection limits. While we confirm expression of *CAMK1D* and *CDC123* in islets and hepatocytes, future studies over-expressing the target gene(s) in these tissues would be necessary to establish the mechanisms by which increased expression leads to diabetes risk.

Discussion

Integration of genome-wide regulatory annotation maps with disease-associated variants identified through GWAS has great potential for elucidation of gene-regulatory variants underlying association signals. In this study, we expand the lexicon of disease-associated functional regulatory variation by examining the type 2 diabetes-association signal at the *CDC123*/*CAMK1D* locus. We prioritized candidate *cis*-regulatory variants and tested whether prioritized variants exhibited allele-specific transcriptional enhancer activity. We provide transcriptional reporter and protein-DNA binding evidence that rs11257655 is part of a *cis*-regulatory complex differentially affecting transcriptional activity. Additionally, we validate FOXA1 and FOXA2 as components of this regulatory complex in human islets.

In recent years, progress has been made in following up mechanistic studies of GWAS type 2 diabetes-association signals (53,72,74,96-101), but challenges remain in sifting through the many associated variants at a locus to identify those influencing disease. We hypothesized that a common variant with modest effect underlies the association at the *CDC123/CAMK1D* locus and evaluated the location of high LD variants ($r^2 \geq .7$; n = 11) at the locus relative to known transcripts and to putative DNA regulatory elements. We identified two variants that overlapped putative islet and/or liver regulatory regions and none located in exons. We did not

assess variants in lower LD ($r^2 < .7$), and additional functional SNPs may exist at this locus acting through alternate functional mechanisms untested in the current study.

Based on our observation of type 2 diabetes-associated SNPs in regions of islet and liver open chromatin, we measured transcriptional activity in two mammalian islet cell models, rat 832/13 and mouse MIN6 insulinoma cells and in one hepatocyte cell model, human HepG2 hepatocellular carcinoma cells. In agreement with our previous observations (53), we found good concordance in allelic transcriptional activity of human regulatory elements across the two rodent islet cell types. Of the two SNPs predicted to be located in predicted enhancer regions, rs11257655 but not rs36062557 demonstrated allele-specific effects in islets and liver, suggesting that rs11257655 is a lead functional candidate. The rs11257655-T allele associated with type 2 diabetes risk displayed increased enhancer activity relative to the C allele, suggesting that increased expression of one or more genes, possibly *CAMK1D* or *CDC123,* may be associated with type 2 diabetes. Our subsequent analysis of protein binding revealed complexes that favored the rs11257655-T allele in 832/13, MIN6 and HepG2 cells. Consistent with predictions that the rs11257655-C allele may disrupt binding to the FOXA1 and FOXA2 transcription factors, we demonstrated that only the T allele of rs11257655 leads to FOXA1- and FOXA2-mediated supershifts. The ChIP enrichment of FOXA1 and FOXA2 in human islets from carriers of the T allele is concordant with EMSAs using nuclear extract from mouse and rat cell lines, further demonstrating the utility of rodent islet cell models to characterize human regulatory elements. Our results suggest that a *cis*-regulatory element surrounding rs11257655 may act in both islet and liver cells. Although we provide evidence that rs11257655 alleles differentially bind *FOXA1* and *FOXA2 in vivo*, it is important to note that this enrichment was detected in isolated human islets. Future experiments will be needed to validate effects of

rs11257655 within a whole organism environment. For example, recently zebrafish have been used to assay the regulatory potential of DNA sequences (102,103).

FOXA1 and FOXA2 are members of the FOXA subclass of the forkhead box transcription factor family and are essential transcriptional activators in development of endodermally-derived tissues including liver and pancreas (104,105). In mature mouse β-cells, ablation of both transcription factors compared to ablation of FoxA2 alone leads to more pronounced impaired glucose homeostasis and insulin secretion, indicating that both factors are important in maintenance of the mature beta cell phenotype (106). In addition, FoxA2 integrates the transcriptional response of mouse adult hepatocytes to a state of fasting (107). FOXA1 and FOXA2 are thought to act as pioneer transcription factors, scanning chromatin for enhancers with forkhead motifs and opening compacted chromatin through DNA demethylation and subsequent induction of H3K4 methylation, epigenetic changes that likely render enhancers transcriptionally competent by allowing subsequent recruitment of transcriptional effectors (108-110). Our data demonstrate increased transcriptional activity and increased binding of FOXA1 and FOXA2 to the rs11257655-T allele, suggesting that rs11257655 may be functioning as part of a transcriptional activator complex. Recent experiments in pancreatic islets support a role for FOXA transcription factors in activation of islet enhancers (111). This same study also showed that FOXA2 binds in pancreatic islets in the T2D-associated region surrounding rs11257655. Further experiments, such as ChIP-seq of additional transcription factors, may identify other key factors present in the activator complex.

Both *CAMK1D* and *CDC123* are candidate transcripts affected by variation at this locus. C*is*-eQTLs in both blood and lung support an effect on *CAMK1D* but not *CDC123*. In blood, initial eQTL evidence for both genes were further analyzed by conditional analyses on the T2D

lead SNP or rs11257655. The conditional analyses abolished the *cis*-eQTL signal for *CAMK1D* but not for *CDC123*, providing evidence that the T2D GWAS signal and the *CAMK1D cis*-eQTL signal are coincident (83). In lung, the GTEx consortium identified an eQTL for *CAMK1D* with rs11257655 as a lead associated variant ($P = 1.1 \times 10^{-7}$); this and other T2D GWAS variants are the strongest cis-eQTLs for *CAMK1D*, while no significant eQTL is observed for *CDC123* (84). For both eQTLs, the rs11257655 type 2 diabetes risk allele is associated with increased *CAMK1D* transcript level, consistent with the direction of transcriptional activity we observed for this allele in islet and liver cells. Many eQTLs are predicted to be shared among tissues (112), and a recent study of the beta cell transcriptome reports good concordance of eQTL direction ($R^2 = .74 - .76$) between beta cells and blood-derived lymphoblastoid cell lines, fat and skin (113), suggesting that the *CAMK1D* eQTL may also exist in islets. Some eQTLs differ across tissues, and evidence of a consistent eQTL in islets would be valuable. Knockout mice provide further evidence supporting *CAMK1D* as a target gene. In *FoxA1/FoxA2* beta cell-specific knockout mice, *Camk1d* expression was reported to be slightly reduced (1.8 fold, $P = 0.13$) (106), consistent with our conclusion that rs11257655 is part of a transcriptional activator complex that includes FOXA1 and FOXA2. Together, these data suggest that *CAMK1D* is a more plausible target for differential regulation by rs11257655 alleles.

The mechanism by which *CAMK1D* may act in type 2 diabetes biology is unclear. CAMK1D is a serine threonine kinase that operates in the calcium-triggered CaMKK-CaMK1 signaling cascade (82,114). In response to calcium influx, *CAMK1D* activates CREB- (cAMP response element-binding protein) dependent gene transcription by phosphorylation (82). CREB is a key beta cell regulator important in glucose sensing, insulin exocytosis and gene transcription and β-cell survival (115), and FOXA2 has been shown to be necessary to mediate

recruitment of CREB in fasting-induced activation of hepatic gluconeogenesis (107). *CAMK1D* also has been reported to regulate glucose in primary human hepatocytes (116). It is important to note that we cannot rule out cell cycle regulator *CDC123* as a target for regulation by rs11257655.

In conclusion, we extend follow up studies of GWAS-identified type 2 diabetes-associated variants to the *CDC123/CAMK1D* locus on chromosome 10. We identify rs11257655 as part of a *cis* regulatory complex in islet and liver cells that alters transcriptional activity through binding FOXA1 and FOXA2. These data demonstrate the utility of experimentally predicted chromatin state to identify regulatory variants for complex traits.

Supplemental Figures and Tables

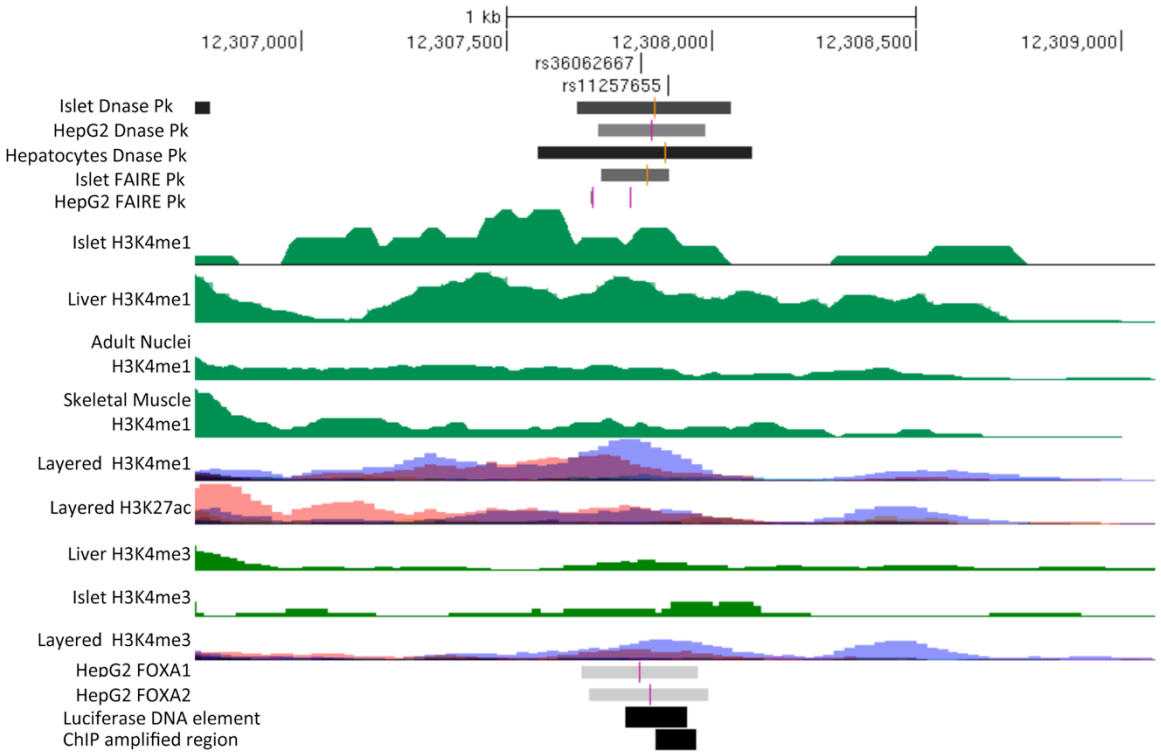

Figure 6: Regulatory potential at rs11257655 and rs36062557. UCSC genome browser (hg18) diagram showing that rs11257655 and rs36062557 overlap regions of open chromatin, detected by DNase hypersensitivity and FAIRE, and histone modifications, including H3K4me1 and H3K9ac in islet, liver, and HepG2 cells. H3K27ac and H3K4me3 histone modifications are also shown. rs11257655 and rs36062557 are also located near to HepG2 ChIP-seq peaks for FOXA1 and FOXA2. DNA sequences amplified to evaluate transcriptional activity in dual-luciferase reporter assays and to evaluate enrichment of binding to FOXA1 and FOXA2 are indicated.

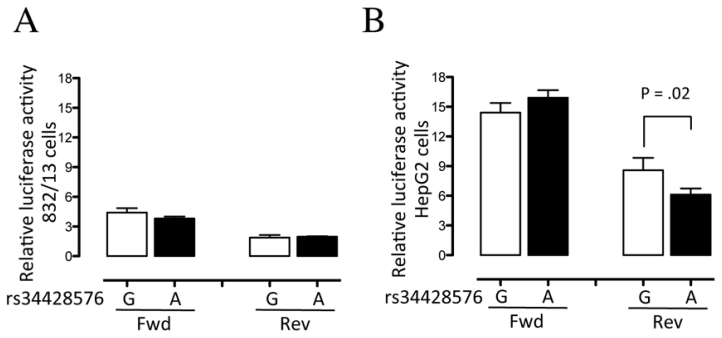

Figure 7: Transcriptional activity at rs34428576. Enhancer activity was measured in 832/13 cells (A) and HepG2 cells (B) for alleles of rs34428576. No difference was observed between alleles in 832/13 cells. In HepG2 cells, moderate allele-specific activity was observed only in the reverse orientation. Error bars represent standard deviation of 4–5 independent clones for each allele. Results are expressed as fold change compared to empty vector control. P values were calculated by a two-sided t-test.

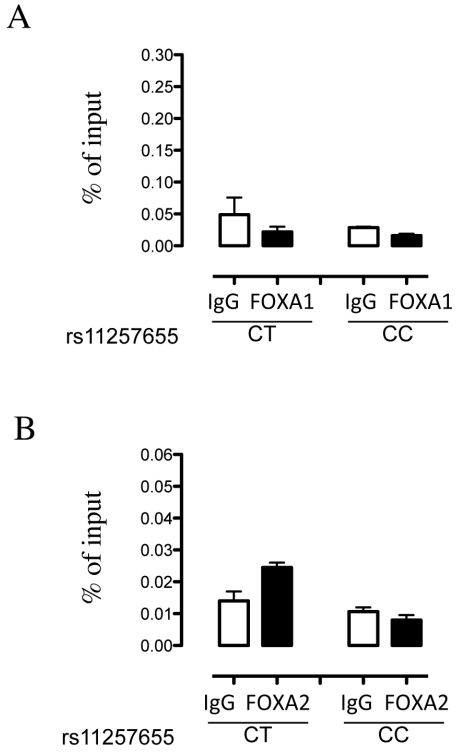

Figure 8: Chromosome 10 region not overlapping open chromatin does not show binding to FOXA1 and FOXA2 in human islets. A negative control region 28 kb downstream of rs11257655 was not substantially enriched in FOXA1- (A) and FOXA2- (B) bound chromatin. Error bars represent standard error of two to three islets for each represented genotype.

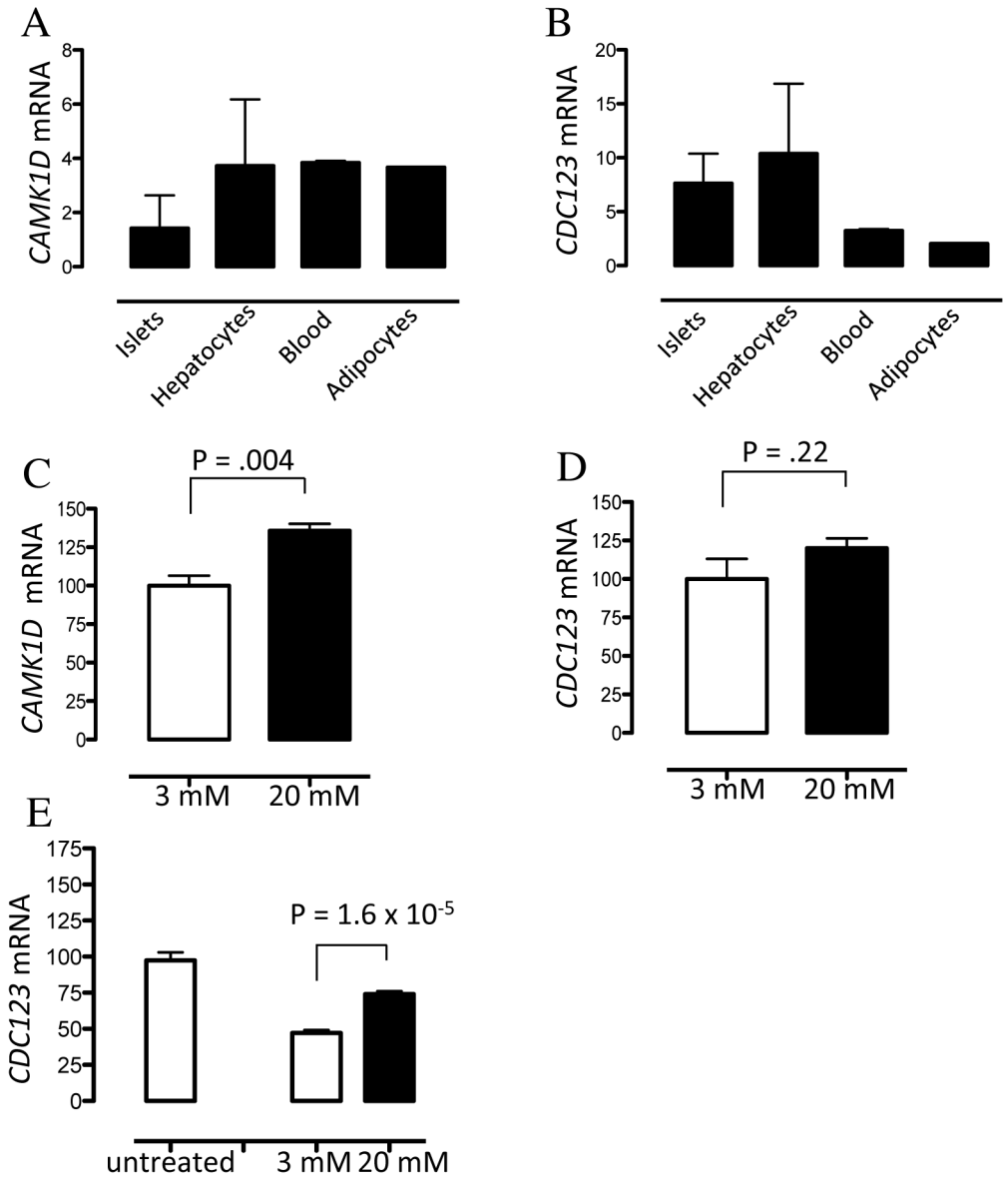

Figure 9: *CDC123* and *CAMK1D* expression and response to glucose. (A, B) Evidence that *CAMK1D* and *CDC123* are expressed in various human tissues. cDNA from human islets, hepatocytes, blood and adipocytes was analyzed by real-time PCR using gene-specific primers for *CAMK1D* (A) and *CDC123* and B2M (B). mRNA level was normalized to B2M. (C, D, E) Effect of glucose stimulus on *CAMK1D* and *CDC123* expression level. 832/13 and MIN6 insulinoma cells were treated with low (3 mM) and high (15 mM) glucose for 16–18 hours. cDNA was analyzed by real-time PCR using TaqMan gene expression assays for *CAMK1D* (C) and CDC123 (D, E). mRNA level was normalized to *RSP9*. High glucose treatment resulted in a significant increase in *CAMK1D* mRNA level (C) but not *CDC123* in MIN6 cells (D). High glucose treatment resulted in increased *CDC123* mRNA level in 832/13 cells. Error bars represent the standard deviation of 4–5 samples for each treatment. *P* values were calculated by a two-sided t-test.

Table 1: DNA sequences amplified for luciferase activity assays

SNP	Chromosome position (hg19)	Sequence 5'-3'
rs11257655	chr10:12,307,791-12,307,941	GGCCCAGAAATGACACAGAA
		AACTGGGTAAGGCTCACTTCC
rs34428576	chr10:12,280,974 -12,281,152	GCGAGACTCTGCCTCAAAAG
		GACAGAGTGAGACCCCATCC

Restriction sites were added to primers for subcloning.

Table 2: PCR primers for quantitative real-time PCR in human tissues

Gene	Sequence 5'-3'
CAMK1D	Fwd – ATCTCCACAGAATGGGCATC
	Rvs – CAGTTCAACGGCTTTGCTGTA
CDC123	Fwd – GCAGCTGGAGGATGAAGAAG
	Rvs –TCATCCCTTCCTGAAACCAC
B2M	Fwd –TGTCTGGGTTTCATCCATCCGACA
	Rvs –TCACACGGCAGGCATACTCATCTT

CHAPTER 3: TRANS-ANCESTRY FINE MAPPING AND MOLECULAR ASSAYS IDENTIFY REGULATORY VARIANTS AT THE ANGPTL8 HDL-C GWAS LOCUS[5,6]

Introduction

To date at least 157 loci have been associated with high-density lipoprotein cholesterol (HDL-C) in genome-wide association studies (GWAS) (38). Analyses of European, African American, Mexican, Pima Indian, and East Asian participants have identified four different lead variants associated with HDL-C located in or near angiopoietin-like protein 8 (*ANGPTL8*) (38,43,66,117,118). *ANGPTL8* is a small protein of 198 amino acids; the gene is located on chromosome 19 within intron 14 of dedicator of cytokinesis 6 (*DOCK6*) and is transcribed in the opposite direction of *DOCK6.* European GWAS meta-analyses identified rs737337 (P=4.6x10^{-17}, N=185,000) (38,119) as the lead variant associated with HDL-C; this variant was also associated with total cholesterol (P=4.1x10^{-5}) but not triglycerides (P=0.12) or low-density lipoprotein cholesterol (LDL-C, P=0.26). rs737337 is located 2.8 kb upstream of the *ANGPTL8* transcription start site (RefSeq NM_018687) and is a synonymous variant in exon 19 of *DOCK6* (Thr723, RefSeq NM_020812). In an African American population, the lead variant rs12979813

[5] This chapter previously appeared as an article in *G3: Genes, Genetics, Genomes*. The citation is: **Cannon ME,** Duan Q, Wu Y, Zeynalzadeh M, Xu Z, Kangas AJ, Soininen P, Ala-Korpela M, Civelek M, Lusis AJ, Kuusisto J, Collins FS, Boehnke M, Tang H, Laakso M, Li Y, Mohlke K. 2017. Trans-ancestry fine mapping and molecular assays identify regulatory variants at the *ANGPTL8* HDL-C GWAS locus. *G3* 7(9):3217-3227.

[6] Maren Cannon designed, performed, and analyzed transcriptional reporter luciferase and electrophoretic mobility shift assays; prioritized candidate variants based on regulatory overlap; analyzed fine-mapping assays; created all figures and wrote the manuscript. Co-authors designed, performed, and analyzed clinical trait measurements; genetic and expression association data; fine-mapping analyses; and reviewed the manuscript.

was only associated with HDL-C (P=1.9x10^{-9}, N=12,157) (43), which is located 7.5 kb upstream of *ANGPTL8* and in intron 22 of *DOCK6*. In a joint analysis of Mexican and Pima Indian samples, the lead variant rs2278426 was associated with HDL-C (P=3.4x10^{-9}, N=4,361) and total cholesterol (P=5.0x10^{-6}) (66,120), which is a nonsynonymous variant in exon 2 of *ANGPTL8* (Arg59Trp) that exhibits relatively high linkage disequilibrium (LD) with the European lead rs737337 (r^2=.74, 1000G Admixed American (AMR)) and moderate LD with the African American lead rs12979813 (r^2=.52 AMR). Finally, in recent lipid associations in East Asian samples, lead variant rs3760782 was the strongest HDL-C-associated variant (P=8.8x10^{-11}) and another variant in strong LD (rs1865063, r^2=0.95) is the lead variant for total cholesterol (P=1.5x10^{-15}) and LDL-C (P=1.8x10^{-8}) (118). rs3760782 is located 3.5 kb upstream of *ANGPTL8* and in intron 20 of *DOCK6.* Given the extensive sharing of GWAS loci across populations (50), we hypothesized that at least one shared variant at the *ANGPTL8* locus affects HDL-C in all of these populations.

ANGPTL8 is a recently defined gene also called *C19orf80, LOC55908,* refeeding-induced fat and liver (*RIFL*)*, TD26,* hepatocellular carcinoma-associated gene*,* lipasin*,* and betatrophin (121-123). Serum ANGPTL8 protein levels have been associated with many metabolic phenotypes including type 1 and type 2 diabetes (124-128), obesity (125,129), and non-alcoholic fatty liver disease (129). The gene is mainly expressed in liver and adipose tissues, despite being entirely contained within one intron of the ubiquitously expressed *DOCK6* (121,130). The precise mechanisms of action of ANGPTL8 remain unclear. ANGPTL8 is secreted into the plasma and is involved in triglyceride storage in adipose tissue; knockout mice gained less fat than wildtype (131) mice and *Angptl8* knockdown in 3T3L1 mouse adipocytes led to decreased triglyceride content (122). *ANGPTL8* expression is increased in response to stress, it up-regulates

early growth response transcription factor and down-regulates adipose triglyceride lipase, suggesting a role in lipid homeostasis (132). ANGPTL8 contains a lipoprotein lipase inhibitory motif and inhibits LPL function when co-expressed with ANGPTL3 (133). Additionally, serum ANGPTL8 protein is inversely associated with HDL-C levels (134) supporting a role of *ANGPTL8* in HDL-C metabolism.

DOCK6 is a member of the dedicator of cytokinesis family of atypical guanine nucleotide exchange factors. *DOCK6* is expressed in many tissues, with highest levels of expression in the lungs, thyroid, and adipose (130). DOCK6 functions as a guanine nucleotide exchange factor for RAC1 and CDC42 and has roles in cytoskeletal remodeling and neurite outgrowth (135). Mutations in *DOCK6* can cause Adams-Oliver syndrome, an actin cytoskeletopathy characterized by limb and skin defects (136). *DOCK6* does not have functions that obviously relate to HDL-C metabolism.

Two coding variants in ANGPTL8 have been proposed to affect HDL-C levels. A rare *ANGPTL8* nonsense variant, rs145464906 (MAF=0.001 in >42,000 individuals of European ancestry) encoding Gln121Ter, is associated with increased HDL-C (P=5.1x10^{-13}) and increased triglycerides (P=0.003) in an exome array-based association analysis (137). Based on conditional analysis (137) and LD (r^2<.01), rs145464906 is independent of the reported *ANGPTL8* GWAS variants. The common Mexican and Pima Indian lead variant, rs2278426 encoding Arg59Trp, has been proposed to increase cleavage of ANGPTL3, leading to decreased LPL activity and lower HDL-C (120). Although coding variants have been identified, the variants responsible for the common HDL-C association signal at this locus remain unclear, and may include regulatory variants. While a simple mechanism involving a single variant that alters

transcription of a single gene is straightforward, multiple variants and/or multiple genes may contribute to the functional consequences of a GWAS signal (54,138,139).

In this study, we describe an association between the HDL-C locus variants and subcutaneous adipose level of *ANGPTL8* RNA. We show that a subset of HDL-C-associated variants overlap regions with strong evidence of regulatory activity (59,60), and we use trans-ancestry fine-mapping and functional assays to identify variants exhibiting allelic differences in regulatory activity at the *ANGPTL8* HDL-C GWAS locus.

Materials and Methods

Study population and phenotypes

The METabolic Syndrome In Men (METSIM) study includes 10,197 men, aged from 45 to 73 years, randomly selected from Kuopio, Eastern Finland, and examined in 2005 – 2010 (140,141). The Ethics Committee of the University of Eastern Finland in Kuopio and the Kuopio University Hospital approved the METSIM study and it was carried out in accordance with the Helsinki Declaration. Triglyceride and lipoprotein characteristics were measured via proton nuclear magnetic-resonance (NMR) or enzymatic assays in 10,079 METSIM participants (142). DNA samples were genotyped on the Illumina OmniExpress and HumanCoreExome arrays, and additional genotypes were imputed using the GoT2D integrated haplotypes reference panel as previously described (47).

We also analyzed a set of 8,421 self-identified African American participants from the Women's Health Initiative SNP Health Association Resource (WHI-SHARe) study. Details of the study design, cohort characteristics, written informed consent, and study approval by local Human Subjects Committees have been described previously (143). Participants who had

consented to genetic research were genotyped on the Affymetrix 6.0 array and additional genotypes were imputed using the 1000 Genomes Project data as a reference panel, as previously described (144).

Association, conditional, and haplotype analysis

For METSIM, we performed a preliminary test for association between ~19 million genetic variants and 72 lipid and lipoprotein subclasses (145). To better observe the relative differences between variants, we use the association of variants with the concentration of phospholipids in medium HDL because the association with this trait is stronger (P=1.9x10^{-7}) than with HDL-C (P=7.7x10^{-4}). We assumed an additive mode of inheritance and accounted for cryptic relatedness between individuals using the EMMAX mixed model (q.emmax test) implemented in EPACTS (146). After accounting for age, BMI, smoking status, and lipid-lowering medication status as covariates trait value residuals were inverse normal transformed. We carried out reciprocal conditional analyses with candidate variants within the associated signal to evaluate the potential presence of a second association signal by adjusting for specific genetic variants including METSIM lead variant rs737337 and WHI lead variant rs4804154. LD in locus plots was calculated from the METSIM imputed genotypes, and all chromosome coordinates correspond to hg19.

For WHI, association analysis was performed under an additive genetic model using linear regression adjusted for covariates. The imputed allelic dosage at each variant was tested via MACH2QTL (147). Genome-wide African ancestry proportion, age, BMI and smoking history were included as covariates. Genome-wide African ancestry proportion was derived from locus-specific ancestry, which has a correlation of 0.99 with the first principal component of

population structure. Ancestry estimation has been described previously (43). To assess the potential presence of multiple, independent variants at the same locus influencing HDL-cholesterol trait, we repeated regression analyses, conditioning on rs4804154. LD in locus plots was calculated from WHI imputed genotypes, and all chromosome coordinates correspond to hg19.

We constructed haplotypes based on five variants at *ANGPTL8* that were previously reported as the GWAS index variants (rs4804154, rs737337, and rs2278426) in studies of different ancestry groups (38,43,66,119,120) or are proxies (rs3745683 and rs3760782) in high LD with both the European and African American lead variants. We performed haplotype analyses using the "haplo.stat" R package (148), estimated haplotypes and haplotype frequencies using the haplo.em function, and tested for association between haplotypes and the concentration of phospholipids in medium HDL (METSIM) or HDL-C (WHI) level using the haplo.glm function. We assumed an additive model in which the regression coefficient represented the expected change in inverse normalized HDL level with each additional copy of the specific haplotype compared with the reference haplotype. The same covariates used for single variant analysis were also used in the haplotype analysis model.

Expression quantitative trait association

RNA from subcutaneous adipose tissue was extracted and the expression levels of probesets were measured using Affymetrix Human Genome U219 Array in 770 METSIM participants (63). The expression quantitative trait locus (eQTL) for *ANGPTL8* was not previously reported because of a gene name difference with RefSeq and Ensembl. Expression data were normalized using Robust Multi-Array Average (RMA) analysis (149). eQTL

associations were performed as previously described (63). Briefly, we applied PEER analysis (150) to account for complex non-genetic factors in the RMA-normalized gene expression levels. We adjusted for 35 inferred confounding factors and inverse normal transformed the PEER-processed residuals. In eQTL analysis, we used genotype dosages imputed using Haplotype Reference Consortium data (151) to test for the variant association with expression levels of all genes within 1 Mb of rs737337. We assumed an additive mode of inheritance and accounted for cryptic relatedness between individuals using the EMMAX mixed model implemented in EPACTS (146). Additionally, we performed conditional analysis on the expression level of *ANGPTL8* by adjusting for rs4804155 and on the expression level of *DOCK6* by adjusting for rs17699089 to assess the presence of multiple, independent associations. We also performed conditional analysis using candidate functional variant rs12463177 to assess the coincidence of GWAS and eQTL signals. To examine the relationship between RMA-normalized expression levels and HDL-C, we adjusted both traits for age and BMI, inverse normal transformed the residuals, and then tested for association in regression analyses.

MANTRA

To fine-map the *ANGPTL8* locus, we performed trans-ancestry meta-analysis using MANTRA (58). MANTRA accounts for the shared similarity in closely related populations using Bayesian partition model assuming the same underlying allelic effect. It models the effect heterogeneity among distant populations by clustering according to the shared ancestry and allelic effects. We conducted the analysis based on the association summary statistics from GWAS in METSIM and WHI and included variants present in both METSIM and WHI. We

constructed a 99% credible set of variants by ranking all variants according to their Bayes factors.

CAVIAR and PAINTOR

To prioritize functional variants, we analyzed variants at the *ANGPTL8* locus using both CAVIAR (56) and PAINTOR (57). CAVIAR estimates the posterior probability of a variant being functional by jointly modeling the *p*-values and beta association statistics, and PAINTOR leverages functional genomic annotation data, in addition to association strength, to prioritize functional variants. Both methods allow for multiple functional variants at the risk locus. Using consistent alleles for METSIM and WHI African Americans, we calculated Z scores based on p-values and the sign of betas from GWAS in these studies. For METSIM, the LD matrix was calculated based on European data from the 1000 Genomes Project Phase 1. For WHI African Americans, LD was calculated based on imputed data as previously described (144). In CAVIAR, we returned the credible set that contains all of the true functional variants with 95% confidence level. The annotation matrix used in PAINTOR contained data from the ENCODE project accessed from the UCSC Genome Browser using the Table Browser tool. The matrix included HepG2 H3K4me1, H3K4me2, H3K4me3, and H3K27ac histone marks, DNase hypersensitivity clusters (ENCODE v3) and transcription factor binding (ENCODE v2) from the Broad Institute. Presence or absence of overlap was determined by the UCSC Table Browser intersection with the signal tracks. In PAINTOR, analysis was performed for METSIM alone (Finns), WHI alone (African Americans), and the two data sets together. We set the number of functional variants as 2, 3, 4 or 5 based on feasible running time. PAINTOR predicted variants to

be functional based on a posterior probability greater than 0.1, a threshold suggested previously (57).

Functional annotation

To identify regulatory overlap of HDL-C-associated variants with histone marks, transcription factor binding, and DNase hypersensitivity sites, we used data from the ENCODE and Roadmap Epigenome projects accessed through the UCSC genome browser (59,60). RNA-seq data were obtained from the Roadmap Epigenome Project (60)adult liver and adipose tissue) or a previous publication(152)HepG2). We evaluated previously published positive correlations of DNase hypersensitivity sites with gene expression in liver cell types (153). LD calculations for selecting candidate variants were based on 1000 Genomes Phase 1 data. The AFR dataset was used for African American LD and EUR for European LD.

Cell culture

HepG2, SW872, SGBS, 293T, 3T3L1, Huh-7, and MIN6 cell lines were maintained at 37ºC with 5% CO_2. HepG2 cells (ATCC, HB-8065) were cultured in MEM-✔ (Gibco) supplemented with 10% FBS and 1 mM sodium pyruvate. SW872 cells (ATCC, HTB92) were cultured in DMEM:F12 (Sigma) supplemented with 10% FBS. SGBS cells (154) were generously provided by Dr. Martin Wabitsch (University of Ulm) and cultured in DMEM:F12 (Sigma) supplemented with 10% FBS and 5% 3.3 mM biotin/1.7 mM panthotenate solution. SW872 and SGBS cells were transfected in the undifferentiated, pre-adipocyte state. 293T cells (ATCC, CRL-3216) were cultured in DMEM (Sigma) supplemented with 10% FBS and 200 mM L-glutamine. 3T3-L1 cells (ATCC, CL-173) were cultured and differentiated as described in the

ATCC protocol. Huh-7 cells (JCRB0403, Japanese Collection of Research Bioresources Cell Bank, National Institute of Biomedical Innovation), were cultured in DMEM with high glucose (Gibco) with 10% FBS, 1 mM sodium pyruvate, 1 mM non-essential amino acids (Sigma), and 1 mM L-glutamine. MIN6 cells(92) were cultured in DMEM (Sigma), supplemented with 10% FBS, 1 mM sodium pyruvate, and 0.1 mM ☌-mercaptoethanol.

Dual luciferase transcriptional reporter assays

We PCR-amplified fragments surrounding each regulatory region or variant with 5 PRIME Mastermix (5 PRIME) or Phusion High-Fidelity Polymerase (New England Biosystems) with the primers listed in Table 5 from DNA of individuals homozygous for alleles associated with increased or decreased HDL-C. Gateway® attB sites were included in primers and Gateway® cloning was used to insert fragments into a Gateway®-compatible pGL4.23 (minimal promoter) or pGL4.10 (promoterless) firefly luciferase reporter vector (Promega). Fragments containing HDL-C-associated variants are designated as 'forward' or 'reverse' based on their orientation with respect to the genome and the *ANGPTL8* gene. The 5-kb and promoter regions were only cloned in the forward orientation (pGL4.10) because they include the *ANGPTL8* promoter. Regions were isolated based on epigenetic marks of promoter and/or enhancer regions surrounding the HDL-C-associated variants. We isolated three to five independent clones (biological replicates) for each allele for each orientation and verified by sequencing. Each clone was cotransfected with *Renilla* luciferase vector in duplicate (HepG2, 293T, Huh-7, SW872) or triplicate (SGBS, 3T3L1) wells (technical replicates) using FUGENE 6 (HepG2, 293T, Huh-7, SW872, 3T3L1; Promega), Lipofectamine 2000 (Min6; Life Technologies), or Lipofectamine 3000 (SGBS; Life Technologies). Twenty-four (SGBS) or forty-eight hours (all other cell types)

after transfection, we collected cell lysates and assayed for luciferase activity using the Dual-Luciferase Reporter Assay System (Promega). Firefly luciferase activity of the clones containing the PCR fragments was normalized to *Renilla* luciferase readings to control for differences in transfection efficiency. We repeated all luciferase transcriptional reporter experiments on independent days and obtained consistent results. Data are reported as fold change in activity relative to an empty pGL4.23 or pGL4.10 vector. We used two-sided Student's t-tests to compare luciferase activity between alleles or haplotypes.

Electrophoretic mobility shift assays (EMSA)

For EMSA, we prepared nuclear cell extracts from HepG2, HuH-7, SGBS, 3T3L1, SW872, and MIN6 cells using the NE-PER nuclear and cytoplasmic extraction kit (Thermo Scientific). Protein concentration was measured with a BCA assay (Thermo Scientific), and lysates were stored at -80ºC until use. We designed 17-19 bp biotin- or IR-Dye® 700-labeled oligos around the HDL-C-associated variants (Table 5) for both alleles. We annealed double-stranded oligos and performed binding reactions as previously described (155). We used 4-6 ug of antibody in supershift assays and 100-200 ng CEBPB purified protein for select EMSAs. Binding reactions were run on non-denaturing PAGE DNA retardation gels in 0.5X TBE (Lonza). For biotin-labeled oligos, we transferred the reactions to Biodyne nylon membranes (Thermo Scientific), cross-linked them on a UV cross linker (Stratagene), and detected DNA-protein complexes by chemiluminescence. For IR-DYE® 700-labeled oligos, we imaged gels on a LiCor Odyssey® CLx Imaging System. We repeated all EMSA experiments on another day with consistent results.

Results

Regulation of tissue-selective expression of ANGPTL8

ANGPTL8 expression is largely restricted to liver and adipose tissues (121,130). To determine drivers of *ANGPTL8* tissue specificity, we tested candidate regulatory regions in human hepatocyte (HepG2, Huh-7), human pre-adipocyte (SGBS), human adipocyte (SW872), mouse adipocyte (3T3-L1), human embryonic kidney (293T), and mouse islet beta (MIN6) cell lines. We tested a region extending 400 bp upstream of the transcription start site (TSS) that spans epigenetic marks characteristic of promoters, and a regulatory region that extends 5 kb upstream of the TSS (regions shown in Figure 10). The 400 bp promoter contains no HDL-C-associated variants; the 5-kb region contains seven HDL-C-associated variants (rs200788077, rs56322906, rs6511729, rs3760782, rs737337, rs737338, and rs3745683). We tested these candidate regions in transcriptional reporter luciferase assays. In comparison to empty vector, the promoter increased transcriptional activity 5-fold in HepG2 ($P<0.0001$) and 1.6-fold in Huh-7 ($P=0.04$), but not in any other cell type ($P>0.5$, Figure 11A). The 5-kb region increased transcriptional activity in HepG2 by 8-fold in HepG2 ($P=0.0007$) and 6-fold in Huh-7 ($P=0.002$) compared to empty vector and increased transcriptional activity 2-fold in pre-adipocyte (SW872, $P=0.0007$) and adipocyte (3T3-L1, $P=0.023$) cells compared to empty vector (Figure 11B). Neither region showed transcriptional activity in MIN6 or 293T cells. These data suggest that the 400-bp region contains promoter regulatory elements contributing to tissue specificity in liver, but may be mediated by additional enhancer elements, especially in adipocytes.

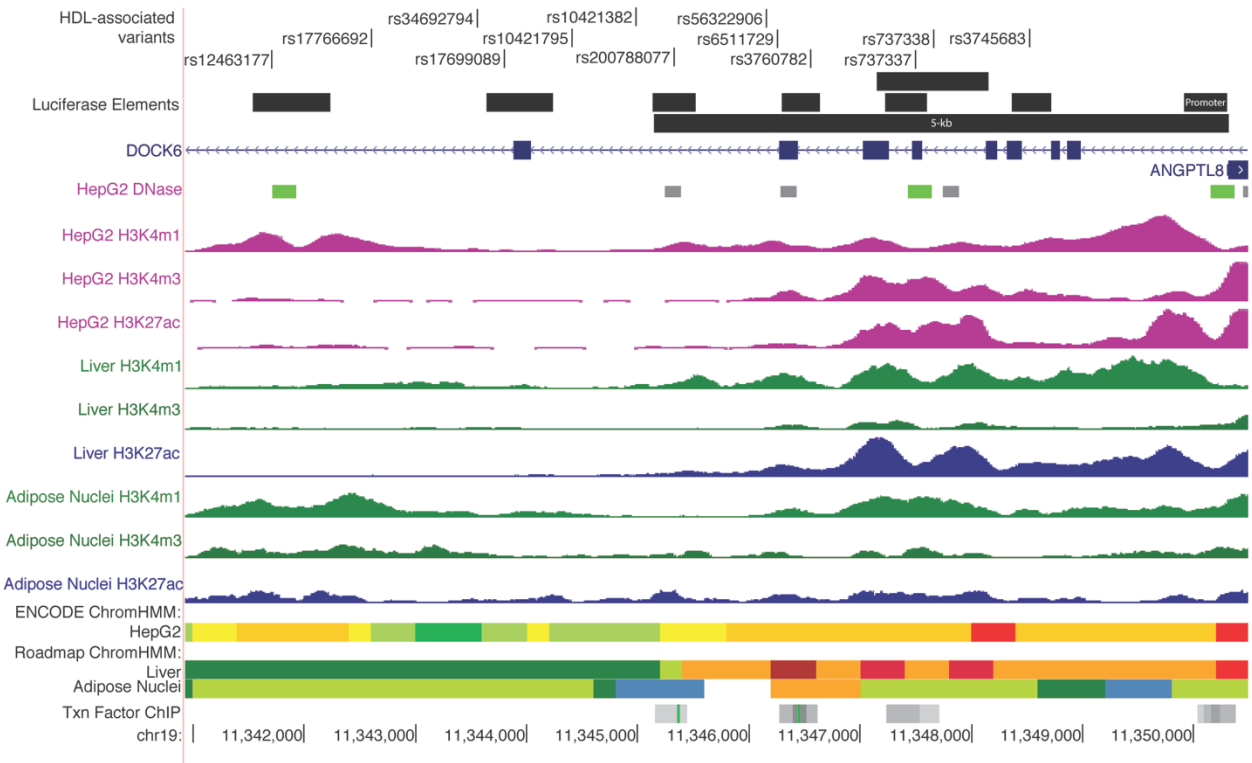

Figure 10: Thirteen variants overlap predicted regulatory regions. 13 variants overlap regulatory regions defined by histone marks, chromatin accessibility, and transcription factor binding. The full set of 42 candidate variants span a 39-kb region (Figure 18). Green rectangles denote DNase hypersensitivity peaks correlated with *ANGPTL8* expression across 112 cell lines in a previous study (153). DNase hypersensitivity correlation with *DOCK6* expression is not indicated because the correlated peaks do not overlap HepG2 or hepatocyte DNase peaks. Gray rectangles represent transcription factor binding; the identities of transcription factors are listed in Table 9. Data was accessed from ENCODE, the Epigenome Roadmap Atlas, and the UCSC Genome Browser. ENCODE ChromHMM: gray, heterochromatin; blue, insulator; green, transcription; yellow and orange, enhancer; red and pink, promoter. Roadmap ChromHMM: orange, enhancer; light green, genic enhancer; dark green, transcription; blue, heterochromatin; red, promoter. Black rectangles represent regions analyzed in transcriptional reporter assays.

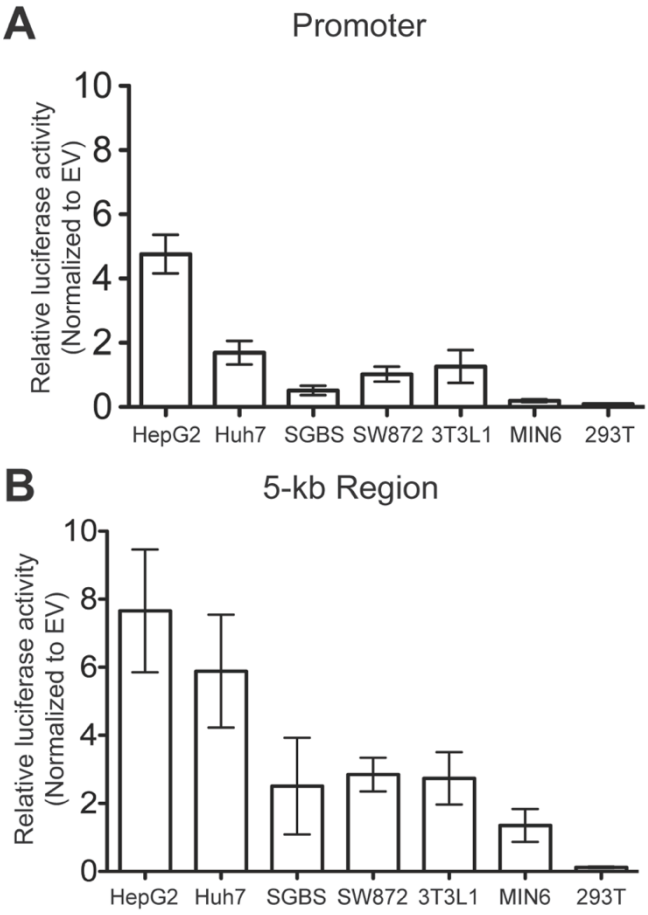

Figure 11: Cell-type specificity of *ANGPTL8* is influenced by nearby regulatory regions. Candidate regulatory regions were tested in a pGL4.10 vector in six cell types to determine drivers of tissue specificity. Reporter activity was normalized to empty vector (EV) in each cell type. Data are represented as the mean ± standard deviation of 10 biological replicates. A) 400-bp promoter B) a 5-kb regulatory region including the promoter. Comparison of the 400-bp promoter vs. 5-kb regulatory region $P<0.05$ for cell types. Cell lines include HepG2 human hepatocellular carcinoma (liver); Huh-7 human hepatocellular carcinoma; SGBS human preadipocyte; SW872 human adipocyte; 3T3L1 mouse adipocyte; MIN6 mouse pancreatic beta cell; and 293T human embryonic kidney.

Characterization of the ANGPTL8 HDL-C GWAS locus

To characterize and determine shared variants of the *ANGPTL8* association signal across populations, we analyzed the METSIM study of Finns (N=8,380) and the African Americans subset (N=8,244) from the Women's Health Initiative (WHI, (43)). In a preliminary METSIM analysis, the variants most strongly associated with HDL-C at the *ANGPTL8* locus were rs737337 ($P=7.7\text{x}10^{-4}$, ß=-0.09) and its LD proxies; these variants were more strongly associated with the concentration of phospholipids in medium HDL particles ($P=1.9\text{x}10^{-7}$, ß=-0.14; Figure 12A, Table 6) than HDL-C ($P=7.7\text{x}10^{-4}$, ß=-0.14) (145). Association analyses conditioned on rs737337 showed no evidence for any additional signals (all $P>8.0\text{x}10^{-4}$, Table 6, Figure 15B). In WHI, the lead variant associated with HDL-C is rs4804154 ($P=8.4\text{x}10^{-17}$), located 13 kb from

rs737337, and we did not observe evidence of any additional signals (P>4.8x10^{-4}, Figure 12B, Table 7, Figure 15D). rs4804154 is in low LD with the reported African American lead variant rs12979813 (r^2=0.18 AFR). rs12979813 is the most strongly associated genotyped variant in the previously reported African American sample (43), which includes the WHI study, but rs4804154 is the strongest variant after imputation in WHI. These two lead variants, rs737337 and rs4804154, are in moderate pairwise LD in Europeans (r^2=.67 EUR) and in low pairwise LD in Africans (r^2=.26 AFR). rs737337 was among the most strongly associated variants in WHI (P=9.1x10^{-10}) and the same was true of rs4804154 in METSIM (P=5.7x10^{-5}). The rare coding variant rs145464906 in *ANGPTL8* (137) is not present in METSIM and was not significantly associated with HDL-C in WHI (P=0.30). Conditioning on the common coding variant rs2278426 attenuates, but does not abolish, the signal in both populations (METSIM rs737337, P=0.03; WHI rs4804154, P=1.82 x 10^{-4}). Conditioning on the lead variant abolished the association with rs2278426 (WHI, P=0.78; METSIM, P=0.86), suggesting that the regulatory variants capture more of the association signal than rs2278426 alone. Because many GWAS loci are shared across populations, the presence of an association signal in both WHI and METSIM supports the hypothesis of at least one shared functional variant across populations.

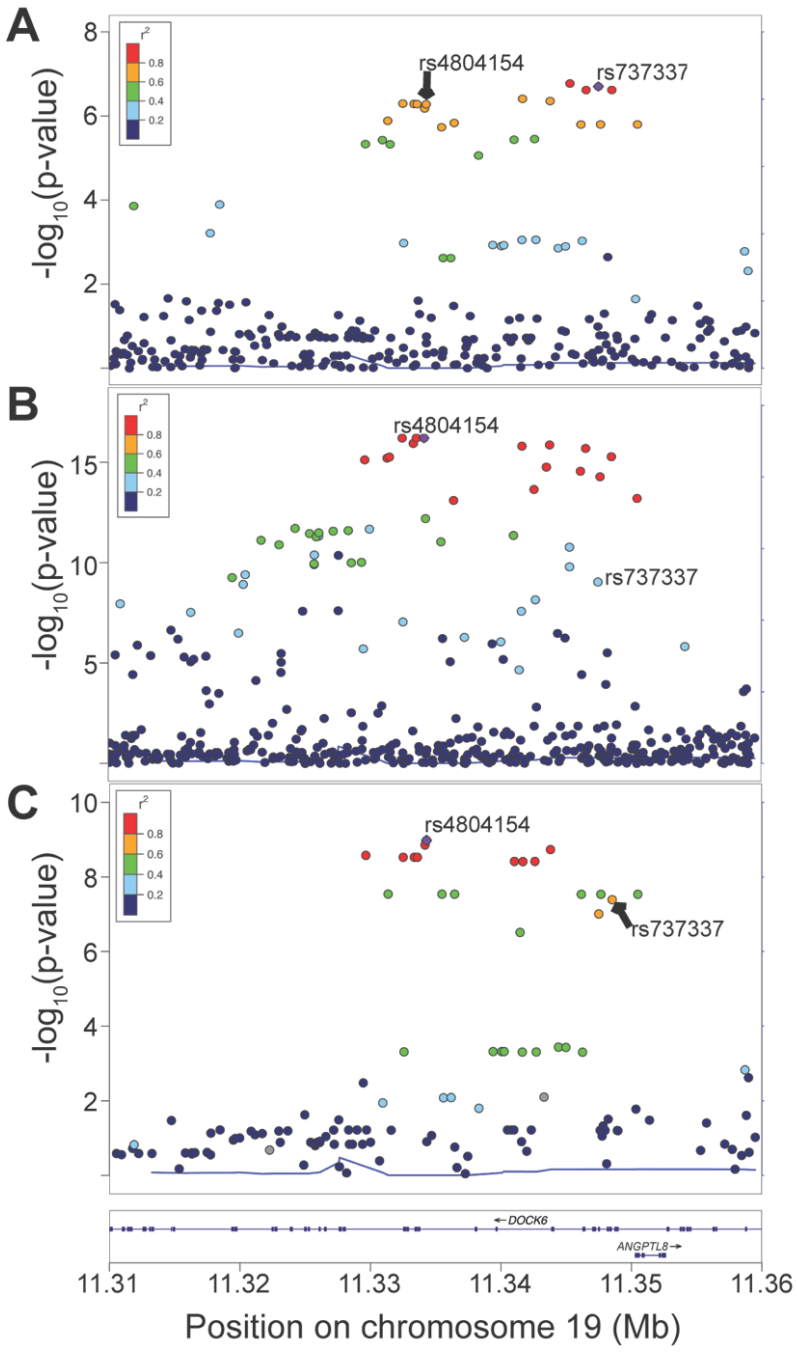

Figure 12: Locus plots of HDL-C and eQTL associations near *ANGPTL8*. A) Concentration of phospholipids in medium HDL in the METSIM study of Finnish individuals (*N*=8,380). Variants are colored according to LD (r^2) with rs737337 (purple), the lead variant in European meta-analyses by GLGC. B) HDL-C association in the WHI subset of African American individuals (*N*=8,244). Variants are colored according to LD (r^2) with rs4804154 (purple). C) HDL-C-associated variants are also associated with *ANGPTL8* expression in 770 subcutaneous adipose samples from the METSIM study. The European and African American lead variants (rs737337 and rs4804154) are among the most significant variants. Variants are colored according to LD (r^2) with rs4804154 (purple).

To further characterize the locus across populations, we conducted haplotype association analyses in METSIM and WHI (Table 3). By comparing variants across haplotypes and populations, we can hypothesize which variants represent the inheritance pattern of variants that have a functional effect. We included five variants in the analyses: the lead variants from European (rs737337), African American (rs4804154), Mexican/Pima Indian (rs2278426), and East Asian (rs3760782) studies and the one variant (rs3745683) in high LD (r^2>.8) with the leads in all four studies. The most common haplotype in both METSIM and WHI contained alleles individually associated with increased HDL-C (haplotype 1, Table 3). In both studies, the haplotype containing all the alleles individually associated with decreased HDL-C (haplotype 3) showed the strongest association with HDL-C ($\hat{\beta}$=-0.044, P<1.0x10^{-22} in WHI and $\hat{\beta}$=-0.146, P=5.3x10^{-7} in METSIM). Haplotypes 1 and 2 differed only for rs737337 alleles. Haplotype 2 is common in African Americans (21% frequency) and nearly absent in Finns (0.02% frequency), which explains the different extent of r^2-based LD in these populations. The small effect sizes of haplotype 2 ($\hat{\beta}$=-0.01 in WHI), 3 ($\hat{\beta}$=-0.044 in WHI, $\hat{\beta}$=-0.146 in METSIM), and 5 ($\hat{\beta}$=-0.009 in WHI, $\hat{\beta}$=-0.049 in METSIM) suggest that rs737337, rs2278426, and rs4804154 may contribute to, but are not alone responsible for, the association signal. These data are consistent with a signal shared across populations driven by one or more functional variants represented by rs3745683 and rs3760782, with potential additional contributions from rs737337, rs2278426, and/or other proxies.

Table 3: Haplotype association analyses in the WHI and METSIM studies.

WHI African American participants							Haplotype 1 reference			Haplotype 3 reference			Haplotype 4 reference		
	rs4804154	rs3760782	rs737337	rs3745683	rs2278426	Freq	Effect	SE	P	Effect	SE	P	Effect	SE	P
1	C	C	T	G	C	0.59	REF	REF	REF	0.042	0.005	8.8E-16	0.019	0.016	0.225
2	C	C	**C**	G	C	0.21	-0.010	0.055	0.051	0.032	0.006	1.5E-07	0.009	0.016	0.564
3	**T**	**T**	**C**	**A**	**T**	0.18	-0.044	0.005	<1E-22	REF	REF	REF	-0.025	0.016	0.126
4	**T**	**T**	**C**	**A**	C	0.01	-0.035	0.019	0.069	0.006	0.020	0.760	REF	REF	REF
5	**T**	C	T	G	C	0.01	-0.009	0.026	0.735	0.040	0.026	0.125	0.014	0.031	0.661
METSIM Finnish participants							**Haplotype 1 reference**			**Haplotype 3 reference**			**Haplotype 4 reference**		
	rs4804154	rs3760782	rs737337	rs3745683	rs2278426	Freq	Effect	SE	P	Effect	SE	P	Effect	SE	P
1	C	C	T	G	C	0.88	REF	REF	REF	0.147	0.029	4.4E-07	0.124	0.056	0.029
2	C	C	**C**	G	C	.0002	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
3	**T**	**T**	**C**	**A**	**T**	0.06	-0.146	0.029	5.3E-07	REF	REF	REF	-0.023	0.063	0.720
4	**T**	**T**	**C**	**A**	C	0.02	-0.120	0.056	0.034	0.027	0.063	0.660	REF	REF	REF
5	**T**	C	T	G	C	0.03	-0.049	0.040	0.210	0.097	0.048	0.041	0.074	0.068	0.280

Haplotype association analyses in 8,244 African Americans in WHI and 8,380 Europeans in METSIM. Alleles shown in bold differ from haplotype 1. In both studies, alleles shown in haplotype 3 were individually associated with decreased HDL-C. Freq, haplotype frequency; SE, standard error; N/A, haplotype was too rare to be analyzed; REF, reference haplotype for interpreting association analyses

eQTL associations with HDL-C-associated variants and nearby genes

To determine which gene(s) the HDL-C-associated variants may affect, we investigated eQTL associations. We observed an eQTL association in subcutaneous adipose tissue from 770 METSIM study participants for both *ANGPTL8* (rs4804154, P=1.0x10^{-9}; Figure 12C, Table 4, Figure 16) and *DOCK6* (rs17699089, P=7.2x10^{-7}; Table 4, Figure 16, Table 8), but not any other gene within 1 Mb of rs737337 (Table 8). The METSIM and WHI HDL-C-associated variants were among the variants most strongly associated with both *ANGPTL8* and *DOCK6* mRNA levels (Figure 12, Table 8), suggesting that the same variants associated with HDL-C may act by affecting expression level of *ANGPTL8* and/or *DOCK6*. Conditional analyses were performed to confirm the coincidence of the GWAS and eQTL signals. For both genes, alleles associated with lower HDL-C levels were associated with lower mRNA levels. In addition, *ANGPTL8* mRNA level was associated with HDL-C level in METSIM samples (P=0.017, Figure 17), whereas *DOCK6* was not (P=0.42, Figure 17). Evidence that the variants most strongly associated with HDL-C are also most strongly associated with *ANGPTL8* mRNA levels suggests that a regulatory mechanism acts at this GWAS locus.

Table 4: Associations with *ANGPTL8* and *DOCK6* expression in subcutaneous adipose tissue.

Variant	Alleles[a]	Gene	Effect[b]	Standard Error	*P*-value	P_{cond}-eSNP	P_{cond}-rs12463177	r^2 with eSNP (EUR)[c]	r^2 with eSNP (AFR)[c]
rs4804155	G/C	*ANGPTL8*	-0.499	0.081	1.04×10^{-9}	-	0.516	-	-
rs737337	C/T	*ANGPTL8*	-0.526	0.098	9.74×10^{-8}	0.244	0.402	0.67	0.40
rs4804154	T/C	*ANGPTL8*	-0.500	0.081	1.38×10^{-9}	0.762	0.270	1.00	0.48
rs12463177	C/G	*ANGPTL8*	-0.479	0.080	3.84×10^{-9}	0.570	-	0.94	0.47
rs17699089	G/A	*DOCK6*	-0.406	0.081	7.21×10^{-7}	-	0.005	-	-
rs737337	C/T	*DOCK6*	-0.298	0.099	2.68×10^{-3}	0.324	0.413	0.74	0.29
rs4804154	T/C	*DOCK6*	-0.398	0.082	1.65×10^{-6}	0.253	0.267	0.94	0.91
rs12463177	C/G	*DOCK6*	-0.382	0.081	3.06×10^{-6}	0.029	-	1.00	0.98

Lead eQTL variants for *ANGPTL8* (rs4804155) and *DOCK6* (rs17699089), lead GWAS variants (rs4804154 and rs737337) and functional candidate variant rs12463177 association with *ANGPTL8* and *DOCK6* expression in 770 primary subcutaneous adipose samples. Conditional analysis on each lead eSNP and the candidate functional variant rs12463177 attenuated both the *ANGPTL8* and *DOCK6* association signals.

[a]The HDL-C-decreasing and eQTL effect alleles are presented first

[b]Effect size is shown as the inverse normal transformed expression levels ($\log_2$(fluorescence intensity)) with each additional copy of the allele.

[c]r^2 is calculated from 1000 Genomes Phase 1 data

Selection of variants to test for allelic differences in regulatory activity

Prioritizing variants at GWAS loci for functional study can be challenging, especially at loci with regulatory mechanisms. To narrow variants that may have regulatory function, we considered three methods: LD with the lead GWAS variants, three fine-mapping algorithms in each of METSIM and WHI, and overlap with predicted regulatory regions (Figure 18). First, we considered variants in LD with the lead GWAS variants in European and/or African American ancestry individuals (1000 Genomes EUR and AFR). Because the variants in Europeans are of moderate allele frequency (MAF~0.07), we considered variants that meet an LD threshold r^2>.5. In total, 42 variants exhibited r^2>.5 with a lead HDL-C-associated variant in at least one study and could be considered as candidates for regulatory function (Table 9, Figure 19). As a second approach to prioritize candidate variants at this locus, we used the MANTRA, CAVIAR, and PAINTOR algorithms to interpret the HDL-C associated variants in the METSIM Finnish and WHI African American ancestry groups (Table 10, 11, 12). Finally, we identified candidate variants based on simple positional overlap with evidence of predicted regulatory regions, as variants overlapping these regions are more likely to be functional (Figure 19, Table 9) (59,60,152,153). Nine variants identified in at least one fine-mapping analysis overlapped regulatory regions (rs12463177, rs17766692, rs17699089, rs200788077, rs56322906, rs3760782, rs737337, rs737338, and rs3745683). We considered these nine variants to represent the most plausible candidate variants based both on fine-mapping and regulatory overlap and examined them for allelic differences on regulatory function. To compare the set of nine most plausible candidate variants to those that only show regulatory overlap, we also examined four additional variants in regulatory regions that were not identified by any fine-mapping analysis (rs34692794, rs10421795, rs10421382, and rs6511729). Thus, we examined a total of 13 variants in assays

examining allelic effects on regulatory function. All 13 variants are located within 9 kb of the *ANGPTL8* transcription start site (Figure 10, Table 9).

Functional characterization of candidate variants

We examined the candidate variants for allelic differences in assays of regulatory function in human liver-derived (HepG2, Huh-7), pre-adipocyte (SGBS), and adipocyte (SW872) cells. We chose these cell types because *ANGPTL8* is expressed in liver and adipose, they had the highest transcriptional activity in our cell type-specificity assays, and they have roles in HDL-C metabolism. We tested variants individually or as a haplotype in luciferase transcriptional reporter assays and/or in electrophoretic mobility shift assays (EMSA, Figure 20, 21). Among the variants analyzed, rs12463177, which was identified in all three fine-mapping analyses and is a candidate at r^2>.5 in both populations (r^2=0.74 with rs737337 EUR, r^2=0.93 with rs4804154 AFR, and is 8 kb upstream from the *ANGPLT8* TSS showed significant (P<.05) allelic differences in two assays of regulatory function (Figure 13). In luciferase transcriptional reporter assays in HepG2, a 697-bp region containing rs12463177-G showed 1.2 to 1.4-fold increased enhancer activity (P=0.02 forward orientation; P=0.004 reverse) compared to rs12463177-C. While modest, this allelic difference was replicated in three independent experiments. The significant effect on transcriptional activity was not replicated in Huh-7, SGBS, and SW872 cells, in which this region showed repressor activity; however, the trend between the alleles is consistent with HepG2, at least in the reverse orientation (Figure 20). In EMSAs using HepG2 and SGBS nuclear extract, the rs12463177-G allele showed increased binding (Figure 13, 21, 22). The direction of effect of rs12463177 transcriptional activity is consistent with the eQTL direction. A functional role for rs12463177 is consistent with

additional evidence that rs12463177 overlaps a DNase hypersensitivity site that was previously correlated to *ANGPTL8* expression (153), green rectangle, Figure 10) and that the eQTL association signal is attenuated when conditioned on rs12463177 (Figure 16, Table 4).

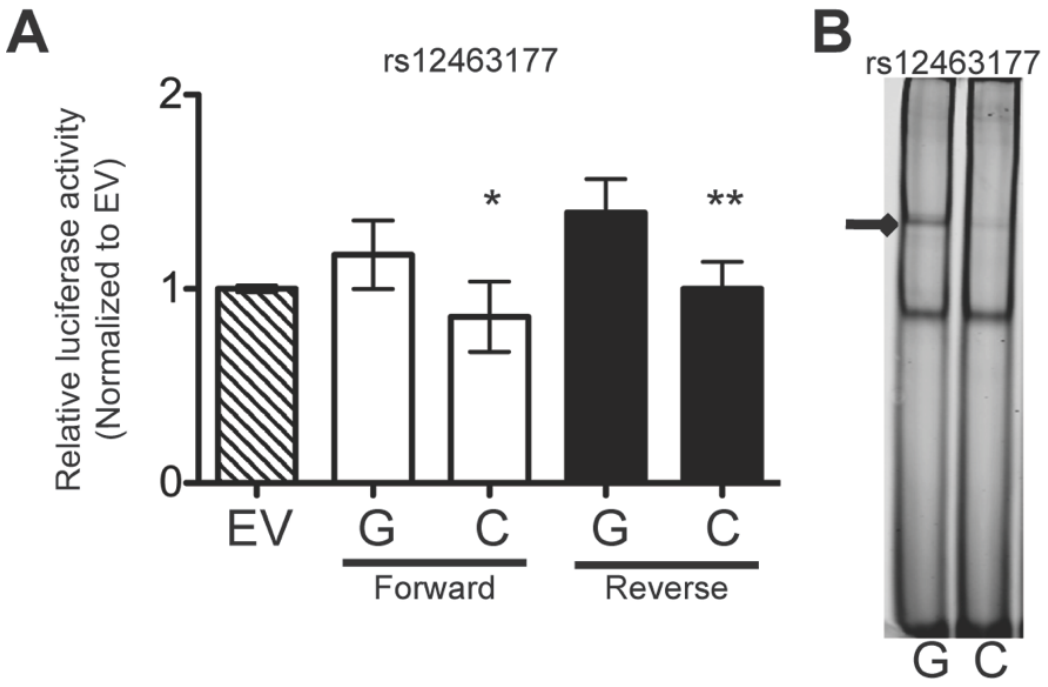

Figure 13: Allelic differences in regulatory assays of rs12463177 A) 698-bp region containing either allele of rs12463177 was cloned into the pGL4.23 vector and transfected in HepG2 cells. Data are represented as the mean ± standard deviation of 3-5 biological replicates. Luciferase activity was normalized to empty vector (EV) and p-values were determined by t-tests. **P*<0.05; ***P*<0.01. B) Differential protein binding was evaluated *in vitro* using EMSA. IR-labeled probes containing either allele of rs12463177 were incubated with 10 ug HepG2 nuclear protein. The arrow shows stronger binding for rs12463177-G. Consistent binding was observed with SGBS nuclear protein (Figure 21). HDL-C-increasing alleles are presented first.

Six additional variants (rs17699089, rs200788077, rs56322906, rs3760782, rs737337, and rs3745683) showed evidence of allelic differences in an assay for regulatory function and two (rs17766692 and rs737338) did not (Figure 20, 21, 22, 23). rs56322906 was only tested in transcriptional reporter assays as part of the 5-kb haplotype described below (Figure 25). rs737337 showed by far the strongest enhancer activity in transcriptional reporter assays, an up to 60-fold increase compared to empty vector in HepG2 (Figure 20), and rs737337-C shows

strong allele-specific binding in liver cell types (Figure 21). Supershift experiments using HepG2 nuclear extract identified RXRα as binding to this probe, although not as part of the allele-specific complex (Figure 24). Interestingly, the four variants that overlap regulatory regions but were not predicted by any fine-mapping analysis (rs34692794, rs10421795, rs10421382, and rs6511729) did not show allelic differences in protein binding (Figure 23). In total, we observed evidence of functional activity for seven of the nine candidate regulatory variants in these assays (Figure 14).

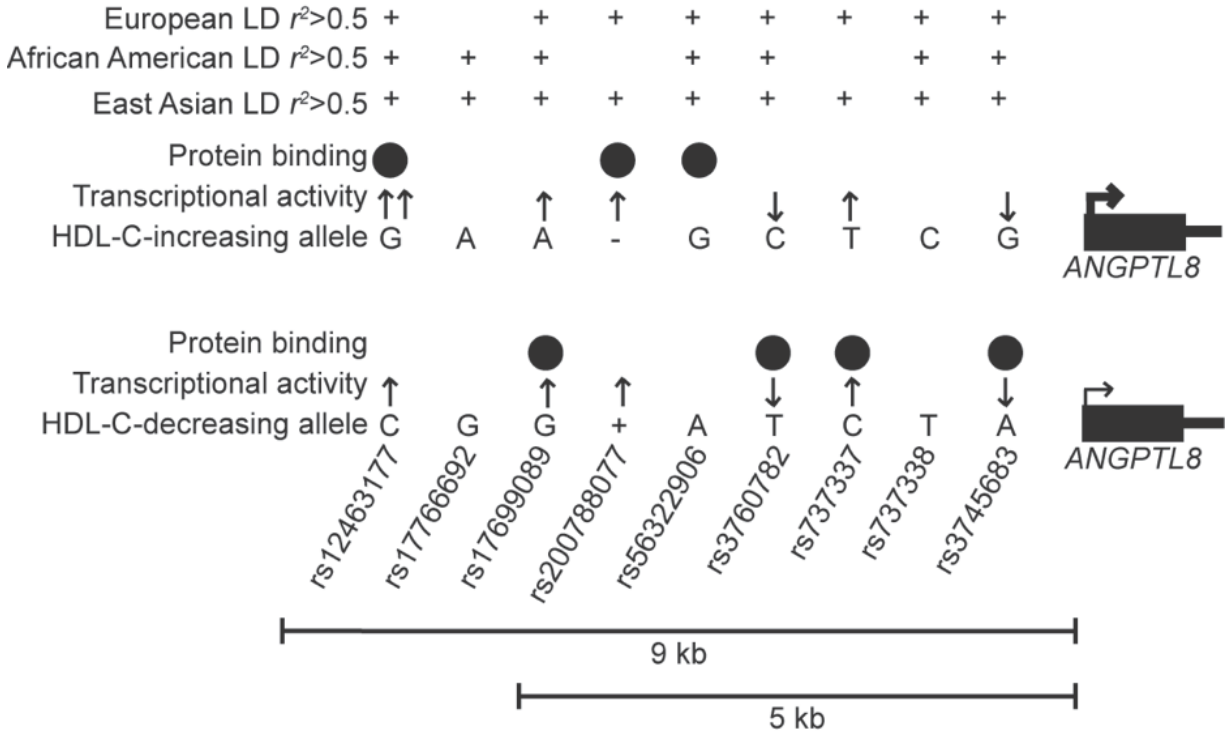

Figure 14: Summary of functional results for seven candidate variants. Summary of shared LD with lead HDL-C-associated variants reported in European (rs737337), African American (rs4804154), and East Asian (rs3760782) populations are shown; variants meeting the r^2>0.5 threshold are marked with a +. Results from transcriptional reporter luciferase assays and electrophoretic mobility shift assays (EMSA) for nine candidate variants are shown. Arrows show the direction of transcriptional activity in reporter assays. Two arrows at rs12463177-G indicate allelic differences in transcriptional activity. Black circles show allele-specific protein binding observed in EMSA experiments. The larger and smaller arrows at the *ANGPTL8* promoter indicate higher and lower expression level from adipose eQTL data. Variants are located within 9 kb of the *ANGPTL8* transcription start site; variant distances are not drawn to scale.

To examine the effect on transcriptional activity of multiple variants together with the *ANGPTL8* promoter, we tested 5-kb haplotypes located immediately upstream of the *ANGPTL8* transcription start site. This region includes five of the seven variants showing allelic differences in protein binding and/or transcriptional activity when examined separately (rs200788077, rs56322906, rs3760782, rs737337, and rs3745683, Figure 20) and two variants that did not show evidence of allelic differences (rs6511729 and rs737338, Figure 23). Transcriptional reporter assays of the smaller segments containing individual variants had shown both activator (e.g. rs737337) and repressor (e.g. rs3760782 and rs3745683) activity. The 5-kb haplotype acted as an enhancer in HepG2, Huh-7, SGBS, and SW872, with intermediate activity observed between the individual segments, but did not show significant differences in transcriptional activity between the two haplotypes (Figure 25). These results suggest a complex regulatory mechanism involving enhancer and repressor regions that work in concert to regulate expression.

Discussion

In this study, we examined the tissue specificity of *ANGPTL8,* reported the first eQTL for variants at the *ANGPTL8* HDL-C GWAS locus in adipose tissue, and identified variants at the GWAS locus that exhibit allelic differences in assays of regulatory function. We found that a 400-bp promoter is the main driver of tissue specificity in liver, and that expression may be mediated by additional enhancer elements within 5 kb upstream of *ANGPTL8*, especially in adipocytes. Of 7 candidate regulatory variants that showed allelic differences in transcriptional activity and/or protein binding, rs12463177 showed the clearest allelic differences consistent with the direction of the *ANGPTL8* eQTL. These data suggest that multiple regions and potentially multiple variants regulate *ANGPTL8* expression in liver and adipose tissue.

ANGPTL8 is a strong candidate gene at this GWAS locus. Although we observed coincident association between the GWAS variants and transcript level for both *ANGPTL8* and *DOCK6* in subcutaneous adipose tissue samples, *ANGPTL8* mRNA level was associated with HDL-C in METSIM, whereas *DOCK6* mRNA level was not (Figure 17). DNase hypersensitivity sites that overlap regulatory variants are correlated to *ANGPTL8* expression (Figure 10), providing further support for *ANGPTL8* as the target gene. *ANGPTL8* protein levels have been shown to be inversely associated with HDL-C (134). This direction is opposite of the association we observed with *ANGPTL8* expression and HDL-C; however, others have suggested that ELISA methods may not consistently quantify serum ANGPTL8 levels (156). Furthermore, ANGPTL8 inhibits lipoprotein lipase when co-expressed with ANGPTL3, giving a direct connection to lipid metabolism (133). While we cannot rule out a role for *DOCK6,* these lines of evidence and a rare coding variant in *ANGPTL8* associated with HDL-C (137) suggest *ANGPTL8* as the most likely target gene at this HDL-C GWAS locus.

The *ANGPTL8* HDL-C GWAS locus exhibited unusual characteristics in fine-mapping because the associated variants defined by LD r^2 with the lead GWAS variants span a larger chromosomal region in African Americans than in Europeans, opposite of most loci in the genome (157). Our haplotype association analyses showed that the LD differences are due to a haplotype with 20% frequency in African Americans from WHI that is essentially not observed in Finns from METSIM (0.02% frequency). At this locus, low-frequency haplotypes (frequency<.03, Table 3) distinguish r^2 thresholds of .8 and .5 in Finns, and the r^2 dependence on allele frequency suggests that a LD threshold of r^2>.8 for selecting candidate variants may be too restrictive and miss potentially functional variants. Here, we considered an initial set of candidate variants based on a more liberal threshold of r^2>.5, resulting in 42 total variants. We

then tested nine variants predicted by MANTRA, CAVIAR, and PAINTOR and that overlapped regulatory datasets in functional assays. Seven of nine variants showed evidence of allelic differences (Figure 14), but only rs12463177 and rs17699089 showed differences consistent with the direction of the eQTL association. Of four additional variants that overlapped regulatory regions but were not predicted to be functional in fine-mapping assays, none showed evidence of regulatory activity. Taken together, these data suggest that the joint use of fine-mapping and regulatory overlap can successfully identify variants exhibiting allelic differences in functional assays.

The mechanisms by which the variants that showed allelic differences in functional assays may work in concert remains unclear. The effect on transcriptional activity of rs12463177 was modest and only significant in one cell type (Figure 13, 20). This marginal effect observed in cell lines may not represent the physiological effect *in vivo.* The magnitude of effect of rs737337 was much greater (50-fold increased transcriptional activity), but did not show significant allelic differences, in contrast to a previous report (158). rs737337 exhibited strong liver-specific allele-specific protein binding, but the specific transcription factor(s) remain unknown. One possible mechanism is that the regulatory region containing rs737337 is a strong enhancer that drives expression of *ANGPTL8,* but that the region containing rs12463177 is important for regulating allele-specific expression. Transcriptional regulators bound at the multiple variants may act together via chromosomal looping with the *ANGPTL8* promoter. Further experiments, especially *in vivo*, are needed to elucidate the precise roles and interactions of the seven variants that showed allelic differences in transcriptional activity and/or protein binding.

In this study, we identified *ANGPTL8* as the target gene at this HDL-C GWAS locus, determined regulatory drivers of tissue specificity, and combined fine-mapping approaches and regulatory overlap with experimental assays to identify variants that may contribute to the HDL-C GWAS signal at *ANGPTL8.* Identifying variants underlying GWAS loci contributes to a growing understanding of target genes, their direction of effect, and metabolic phenotypes. Our results are consistent with previously described results at other GWAS loci where multiple common regulatory variants act together (54,138,139) and continued work on the *ANGPTL8* locus will further clarify the complex mechanism of these variants.

Supplemental Figures and Tables

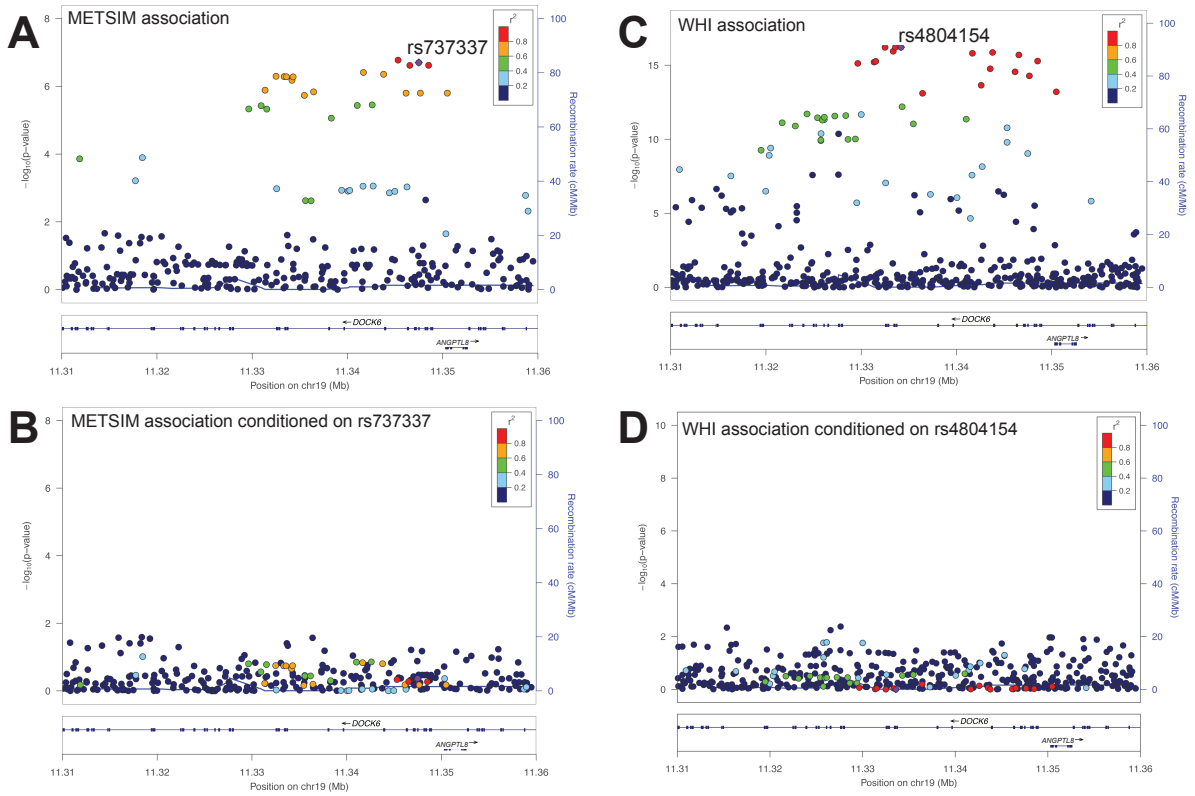

Figure 15: HDL association and conditional analysis. A) Variant association with concentration of phospholipids in medium HDL in the METSIM study of Finnish individuals (N=8380). rs737337 (purple) was among the most significantly associated variants. Variants are colored according to LD (r^2) with rs737337. B) Conditional analysis on rs737337 attenuated the association signal. C) HDL-C association in the WHI study of African American individuals (N=8244). Variants are colored according to LD (r^2) with rs4804154 (purple). D) Conditional analysis on rs4804154 attenuated the signal.

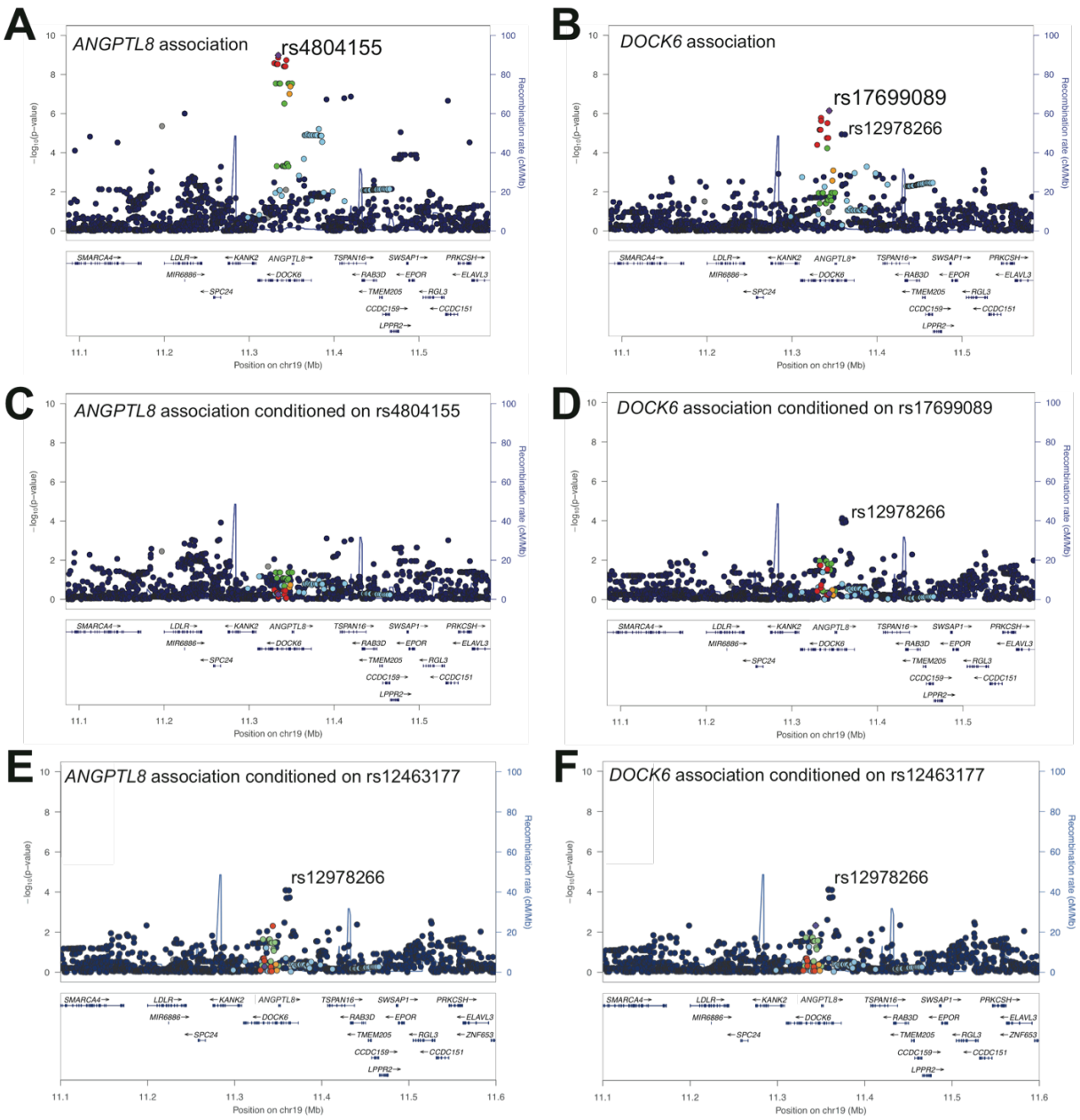

Figure 16: eQTL association in subcutaneous adipose from 770 individuals in the METSIM study. A) HDL-C-GWAS variants are associated with *ANGPTL8* expression. B) HDL-C GWAS variants are associated with *DOCK6* expression. C) Conditional analysis on the top variant associated with *ANGPTL8* expression, rs4804155, attenuated the signal. D) Conditional analysis on the top variant associated with *DOCK6* expression, rs17699089, revealed a secondary association signal with *DOCK6* expression (rs12978266, P=7.33x10-5). E) Conditional analysis on candidate functional variant rs12463177 also attenuated the *ANGPTL8* association signal and reveals the secondary association of rs12978266. F) Conditional analysis on candidate functional variant rs12463177 also attenuated the *DOCK6* association signal.

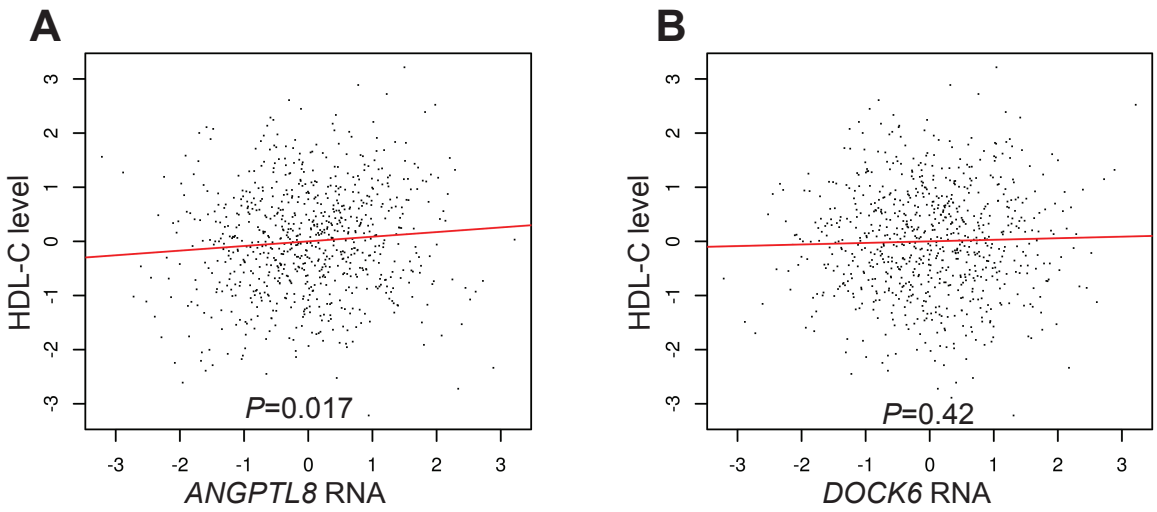

Figure 17: RNA associations with HDL-C in METSIM. *ANGPTL8* RNA levels (A) are associated with HDL-C level in 770 Finnish individuals from METSIM. *DOCK6* RNA levels (B) are not associated with HDL-C. To examine the relationship between RMA-normalized expression levels and HDL-C, we adjusted both traits for age and BMI, inverse normal transformed the residuals, and then tested for association in regression analysis. Correlation coefficients: R=0.086 (*ANGPTL8*), R=0.029 (*DOCK6*).

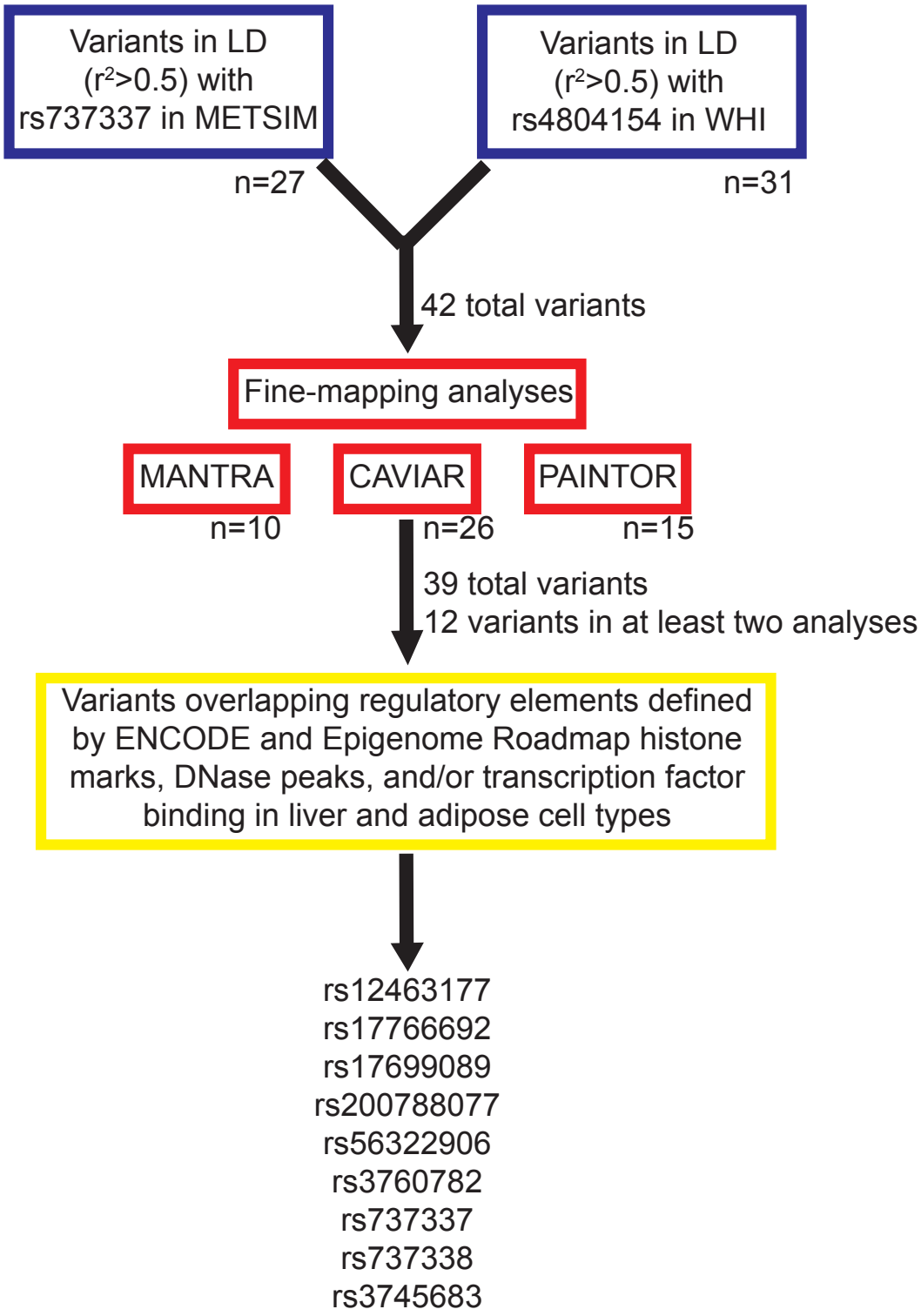

Figure 18: Flow chart describing selection of variants to test in functional experiments.

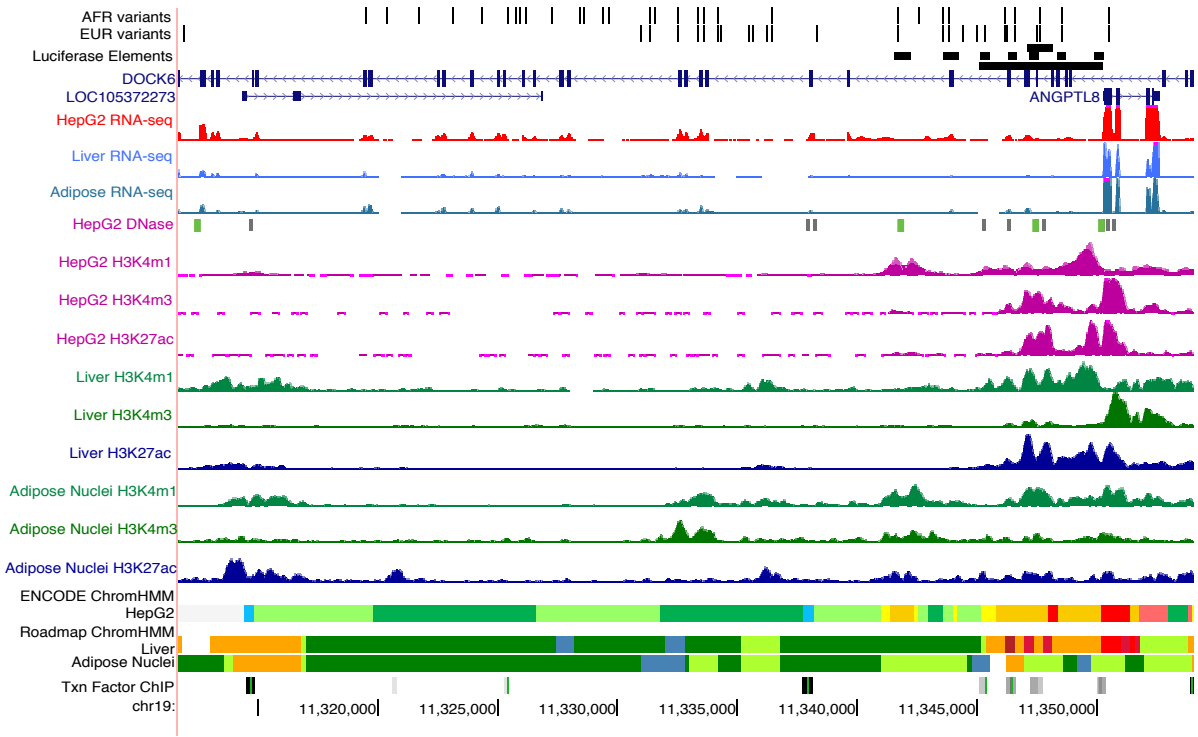

Figure 19: Candidate variants relative to predicted regulatory regions. 27 variants exhibited r^2>0.5 with rs737337 in METSIM (EUR variants) and 31 variants with rs4804154 in WHI (AFR variants). These variants span a 39-kb window within *DOCK6*. 13 of 42 total variants overlap regulatory regions defined by histone marks, chromatin accessibility, and transcription factor binding (Figure 10). Green rectangles represent DNase hypersensitivity sites correlated with *ANGPTL8* expression (153). Consistent with our tissue-specificity experiments, *ANGPTL8* is highly expressed in liver and adipose RNA-seq datasets (59,60,130). Data was accessed from ENCODE, the Epigenome Roadmap Atlas, and the UCSC Genome Browser. Black rectangles represent regions analyzed in transcriptional activity assays.

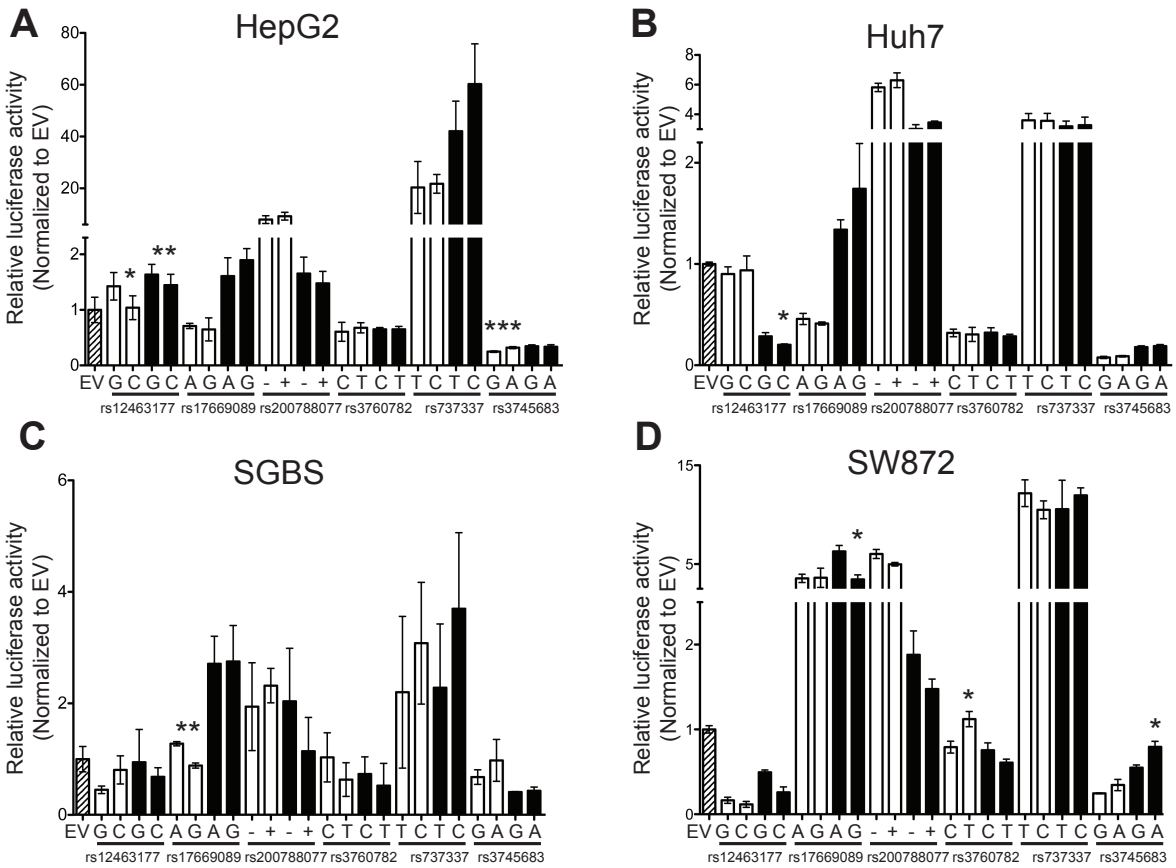

Figure 20: Variants tested in transcriptional reporter luciferase assays. Transcriptional activity was evaluated for six variants (rs56322906 was evaluated in a 5-kb haplotype, Figure 25) in HepG2 (A), Huh7 (B), SGBS (C), and SW872 (D) cells. Data are represented as the mean ± standard deviation of 3-5 biological replicates. Luciferase activity was normalized to empty vector (EV). White bars represent the forward orientation with respect to the genome; black are reverse. **P*<0.05, ***P*<0.01 ****P*<0.001 HDL-C-increasing alleles are presented first.

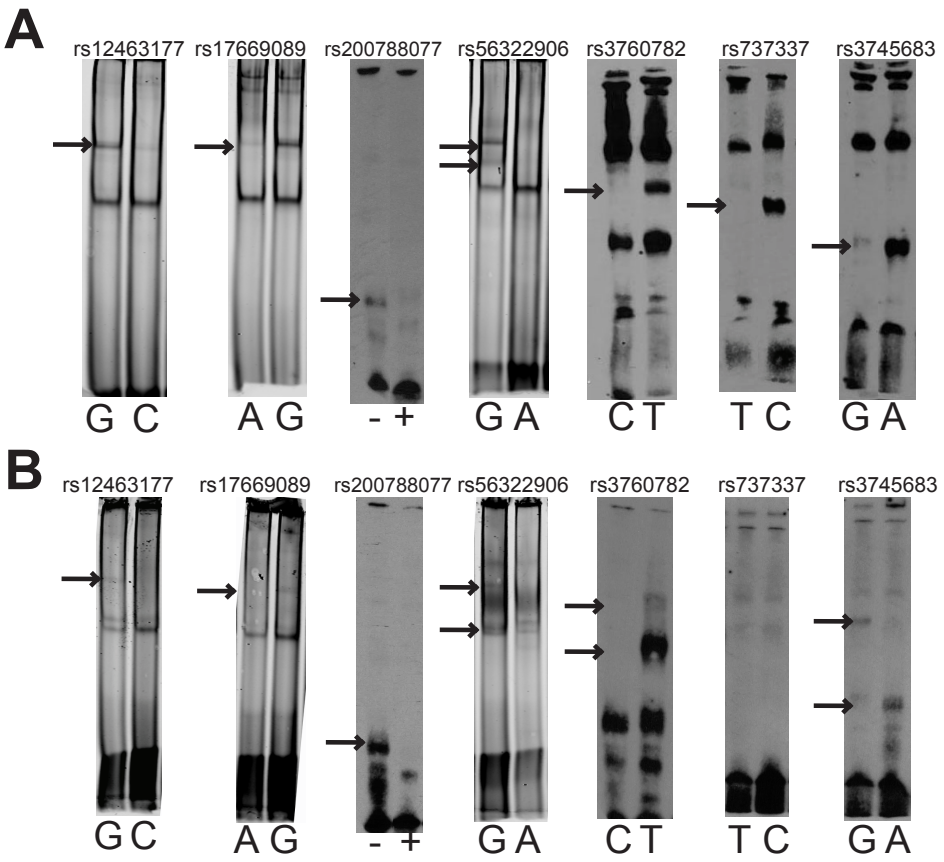

Figure 21: Seven variants show differential protein binding in EMSAs.
Allelic differences in protein binding were observed for all seven variants with nuclear extract from HepG2 cells (A). Six variants (except rs737337) showed allelic differences in protein binding with nuclear extract from SGBS cells (B). Arrows show allelic differences.

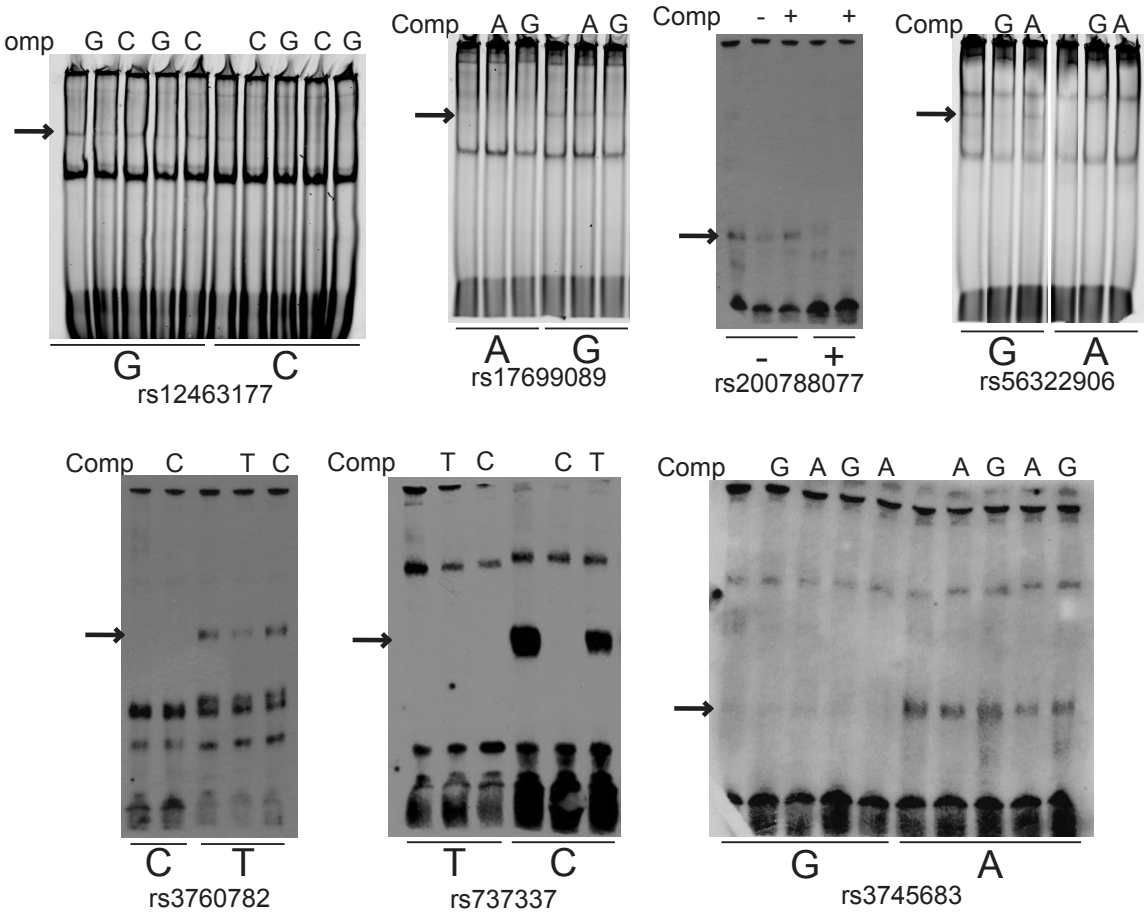

Figure 22: Competition EMSAs confirm allele-specific effects.
Competition EMSA experiments using HepG2 nuclear extract were conducted with unlabeled competitor probes for each allele. rs12463177 is competed with 100x competition (lanes 2, 3, 7, 8) and 200x (lanes 4, 5, 9,10) compared to labeled probe. rs17699089 is competed with 100x competition, rs200788077 is competed with 100x competition, rs56322906 is competed with 100x competition, rs3760782 is competed with 192x competition, rs737337 is competed with 269x competition, and rs3745683 is competed with 50x (lanes 2, 3, 7, and 8) and 100x (lanes 4, 5, 9, and 10) competition. HDL-C-increasing alleles are presented first.

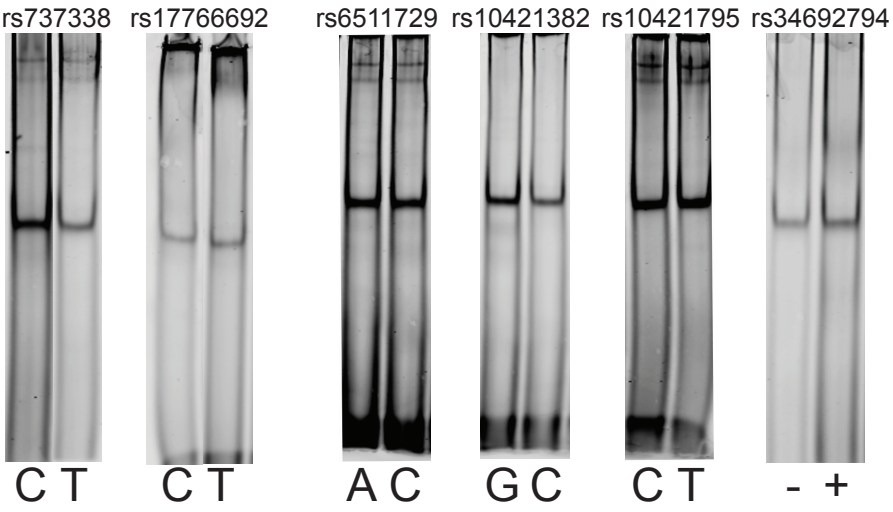

Figure 23: Six variants overlapping regulatory regions did not alter protein binding. EMSAs were performed with IR-labeled probes for each allele and incubated with 10 ug of HepG2 nuclear protein. rs737338 and rs17766692 were predicted in the CAVIAR causal set. The remaining four variants were not predicted in fine-mapping analyses. The prominent band in the middle of all gels represents non-specific binding that is observed in all IR-labeled EMSAs. HDL-C-increasing alleles are presented first. No allelic differences were observed.

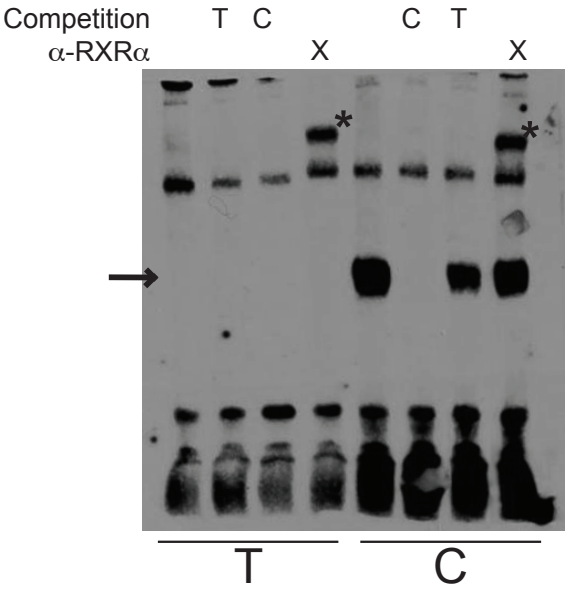

Figure 24: RXRα may bind as part of a complex at rs737337. Supershift EMSA assays were performed using HepG2 nuclear extract (NE). The allele-specific band (arrow) is not disrupted when RXRα antibody is added to the reaction; however, there is a supershift in both alleles (asterisks). No disruption or supershift is observed with 36 other transcription factor antibodies (data not shown; Table 9). Competitor lanes are competed with 269x unlabeled probe. HDL-C-increasing alleles are presented first.

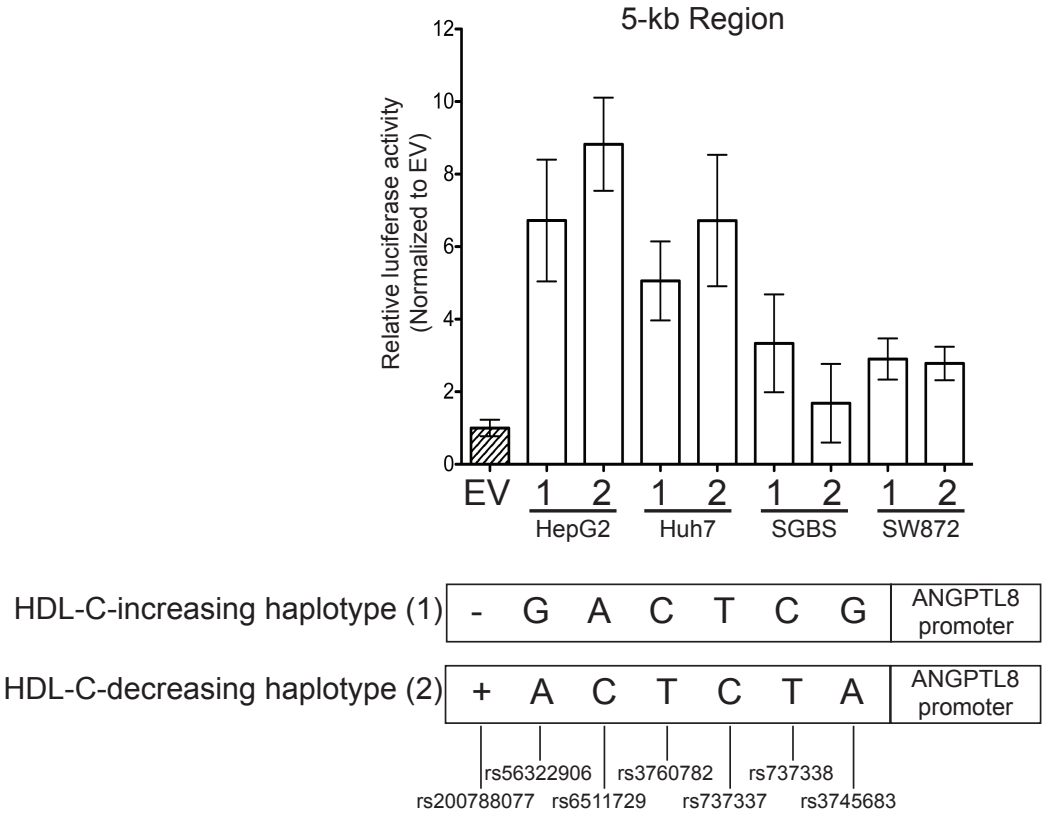

Figure 25: A 5-kb haplotype did not show allelic differences in transcriptional activity. Transcriptional activity was evaluated for a 5-kb haplotype containing 7 variants in HepG2, Huh7, SGBS, and SW872 cells. Data are represented as the mean ± standard deviation of 3-5 biological replicates. Luciferase activity was normalized to empty vector (EV). *P*>0.07 The HDL-C-increasing haplotype is haplotype 1.

Table 5: Primer and probe sequences for functional assays

Primer sequences for luciferase assays	5'- 3' Sequence	Chromosome Position (hg19)
rs737337_F	gcaccagggtgaagaatttg	chr19:11347220-11347605
rs737337_R	atcagtcagggagggctga	
rs737337_long_F	tcagcacagtgtccttgagc	chr19:11347169-11348145
rs737337_long_R	tgctcacacccgatgtatgt	
rs3745683_F	ctggcagctgacatggtaga	chr19:11348359-11348719
rs3745683_R	tatgtaggggacacgtgag	
rs3760782_F	agtgccaggaaggcgaaag	chr19:11346293-11346640
rs3760782_R	aggtgacagtgagccgagat	
rs200788077_F	ccctgagaataatgcctgaca	chr19:11345134-11345515
rs200788077_R	aatgttttgcccacttttgc	
ANGPTL8prom_F	ggaggggaacaagagcagat	chr19:11349914-11350304
ANGPTL8prom_R	tctaaggtatagccacagcac	
rs12463177_F	gctggtagggggtgagg	chr19:11341542-11342239
rs12463177_R	tgtgcttgagttagggctga	
rs17699089_F	ttgttcagccacgccaag	chr19:11343636-11344234
rs17699089_R	cctggcctattctcagtttttc	
5kb_F	ccctgagaataatgcctgaca	chr19:11345134-11350304
5kb_R	tctaaggtatagccacagcac	
Probe sequences for EMSA	**5'- 3' Sequence**	
rs737337_T	gacacggctgtgagctc	
rs737337_C	gacacggccgtgagctc	
rs3760782_C	aggggtcacaaattttt	
rs3760782_T	aggggtcataaattttt	
rs200788077_+	aggaaaaaacaggctca	
rs200788077_-	aggaaaaacaggctca	
rs3745683_G	tcacctctgccatgcca	
rs3745683_A	tcacctctaccatgcca	
rs12463177_G	tgtgcaccgtgagggcc	
rs12463177_C	tgtgcaccctgagggcct	
rs56322906_G	cgaactcctgacctcaaat	
rs56322906_A	cgaactcctaacctcaaat	
rs6511729_A	ctcgtctcacagggatt	
rs6511729_C	ctcgtctcccagggatt	
rs10421382_G	tgcctggcgtattttat	
rs10421382_C	tgcctggcctattttat	
rs10421795_C	ccagttacctggggagg	
rs10421795_T	ccagttacttggggagg	
rs17699089_A	tgtttcccatgcttcat	
rs17699089_G	tgtttcccgtgcttcat	
rs34692794_-	atatgcatgggggtg	
rs34692794_G	atatgcatgggggggtg	
rs17766692_C	tggatttgcacttcgtt	
rs17766692_T	tggatttgtacttcgtt	
rs737338_C	gtgtagcccggtctggg	
rs737338_T	gtgtagcctggtctggg	

Table 6: Association of 100 ANGPTL8 locus variants with concentration of phospholipids in medium HDL in METSIM

				Unconditioned			Conditioned on rs737337		
Variant	Position	Alleles[a]	MAF	Effect	Std Error	P-value	Effect	Std Error	P-value
rs737337	19:11347493	T/C	0.08	-0.137	0.026	1.99E-07	-	-	-
rs112108870[b]	19:11345315	G/GA	0.08	-0.138	0.026	1.68E-07	-0.369	0.499	0.460
rs3745683	19:11348521	G/A	0.08	-0.137	0.026	2.39E-07	0.359	0.576	0.533
rs3760782	19:11346550	C/T	0.08	-0.137	0.026	2.40E-07	0.367	0.576	0.525
rs12463177	19:11341680	G/C	0.12	-0.115	0.023	3.88E-07	-0.056	0.039	0.146
rs17699089	19:11343795	A/G	0.12	-0.114	0.023	4.39E-07	-0.055	0.039	0.156
rs3826815	19:11332505	C/T	0.12	-0.114	0.023	5.03E-07	-0.053	0.039	0.178
rs72994363	19:11333358	G/T	0.12	-0.114	0.023	5.14E-07	-0.053	0.039	0.180
rs12974173	19:11333359	A/T	0.12	-0.114	0.023	5.16E-07	-0.052	0.039	0.180
rs3810308	19:11333596	T/C	0.12	-0.114	0.023	5.19E-07	-0.052	0.039	0.181
rs4804155	19:11334295	C/G	0.12	-0.114	0.023	5.24E-07	-0.052	0.039	0.179
rs4804154	19:11334179	C/T	0.12	-0.114	0.023	6.61E-07	-0.049	0.040	0.222
rs4804576	19:11331354	G/T	0.06	-0.142	0.029	1.30E-06	-0.033	0.063	0.599
rs66466742	19:11336444	C/T	0.06	-0.141	0.029	1.45E-06	-0.030	0.063	0.635
rs737338	19:11347657	C/T	0.06	-0.141	0.029	1.58E-06	-0.028	0.063	0.661
rs2278426	19:11350488	C/T	0.06	-0.141	0.029	1.58E-06	-0.028	0.063	0.661
rs56322906	19:11346155	G/A	0.06	-0.141	0.029	1.59E-06	-0.027	0.063	0.662
rs8101801	19:11335477	C/A	0.07	-0.139	0.029	1.84E-06	-0.025	0.062	0.689
rs17766692	19:11342599	C/T	0.10	-0.112	0.024	3.52E-06	-0.049	0.033	0.138
rs1865063	19:11341029	C/T	0.10	-0.112	0.024	3.65E-06	-0.049	0.033	0.140
rs17699030	19:11330942	A/G	0.05	-0.158	0.034	3.72E-06	-0.057	0.051	0.272
rs4804575	19:11329641	G/A	0.10	-0.110	0.024	4.63E-06	-0.046	0.033	0.157
rs4804153	19:11331531	C/T	0.10	-0.111	0.024	4.68E-06	-0.046	0.033	0.166
rs138572354	19:11338309	C/A	0.05	-0.149	0.033	8.66E-06	-0.036	0.053	0.496
rs143466522	19:11318472	G/A	0.02	-0.213	0.056	1.27E-04	-0.103	0.062	0.097
rs79846490	19:11311885	G/C	0.04	-0.146	0.038	1.38E-04	-0.023	0.052	0.661
rs111279811	19:11298369	C/T	0.03	-0.158	0.042	1.49E-04	-0.046	0.052	0.375
rs56048141	19:11317744	C/T	0.02	-0.173	0.051	6.10E-04	-0.056	0.058	0.342
rs12979813	19:11342703	A/G	0.18	-0.062	0.019	8.72E-04	-0.003	0.024	0.883
rs10406522	19:11341635	T/C	0.18	-0.062	0.019	8.81E-04	-0.003	0.024	0.886
rs6511729	19:11346252	A/C	0.18	-0.062	0.019	9.28E-04	-0.003	0.024	0.891
rs3810307	19:11332570	T/A	0.18	-0.061	0.019	1.05E-03	-0.002	0.024	0.941
rs2043302	19:11339396	T/C	0.18	-0.061	0.019	1.16E-03	-0.001	0.024	0.972
rs10418759	19:11340242	A/G	0.18	-0.061	0.019	1.18E-03	-0.001	0.024	0.976
rs7247404	19:11268556	G/A	0.34	0.050	0.015	1.22E-03	0.037	0.016	0.018
rs17001244	19:11340057	A/G	0.18	-0.060	0.019	1.23E-03	0.000	0.024	0.989
rs11672123	19:11194823	G/A	0.09	-0.080	0.025	1.24E-03	-0.075	0.025	0.002
rs10421382	19:11344973	G/C	0.18	-0.060	0.019	1.25E-03	-0.001	0.024	0.980
rs776487142	19:11013429	C/T	0.00	-5.925	1.842	1.31E-03	-5.973	1.840	0.001
rs10409274	19:11273179	G/A	0.29	0.051	0.016	1.34E-03	0.034	0.016	0.035
rs11671937	19:11264514	C/T	0.34	0.049	0.015	1.36E-03	0.036	0.016	0.020
rs7408517	19:11264063	C/T	0.34	0.049	0.015	1.39E-03	0.036	0.016	0.020
rs10421795	19:11344406	C/T	0.18	-0.060	0.019	1.39E-03	0.000	0.024	0.991
rs11670169	19:11266015	T/C	0.34	0.049	0.015	1.41E-03	0.036	0.016	0.020
rs892115	19:11263650	G/T	0.34	0.049	0.015	1.42E-03	0.036	0.016	0.020
rs934424	19:11267613	G/A	0.34	0.049	0.015	1.49E-03	0.036	0.016	0.021
rs9749459	19:11270016	T/C	0.29	0.050	0.016	1.55E-03	0.034	0.016	0.039

rs4804148	19:11270867	C/T	0.29	0.050	0.016	1.60E-03	0.034	0.016	0.040
rs9749257	19:11269960	G/T	0.29	0.050	0.016	1.61E-03	0.033	0.016	0.040
rs17248748	19:11206040	C/T	0.02	-0.175	0.056	1.62E-03	-0.110	0.057	0.056
rs4804579	19:11358700	T/C	0.18	-0.059	0.019	1.65E-03	0.004	0.024	0.882
rs8104261	19:11272585	C/T	0.29	0.050	0.016	1.71E-03	0.033	0.016	0.042
rs9749350	19:11257299	C/G	0.18	-0.059	0.019	1.98E-03	-0.044	0.019	0.023
rs934425	19:11267503	C/T	0.29	0.050	0.016	2.02E-03	0.042	0.016	0.011
rs551841239	19:11671580	G/C	0.00	-1.435	0.465	2.05E-03	-1.371	0.465	0.003
rs150205856	19:11391990	C/T	0.02	-0.180	0.059	2.18E-03	-0.062	0.065	0.338
rs200600677	19:11369440	TC/T	0.23	-0.053	0.017	2.19E-03	-0.022	0.019	0.240
rs17001264	19:11348212	C/T	0.02	-0.181	0.059	2.25E-03	-0.062	0.065	0.340
rs6511728	19:11335597	A/G	0.15	-0.062	0.020	2.37E-03	0.026	0.029	0.361
rs8101802	19:11336182	G/C	0.15	-0.062	0.020	2.38E-03	0.026	0.029	0.360
rs139606057	19:11393073	C/T	0.00	-1.089	0.359	2.40E-03	-1.100	0.358	0.002
rs112550373	19:11071469	G/A	0.01	0.277	0.091	2.41E-03	0.299	0.091	0.001
rs147629608	19:11411726	C/T	0.00	0.862	0.284	2.41E-03	0.953	0.284	0.001
rs34098	19:11539681	A/T	0.10	0.079	0.026	2.45E-03	0.057	0.027	0.031
rs186292971	19:11435015	A/C	0.00	1.608	0.531	2.47E-03	1.598	0.530	0.003
rs187416509	19:11528292	C/T	0.00	-1.387	0.462	2.66E-03	-1.414	0.461	0.002
rs145277768	19:11440420	C/T	0.00	1.595	0.532	2.69E-03	1.586	0.531	0.003
rs541868466	19:11543169	C/T	0.00	-1.384	0.462	2.73E-03	-1.411	0.461	0.002
rs181565096	19:11232702	C/G	0.02	-0.163	0.054	2.75E-03	-0.098	0.056	0.080
rs10423399	19:11273603	T/G	0.29	0.048	0.016	2.78E-03	0.031	0.016	0.057
rs565352617	19:11992491	CAA/C	0.00	-1.488	0.498	2.79E-03	-1.445	0.497	0.004
rs570249721	19:11252163	G/A	0.00	-9.531	3.225	3.13E-03	-9.358	3.221	0.004
rs191370629	19:11569128	C/T	0.00	-1.366	0.463	3.16E-03	-1.394	0.462	0.003
rs182210127	19:11659265	G/A	0.00	-1.352	0.458	3.16E-03	-1.377	0.457	0.003
rs112541805	19:11221946	G/GA	0.03	-0.128	0.045	4.12E-03	-0.087	0.045	0.054
rs386474655	19:11026074	G/C	0.00	-0.486	0.170	4.12E-03	-0.407	0.170	0.017
rs34254024	19:11369433	TTC/T	0.24	-0.049	0.017	4.14E-03	-0.019	0.018	0.311
rs13306513	19:11218226	G/A	0.03	-0.129	0.045	4.25E-03	-0.088	0.046	0.054
rs17242899	19:11216768	T/C	0.03	-0.129	0.045	4.25E-03	-0.088	0.046	0.054
rs145446845	19:11299431	G/A	0.02	-0.164	0.057	4.26E-03	-0.046	0.063	0.467
rs140898392	19:11218361	G/A	0.03	-0.129	0.045	4.28E-03	-0.088	0.046	0.055
rs146559752	19:11264537	A/C	0.03	0.118	0.042	4.42E-03	0.109	0.041	0.008
rs528191740	19:11448303	T/G	0.00	2.902	1.024	4.62E-03	2.889	1.023	0.005
rs200345643	19:11286646	G/C	0.00	-0.619	0.219	4.79E-03	-0.499	0.221	0.024
rs72996217	19:11358966	G/A	0.16	-0.055	0.020	4.81E-03	-0.007	0.023	0.764
rs11085768	19:11370653	G/A	0.24	-0.048	0.017	4.86E-03	-0.017	0.018	0.346
rs7248924	19:11372077	T/C	0.24	-0.047	0.017	5.55E-03	-0.017	0.018	0.352
rs142159985	19:11516368	A/C	0.00	2.395	0.865	5.63E-03	2.367	0.864	0.006
rs199636757	19:11378355	CAG/C	0.24	-0.047	0.017	5.89E-03	-0.017	0.018	0.363
.	19:11798665	A/G	0.00	-3.388	1.231	5.92E-03	-3.429	1.229	0.005
rs2043303	19:11368648	C/T	0.24	-0.047	0.017	5.98E-03	-0.017	0.018	0.366
rs7258016	19:11367353	A/C	0.24	-0.047	0.017	5.98E-03	-0.016	0.018	0.367
rs12462741	19:11365281	C/T	0.24	-0.047	0.017	6.02E-03	-0.016	0.018	0.369
rs2043301	19:11365650	C/A	0.24	-0.047	0.017	6.02E-03	-0.016	0.018	0.369
rs556896609	19:11444804	A/G	0.00	1.369	0.499	6.04E-03	1.358	0.498	0.006
rs322135	19:11379717	A/G	0.24	-0.047	0.017	6.08E-03	-0.016	0.018	0.368
rs7247840	19:11267678	T/C	0.26	0.046	0.017	6.09E-03	0.039	0.017	0.019
rs10422673	19:11265408	C/G	0.26	0.046	0.017	6.09E-03	0.039	0.017	0.019
rs396460	19:11374916	C/T	0.24	-0.047	0.017	6.10E-03	-0.016	0.018	0.368
rs416231	19:11374675	C/T	0.24	-0.047	0.017	6.11E-03	-0.016	0.018	0.369

Evidence of association with the concentration of phospholipids in medium HDL in 8,380 individuals in the METSIM study. Effect represents the change in standard-normalized residuals of phospholipids in medium HDL. Conditioning on variant rs737337 attenuated the signal. MAF, minor allele frequency

[a]Non-effect allele/effect allele

[b]rs112108870 is also known as rs200788077

Table 7: Association of 100 ANGPTL8 locus variants with HDL-C in WHI

					Unconditioned			Conditioned on rs48041545		
Variant	Position	NEA	EA	MAF	Effect	Std Error	P-value	Effect	Std Error	P-value
rs4804154	19:11334179	C	T	0.19	-0.042	0.005	6.15E-17	-	-	-
rs3810308	19:11333596	T	C	0.19	-0.042	0.005	6.11E-17	0.000	0.005	0.999
rs3826815	19:11332505	C	T	0.19	-0.042	0.005	6.20E-17	0.000	0.005	0.996
rs12974173	19:11333359	A	T	0.17	-0.046	0.006	1.14E-16	-0.001	0.006	0.923
rs72994363	19:11333358	G	T	0.17	-0.046	0.006	1.14E-16	-0.001	0.006	0.923
rs17699089	19:11343795	A	G	0.20	-0.043	0.005	1.36E-16	0.000	0.005	0.998
rs12463177	19:11341680	G	C	0.20	-0.042	0.005	1.56E-16	0.000	0.005	0.985
rs3760782	19:11346550	C	T	0.19	-0.044	0.005	2.01E-16	-0.001	0.005	0.892
rs3745683	19:11348521	G	A	0.19	-0.044	0.005	5.22E-16	0.000	0.005	0.955
rs4804153	19:11331531	C	T	0.18	-0.044	0.005	5.43E-16	0.001	0.005	0.920
rs4804576	19:11331354	G	T	0.18	-0.043	0.005	6.14E-16	0.000	0.005	0.977
rs4804575	19:11329641	G	A	0.19	-0.043	0.005	7.50E-16	0.001	0.005	0.861
rs34692794	19:11343547	R	I	0.21	-0.041	0.005	1.72E-15	0.001	0.005	0.875
rs56322906	19:11346155	G	A	0.18	-0.043	0.005	2.76E-15	0.000	0.005	0.971
rs737338	19:11347657	C	T	0.18	-0.043	0.005	5.20E-15	0.001	0.005	0.917
rs17766692	19:11342599	C	T	0.20	-0.040	0.005	2.24E-14	0.002	0.005	0.736
rs2278426	19:11350488	C	T	0.19	-0.042	0.006	6.22E-14	0.002	0.006	0.780
rs66466742	19:11336444	C	T	0.17	-0.040	0.005	7.83E-14	0.002	0.005	0.640
rs4804155	19:11334295	C	G	0.31	-0.030	0.004	6.33E-13	-0.004	0.004	0.300
rs35472533	19:11324312	G	A	0.34	-0.032	0.004	1.93E-12	-0.005	0.004	0.313
rs62129150	19:11330005	G	A	0.52	0.030	0.004	2.12E-12	0.010	0.004	0.017
rs113441245	19:11328383	G	A	0.35	-0.031	0.004	2.50E-12	-0.004	0.004	0.356
rs11085764	19:11327227	G	C	0.35	-0.031	0.004	2.68E-12	-0.004	0.004	0.350
rs2304154	19:11326125	C	T	0.35	-0.031	0.004	3.22E-12	-0.004	0.004	0.357
rs2163830	19:11325417	A	G	0.35	-0.031	0.004	3.51E-12	-0.004	0.004	0.359
rs1865063	19:11341029	C	T	0.31	-0.030	0.004	4.35E-12	-0.005	0.004	0.256
rs2304155	19:11326119	G	A	0.52	0.030	0.004	4.72E-12	0.010	0.004	0.017
rs11673129	19:11325924	C	G	0.35	-0.031	0.004	5.06E-12	-0.004	0.004	0.375
rs8113156	19:11321705	T	G	0.35	-0.030	0.004	7.63E-12	-0.004	0.004	0.317
rs8101801	19:11335477	C	A	0.30	-0.029	0.004	9.08E-12	-0.005	0.004	0.278
rs11666686	19:11323085	T	C	0.35	-0.030	0.004	1.26E-11	-0.004	0.004	0.387
rs112108870	19:11345315	R	I	0.40	-0.029	0.004	1.67E-11	-0.008	0.004	0.053
rs2116873	19:11325784	A	T	0.55	0.030	0.004	4.11E-11	0.008	0.004	0.062
rs4804152	19:11327626	G	A	0.39	0.030	0.004	4.28E-11	0.013	0.004	0.004
rs59389322	19:11329394	G	A	0.34	-0.029	0.005	9.54E-11	-0.002	0.004	0.578
rs67076391	19:11328617	C	T	0.34	-0.029	0.005	1.01E-10	-0.002	0.004	0.582
rs2116875	19:11325764	A	G	0.30	-0.032	0.005	1.13E-10	-0.001	0.005	0.777
rs2116874	19:11325767	T	C	0.43	-0.030	0.005	1.23E-10	-0.011	0.005	0.017
rs200788077	19:11345320	R	I	0.37	-0.028	0.004	1.60E-10	-0.009	0.004	0.051
rs111705028	19:11320494	C	T	0.25	-0.036	0.006	3.91E-10	-0.007	0.006	0.215
rs8409	19:11319491	G	A	0.37	-0.027	0.004	5.46E-10	-0.003	0.004	0.515
rs737337	19:11347493	T	C	0.41	-0.026	0.004	9.09E-10	-0.006	0.004	0.169
rs12981155	19:11320339	G	C	0.27	-0.036	0.006	1.20E-09	-0.006	0.006	0.309
rs12979813	19:11342703	A	G	0.50	-0.022	0.004	7.01E-09	-0.006	0.004	0.102
rs7252965	19:11309160	G	C	0.53	0.024	0.004	8.49E-09	0.006	0.004	0.149
rs8101345	19:11310920	T	C	0.45	-0.023	0.004	1.10E-08	-0.005	0.004	0.195
rs4804151	19:11327608	C	T	0.36	0.026	0.005	2.44E-08	0.010	0.005	0.033
rs12609620	19:11324890	C	T	0.36	0.026	0.005	2.58E-08	0.010	0.005	0.031
rs10406522	19:11341635	T	C	0.51	-0.022	0.004	2.62E-08	-0.006	0.004	0.138
rs8110433	19:11316317	A	C	0.55	0.023	0.004	2.97E-08	0.005	0.004	0.208
rs138111115	19:11307572	R	D	0.50	0.023	0.004	4.66E-08	0.006	0.004	0.171

rs3810307	19:11332570	T	A	0.49	-0.022	0.004	8.85E-08	-0.004	0.004	0.277
rs149928810	19:11308475	R	D	0.49	0.022	0.004	2.24E-07	0.005	0.004	0.255
rs764304127	19:11314807	D	R	0.46	0.021	0.004	2.27E-07	0.006	0.004	0.166
rs11878417	19:11319978	A	G	0.52	0.021	0.004	3.19E-07	0.002	0.004	0.615
rs10421795	19:11344406	C	T	0.62	-0.022	0.004	3.30E-07	-0.008	0.004	0.075
rs113535966	19:11337269	G	A	0.06	-0.051	0.01	5.22E-07	-0.002	0.01	0.822
rs10421382	19:11344973	G	C	0.61	-0.021	0.004	5.61E-07	-0.007	0.004	0.100
rs6511728	19:11335597	A	G	0.61	-0.021	0.004	5.95E-07	-0.007	0.004	0.086
rs7252976	19:11315343	G	A	0.20	0.025	0.005	6.36E-07	0.014	0.005	0.005
rs17001244	19:11340057	A	G	0.54	-0.019	0.004	8.72E-07	-0.004	0.004	0.298
rs2043302	19:11339396	T	C	0.61	-0.021	0.004	1.09E-06	-0.006	0.004	0.162
rs35248735	19:11312238	G	A	0.42	0.019	0.004	1.27E-06	0.005	0.004	0.181
rs56865998	19:11354146	C	G	0.07	-0.047	0.01	1.49E-06	-0.002	0.01	0.872
rs59175057	19:11329534	R	D	0.44	-0.021	0.004	1.94E-06	-0.004	0.004	0.368
rs184781818	19:11317340	A	G	0.00	-1.927	0.413	2.99E-06	-0.492	0.411	0.231
rs114281937	19:11348208	C	T	0.02	-0.084	0.018	3.04E-06	-0.034	0.018	0.060
.	19:11323240	D	R	0.38	0.020	0.004	3.29E-06	0.005	0.004	0.243
rs7249565	19:11302807	G	A	0.60	0.018	0.004	3.44E-06	0.004	0.004	0.300
rs12980863	19:11309871	C	T	0.41	0.018	0.004	3.52E-06	0.005	0.004	0.214
rs3745681	19:11303943	A	G	0.59	0.018	0.004	3.63E-06	0.003	0.004	0.447
rs7246614	19:11310538	C	T	0.24	0.021	0.005	3.87E-06	0.010	0.005	0.024
rs3745682	19:11313256	G	A	0.24	0.021	0.005	4.20E-06	0.010	0.004	0.025
rs4804574	19:11317482	A	G	0.35	0.019	0.004	4.55E-06	0.007	0.004	0.118
rs7250652	19:11302606	G	A	0.40	-0.018	0.004	4.63E-06	-0.004	0.004	0.255
rs145352947	19:11306346	R	D	0.58	0.019	0.004	4.86E-06	0.003	0.004	0.478
rs6511727	19:11315817	G	T	0.23	0.021	0.005	4.97E-06	0.011	0.005	0.021
rs10421221	19:11316547	T	C	0.32	0.019	0.004	6.41E-06	0.007	0.004	0.080
rs10418759	19:11340242	A	G	0.64	-0.019	0.004	6.54E-06	-0.005	0.004	0.196
rs8101802	19:11336182	G	C	0.62	-0.019	0.004	8.49E-06	-0.006	0.004	0.187
rs8110823	19:11316315	G	A	0.23	0.021	0.005	8.52E-06	0.011	0.005	0.019
.	19:11323239	D	R	0.38	0.019	0.004	9.08E-06	0.004	0.004	0.341
rs7250778	19:11306265	G	A	0.21	0.023	0.005	1.50E-05	0.012	0.005	0.018
rs61045132	19:11303068	C	T	0.22	0.020	0.005	1.73E-05	0.013	0.005	0.005
rs58543390	19:11342434	C	T	0.00	-1.187	0.279	2.07E-05	-0.328	0.278	0.237
rs73506605	19:11307564	G	A	0.44	0.018	0.004	2.10E-05	0.003	0.004	0.442
rs34757881	19:11341462	C	T	0.07	-0.039	0.009	2.18E-05	0.012	0.009	0.190
rs200384092	19:11323225	R	D	0.41	-0.023	0.005	2.93E-05	-0.006	0.005	0.258
rs79846490	19:11311885	G	C	0.01	-0.115	0.028	3.71E-05	-0.030	0.028	0.280
rs6511729	19:11346252	A	C	0.65	-0.018	0.004	3.72E-05	-0.004	0.004	0.336
rs57681847	19:11300648	G	T	0.21	0.019	0.005	4.02E-05	0.012	0.005	0.009
rs58495388	19:11300312	G	C	0.18	0.021	0.005	6.76E-05	0.014	0.005	0.008
rs139048611	19:11321312	R	D	0.02	-0.067	0.017	7.34E-05	-0.023	0.017	0.167
rs8111456	19:11301147	A	G	0.22	0.018	0.005	7.44E-05	0.013	0.005	0.007
rs147045092	19:11300357	G	C	0.00	14.200	3.656	0.000103	10.115	3.641	0.005
rs34301174	19:11348098	G	A	0.11	0.034	0.009	0.000115	0.017	0.009	0.054
rs2278013	19:11305429	C	A	0.15	0.024	0.006	0.000168	0.014	0.006	0.032
rs148312284	19:11358858	G	C	0.02	-0.071	0.019	0.000191	-0.025	0.019	0.182
rs199653227	19:11317508	D	R	0.25	0.017	0.005	0.000234	0.007	0.005	0.134
rs73506665	19:11358644	C	T	0.05	-0.036	0.01	0.000263	0.000	0.01	0.970

Evidence of association with HDL-C in 8,244 individuals in the WHI study. Effect represents the change in standard-normalized residuals of phospholipids in medium HDL. Conditioning on lead variant rs4804154 attenuated the signal. NEA, non-effect allele; EA, effect allele; MAF, minor allele frequency

Table 8: Variant associations with gene expression levels in subcutaneous adipose tissue

	eQTL p-value					
Gene	**rs4804155**	**rs17699089**	**rs4804154**	**rs737337**	**rs12463177**	**Probeset ID**
ANGPTL8	1.04E-09	1.84E-09	1.38E-09	9.74E-08	3.84E-09	11756040_a_at
DOCK6	2.37E-06	7.21E-07	1.65E-06	2.68E-03	3.06E-06	11719400_a_at
CCDC159	0.071	0.066	0.085	0.092	0.054	11744239_a_at
KANK2	0.056	0.073	0.048	0.565	0.116	11726538_x_at
KRI1	0.052	0.062	0.036	0.348	0.053	11725949_x_at
LDLR	0.303	0.300	0.258	0.616	0.349	11720028_x_at
LPPR2	0.059	0.056	0.054	0.215	0.070	11722506_a_at
S1PR5	0.152	0.202	0.157	0.014	0.247	11752664_a_at
SLC44A2	0.028	0.029	0.028	0.178	0.036	11740973_s_at
TSPAN16	0.050	0.032	0.043	0.288	0.034	11761297_x_at

Gene expression was measured in 770 subcutaneous adipose samples. eQTL data are reported for genes within 1 Mb of rs737337 that have at least one variant with $p<0.05$. Lead eQTL variants for *ANGPTL8* (rs4804155) and *DOCK6* (rs17699089), lead GWAS variants (rs4804154 and rs737337) and functional candidate variant rs12463177 are shown. Data are shown for the most strongly associated probeset.

Table 9: Variants associated with HDL-C in METSIM and/or WHI

SNP[a]	chr19 position	r^2 with rs737337 (EUR)[c]	r^2 with rs4804154 (AFR)[c]	Dnased	FAIRE[d]	H3K4me1[d]	H3K27ac[d]	H3K4me3[d]	H3K9ac[d]	H3K4me2[d]	Transcription Factor Binding (ChIP-seq)[d]	Posterior Probability MANTRA	Posterior Probability CAVIAR (Finnish)[e]	Predicted in at least one PAINTOR analysis[f]	Antibodies tested in EMSA
rs79846490	11311884	0.58				L									
rs8409	11319491		0.66	NM											
rs8113156	11321705		0.72								B: ZEB1			Yes	
rs11666686	11323085		0.78												
rs35472533	11324312		0.81												
rs2163830	11325417		0.8												
rs2116875	11325764		0.68												
rs11673129	11325924		0.8												
rs2304154	11326125		0.82												
rs11085764	11327227		0.82											Yes	
rs113441245	11328383		0.78												
rs67076391	11328617		0.78												
rs59389322	11329394		0.79												
rs4804575	11329641	0.34	0.98									0.240	0.021		
rs17699030	11330942	0.45											0.029	Yes	
rs4804576	11331354	0.46	0.95									0.450	0.037		
rs4804153	11331531	0.3	0.95									0.240			
rs3826815	11332505	0.67	0.99					A					0.050	Yes	
rs12971537	11333358							A							
rs12974173	11333359	0.61	0.81					A					0.049		
rs3810308	11333596	0.67	0.98					A				0.290	0.049	Yes	
rs4804154	11334179	0.67	1								L: USF1		0.037		
rs4804155	11334295	0.67	0.48										0.050	Yes	
rs8101801	11335477	0.46	0.47										0.024		
rs6511728	11335597	0.48													
rs8101802	11336182	0.53					A								
rs66466742	11336444	0.46	0.92				A						0.028		
rs138572354	11338309	0.72		ONEBL									0.023		
rs12463177	11341680	0.74	0.93			AL	AL			L		0.320	0.065	Yes	
rs17766692	11342599	0.34	0.84			AL	L			L	B: OCT2, POU2f2		0.028		
rs34692794	11343547	0.74	0.86			AL									
rs17699089	11343795	0.74	0.91			AL	A					0.290	0.061	Yes	
rs10421795	11344406	0.47				AL									
rs10421382	11344973	0.47				L	A								
rs200788077[b]	11345321	0.96	0.22	L		L	A				L: JUND, cJUN		0.066		SMAD4, HNF4G, RXRA
rs56322906	11346155	0.54	0.86			AL	AL	L	L	L		0.420	0.026		
rs6511729	11346252	0.48		M		AL	AL	L	L	L					
rs3760782	11346550	0.98	0.87			AL	AL	L	L	L	L:HDAC2	0.430	0.051	Yes	SP1, RXRA, HNF4A, HNF4G, SREBP1, SREBP2
rs737337	11347493	1	0.26	L		AL	AL	L	L	L	L: USF1, TAF1, RXRA, ELF1		0.058		RXRA, AP2A, HIF1A, CEBPB, LXRA, CHREBP, THR, SP1, USF1, NF1, PAX4, FOXA1, FOXA2, FOXO3, PPARG, CEBPB, YY1, PAX6, PARP1, SMAD4, HEY1, FOSL1, ELF1, TAF1, HNF4G, SREBP1, SREBP2, CEBPA, AHR, ARNT, PPARA, EGR1, CREB, HNF4A, PXR, LRH1, CAR
rs737338	11347657	0.54	0.86			AL	AL	L	L	L	L: CEBPb, TAF1, ELF1, HNF4g	0.420	0.026		
rs3745683	11348521	0.98	0.82	N		AL	AL	L	L	L		0.420	0.051		YY1, SREBP1, RXRA, PPARG
rs2278426	11350488	0.54	0.78	L		AL	AL	L	L	L	L: POL2		0.026		

[a]Variants are ordered by position, hg19. Nearby genes: *ANGPTL8* is located at position 11350295-11352619, *DOCK6* position 11309969-11373168. The thirteen candidate variants evaluated in functional assays are bolded.

[b]rs200788077 is also known as rs112108870

[c]r^2-Haploreg v4.1 (1000 Genomes Phase I)

[d]DNase, FAIRE, histone marks, and ChIP-seq data are from ENCODE and Roadmap Epigenomics

[e]Three additional variants are identified in the Finnish CAVIAR analysis and two in the African American CAVIAR analysis that are not in LD with either lead.

[f]Nine additional variants are identified in at least one PAINTOR analysis that are not in LD with either lead.

O=Bone: ENCODE Osteoblast

B= Blood: ENCODE GM19240, GM19239, GM19238, GM190999, GM18951, GM18526, GM18507, GM15510, GM12892, GM12891, GM12878, GM12875, GM12874, GM12873, GM12872, GM12865, GM12864, GM12802, GM12193, GM06990

L=Liver: Roadmap Epigenomics Adult Liver; ENCODE Hepatocytes, HepG2, Huh-7

E=Endothelial: ENCODE HUVEC

M=Muscle: Roadmap Epigenomics Skeletal Muscle; ENCODE Myocyte, PSOAS muscle

N=Brain: Roadmap Epigenomics Brain Anterior Caudate, Brain Mid Frontal Lobe, Brain Substantia Nigra; ENCODE Astrocytes, Cerebellum, Cerebral Frontal, Frontal Cortex

A=Adipose: Roadmap Epigenomics Adipose Nuclei, Adipose Tissue

Table 10: Fine-mapping analysis using MANTRA

Variant	Position	EA	NEA	logBF	PP	Direction	Bayes Factor	Cum(BF)	credible set
rs3760782	11,346,550	T	C	18.9	0.43	--	7.57E+18	7.57E+18	0.331
rs3745683	11,348,521	A	G	18.8	0.42	--	6.04E+18	1.36E+19	0.595
rs4804153	11,331,531	T	C	18.6	0.24	--	3.72E+18	1.73E+19	0.757
rs17699089	11,343,795	G	A	18.4	0.29	--	2.28E+18	1.96E+19	0.856
rs4804576	11,331,354	T	G	17.9	0.45	--	7.56E+17	2.04E+19	0.889
rs56322906	11,346,155	A	G	17.8	0.42	--	6.60E+17	2.10E+19	0.918
rs4804575	11,329,641	A	G	17.8	0.24	--	6.48E+17	2.17E+19	0.947
rs737338	11,347,657	T	C	17.8	0.42	--	5.95E+17	2.23E+19	0.973
rs3810308	11,333,596	C	T	17.5	0.29	--	3.14E+17	2.26E+19	0.986
rs12463177	11,341,680	C	G	17.3	0.32	--	2.00E+17	2.28E+19	0.995

MANTRA analysis was conducted in 16624 individuals from METSIM and WHI. Credible set values are the cum(BF) divided by the total cumulative Bayes Factor. EA, effect allele; NEA, non-effect allele; logBF, log(Bayes Factor); PP, posterior probability

Table 11: Fine-mapping analysis using CAVIAR

Finnish		African American	
Posterior probability	Variant	Posterior probability	Variant
0.066	rs200788077	0.499	rs2116874
0.065	rs12463177	0.499	rs2304155
0.061	rs17699089		
0.058	rs737337		
0.051	rs3745683		
0.051	rs3760782		
0.050	rs3826815		
0.050	rs4804155		
0.049	rs72994363		
0.049	rs12974173		
0.049	rs3810308		
0.037	rs4804154		
0.031	rs4804576		
0.029	rs17699030		
0.028	rs66466742		
0.028	rs143466522		
0.028	rs17766692		
0.027	rs1865063		
0.026	rs737338		
0.026	rs2278426		
0.026	rs56322906		
0.024	rs8101801		
0.023	rs138572354		
0.021	rs4804575		

Variants shown are in the 95% causal set.

Table 12: Fine-mapping analysis using PAINTOR

METSIM Finnish

2 causal variants			3 causal variants			4 causal variants			5 causal variants		
Position	**Variant**	**Probability**	**Position**	**Variant**	**Probability**	**Position**	**Variant**	**Probability**	**Position**	**Variant**	**Probability**
11327571	rs4804150	0.42	11327571	rs4804150	1.00	11327571	rs4804150	1.00	11327571	rs4804150	1.00
11341680	rs12463177	0.29	11313256	rs3745682	1.00	11313256	rs3745682	0.96	11317770	rs56034303	1.00
11343795	rs17699089	0.29	11343795	rs17699089	0.66	11318375	rs116504889	0.86	11313256	rs3745682	0.97
11330942	rs17699030	0.25	11332505	rs3826815	0.34	11343795	rs17699089	0.65	11318375	rs116504889	0.88
11332505	rs3826815	0.23				11332505	rs3826815	0.30	11343795	rs17699089	0.62
11346550	rs3760782	0.20							11332505	rs3826815	0.26
									11318235	rs114277401	0.11

WHI African American

2 causal variants			3 causal variants			4 causal variants			5 causal variants		
Position	**Variant**	**Probability**	**Position**	**Variant**	**Probability**	**Position**	**Variant**	**Probability**	**Position**	**Variant**	**Probability**
11330005	rs62129150	1.00	11334295	rs4804155	1.00	11350086	rs115758240	1.00	11340498	rs73506650	1.00
11321705	rs8113156	1.00	11330005	rs62129150	1.00	11334295	rs4804155	1.00	11340057	rs17001244	1.00
			11327571	rs4804150	1.00	11330005	rs62129150	1.00	11334295	rs4804155	1.00
						11327571	rs4804150	1.00	11330005	rs62129150	1.00
									11327571	rs4804150	1.00

METSIM Finnish and WHI African American

2 causal variants			3 causal variants			4 causal variants			5 causal variants		
Position	**Variant**	**Probability**	**Position**	**Variant**	**Probability**	**Position**	**Variant**	**Probability**	**Position**	**Variant**	**Probability**
11330005	rs62129150	1.00	N/A	N/A	N/A	11334295	rs4804155	1.00	N/A	N/A	N/A
11321705	rs8113156	1.00				11333596	rs3810308	1.00			
						11330005	rs62129150	1.00			
						11327227	rs11085764	1.00			

PAINTOR analysis was performed in METSIM, WHI, and METSIM/WHI combined. In each dataset, 2, 3, 4, or 5 causal variants were assumed. Posterior probabilities are presented. Positions are on chromosome 19 and hg19. N/A, no variants are predicted.

CHAPTER 4: OPEN CHROMATIN PROFILING IN ADIPOSE MARKS GENOMIC REGIONS WITH FUNCTIONAL ROLES IN CARDIOMETABOLIC TRAITS

Introduction

Cardiometabolic diseases, including cardiovascular disease, type 2 diabetes (T2D), and obesity, are a significant health burden in many populations (9-11). Each of these diseases and related traits is regulated by genetic and environmental risk factors (25). Genome-wide association studies (GWAS) have identified hundreds of loci associated with cardiometabolic diseases and risk factors (www.ebi.ac.uk/gwas/). Many associated variants are located in noncoding regions, suggesting a regulatory mechanism of action. The causal variant(s) and target gene(s) at noncoding GWAS loci have yet to be fully delineated. A mechanistic understanding of loci is necessary to inform the development of therapies for the appropriate genes (1) and direction of effect (2), especially given the increasing recognition of allelic heterogeneity at GWAS loci (3-8).

Adipose tissue, especially in subcutaneous depots, is involved in cardiometabolic traits and diseases. Subcutaneous adipose tissue serves as a buffering system for lipid energy balance, particularly fatty acids (159-161), and may play a protective role in cardiometabolic risk (162). Accumulation of fat, particularly in the central abdomen (163), and specifically in the subcutaneous depot (164), confers an elevated risk of cardiometabolic diseases and mortality. In addition, subcutaneous adipose expression quantitative trait loci (eQTL) studies have identified genes involved in central obesity and metabolic traits (63,165-168), and specific GWAS loci for

type 2 diabetes, lipid levels, measures of obesity, and adiponectin colocalized with subcutaneous adipose eQTLs (4-7). In addition, a recent GWAS meta-analysis for waist-hip ratio, a measure of obesity, identified loci that were enriched both for putative regulatory elements in adipocyte nuclei and for genes expressed in subcutaneous adipose tissue (6), many of which are now linked to adipose function (169).

Adipose tissue is highly complex and composed of many cell types, including adipocytes, preadipocytes, nerve cells, immune cells, and vascular cells (170). Characterization of heterogeneous whole adipose tissue and its component cell types are both needed to fully delineate the role of adipose tissue in cardiometabolic disease. A benefit of using human adipose tissue samples is to identify differences in chromatin accessibility due to genotype and link them to cardiometabolic traits; however, inherent differences exist between samples due to site of tissue extraction, sample handling, tissue storage conditions, genotype, and/or environmental contributions such as obesity or T2D. Although cell models do not fully replicate cells within a complex tissue, they provide a means to characterize individual cell types and have consistent growth, storage, and environmental conditions. The Simpson Golabi-Behmel Syndrome (SGBS) human cell strain is an ideal model for studying adipocyte and preadipocyte biology because the cells are diploid, easy to grow in culture, can be differentiated to mature adipocytes (171). Additionally, using SGBS cells rather than primary human adipocytes from multiple individuals decreases experimental variation due to genotype or sample collection differences.

Variability in chromatin accessibility is heritable and mediates the effects of gene expression (172-174). Adipose and adipocytes are poorly represented in open chromatin datasets because the high lipid content makes experimental assays challenging. For human adipose and adipocytes, three DNase-seq datasets exist: one from *in vitro* differentiated adipocytes

(ENCODE), one SGBS adipocytes (175), and a third from adipose-derived differentiated stem cells (176). Accessible chromatin can be assayed in small amounts of tissue or cells using the assay for transposase-accessible chromatin followed by sequencing (ATAC-seq) (62). To date, only three ATAC-seq datasets have been generated from human adipose tissue or cells: one from adipocyte nuclei from subcutaneous adipose tissue (61,177), one from subcutaneous adipose tissue (ENCODE), and one from omental fat tissue (ENCODE). In addition to chromatin accessibility, regulatory histone marks have been characterized in adipose nuclei from subcutaneous adipose tissue and differentiated adipocytes from mesenchymal stem cells in the Roadmap Epigenomics Project. The collection of histone marks were integrated into chromatin states (178), which predict regulatory function (i.e. promoter or enhancer). Previous studies have found that regions of chromatin accessibility occur preferentially in promoters and enhancers compared to transcribed and repressed regions of the corresponding cell type (178,179). Adipose chromatin states based on histone marks will be strengthened by the addition of open chromatin datasets.

Robust chromatin accessibility data from adipose tissue and adipocytes can identify candidate variants at GWAS loci. Allelic differences have been found in accessible chromatin and histone marks of chromatin state (172,180-184), and these differences have provided a functional context for interpreting GWAS loci for blood and autoimmune diseases (8,173). Chromatin accessibility data can also identify sites of transcription factor binding (61). Transcription factor motifs predict where a transcription factor may bind; however, only a small fraction of predicted motifs show factor binding (185). Transcription factor footprints can potentially identify the factor(s) bound at a given site. Cell models such as SGBS may provide a more pure population of cells to prioritize candidate variants and identify candidate transcription

factors because chromatin accessibility in these cells can identify preadipocyte-specific regulatory elements and may help characterize cell-type specificity within heterogeneous adipose tissue. Taken together, chromatin accessibility data in adipose tissue and adipocytes will improve annotation of candidate regulatory variants and candidate transcription factors at GWAS loci.

In this study, we use ATAC-seq data from frozen clinical subcutaneous adipose tissue needle biopsy samples and SGBS preadipocytes to identify accessible chromatin. We compare adipose tissue and preadipocyte open chromatin profiles and identify regulatory regions specific to each. We identify cardiometabolic GWAS and transcription factor enrichment within ATAC peaks. Finally, we use the ATAC-seq annotations to prioritize candidate variants at cardiometabolic GWAS loci with colocalized eQTL associations. Based on ATAC-seq overlap, we test candidate variants at two cardiometabolic GWAS loci and identify functional variants.

Materials and Methods

METSIM study participants

Subcutaneous adipose tissue needle biopsies were obtained from METabolic Syndrome in Men (METSIM) participants as previously described (63). We use three adipose tissue needle biopsy samples for ATAC-seq (Table 14). The METSIM study includes 10,197 men, aged from 45 to 73 years, randomly selected from Kuopio, Eastern Finland, and examined in 2005 – 2010 (140,141). The Ethics Committee of the University of Eastern Finland in Kuopio and the Kuopio University Hospital approved the METSIM study and it was carried out in accordance with the Helsinki Declaration. DNA samples were genotyped on the Illumina OmniExpress and HumanCoreExome arrays and imputed using the Haplotype Reference Consortium as previously described (151).

Sample processing and ATAC-seq library preparation

Adipose tissue was flash frozen and stored at -80º until use. SGBS cells (154) were generously provided by Dr. Martin Wabitsch (University of Ulm) and cultured as previously described (186). For adipose tissue samples 1 and 3, we generated libraries using nuclei isolation buffers that contained detergent (1% NP-40) or did not contain detergent. For tissue sample 2, we generated libraries using ~12 mg and ~36 mg of tissue. For SGBS, we have generated profiles using as few as 50,000 cells (data not shown); however, in this study we generated profiles with 800,000 cells (SGBS 1) and 400,000 cells (SGBS 2), both with detergent (Table 15). Tissue was pulverized in liquid nitrogen using a Cell Crusher homogenizer (cellcrusher.com). For SGBS cells, cells were trypsinized and pelleted. The tissue powder or cell pellet was resuspended in nuclei isolation buffer (20 mM Tris-HCl, 50 mM EDTA, 60 mM KCl, 40% glycerol, 5 mM spermidine, 0.15 mM spermine, 0.1% mercaptoethanol, 1% NP-40). Tubes were rotated at 4º for 5 minutes. The solution was completely homogenized using a tight homogenizer (Wheaton) for 10 strokes. Following homogenization, the solution was centrifuged at 1500 x g for 10 minutes at 4º. Following removal of the lipid layer from the adipose tissue and supernatant, the pellet was resuspended in resuspension buffer (10 mM Tris-HCl, 10 mM NaCl, 3 mM $MgCl_2$) and centrifuged at 1200 x g for 10 minutes at 4º. The supernatant was removed and the pellet was used for the transposase reaction as previously described (62). For SGBS libraries, we used 5 ul Tn5 transposase. Following library PCR amplification, we removed primer dimers using Ampure Beads (Agencourt) with a 1:1.2 ratio of library to beads. Libraries were visualized and quantified using a TapeStation or Bioanalyzer and sequenced at the Duke University Genome Sequencing Center (adipose tissue single-end sequencing) or the University of North Carolina High-Throughput Sequencing Facility (SGBS paired-end sequencing).

ATAC-seq alignment and peak calling

We obtained previously published adipose ATAC-seq datasets from subcutaneous adipose tissue (ENCODE ENCSR540BML) and mature adipocytes (177). The mature adipocyte ATAC-seq data was shared by the McGill Epigenomics Mapping Centre and is available from the European Genome-phenome Archive of the European Bioinformatics Institute (dataset EGAD00001001300). To perform alignments consistently across samples, we merged the mate pair fastq files and trimmed reads to 50 nucleotides for each paired-end SGBS, mature adipocyte, and ENCODE ATAC-seq samples. We removed sequencing adapters from raw ATAC-seq sequence reads using Tagdust (187) with a false discovery rate of 0.001 and selected high quality reads with a Phred score of at least 20 for at least 90% of bases using the FASTX toolkit (http://hannonlab.cshl.edu/fastx_toolkit). We aligned filtered reads to the hg19 human genome using bowtie2 (188), penalizing ambiguous bases as mismatches. We removed any alignments with mapping quality less than 20, mitochondrial reads, or blacklisted regions (189,190) and shifted the resulting alignments by +4 on the + strand and -5 on the – strand so that the 5' base of each alignment corresponded to the center of the binding site of the Tn5 transposase (62,191). We verified sample identity using verifyBamID (192) using genotyped variants with at least 10 ATAC-seq reads in the sample with the lowest read depth (Tissue 2; 8,683 variants), minimum minor allele frequency of 0.01, and call rate of at least 0.5; we used the best-matched genotypes for each sample. We called peaks using MACS2 (193) with no background dataset, smoothing ATAC signal over a 200 bp window centered on the Tn5 binding site, allowing no duplicates, and a 0.05 false discovery rate.

Comparison of adipose ATAC profiles

We compared ATAC peak genomic positions between our ATAC-seq samples and subcutaneous adipose ATAC-seq profiles from ENCODE (ENCSR540BML), and mature adipocytes (177) using two metrics in BEDTools: the Jaccard index, which evaluates similarity between a pair of samples as the intersection divided by the union of sample set contents, and the percent of peak bases in one sample covered by another sample. When comparing peaks between library preparations, we used the top 10,000 peaks ranked by peak *p* value. We used two methods to compare peaks between our adipose tissue, SGBS, ENCODE, and mature adipocyte.samples. We compared the top 25,000 peaks ranked by peak p value in each individual sample and the top 50,000 peaks between the union of our tissue samples, the union of the SGBS samples, the union of the mature adipocyte samples, and the ENCODE sample. Results were similar between the individual and union comparisons. The p values of overlapping peaks were averaged when merging peaks between samples. We generated Venn diagrams showing the overlap of the union samples using the Vennerable R package (194).

Chromatin state analyses

We obtained chromatin states for the expanded 18-state model consisting of data for 98 cell and tissue types and 6 histone marks (H3K4me1, H3K4me3, H3K36me3, H3K27me3, H3K9me3, and H3K27ac) from the Roadmap Epigenomics Consortium (178). We generated the following combined states by merging states of similar genomic context: promoter (1_TssA, 2_TssFlnk, 3_TssFlnkU, 4_TssFlnkD, 14_TssBiv), transcribed (5_Tx, 6_TxWk), enhancer (7_EnhG1, 8_EnhG2, 9_EnhA1, 10_EnhA2, 11_EnhWk, 15_EnhBiv), and polycomb repressed (16_ReprPC, 17_ReprPCWk). Using BEDTools (46), we calculated the number of ATAC peak

bases overlapped by each chromatin state in each individual ATAC sample, the union of the adipose tissue samples, and the union of the SGBS samples. To control for differences in peak number and width, we performed overlap using two sets of normalized peaks in the union samples: the top 50 thousand peaks ranked by peak p value and the central 200 bases of the top 50 thousand peaks. We generated stacked bar plots of the percent of ATAC peak bases covered by each chromatin state in each cell type using R (195) and ranked the ATAC peak coverage of each chromatin state in adipose nuclei (Roadmap epigenome ID = E063) relative to all other cell types, where a rank of 1 corresponds to higher coverage than all other cell types.

Comparison of adipose tissue and SGBS specific ATAC peaks

We classified adipose tissue- or SGBS-specific peaks as those found in at least one tissue sample and no SGBS sample, or vice versa. To correct for the greater number of peaks in SGBS, we compared the top 10,000 (of 19,757) tissue-specific peaks to the top 10,000 (of 162,414) SGBS-specific peaks ranked by peak p value. We identified TF motifs enriched in peaks specific to tissue or SGBS relative to peaks shared between tissue and SGBS using 319 vertebrate motifs in Homer (188). We classified motifs with an enrichment p value less than $2*10^{-4}$ (0.05/319) as enriched. For shared peaks, we required that at least 50% of a tissue peak was overlapped by an SGBS peak and vice versa (n=45,120). We performed gene ontology enrichment using GREAT with default parameters (196) of METSIM and SGBS-specific peaks using the Gene Ontology Biological Process ontology (197). We identified transcription start sites (TSSs) of 78,589 unique TSSs at 43,527 genes TSSs (GENCODE 24lift37 Basic Set) specific to tissue as those overlapping one of the top 10,000 tissue-specific peaks and at least 1 kb or 5 kb from any SGBS

peak. We defined SGBS-specific TSSs as those overlapping one of the top 10,000 SGBS-specific peaks and at least 1 kb or 5 kb from any tissue peak.

Transcription factor motif scanning and footprinting

We scanned the hg19 human genome (SGBS) or haplotypes of personalized reference genomes constructed from individual genotypes (adipose tissue samples) for 519 transcription factor binding motifs from the JASPAR core 2016 vertebrates database using FIMO (198,199). If two motifs for the same factor existed at the exact same genomic coordinates and on the same strand on each haplotype, we used the motif with the highest motif score.

We performed transcription factor footprinting for 35 transcription factor motifs corresponding to 34 unique adipose-related transcription factors (Table 16). The 34 transcription factors included 21 described as adipose core transcription factors (200), six dimer motifs that contained a core transcription factor, CEBPA, CEBPB, CEBPD, ZEB1, SPI1, SPIB, and CTCF. For the resulting motifs, we generated windows containing the genomic coordinates of the motif and 100 bp flanking both motif edges. We removed motif windows where less than 90% of bases uniquely mapped and windows that overlapped blacklisted regions (47). We constructed matrices of the number of Tn5 transpositions across the remaining motif windows and predicted which motifs were likely bound using CENTIPEDE (185). We used motif scores calculated by FIMO for CENTIPEDE priors and classified a motif with a CENTIPEDE posterior binding probability greater than 0.99 as bound and less than 0.5 as unbound.

To determine which transcription factors exhibited an aggregate footprint profile, we calculated the average transposition probability at each window position separately for bound and the top 10,000 unbound sites to obtain aggregate bound and unbound profiles. We then

calculated the transposition probability ratio (TPR) by dividing each position in the bound profiles by the corresponding position in the unbound profiles and calculated the average TPR across the motifs (mTPR) and the 100 bp flanking regions (fTPR). We considered transcription factor motifs to display an aggregate footprint profile if mTPR was less than fTPR.

Enrichment of GWAS variants in ATAC peaks and Roadmap chromatin states

We tested for enrichment of cardiometabolic GWAS loci in ATAC peaks and Roadmap chromatin states using GREGOR with an LD r^2 threshold of 0.8 and a window size of 1 Mb (158). We used loci from 12 trait categories from the GWAS catalog (December 2016): type 2 diabetes, insulin, glucose, cardiovascular outcomes, blood pressure traits, low-density lipoprotein cholesterol (LDL-C), high-density lipoprotein cholesterol (HDL-C), triglycerides, total cholesterol, body mass index (BMI), waist-hip ratio adjusted for BMI (WHR), and adiponectin. Loci that were associated with multiple traits were assigned to each trait. We classified loci as enriched in a given region if the enrichment *p*-value was less than the Bonferroni-corrected threshold of 5×10^{-4} (0.05/(8 regions*12 traits)). To compare enrichment magnitudes between regions and traits, we calculated an enrichment z-score:

$$\text{z-score} = \frac{\text{observed overlaps-expected overlaps}}{\text{standard deviation}}$$

The expected overlaps and standard deviation were estimated from matched control loci using GREGOR. We visualized the enrichment results using the heatmap.2 function in the gplots R package (195,201). We computed the percent of loci and proxies for each cardiometabolic trait that overlapped adipose tissue and SGBS ATAC peaks using the LD proxies calculated by GREGOR (1000 Genomes Phase I) (57). Of the loci and proxies overlapping ATAC peaks, we calculated the fraction that overlapped transcription factor motifs from JASPAR (54).

Overlap of GWAS-eQTL colocalized loci with ATAC peaks

eQTL mapping in 770 subcutaneous adipose samples and determination of GWAS-coincident eQTLs was described previously (63,186). We identified overlap of ATAC peaks with any variant in LD (r^2>0.8) with the eQTL lead variant at 110 loci using BEDTools (46). LD was calculated using the 770 METSIM individuals included in the eQTL analysis. The 110 GWAS-eQTL colocalized loci contain 6,746 total LD proxies.

Transcriptional reporter luciferase assays

SGBS and SW872 liposarcoma cells were maintained and transcriptional reporter luciferase assays were performed as previously described (186). Table 17 contains primers used for amplifying ATAC peaks overlapping the variant of interest. Fragments containing potential enhancers are designated as 'forward' or 'reverse' based on their orientation with respect to the genome. Regions were designed to include the entire ATAC peak overlapping the variant of interest. Three to five independent clones were cotransfected with *Renilla* luciferase vector in triplicate (SGBS) or duplicate (SW872) wells using Lipofectamine 3000 (SGBS, Life Technologies) or FUGENE 6 (SW872, Promega). Firefly luciferase activity of the clones containing the PCR fragments was normalized to *Renilla* luciferase readings to control for differences in transfection efficiency. We repeated all luciferase transcriptional reporter experiments on independent days and obtained consistent results. Data are reported as fold change in activity relative to an empty pGL4.23 vector. We used two-sided Student's t-tests to compare luciferase activity.

Electrophoretic mobility shift assays (EMSA)

For EMSA, we prepared nuclear cell extracts from SGBS and SW872 cells using the NE-PER nuclear and cytoplasmic extraction kit (Thermo Scientific) as previously described (155). Double-stranded oligos (Table 17) were incubated with SGBS or SW872 nuclear extract or 100 ng purified PU.1 protein (Creative Biomart SPI1-172H) and DNA-protein complex visualization as previously described (155). A positive control oligo contained the PU.1 motif from JASPAR and a negative control did not contain the motif (Table 17). We repeated all EMSA experiments on independent days and obtained consistent results.

Personal genome construction and ATAC-seq allelic imbalance

We constructed personalized reference genomes with the –create_reference option in the AA-ALIGNER pipeline (202). We aligned reads to personalized genomes using the allele-aware aligner GSNAP allowing two mismatches, no INDELs, and treating ambiguous bases (encoded as Ns) as mismatches (203).We extracted unique alignments and filtered alignments for mitochondrial alignments and blacklisted regions (190,204). Using WASP (205), we removed alignments that did not uniquely map to each allele at heterozygous sites. Allele count pileup files were generated at heterozygous sites with a minimum base quality Phred score of 30 to minimize the impact of sequencing errors. We removed heterozygous loci with aligned bases other than the two genotyped alleles and selected heterozygous sites with at least 10 total counts and at least 1 count per allele. To account for residual biases, we fit allele counts to a beta-binomial distribution with the probability of success (reference allele ratio) and dispersion estimated using maximum likelihood separately for each sample using the VGAM R package (195,206). We performed beta-binomial tests of allelic imbalance using VGAM and multiplied

the resulting one-tailed *p*-values by 2 to obtain estimated two-tailed imbalance *p*-values. We tested the significance of overlap between variants exhibiting nominal imbalance (beta-binomial $p<0.05$) with subcutaneous adipose GWAS-coincident eQTLs using the chi square test using R (53).

Results

Chromatin accessibility in frozen adipose tissue and SGBS preadipocytes

We generated ATAC-seq open chromatin profiles from three frozen subcutaneous adipose tissue needle biopsy samples (Table 14) and two preparations of SGBS preadipocytes. We generated ~100-160 million raw ATAC-seq reads for each ATAC profile and ~26-70 million high quality aligned reads (Table 13, Methods). Using MACS2 (193) and an FDR of 5%, we identified ~29,000-58,000 peaks in the tissue samples and ~180,000 peaks in the SGBS preadipocytes (Table 13). We evaluated the use of detergent in library preparation and found that including detergent resulted in a greater number of peaks and higher peak similarity compared to no detergent (Table 14, 18). The three detergent-treated tissue profiles were similar with 64-80% of bases overlapping a peak in another sample and a mean pairwise Jaccard index of 0.58 (Table 19). The two SGBS replicates showed strong similarity with 82-88% of bases in one replicate overlapped the other replicate and a Jaccard index of 0.73 (Table 19). Differences between adipose tissue samples may be due to individual variation in tissue collection and storage, genotype, or cellular environment. These data demonstrate that ATAC-seq open chromatin profiles can be obtained from small amounts (12-36 mg, one-third to two-thirds of a needle biopsy) of frozen clinical subcutaneous adipose tissue samples and from SGBS preadipocytes.

Table 13: ATAC-seq alignment metrics of human adipose tissue and SGBS preadipocytes

Sample	Total reads	Aligned reads	Percent Mitochondrial Reads	Nuclear alignments	Remaining reads after blacklist filtering	Remaining reads after duplicates removed	Number of peaks
Tissue 1	129.5	87.4	8.5	80.0	79.0	70.6	58,279
Tissue 2	131.5	83.6	12.8	72.9	71.8	60.6	36,612
Tissue 3	119.3	70.5	11.9	62.2	61.3	57.1	49,631
SGBS 1[a]	194.9	168.4	14.2	144.4	126.2	40.5	196,211
SGBS 2[a]	105.5	93.3	42.7	53.4	42.6	25.6	173,084

Reads are reported in millions of reads

[a]These samples were sequenced using paired-end reads, but processed as single-end reads

Comparison of adipose tissue, adipocyte, and preadipocyte ATAC profiles

To determine the differences in genomic distribution and regulatory overlap of our adipose tissue, SGBS preadipocyte, and previously published adipose ATAC-seq datasets (ENCODE, (177)), we compared the ATAC profiles using three methods: overlap of peak base positions, the distribution of ATAC peaks across Roadmap adipose nuclei regulatory regions, and the distribution of ATAC peaks across all Roadmap cell types. The ATAC peak bases were most similar between adipose tissue and ENCODE or mature adipocytes (mean 52% overlapped bases; Jaccard index = 0.33, 0.34 respectively; Figure 26A, Table 19). SGBS preadipocytes showed less similarity with mean 42%, 41%, and 48% overlapped bases for adipose tissue, mature adipocytes, and ENCODE, respectively (Jaccard index = 0.24, 0.24, and 0.31, respectively, Figure 26A, Table 19)

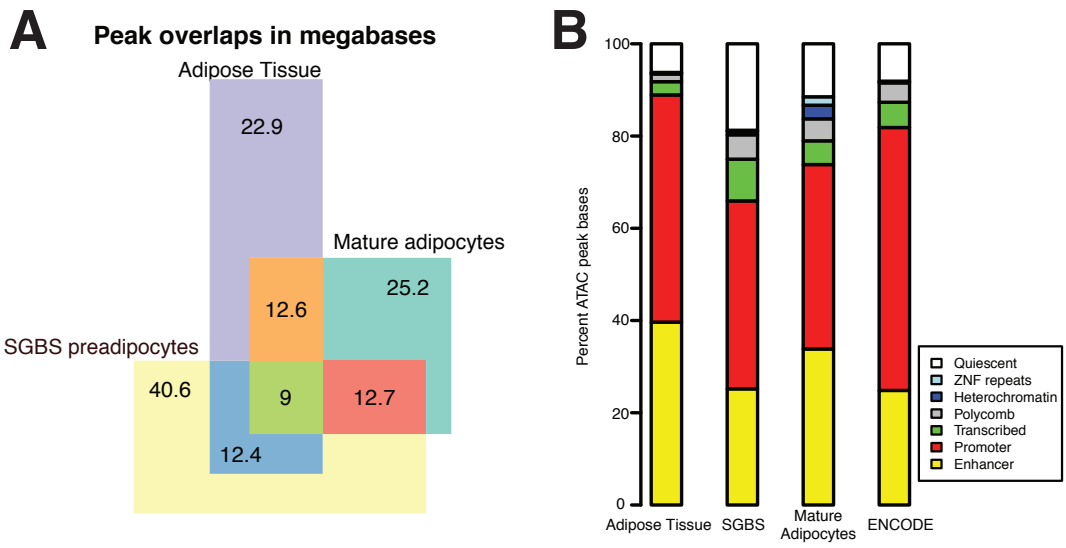

Figure 26. Differences in ATAC-seq peaks and overlap with Roadmap adipose nuclei chromatin states. A) Venn diagram showing overlap of peak megabases between adipose tissue, SGBS preadipocytes, and mature adipocyte (177). B) Overlap of ATAC peaks with Roadmap adipose nuclei chromatin states. Adipose tissue ATAC peak bases were mostly located in adipose nuclei promoter (49% of total peak bases) and enhancer (40%) chromatin state regions. SGBS preadipocyte ATAC peak bases showed less overlap with adipose nuclei promoter (41% of total peak bases) and enhancer (25%) states. The ENCODE peaks showed strong overlap with promoter states (57%) and less with enhancer (25%) states. Mature adipocytes also showed strong overlap with regulatory regions (40% promoter, 34% enhancer).

We next compared the ATAC peak locations to chromatin states in adipose nuclei from subcutaneous adipose tissue from the Roadmap Epigenomics Consortium (178). For all four ATAC profiles, the majority of peaks were located in adipose nuclei promoter and enhancers states with minimal overlap of regions associated with closed chromatin (Figure 26B, Table 19). Our adipose tissue samples showed the strongest overlap of peaks with the promoter (49% total union bases) and enhancer (40%) chromatin states of adipose nuclei (Figure 26B, Table 20). More promoter (57%) and less enhancer (25%) overlap was observed for ENCODE peaks and less overlap for mature adipocytes (40% promoter, 34% enhancer) compared to our adipose tissue. SGBS preadipocytes showed the least overlap with adipose nuclei promoters (41%) and enhancers (25%), suggesting that SGBS preadipocyte ATAC peaks may overlap novel

preadipocyte-specific regulatory regions that are not identified in the adipose nuclei chromatin states. Multiple peak normalization strategies confirm that results are not due to differences in peak number and width between samples (Table 20). The strong overlap of our tissue, ENCODE, and mature adipocyte peaks with adipose nuclei promoters and enhancers is surprising given the modest peak base overlap between them (Figure 26A, Table 19).

Finally, we identified which cell types showed the strongest epigenomic similarity to each ATAC profile by identifying percent overlap and rank of 98 tissues and cell type chromatin states from Roadmap (178). For adipose tissue peaks, enhancers in adipose nuclei rank first of all 98 cell types and promoters rank 4th behind stomach smooth muscle, chondrocyte cultured cells, and foreskin fibroblast cells (Figure 31, Table 20, 21). For SGBS preadipocyte peaks, adipose nuclei promoters rank 11th and enhancers rank 25th. 17 of the 24 cell types ranking above adipose nuclei in enhancer overlap correspond to cell lines or cultures (Figure 32, Table 20, 21) and include five types of fibroblast cells, which may reflect the fibroblast-like nature of preadipocytes (207,208). For ENCODE adipose peaks, adipose nuclei promoters rank 8th and enhancers rank 18th (Figure 33, Table 20, 21), further demonstrating that differences may exist due to adipose tissue collection and processing. For mature adipocyte peaks, adipose nuclei promoters rank 7th and enhancers rank 1st (Figure 34, Table 20, 21). Taken together, we identify important differences between adipose ATAC profiles, which are likely due to tissue collection and storage differences, cell type heterogeneity, state of differentiation, cellular environment, and/or genotype.

Open chromatin regions selective to adipose tissue or SGBS preadipocytes exhibit different regulatory signatures

Motivated by the widespread differences in adipose tissue and SGBS peak locations and Roadmap adipose nuclei regulatory overlap, we tested if peaks specific to tissue or SGBS differed in transcription factor (TF) motif enrichment, biological process enrichment, and location near transcription start sites (TSSs). We identified 57 TF motifs enriched in tissue-specific peaks (Table 22), including TFs known to promote adipogenesis (*ie:* CEBP family members) and adipose core TFs RXR, FLI1, ETS1, FOXO1, and IRF1. We observed 35 motifs enriched in SGBS-specific peaks including adipogenesis inhibitors GATA2 and GATA3 (209,210). Interestingly, SGBS-specific peaks also showed enrichment for AP-1 family members, which promote adipogenesis (210). Enhancers for multiple lineages can develop in a poised state and poised enhancers are enriched for motifs involved in development of multiple endodermal tissues (211), suggesting that a subset of SGBS ATAC peaks may mark poised regulatory elements involved in adipogenesis.

We next tested if peaks specific to tissue or SGBS were enriched for different biological processes. Adipose tissue peaks showed significant enrichment for five processes, including both adipocyte (brown fat cell differentiation) and endothelial (blood vessel endothelial cell differentiation) processes (Figure 35), reflecting the heterogeneous nature of adipose tissue. SGBS peaks showed enrichment for eight processes, including processes relevant to cultured cells, such as chemotaxis and hippo signaling (Figure 35). However, hippo signaling has also been shown to regulate adipogenesis (212). These data suggest that ATAC peaks are marking nearby genes with relevant functions to the source cells or tissue.

Finally, we identified TSSs with differential chromatin accessibility. We evaluated 78,589 unique TSSs at 43,527 genes in GENCODE version 24 for overlap with ATAC peaks. We identified 346 TSSs at 248 unique genes specific to tissue (Table 23) and 338 TSSs at 247 genes specific to SGBS. Only two genes (*EPS8* and *RGS3*) with multiple TSSs have both a tissue-specific and SGBS-specific TSS. We identified tissue-specific TSSs at *ADIPOQ,* which encodes adiponectin, a vital metabolic hormone secreted by adipocytes. Adipose-selective ATAC peaks overlap the TSS and parts of previously described regulatory elements upstream and in intron 1 of *ADIPOQ* that showed increased transcriptional activity in reporter assays (213,214), and additional peaks may mark novel regulatory elements (Figure 27). Consistently, there is also an ATAC peak at the *ADIPOQ* TSS in mature adipocytes (Figure 27) (61,177). Interestingly, strong peaks exist in SGBS preadipocytes at the promoter of an isoform of *ADIPOQ-AS1,* the antisense transcript, suggesting that the antisense may be upregulated in SGBS preadipocytes. We also observed tissue-specific TSSs at von Willebrand factor (*VWF*), which is expressed in endothelial cells and involved in platelet formation in blood. Notably, ATAC peaks are absent in preadipocytes and mature adipocytes, suggesting that *VWF* is functioning in vascular cells within adipose tissue (Figure 36). The presence of tissue-specific peaks at *VWF* and *ADIPOQ* is consistent with the heterogeneous nature of adipose tissue (25) and suggests that open chromatin profiles in heterogeneous tissue can identify regulatory regions specific to component cell types.

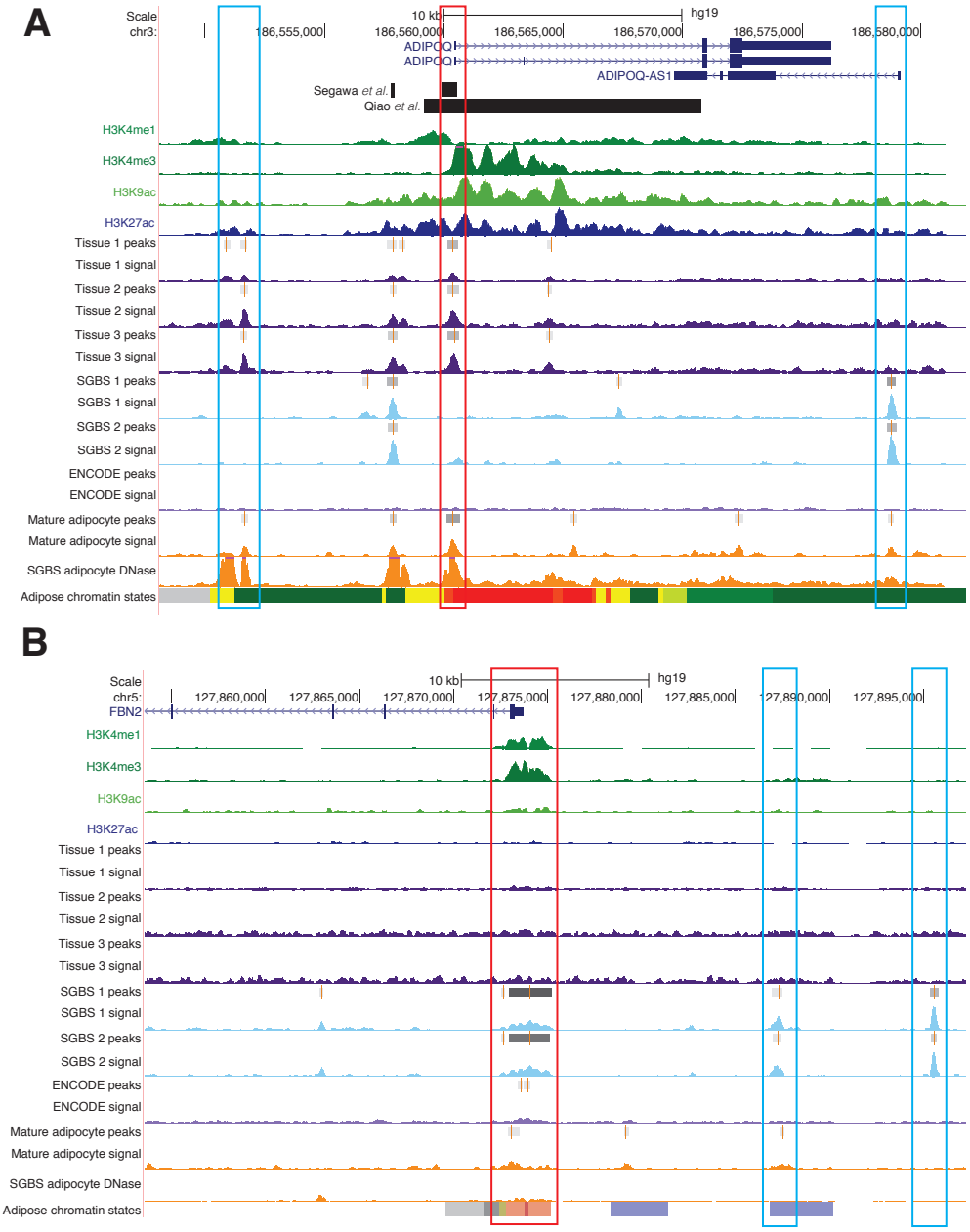

Figure 27. *ADIPOQ* and *FBN2* TSSs show differential ATAC-seq peaks. UCSC genome browser image showing the *ADIPOQ* (A) and *FBN2* (B) gene regions. Histone marks from the Roadmap Epigenomics project adipose nuclei are shown in green and blue. The three adipose tissue (purple) and two SGBS (light blue) ATAC-seq signal tracks and peaks (gray) are shown. ATAC-seq signal and peaks from the ENCODE adipose ATAC-seq (light purple, ENCODE ID: ENCSR540BML), ATAC-seq from mature adipocytes (177), and signal from SGBS adipocyte DNase (175). Chromatin states from the Roadmap Epigenomics Project are shown for adipose nuclei. Yellow, enhancer; green, transcribed; orange and red, promoter; light green, genic enhancer. Red boxes show adipose tissue-selective (A) or SGBS-selective (B) ATAC-seq peaks at TSSs. Blue boxes show intronic ATAC-seq peaks also showing specificity. Data was visualized on the hg19 UCSC genome browser.

We identified SGBS-specific TSSs at fibrillin 2 (*FBN2*) and lumican (LUM) (Figure 27, 36), which are involved in the construction and maintenance of the extracellular matrix, consistent with previous findings that extracellular matrix genes are upregulated in preadipocytes (215). FBN2 is a large glycoprotein involved in creating microfibrils in connective tissues and lumican is a extracellular matrix proteoglycan. Another fibrillin family member, *FBN1,* is involved in the transition to adipogenesis by assisting the changing extracellular matrix (216). The precise roles of *FBN2* and *LUM* in preadipocytes are unknown, but our data suggests that they may be expressed in preadipocytes, perhaps working to maintain or alter the extracellular matrix. Taken together, SGBS-specific peaks show both preadipocyte and cell culture regulatory signatures, and adipose tissue-specific peaks show regulatory signatures of mature adipocytes and non-adipocyte cell types present in adipose tissue.

Cardiometabolic GWAS loci annotation in ATAC peaks

To identify cardiometabolic traits with a strong genetic link to adipocyte regulatory elements, we annotated the ATAC-seq profiles by testing for enrichment of GWAS loci for 12 cardiometabolic traits (Table 24) and identifying TF motifs and footprints within ATAC peaks at these loci. Adipose tissue peaks showed significant enrichment ($P<5x10^{-4}$) in 9 traits (WHR, T2D, insulin traits, triglycerides, HDL-C, LDL-C, total cholesterol, cardiovascular outcomes, and blood pressure traits) (Figure 28, 37, Table 25). SGBS preadipocyte peaks showed significant enrichment in 5 traits (WHR, adiponectin, T2D, HDL-C, and total cholesterol) (Figure 28, 37, Table 25). The enrichment of circulating adiponectin GWAS loci (n=21) in SGBS but not tissue peaks is likely due to low power; adipose tissue peaks showed a strong enrichment z score (3.99), but the p value ($8*10^{-4}$) was slightly above the significance threshold

($5*10^{-4}$) (Figure 28, 37, Table 25). Although adipose tissue peaks showed enrichment for GWAS loci for more traits than SGBS peaks, SGBS peaks overlap more loci for all tested traits (Figure 28, Table 26), likely due to the presence of many more peaks in SGBS preadipocytes than adipose tissue.

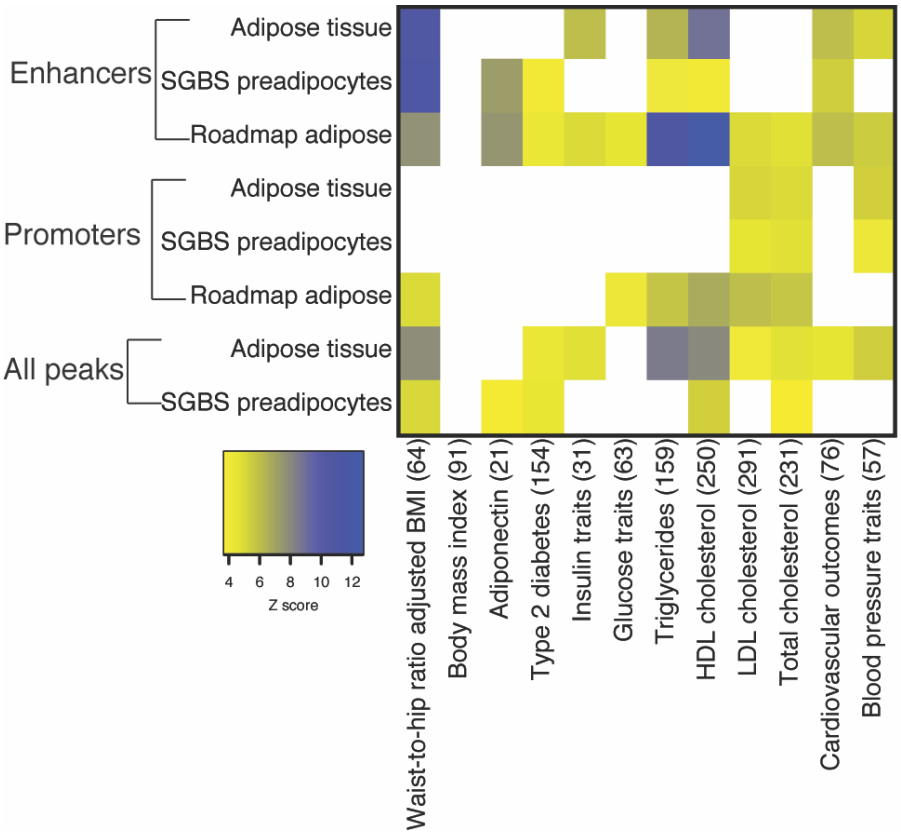

Figure 28. Cardiometabolic GWAS loci are enriched in ATAC-seq peaks. Heatmap shows enrichment of ATAC-seq peaks with cardiometabolic GWAS loci by z-score. We compare all peaks in adipose tissue and SGBS preadipocytes, peaks within Roadmap adipose promoters and enhancers, and Roadmap adipose promoters and enhancers themselves. Cells with a non-significant *p*-value are labeled white. The heatmap colored by *p*-value is shown in Figure 37.

We next tested if enrichment of cardiometabolic GWAS loci in open chromatin varies based on chromatin context by testing for enrichment of GWAS loci in promoter and enhancer states from Roadmap adipose nuclei (35) and in ATAC peaks overlapping these promoters and enhancers. Roadmap enhancers showed enrichment for loci in 11 traits and promoters in 6 traits. Enhancers showed stronger enrichment magnitudes than promoters for all but two traits (total

cholesterol and LDL-C). In ATAC peaks overlapping adipose nuclei promoters and enhancers, we found that 4 trait loci (WHR, triglycerides, HDL-C, cardiovascular outcomes) were enriched in adipose tissue and SGBS preadipocyte enhancer peaks but not promoter peaks, whereas loci for 2 traits (total cholesterol, LDL-C) were enriched in tissue and SGBS only in promoter peaks. We observed stronger enrichment of WHR loci in tissue (z score=10.6, $p=6.4*10^{-13}$) and SGBS (z score=11.1, $p=1.2*10^{-15}$) enhancer peaks than in all peaks. Interestingly, loci for WHR and insulin and for WHR alone showed stronger enrichment in adipose tissue and SGBS enhancer peaks, respectively, compared to all Roadmap enhancers, suggesting that ATAC peaks can sometimes pinpoint disease-relevant subsets of enhancers. Taken together, integrating chromatin accessibility with chromatin states revealed that enrichment of GWAS loci in chromatin states and chromatin accessibility can vary by trait, which may provide trait-specific mechanistic insights.

We next identified transcription factor motifs and footprints within ATAC peaks at cardiometabolic GWAS loci to inform functional regulatory assays. The percent of cardiometabolic GWAS loci and variants in ATAC peaks that overlapped TF motifs varied widely between traits (Table 26). Across all traits, 242 variants at 257 loci overlapped TF motifs and ATAC peaks in adipose tissue and 631 variants at 476 loci overlapped TF motifs and ATAC peaks in SGBS. We performed transcription factor footprinting using CENTIPEDE to predict binding sites for 35 motifs (Table 16) in each of the adipose tissue samples and in one SGBS sample, due to sequence duplication (SGBS 2). Using a stringent definition (see Methods), we identified aggregate footprint profiles for 9 of 35 tested transcription factor motifs in all three adipose tissue samples: CEBPA, CREB3L2, FOXO1, MEF2D, NFIA, SPI1, STAT3, STAT5A::STAT5B, and TCF7L2 (Figure 38-44, Table 16). Aggregate footprint profiles for

IRF1 and PBX1 were identified in two adipose tissue samples and for ETS1 only in adipose tissue sample 2 (Figure 42, Table 16). We identified 17 aggregate footprint profiles in SGBS, including all 9 identified in the adipose tissue samples and CEBPB, CEBPD, CTCF, IRF1, PBX1, RARA, RREB1, and RXRA::VDR (Figure 38-44, Table 16). Identification of more aggregate footprint profiles in SGBS is likely a reflection of deeper read depth in SGBS than the adipose tissue samples. 53 cardiometabolic GWAS variants at 46 loci overlap at least one footprint (Table 16). Collectively, these results will serve as a useful resource for prioritization of variants at GWAS loci and will provide insights into molecular mechanisms of cardiometabolic disease.

Functional evaluation of variants at cardiometabolic GWAS with colocalized eQTLs overlapping ATAC peaks

We further examined ATAC-seq signals at the subset of GWAS loci that are also associated with subcutaneous adipose gene expression and tested candidate variants in functional assays. We evaluated 110 cardiometabolic GWAS loci previously described to have colocalized eQTLs (see Methods) (63,186); these loci consist of 6,702 variants (LD r^2>0.8 with lead GWAS variants). For tissue peaks, 52 loci have at least one proxy variant overlapping an ATAC peak in at least one sample (Table 27); 147 of 4,538 total candidate variants overlap an ATAC peak. If we more stringently require GWAS variants to overlap an ATAC peak in all three adipose tissue samples, then we identify 81 variants at 39 loci (Table 27), of which 13 have only one variant that overlaps an ATAC peak, suggesting these variants as strong candidates for a regulatory function at these loci. For SGBS, 381 proxies at 86 loci overlap at least one SGBS peak. Importantly, 124 of the 147 proxies overlapping tissue peaks also overlap SGBS peaks. Proxies

overlapping both tissue and SGBS peaks are more likely to act through regulatory elements present in adipose and are not present simply due to tissue heterogeneity or cell culture. We identified a previously implicated functional variant at the *ANGPTL8* HDL-C GWAS locus (186). Of the 147 variants overlapping ATAC peaks in at least one sample, 97 (66%) overlap at least one TF motif, two of which overlap a TF footprint (rs7187776 and rs174538).

rs1534696, located in the second intron of *SNX10* (encoding sorting nexin 10) is associated with waist-hip ratio (P=2x10^{-8}, ß=0.027, in women) (6) and exhibits a colocalized eQTL for *SNX10* (P=3.4x10^{-150}, ß=1.12) and *CBX3* (P=1.1x10^{-13}, ß=0.39) in adipose tissue (63). rs1534696 is the only candidate variant for both the GWAS and eQTL associations based on LD (r^2>0.8) and overlaps a MEF2C motif. Interestingly, is not located in a predicted regulatory region as defined by any adipose-related Roadmap chromatin states. We tested a 250-bp region containing either allele of rs1534696 and encompassing the entirety of the surrounding ATAC peak in transcriptional reporter luciferase assays in SGBS preadipocytes and SW872 liposarcoma cells (Figure 29, 45). In SGBS preadipocytes, the construct containing rs1534696-A showed higher transcriptional activity than rs153696-C (P=0.007) in the reverse orientation with respect to the genome (Figure 30), and in SW872 cells, the construct containing rs1534696-A showed a similar trend (P=.065) in the forward orientation (Figure 45). In EMSA, we observed increased protein binding for rs1534696-A using nuclear extract from SGBS preadipocytes but not with SW872 nuclear extract (Figure 29, 45). Taken together, these data suggest that rs1534696-A may increase transcription factor binding and transcriptional activity of *SNX10* and/or *CBX3* and contribute to the molecular mechanism at this WHR GWAS locus.

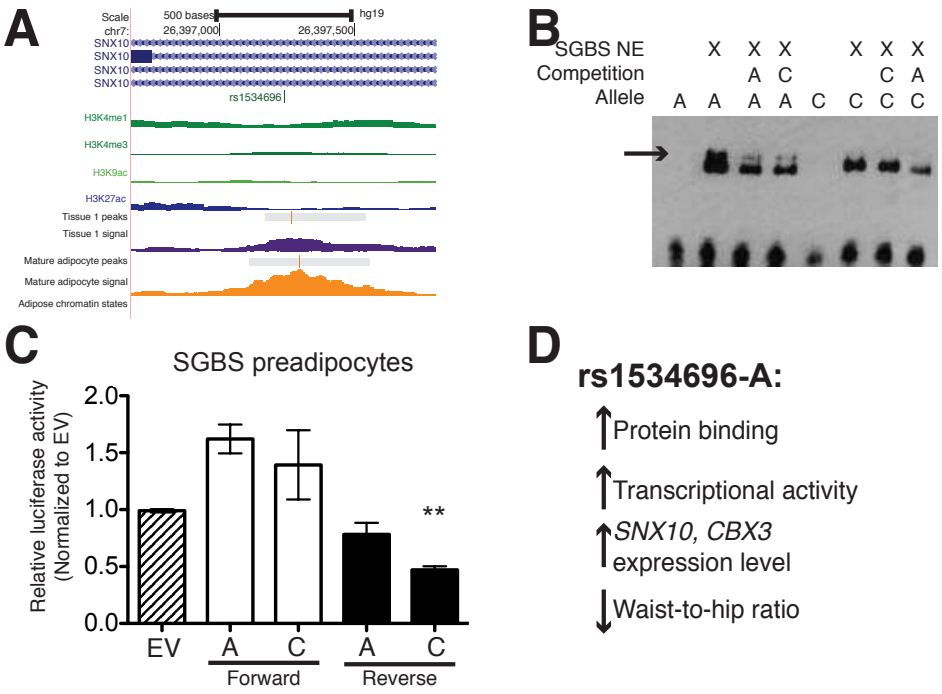

Figure 29. A variant at the *SNX10* WHR GWAS locus alters transcriptional activity and protein binding. A) rs1534696 overlaps an ATAC-seq peak and is located in intron 2 of *SNX10* but is not in a predicted regulatory region based on Roadmap chromatin states or histone marks. B) The genomic region containing rs1534696-A shows increased transcriptional activity and allelic differences in the reverse orientation (**P=0.007). Bars represent the mean ±standard deviation of three independent clones. C) rs1534696-A shows increased protein binding in EMSA using SGBS nuclear extract. Arrow shows allelic differences in protein binding. D) Summary of the direction of effect of rs1534696-A. Additional regulatory assays are shown in Figure 45.

rs7187776 is located in the 5'UTR of a long isoform of *SH2B1* (encoding SH2B adaptor protein 1) and is in strong LD (r^2 > 0.8) with the lead variant associated with BMI (rs3888190, P=3.14x10^{-23}, ß=0.031)(7), which exhibits a colocalized eQTL for *SH2B1* (P=4.7x10^{-15}, ß=-0.39) and *ATXN2L* (P=2.5x10^{-11}, ß=-0.34) in adipose tissue (63). rs7187776 is one of 124 candidate variants based on LD (r^2>0.8) with the lead GWAS and eQTL variants and one of five variants that overlap adipose tissue ATAC peaks at this locus (Table 27). rs7187776 overlaps a SPI1 (encodes PU.1 protein) motif and footprint (Figure 30). We tested a 456-bp region containing either allele of rs7187776 and encompassing the entirety of the surrounding ATAC

peak in transcriptional reporter luciferase assays in SGBS preadipocytes and SW872 cells (Figure 30, 46). In both cell types, we observed extremely strong transcriptional activity (>200-fold compared to background) but no allelic differences between rs7187776 alleles (Figure 46). Allelic differences in transcriptional activity may have been masked by the massive transcriptional effect of this region. In EMSA, we observed allele-specific binding of rs7187776-G to purified PU.1 protein and similar binding using nuclear extract from SW872 cells, consistent with the predicted motif (Figure 30, Figure 46). One additional variant that overlaps an ATAC-seq peak showed allelic differences in protein binding using nuclear extract from SW872 cells (Figure 46). These data suggest that rs7187776-G increases PU.1 binding, and may contribute to the molecular mechanism at the *ATP2A1-SH2B1* BMI GWAS locus.

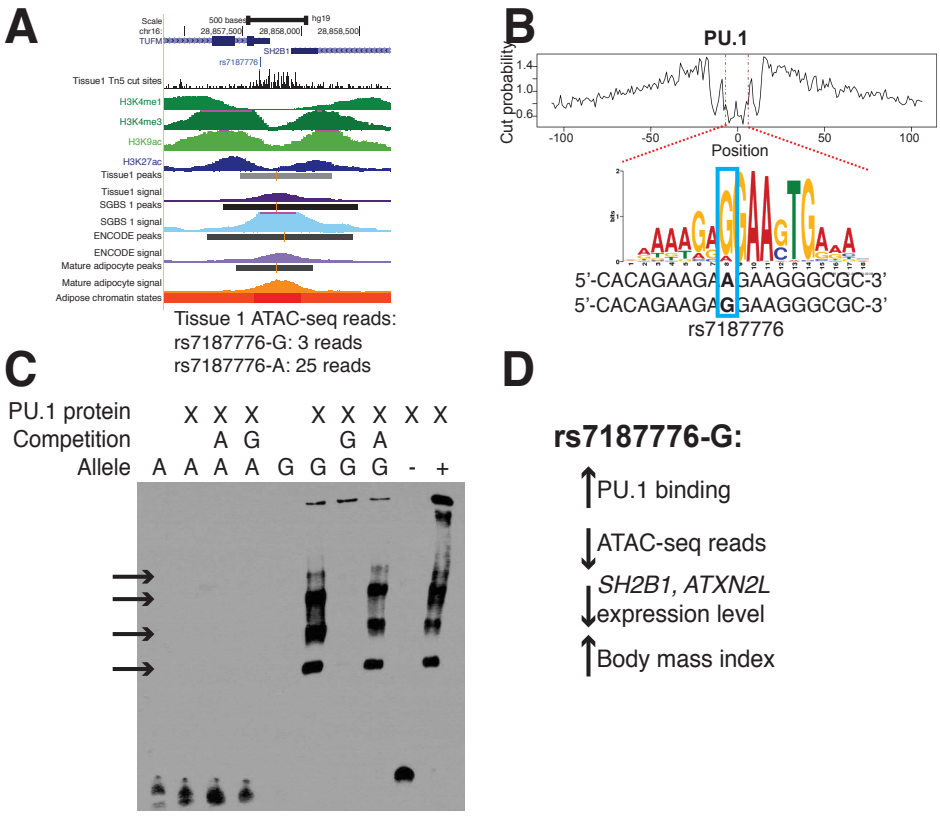

Figure 30. A variant at the *ATP2A1-SH2B1* BMI GWAS locus alters chromatin accessibility and PU.1 binding. A) rs7187776 is located in the promoter of a long *SH2B1* isoform and the 5'-UTR of *TUFM.* We observe more ATAC-seq reads for rs7187776-A (Table 29). B) rs7187776-G creates a PU.1 binding motif (JASPAR motif) and we identify a PU.1 footprint in ATAC-seq reads. C) rs7187776-G shows increased protein binding to purified PU.1 in EMSA, similar to a positive control containing the PU.1 motif (+). A negative control (-) and rs7187776-A show no binding to PU.1. Arrows show allele-specific protein binding. Similar protein binding patterns were observed using SW872 nuclear extract (Figure 46). D) Summary of the direction of effect of rs7187776-G.

Allelic imbalance in ATAC-seq reads

Allelic imbalance in ATAC-seq reads, especially at a colocalized GWAS-eQTL locus, can suggest that a variant alters chromatin accessibility and may contribute to the underlying mechanism at the locus. To find genetic variants that exhibit allelic differences in chromatin accessibility, we identified heterozygous sites in adipose tissue samples that exhibited allelic

imbalance in ATAC-seq reads with at least 10 total reads and at least one read per allele. We identified 812 sites showing nominal allelic imbalance (beta-binomial $P<0.05$) in at least one sample (Table 28). Only 40 GWAS-eQTL variants were heterozygous in at least one tissue sample and met our threshold of 10 total reads and 1 read per allele. Of these 40 variants, 8 showed nominal evidence of ATAC allelic imbalance (Table 29), including rs7187776 at the *SH2B1* locus (chi square p=0.0044 Figure 30). We confirmed that the overlap of nominally imbalanced loci with GWAS-eQTLs is still greater than expected by chance when removing duplicated reads from the imbalance calculation (chi square p=0.043). These results suggest that increased power through higher read depth will enable identification of many more disease-associated loci exhibiting allelic imbalance in chromatin accessibility. Open chromatin profiling in additional samples will also allow the identification of chromatin quantitative trait loci (cQTLs), which are not restricted to heterozygous variants. Collectively, these analyses will help identify disease-associated variants that mediate their effects on disease through chromatin accessibility.

Discussion

In this study, we generated high quality ATAC-seq open chromatin profiling from three frozen clinical adipose samples and two preparations of SGBS preadipocytes. We identified differences between adipose tissue, preadipocyte, and adipocyte open chromatin profiles, including adipose tissue-selective peaks at *ADIPOQ* and *VWF* and preadipocyte-selective peaks at extracellular matrix genes. Adipose tissue open chromatin profiles largely overlapped Roadmap adipose nuclei chromatin states and SGBS preadipocytes showed more overlap with fibroblast-type cells. Transcription factor motifs and footprints in ATAC peaks overlap GWAS

variants and GWAS traits are enriched in ATAC peaks within enhancers. Finally, we used the ATAC-seq profiles to annotate potential regulatory variants at GWAS-eQTL colocalized loci and provided experimental evidence of allelic differences in regulatory activity for variants at the *SNX10* and *ATP2A1-SH2B1* GWAS loci. Taken together, these data are the deepest characterization of chromatin accessibility in adipose tissue and preadipocytes to date.

Important differences exist between adipose tissue, preadipocyte, and mature adipocyte ATAC-seq profiles. Explanations for these differences include cell-type composition/heterogeneity, the differentiation state of adipocytes, the cultured nature of SGBS preadipocytes, and technical differences that arose when generating the ATAC-seq data, such as sequencing depth. At the TSS of *VWF,* adipose tissue-selective ATAC peaks are present at the TSS and very little signal is present in preadipocytes or mature adipocytes (Figure 32), suggesting that *VWF* functions within vascular cells in adipose tissue. At the *ADIPOQ* gene we observed adipose tissue-selective and mature adipocyte ATAC peaks and ATAC peaks specific to SGBS preadipocytes at the TSS of *ADIPOQ-AS1,* the antisense transcript of *ADIPOQ.* These data led us to hypothesize that *ADIPOQ-AS1* may be expressed in preadipocytes to stop production of *ADIPOQ,* as antisense transcripts are traditionally thought to interfere with production of the sense transcript (217). The accessibility pattern of *ADIPOQ* is consistent with a previous finding that the *ADIPOQ* promoter is inaccessible until differentiation (218) and it's role in adipocyte differentiation (219-221). Among 98 Roadmap tissue and cell types, SGBS preadipocyte ATAC profiles were more similar to fibroblast-like cells and cell lines than to adipose nuclei, reflecting differences due to the cultured nature of SGBS preadipocytes. Additionally, we identified 17 transcription factor footprints in SGBS, but only 9 in adipose tissue due to more read depth in SGBS. We also observed differences between the previously

published ENCODE adipose tissue and our adipose tissue ATAC-seq profiles; these differences are likely due to technical variation such as such as biopsy location, freezing method, or storage conditions could contribute to these differences.

Adipose ATAC-seq profiles provide insight into the mechanisms of cardiometabolic GWAS loci. We found that ATAC peaks can sometimes pinpoint disease-relevant subsets of enhancers and that adipose ATAC peaks are enriched in WHR, but not BMI GWAS loci. This enrichment is consistent with recent findings that WHR loci are enriched in adipose transcriptional regulatory elements (10) and that BMI GWAS loci are enriched in pathways involved in central nervous system biology (11). We also identify enrichment of other cardiometabolic traits including insulin traits, lipids, cardiovascular outcomes, and blood pressure traits that have historically been enriched in liver, islet, and blood cell types more than adipose regulatory datasets. Identifying the transcription factor(s) bound to a regulatory variant is a challenging part of defining the molecular mechanisms underlying cardiometabolic GWAS loci; transcription factor footprints identify factors that are physically occupying the DNA at variants or loci with more accuracy than the simple overlap of binding motifs (185). We successfully generated transcription factor footprints for 17 transcription factors (Figures 38-44), which can be used to identify transcription factor that bind sites overlapping candidate variants at cardiometabolic GWAS loci.

We show two example loci where ATAC peaks helped prioritize candidate variants. At the *SNX10* waist-hip ratio GWAS and adipose eQTL locus, we identified a potentially functional variant. rs1534696 is not located in a predicted regulatory region based on existing chromatin state data. However, rs1534696 overlaps an ATAC peak in adipose tissue and shows allelic differences in transcriptional reporter and protein-binding assays. At a second example, the

ATP2A1-SH2B1 BMI GWAS and adipose eQTL locus, we identified a PU.1 binding motif and footprint at rs7187776, as well as allelic imbalance in ATAC-seq reads, and confirmed the allelic differences in PU.1 binding *in vitro.* However*,* further experiments are needed to confirm that PU.1 and/or other ETS family members are binding at this locus in vivo. These data provide an excellent example of how to integrate GWAS, eQTL, and ATAC-seq data to identify functional variants at GWAS loci. Further experiments are needed to determine if these variants are the sole functional variant at each locus, and which gene(s) are contributing to obesity risk.

ATAC-seq open chromatin profiling can be further improved and future experiments are needed to fully characterize chromatin accessibility in adipose. The library preparation methods can be further optimized; we considered multiple library preparation conditions and chose detergent-treated samples to continue with analyses; however, we note that our sample size is too small to make definitive conclusions regarding these choices. The ATAC-seq method is relatively new and continues to improve, especially for frozen tissues and cells; a recent update to the original protocol, called Omni-ATAC, includes additional detergents for frozen tissues (222). Future chromatin profiles from additional individuals will enable detection chromatin quantitative trait loci (cQTL) associations, which could provide evidence that a variant allele alters chromatin accessibility and would help to prioritize variants for functional followup. Cellular environment can alter chromatin state, especially for cardiometabolic phenotypes, where both genetic and environmental factors can affect phenotypic outcome. Future work treating the cells with various metabolic stimuli and in in different cellular states (e.g. insulin or glucose treatment and stages of adipocyte differentiation) will provide needed insight to how chromatin changes with metabolic environmental triggers.

Here, we present ATAC-seq open chromatin profiles for frozen adipose tissue and cultured preadipocytes. We showed the utility of open chromatin profiles in multiple tissue samples and across cell types within heterogeneous tissue. Together, these data add to the growing understanding of gene regulation in adipose and the complex genetic mechanisms of cardiometabolic traits and diseases.

Supplemental Figures and Tables

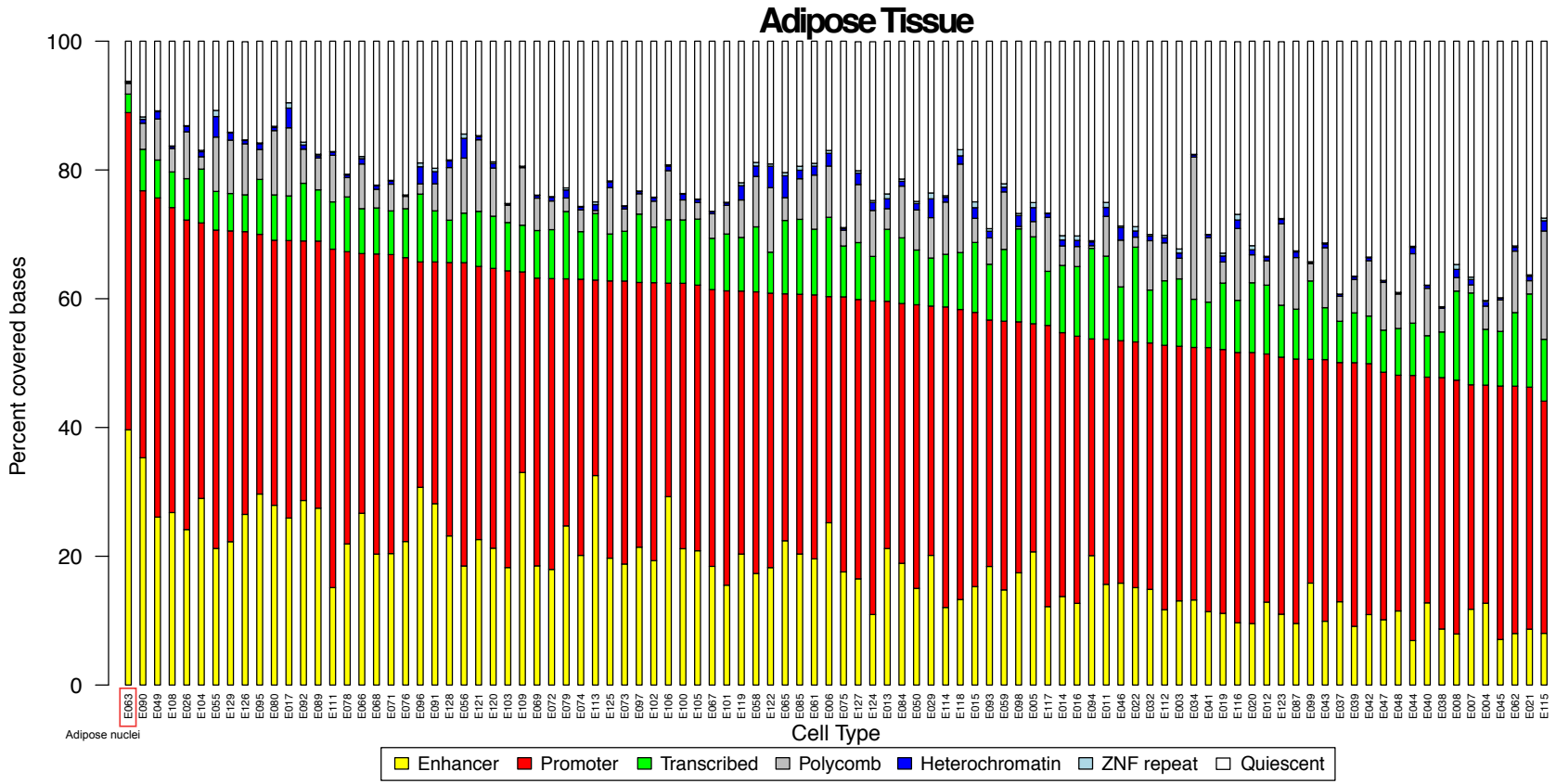

Figure 31. Adipose tissue ATAC peak overlap of all chromatin states. ATAC-seq peak base overlap was identified in 98 tissue and cell types chromatin states from the Epigenome Roadmap Project (178). Adipose nuclei is boxed in red. Bars are ordered based on promoter and enhancer overlap. Cell type IDs are listed in Table 21.

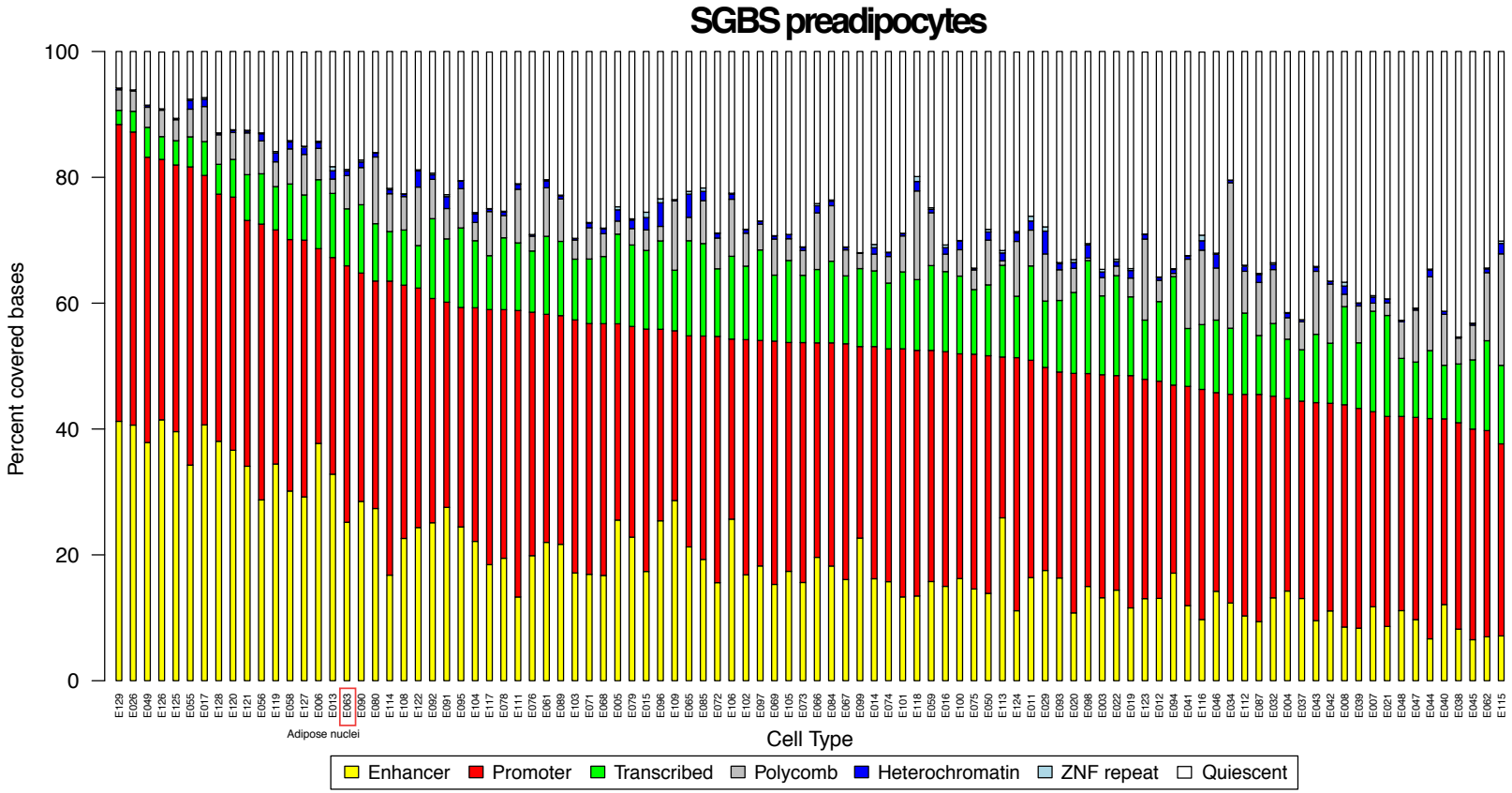

Figure 32. SGBS ATAC peak overlap of all chromatin states. ATAC-seq peak base overlap was identified in 98 tissue and cell types chromatin states from the Epigenome Roadmap Project (178). Adipose nuclei is boxed in red. Bars are ordered based on promoter and enhancer overlap. Cell type IDs are listed in Table 21.

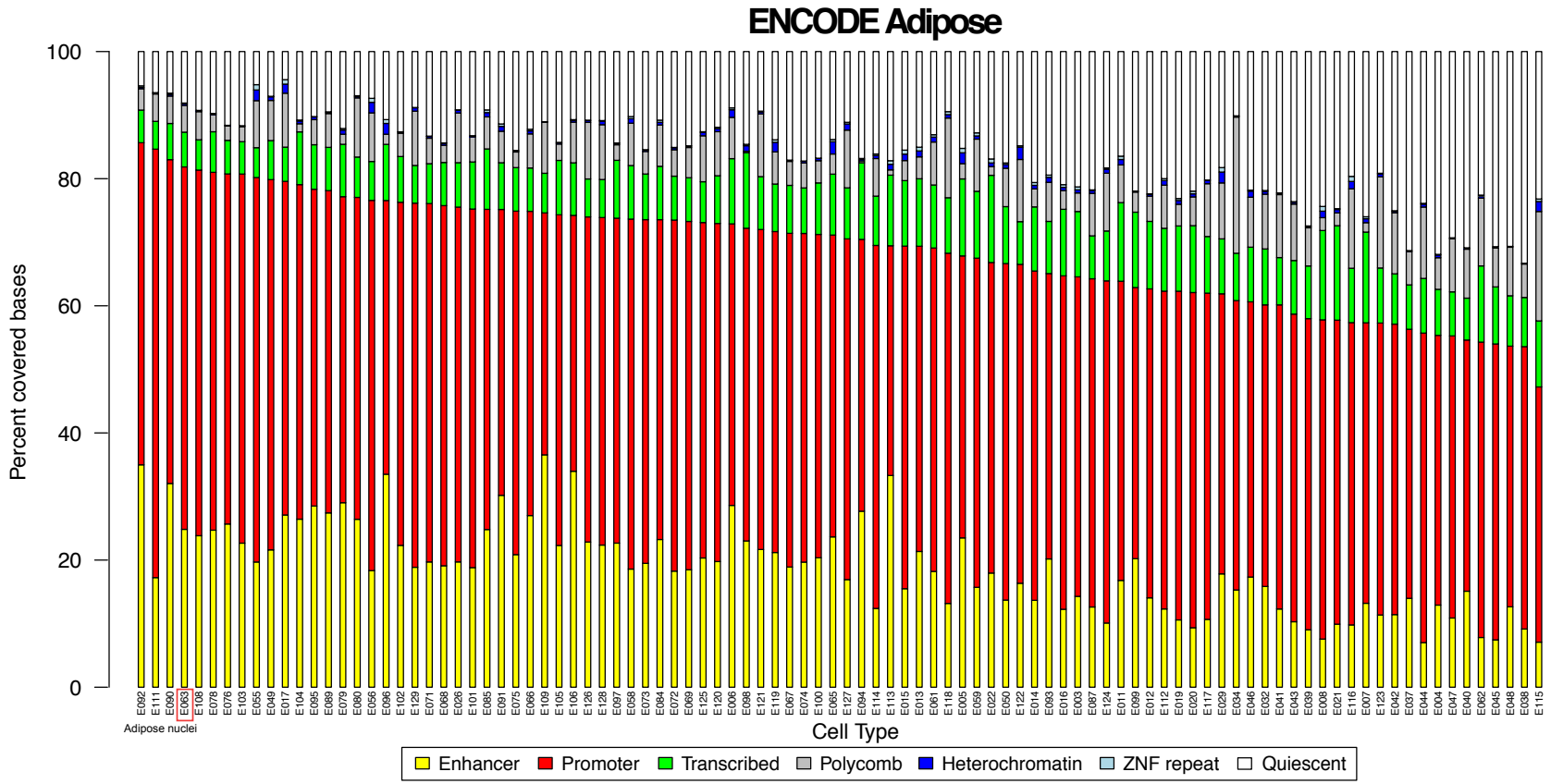

Figure 33. ENCODE ATAC peak overlap of chromatin states. ATAC-seq peak base overlap was identified in 98 tissue and cell types chromatin states from the Epigenome Roadmap Project (178). Adipose nuclei is boxed in red. Bars are ordered based on promoter and enhancer overlap. Cell type IDs are listed in Table 21.

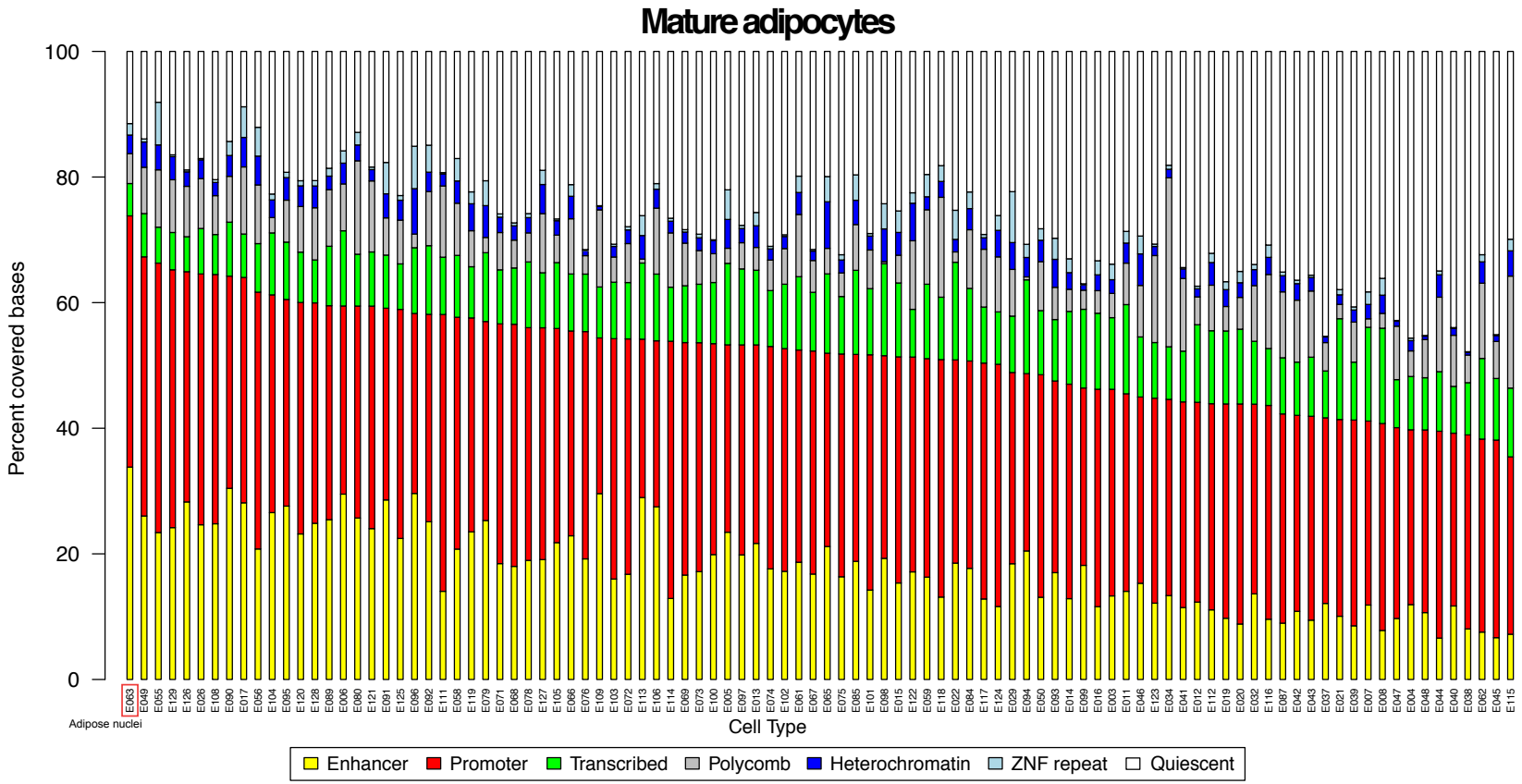

Figure 34. Mature adipocyte ATAC peak overlap of all chromatin states. ATAC-seq peak base overlap was identified in 98 tissue and cell types chromatin states from the Epigenome Roadmap Project (178). Adipose nuclei is boxed in red. Bars are ordered based on promoter and enhancer overlap. Cell type IDs are listed in Table 21.

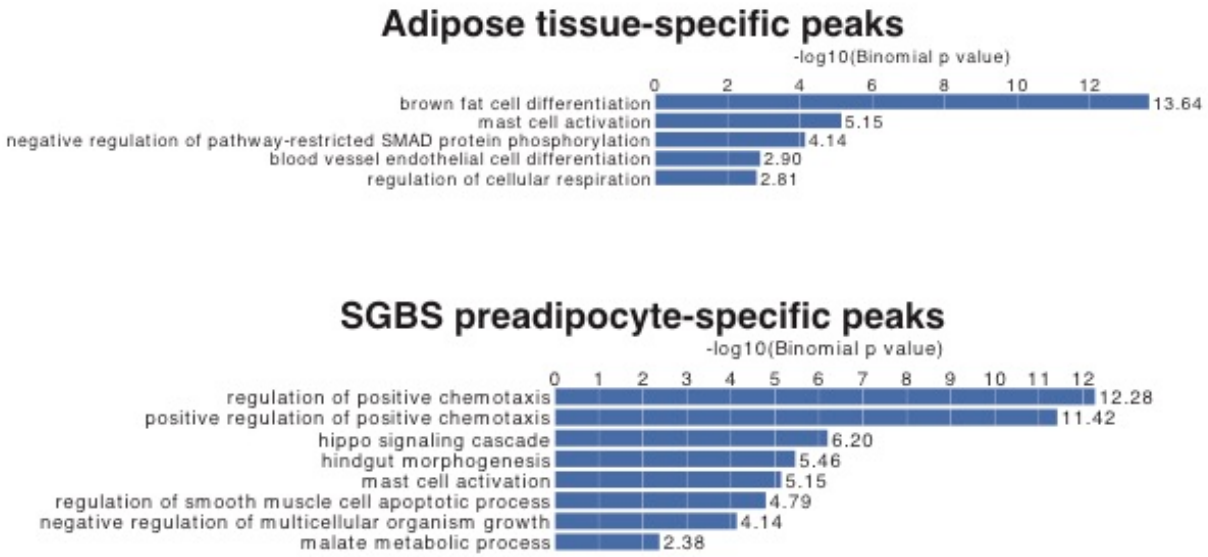

Figure 35. Gene ontology enrichment of adipose tissue- and SGBS- specific ATAC peaks. Gene ontology enrichment of genes near adipose tissue- or preadipocyte-specific peaks was performed using GREAT with default parameters.

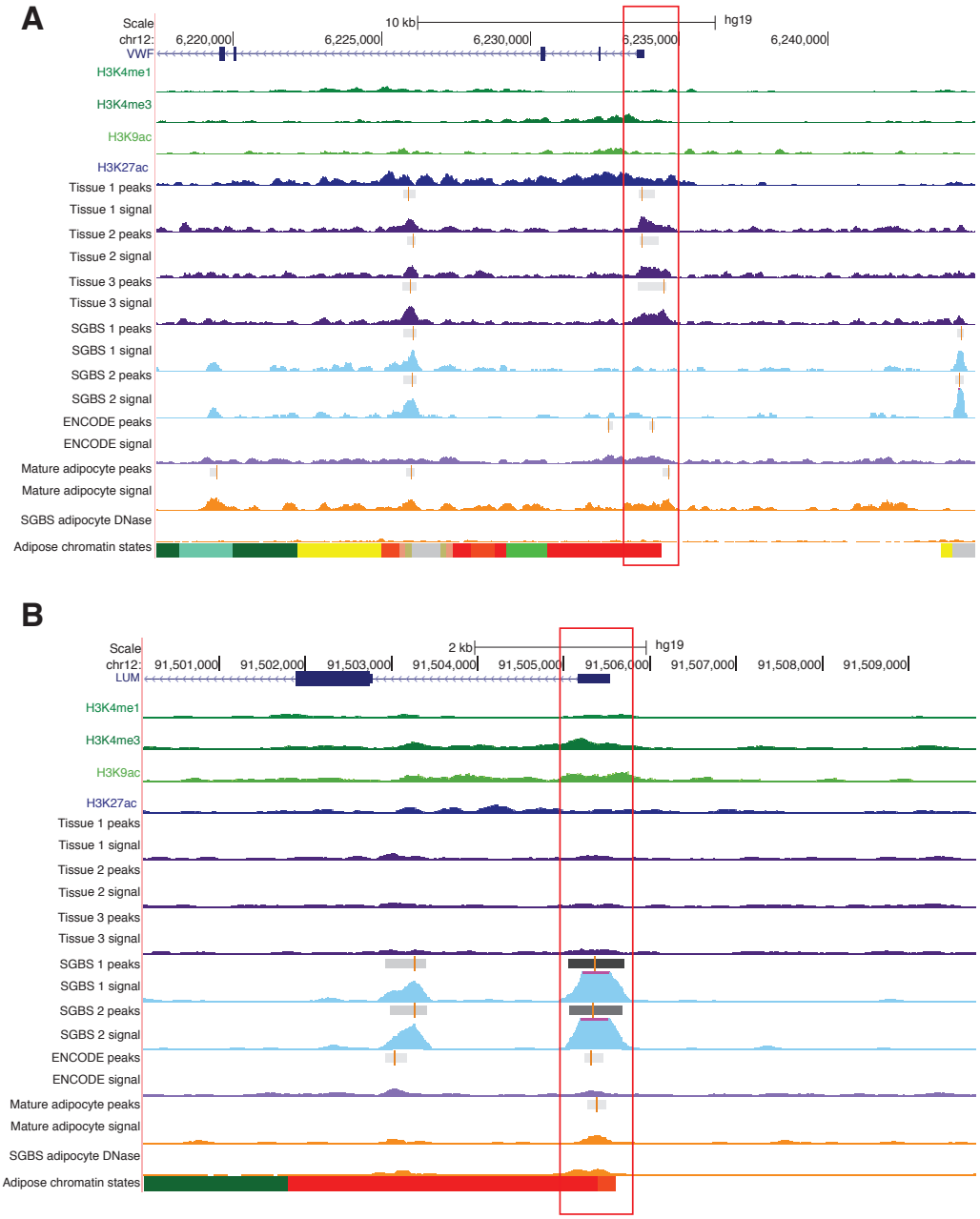

Figure 36. *VWF* and *LUM* TSSs show differential ATAC-seq peaks. UCSC genome browser image showing the *VWF* (A) and *LUM* (B) gene regions. Histone marks from the Roadmap Epigenomics project adipose nuclei are shown in green and blue. The three adipose tissue (purple) and two SGBS (light blue) ATAC-seq signal tracks and peaks (gray) are shown. ATAC-seq signal and peaks from the ENCODE adipose ATAC-seq (light purple, ENCODE ID: ENCSR540BML), ATAC-seq from mature adipocytes (177), and signal from SGBS adipocyte DNase (175). Chromatin states from the Roadmap Epigenomics Project are shown for adipose nuclei. Yellow, enhancer; green, transcribed; orange and red, promoter; light green, genic enhancer. Red boxes show adipose tissue-selective (A) or SGBS-selective (B) ATAC-seq peaks at TSSs. Data was visualized on the hg19 UCSC genome browser.

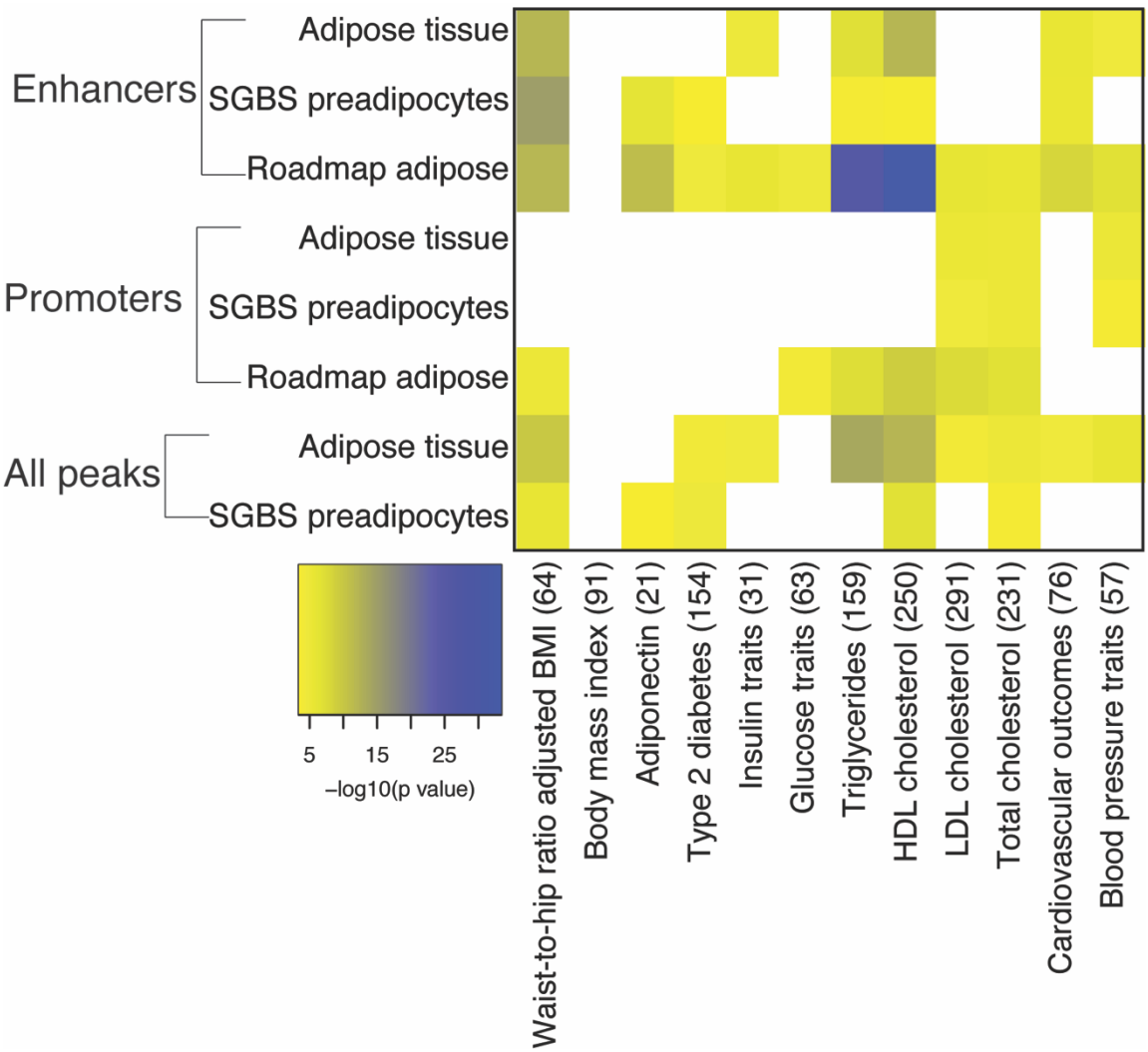

Figure 37. GWAS loci are enriched in ATAC-seq peaks. Heatmap shows enrichment of ATAC-seq peaks with cardiometabolic GWAS loci by *p*-value. We compare all peaks in adipose tissue and SGBS preadipocytes, peaks within Roadmap adipose promoters and enhancers, and Roadmap adipose promoters and enhancers themselves. Cells with a non-significant *p*-value are labeled white.

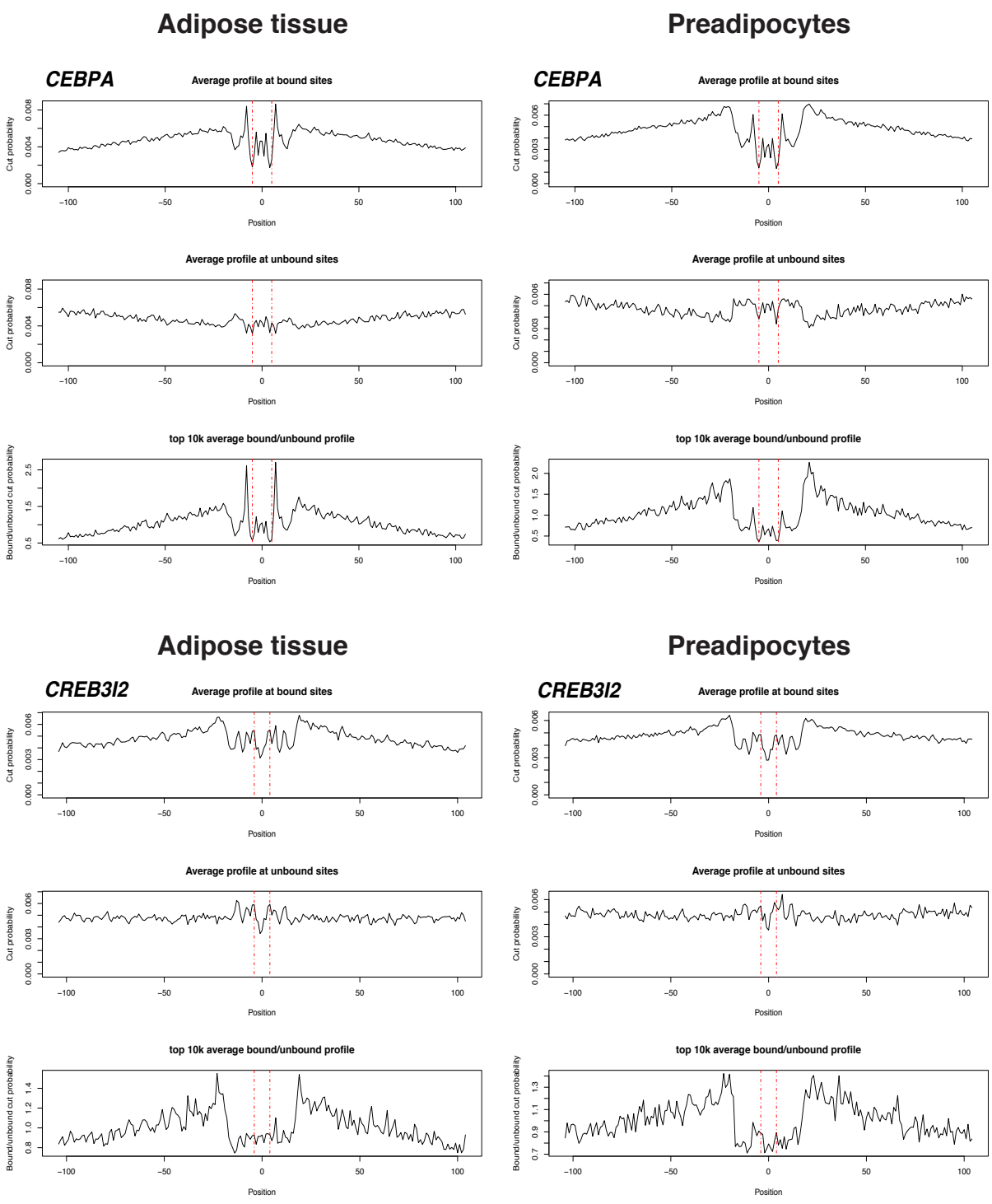

Figure 38. Aggregate footprint profiles for CEBPA and CREB312 in adipose tissue samples and SGBS preadipocytes. Footprints are shown for adipose tissue sample 1, but results were similar for the tissue Sample 2 and 3. Bound, unbound, and normalized profiles are shown for each factor.

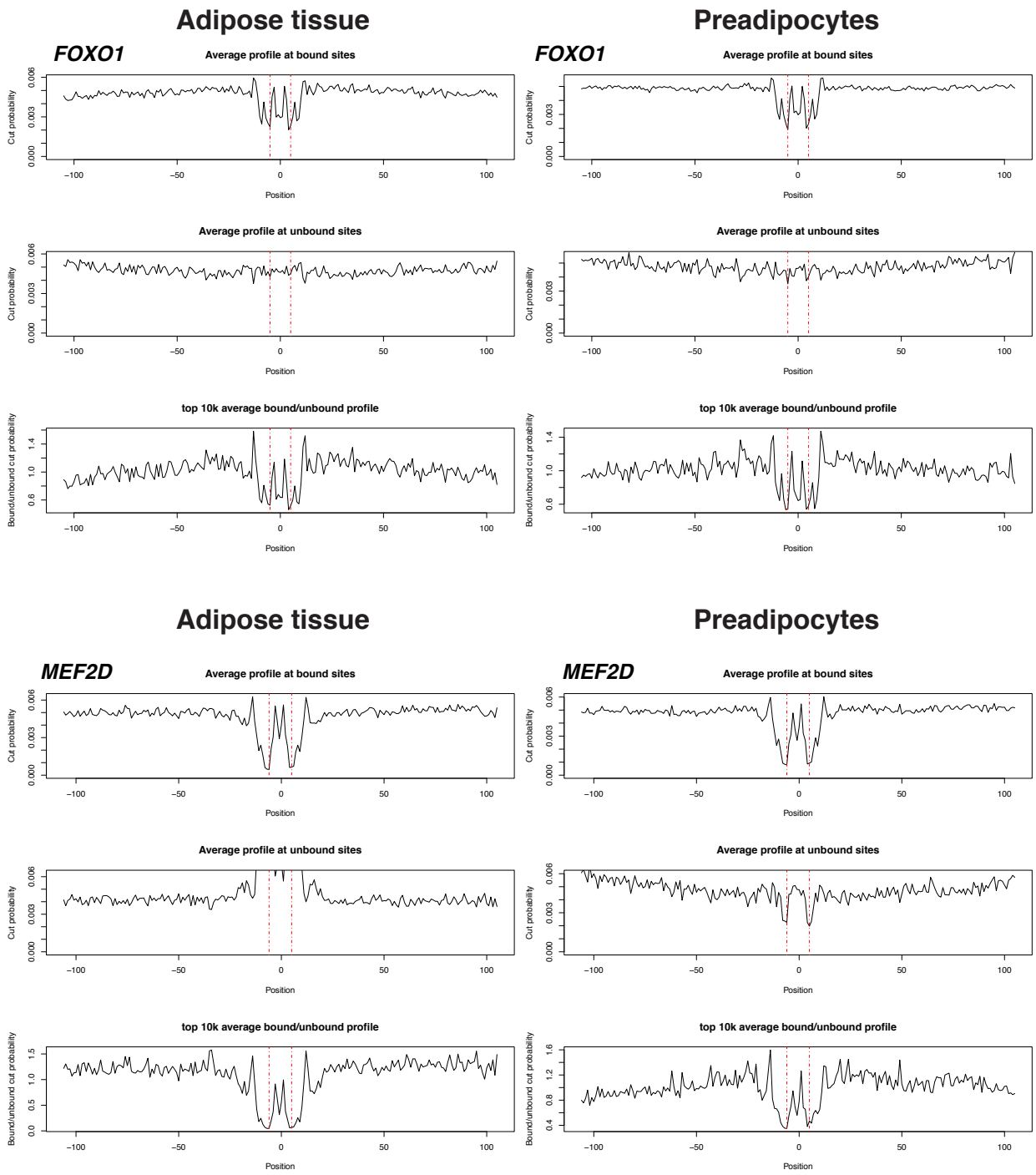

Figure 39: Aggregate footprint profiles for FOXO1 and MEF2D in adipose tissue samples and SGBS preadipocytes. Footprints are shown for adipose tissue sample 1, but results were similar for the tissue Sample 2 and 3. Bound, unbound, and normalized profiles are shown for each factor.

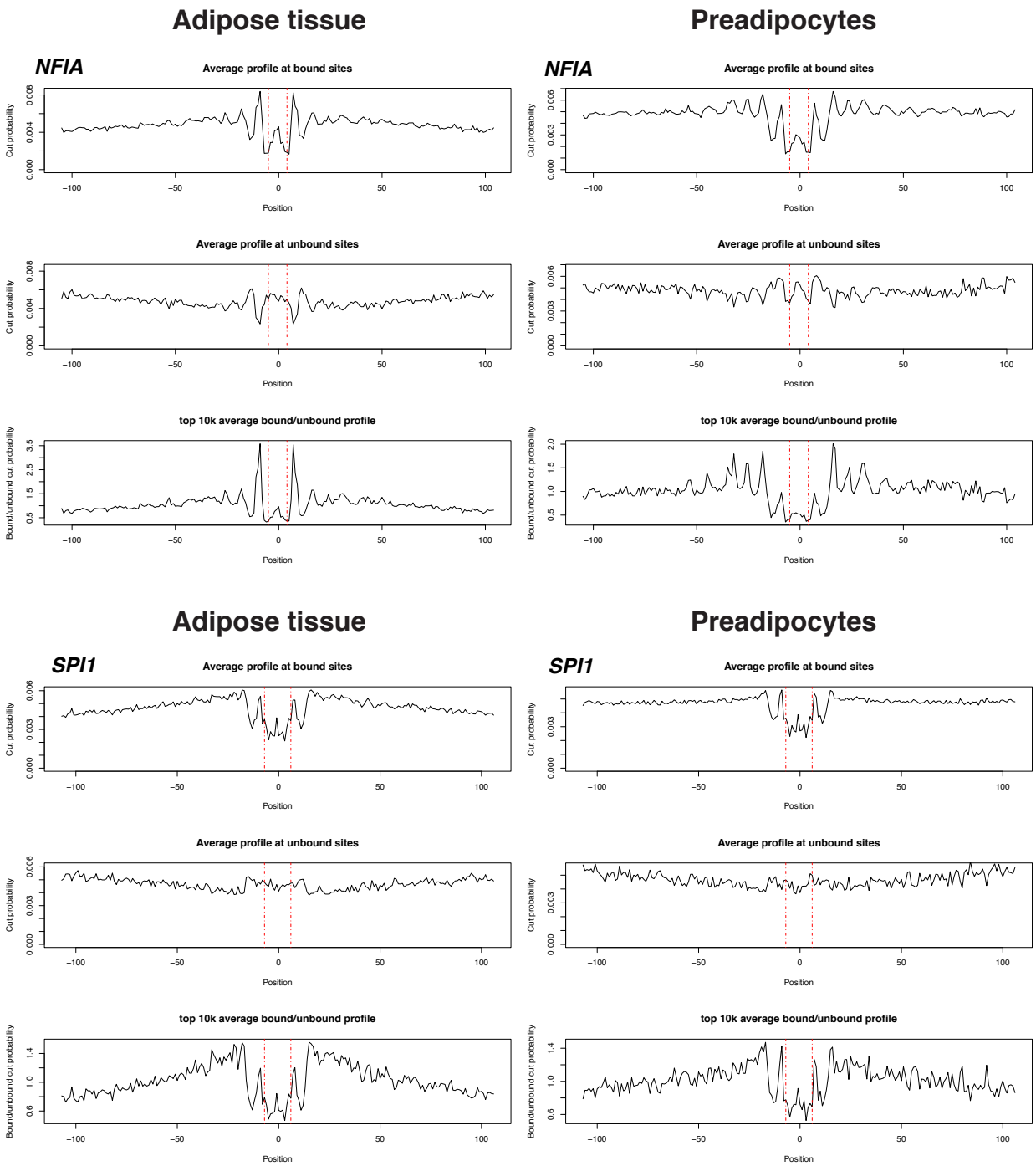

Figure 40: Aggregate footprint profiles for NFIA and SPI1 in adipose tissue samples and SGBS preadipocytes. Footprints are shown for adipose tissue sample 1, but results were similar for the tissue Sample 2 and 3. Bound, unbound, and normalized profiles are shown for each factor.

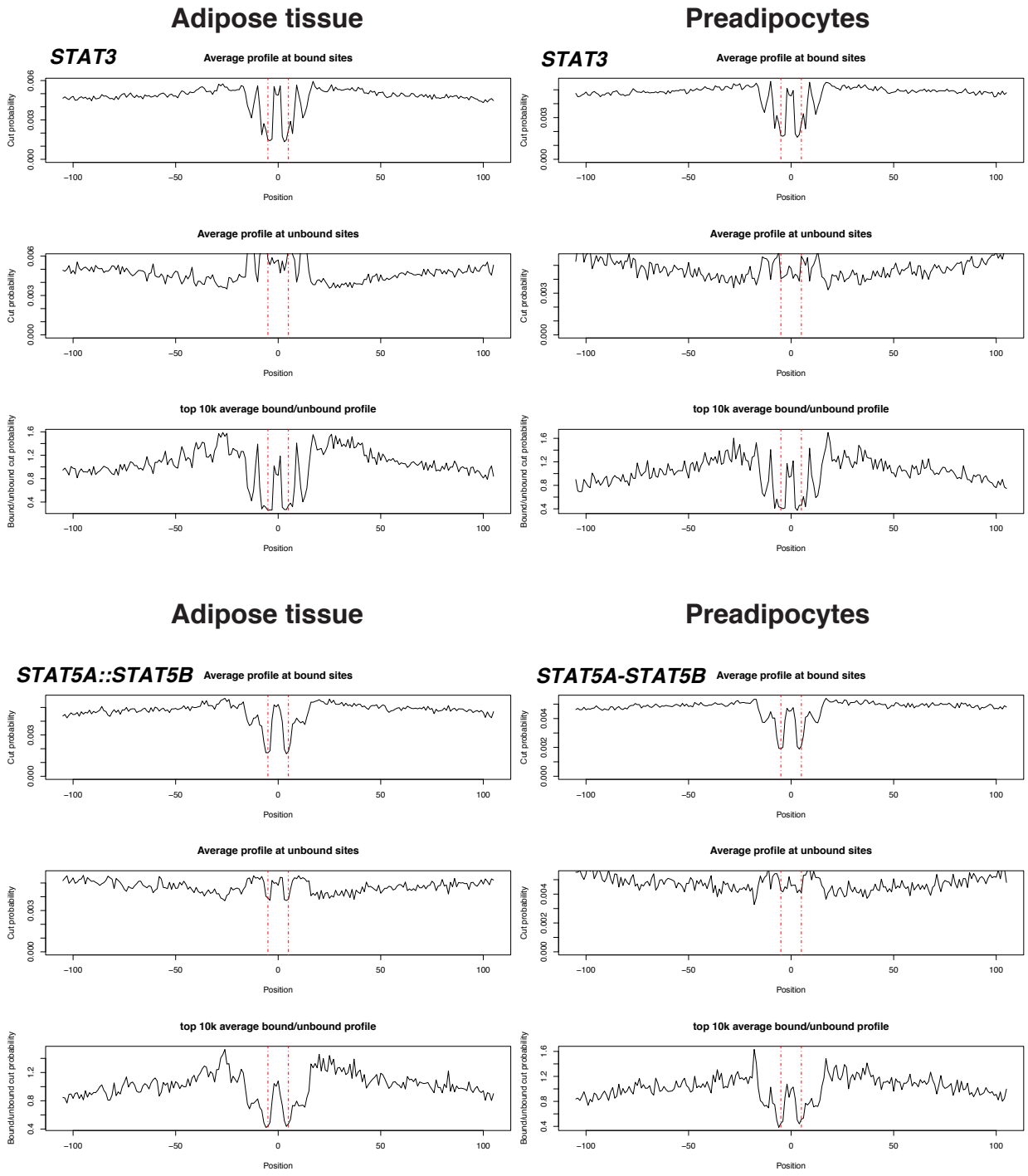

Figure 41: Aggregate footprint profiles for STAT3A and STAT5A::STAT5B in adipose tissue samples and SGBS preadipocytes. Footprints are shown for adipose tissue sample 1, but results were similar for the tissue Sample 2 and 3. Bound, unbound, and normalized profiles are shown for each factor.

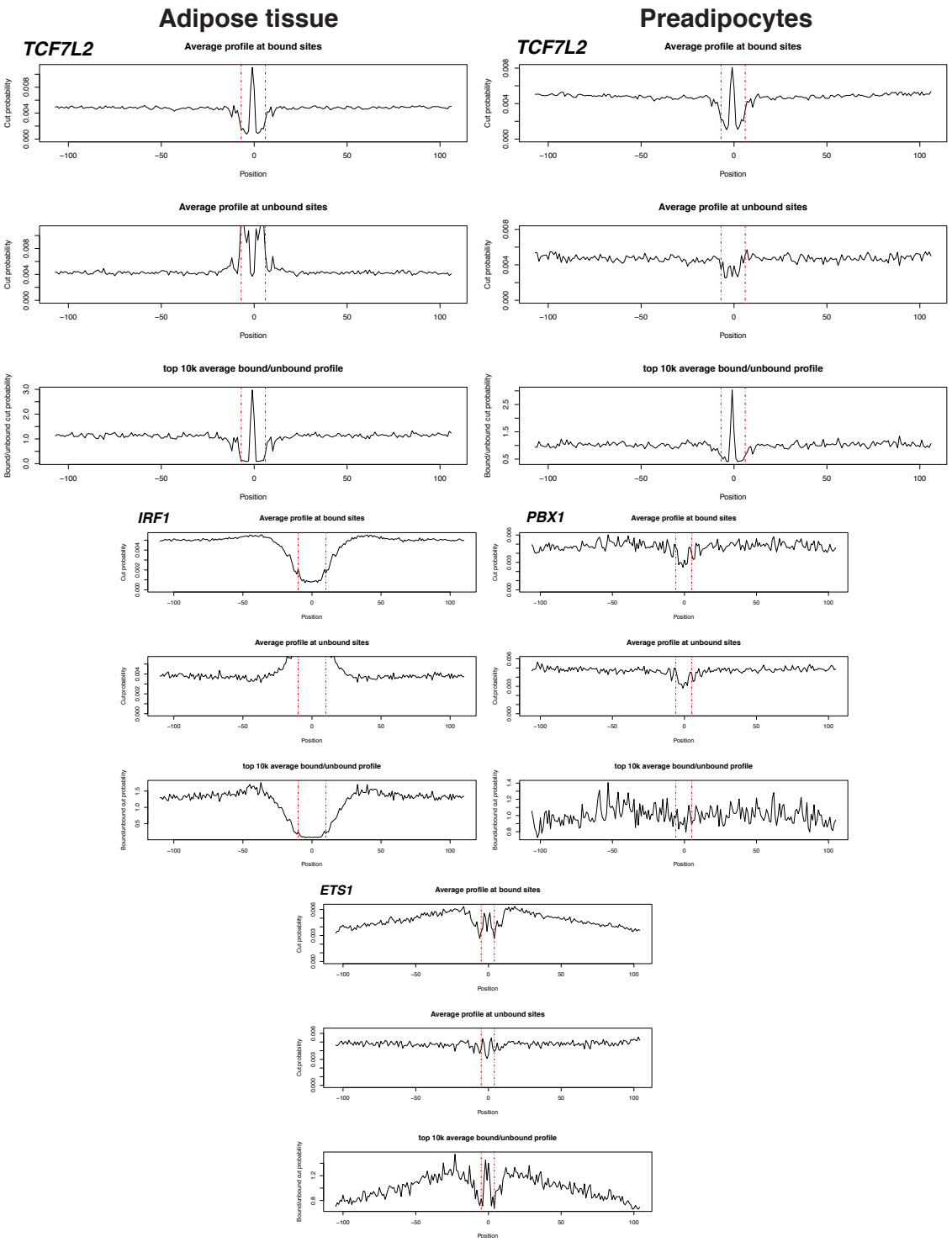

Figure 42: Aggregate footprints for TCFL7, IRF1, PBX1, and ETS1. Aggregate footprints for *TCFL7* were observed in all three adipose tissue samples and SGBS preadipocytes. Aggregate footprint for *IRF1* was observed in adipose samples 1 and 3; *PBX1* in adipose samples 1 and 2; and *ETS1* in adipose sample 2. For footprints in more than one sample, profiles are shown for adipose sample 1. Bound, unbound, and normalized profiles are shown for each factor.

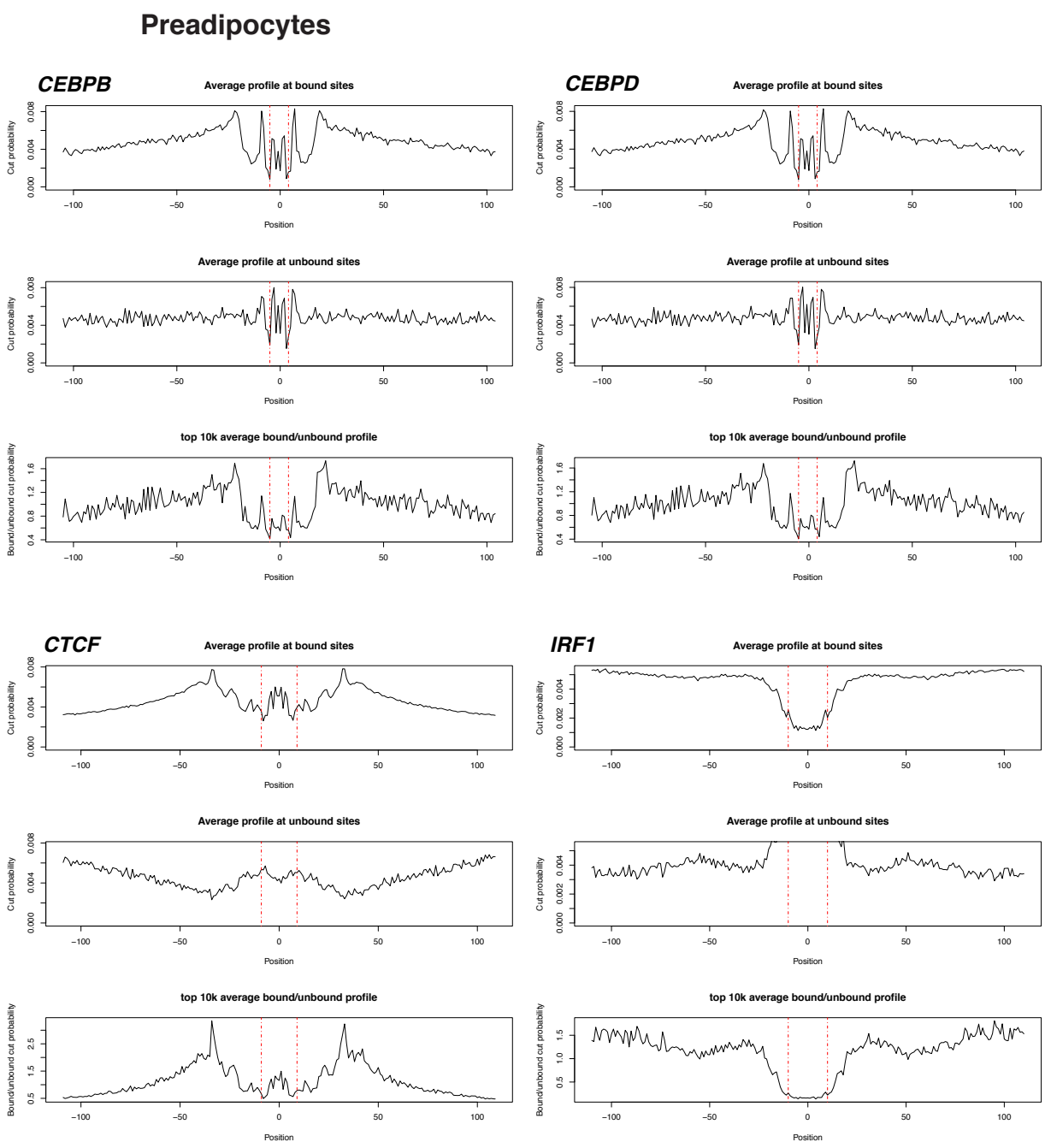

Figure 43: Aggregate footprint profiles for CEBPB, CEBPD, CTCF, and IRF1 in SGBS preadipocytes. Footprints are shown for SGBS 2. Bound, unbound, and normalized profiles are shown for each factor.

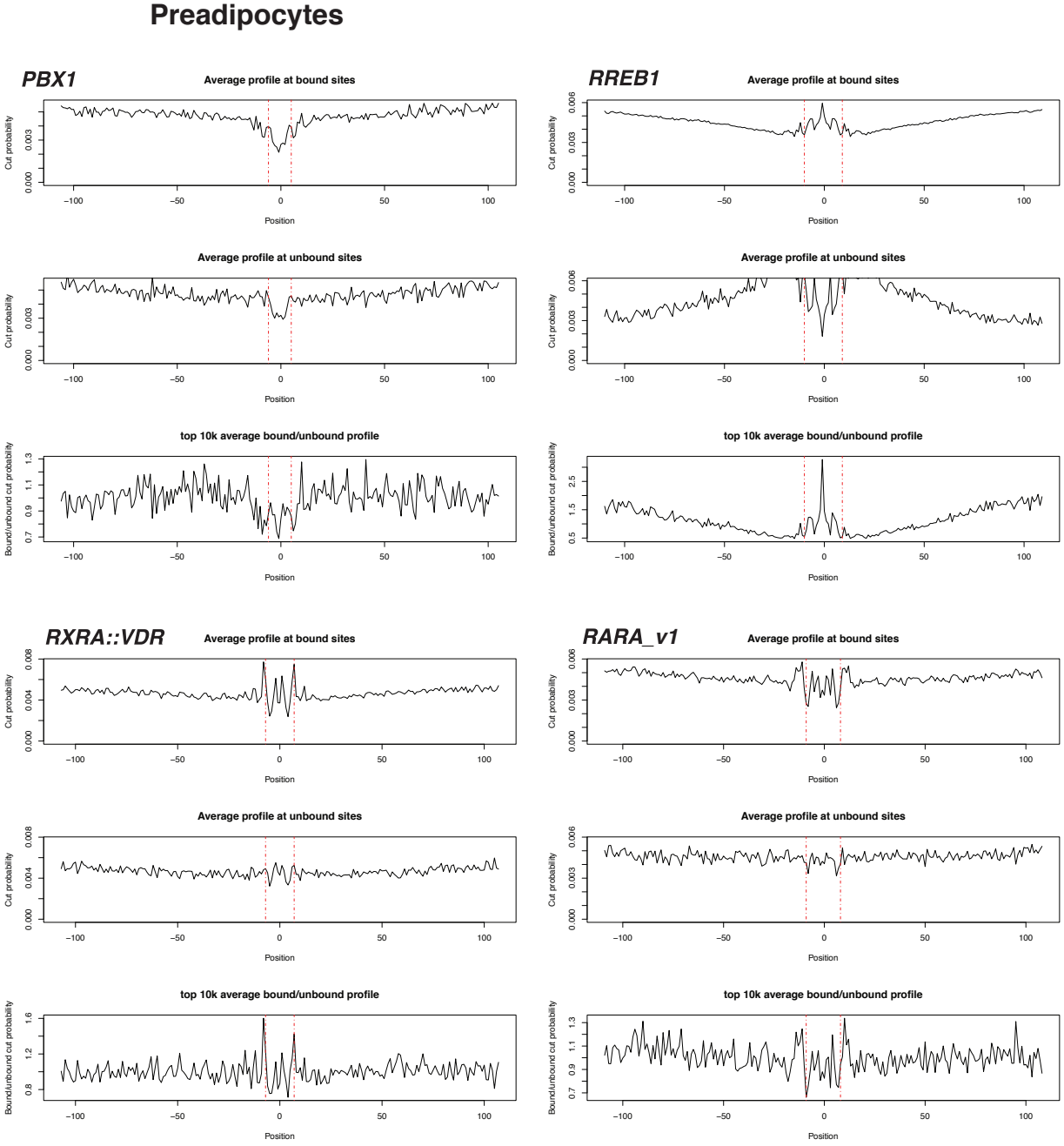

Figure 44: Aggregate footprint profiles for PBX1, RREB1, RXRA::VDR, and RARA in SGBS preadipocytes. Footprints are shown for SGBS 2. Bound, unbound, and normalized profiles are shown for each factor.

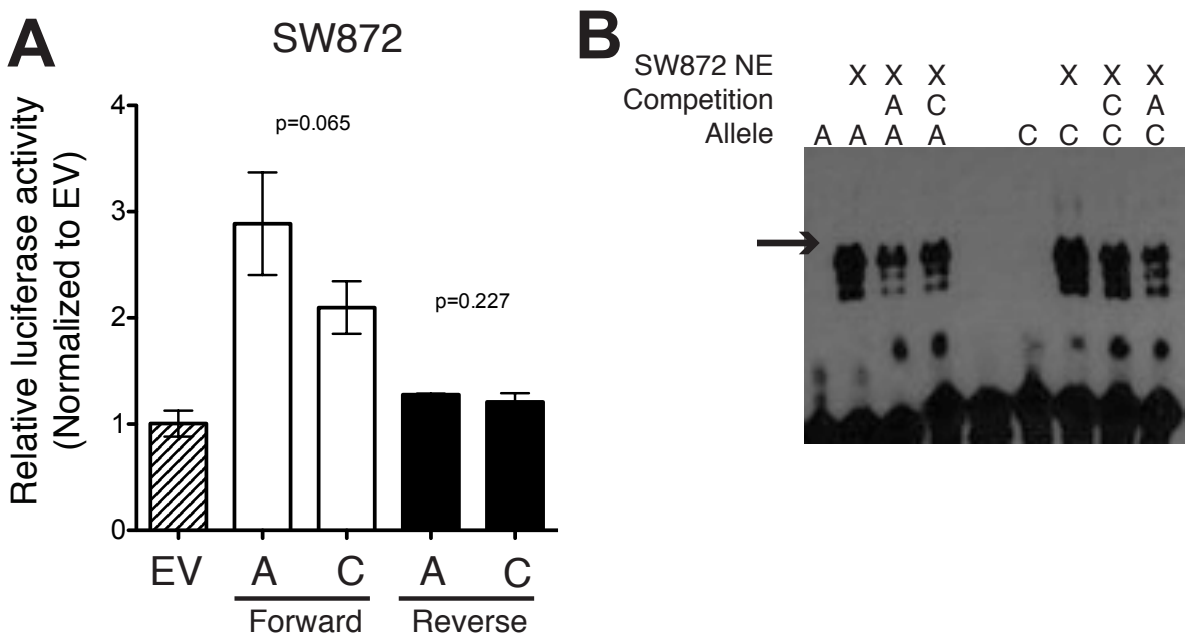

Figure 45. Additional regulatory assays for rs1534696. A) The region containing rs1534696 was tested in transcriptional reporter luciferase assays in SW872 cells. Each bar represents mean ±standard deviation of three independent clones. Transcriptional activity is greater in SW872, and the trend of higher transcriptional activity for rs1534696-A is consistent with results in SGBS (Figure 29). B) Allelic differences in protein binding with SW872 nuclear extract are not present as with SGBS nuclear extract (Figure 29). Arrow shows where the allele-specific protein complex appears with SGBS nuclear extract.

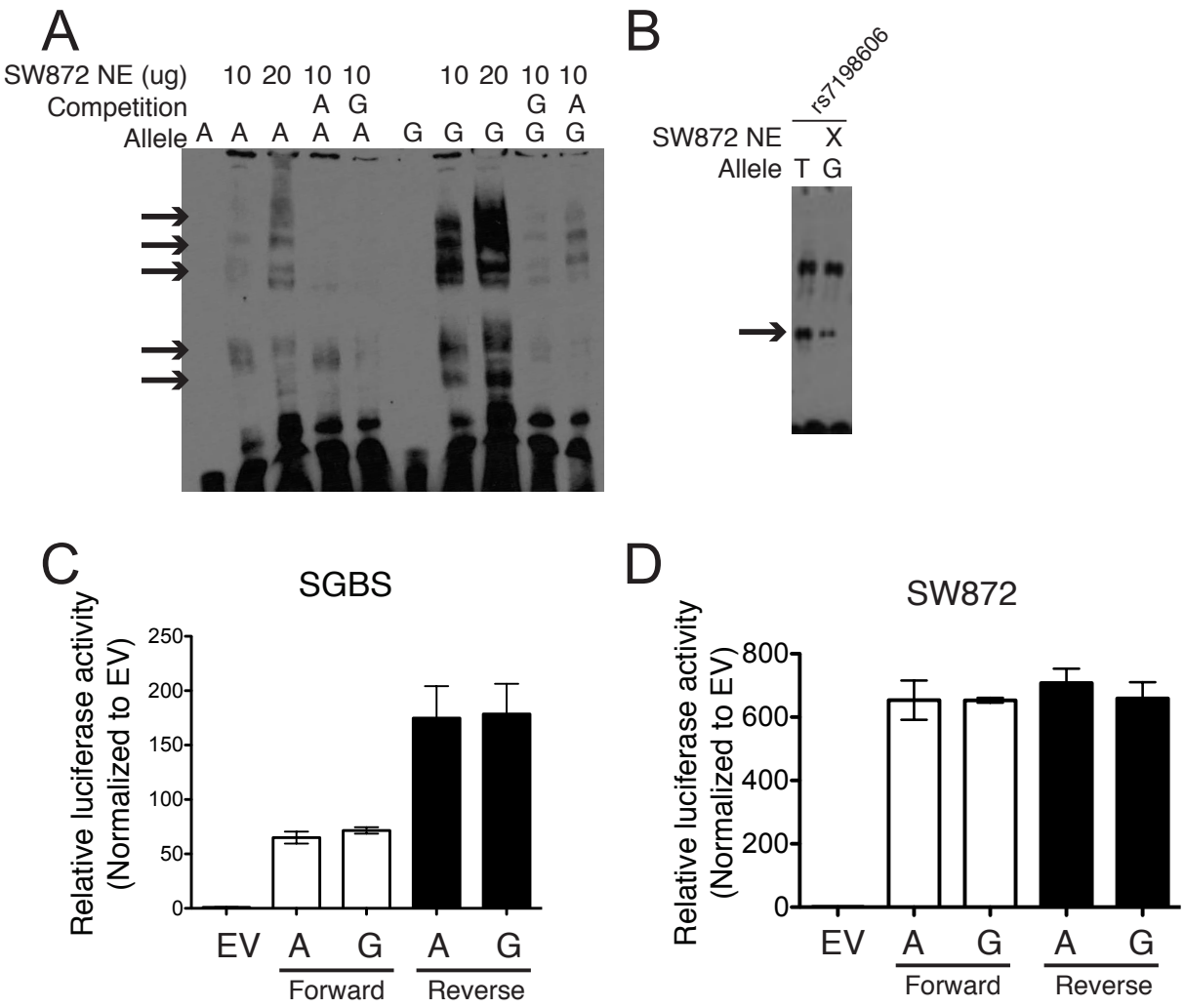

Figure 46. Additional regulatory assays for rs7187776. A) rs7187776-G shows increased protein binding using SW872 nuclear extract in EMSA. The banding pattern of protein binding is similar to protein binding to purified PU.1 (Figure 26). Arrows show allelic differences. B) One of four additional variants (rs7198606) overlapping ATAC-seq peaks showed allelic differences in protein binding (arrow) using SW872 nuclear extract in EMSA. The remaining three variants did not show consistent results across EMSA experiments. C-D) The region containing rs7187776 shows extremely strong transcriptional activity in reporter assays in SGBS (C) and SW872 (D) cells. Bars represent the mean ± standard deviation of 3-5 independent clones.

Table 14: Characteristics of METSIM individuals

Sample ID	Body Mass Index	Waist (cm)	Waist-to-hip ratio	Age (yrs)	Glucose Tolerance
Tissue 1	28.4	106	1.03	62	impaired
Tissue 2	24.0	91.5	0.94	51	normal
Tissue 3	24.9	100	1.09	55	normal

Table 15: Alignment metrics of all samples and conditions

Sample	Tissue Size (ug)/Cell count	Detergent treatment (1% NP-40)	Amount of Tn5 Transposase (ul)	Total reads	Uniquely aligned	Percent Mitochondrial Reads	Nuclear alignments	Remaining reads after blacklist filtering	Remaining reads after duplicates removed	Number of peaks
Tissue 1	36	No	2.5	160,456,871	103,884,007	4.2	99,514,873	98,503,345	87,261,270	13,073
Tissue 1[a]	36	Yes	2.5	129,550,842	87,423,735	8.5	80,036,369	79,034,093	70,607,644	58,279
Tissue 2[a]	12	Yes	2.5	131,596,088	83,656,351	12.8	72,952,105	71,836,642	60,612,852	36,612
Tissue 2	36	Yes	2.5	152,103,576	89,636,408	17.0	74,406,166	73,172,989	66,897,385	29,409
Tissue 3	36	No	2.5	156,251,262	97,752,018	10.8	87,234,180	86,077,206	76,890,324	21,016
Tissue 3[a]	36	Yes	2.5	119,284,302	70,587,237	11.9	62,214,172	61,339,160	57,119,081	49,631
SGBS 1[b]	800,000	Yes	5.0	194,938,464	168,498,157	14.2	144,425,078	126,236,487	40,598,393	196,211
SGBS 2[b]	400,000	Yes	5.0	105,391,646	93,389,720	42.7	53,460,836	42,620,701	25,646,124	173,084

[a]These samples were used in the main table and in all analyses.

[b]These samples were sequenced using paired-end reads, but processed as single-end reads

Table 16: Footprinting of 35 transcription factors

Transcription Factor	JASPAR motif ID	Aggregate footprint in adipose tissue samples	Aggregate footprint in SGBS 2	Cardiometabolic GWAS variants overlapping aggregate footprint in adipose tissue	Cardiometabolic GWAS variants overlapping aggregate footprint in SGBS 2	Number of footprints in Tissue 1	mTPR/fTPR Tissue 1	Number of footprints in Tissue 2	mTPR/fTPR Tissue 2	Number of footprints in Tissue 3	mTPR/fTPR Tissue 3	Number of footprints in SGBS 2	mTPR/fTPR SGBS 2
CEBPA	MA0102.3	1,2,3	Yes	rs174538, rs10495712	rs1553832, rs4812488	10010	0.79	7074	0.53	9343	0.87	16165	0.51
CEBPB	MA0466.2		Yes									7062	0.62
CEBPD	MA0826.1		Yes									7000	0.62
CREB3I2	MA0608.1	1,2,3	Yes			3972	0.90	2799	0.70	3889	0.94	8033	0.78
CTCF	MA0139.1		Yes		rs7131882, rs11838776, rs7251881, rs79813245, rs76226186, rs6853156, rs6830765, rs1042701							54970	0.79
EBF1	MA0154.3												
ELK3	MA0759.1												
ETS1	MA0098.3	2						5535	0.99				
FLI1	MA0475.2												
FOXO1	MA0480.1	1,2,3	Yes		rs4149272	8609	0.73	5522	0.51	7622	0.86	19405	0.75
IRF1	MA0050.2	1,3	Yes		rs12128131, rs662799, rs7104207, rs9913522, rs17242395, rs62233075, rs150111048, rs55875205, rs1401419, rs1800759	65530	0.09			48881	0.09	70371	0.14
MEF2D	MA0773.1	1,2,3	Yes		rs35581848	10321	0.35	3547	0.37	6785	0.42	14961	0.66
NFIA	MA0670.1	1,2,3	Yes			10166	0.54	7200	0.42	9085	0.72	15851	0.46
NR1H2::RXRA	MA0115.1												
NR1H3::RXRA	MA0494.1												
NR3C1	MA0113.3												
PBX1	MA0070.1	1,2	Yes			1694	0.93	1112	0.95			6849	0.85
PPARG::RXRA	MA0065.2												
PPARG	MA0066.1												
RARA::RXRA	MA0159.1												
RARA_v1	MA0729.1		Yes		rs2493410, rs4722530							11344	0.89
RARA_v2	MA0730.1												
RFX2	MA0600.2												
RREB1	MA0073.1		Yes		rs11102965, rs10750216, rs8027181, rs12051272, rs148222624,rs72838036, rs72838060, rs2795159, rs2122031, rs11064, rs1936801, rs6929846, rs1294404, rs57246313, rs1004558, rs118146489							100406	0.98
RUNX1	MA0002.2												
RXRA::VDR	MA0074.1		Yes		rs6129801, rs73069965							8268	0.98
RXRA	MA0512.1												
SMAD2::SMAD3::SMAD4	MA0513.1												
SMAD3	MA0795.1												
SPI1	MA0080.4	1,2,3	Yes	rs7187776, rs1401419	rs1401419, rs1800759, rs10790517	10833	0.63	6171	0.38	7982	0.73	20323	0.69
SPIB	MA0081.1												
STAT3	MA0144.2	1,2,3	Yes	rs11097198		16627	0.50	13479	0.36	13079	0.70	32195	0.63
STAT5A::STAT5B	MA0519.1	1,2,3	Yes	rs11097198, rs6689393, rs7178051	rs11041999, rs6776159	11542	0.70	8490	0.47	9997	0.83	27291	0.71
TCF7L2	MA0523.1	1,2,3	Yes		rs67248872, rs2243312	11090	0.51	5200	0.46	7300	0.65	16950	0.80
ZEB1	MA0103.2												

Table 17: Primer and probe sequences for functional assays

Primer sequences for luciferase assays	5'- 3' Sequence	Chromosome Position (hg19)
SNX10_F	CAGAAGGAAAGCAACATCATCA	chr7:26397159-26397408
SNX10_R	AGAAACCCAGTTCTCCTAGACC	
SH2B1_F	GGGAGCCGGGACCAGGA	chr16:28857631-28858086
SH2B1_R	GGGACTCGGGGTCTCTCT	
Probe sequences for EMSA	**5'- 3' Sequence**	**Chromosome Position (hg19)**
rs1534696_A	TATGGGCCC**A**GAAATAAAT	chr7:26396980-26396999
rs1534696_C	TATGGGCCC**C**GAAATAAAT	
rs7187776_A	CACAGAAGA**A**GAAGGGCGC	chr16:28857636-28857655
rs7187776_G	CACAGAAGA**G**GAAGGGCGC	
rs11864750_A	TGGGGGCGC**A**GGGAGCGGG	chr16:28875195-28875214
rs11864750_T	TGGGGGCGC**T**GGGAGCGGG	
rs7198606_T	TCTATGGTC**T**CTTCCTTCA	chr16:28875112-28875131
rs7198606_G	TCTATGGTC**G**CTTCCTTCA	
rs8055138_C	TGGCCTTAG**C**CCTTCCCCG	chr16:28891456-28891475
rs8055138_T	TGGCCTTAG**T**CCTTCCCCG	
rs148099387_-	TGGCAGCCCGTCAGCCTT	chr16:28,851,369-28,851,387
rs148099387_CT	TGGCAGCCC**CT**GTCAGCCTT	
PU.1_positive control*	CAGAAAAGAGGAAGTGAAACCG	N/A
PU.1_negative control	TGAGGGGAGAAGGACAGGGTTA	N/A

*PU.1 positive control sequence was constructed from the JASPAR motif

Table 18: Comparison of ATAC-seq peak overlap between tissue samples

Percent overlap of bases in peaks

Sample	Tissue 1 36 mg no det	Tissue 1 36 mg det	Tissue 2 36 mg det	Tissue 2 12 mg det	Tissue 3 36 mg no det	Tissue 3 36 mg det
Tissue 1 36 mg no det	100.0	74.3	60.4	65.6	73.5	69.7
Tissue 1 36 mg det	38.7	100.0	56.6	64.4	57.8	71.6
Tissue 2 36 mg det	43.1	77.7	100.0	77.9	61.9	77.2
Tissue 2 12 mg det	43.1	81.3	71.7	100.0	63.6	78.0
Tissue 3 36 mg no det	56.9	85.9	67.1	74.9	100.0	81.4
Tissue 3 36 mg det	42.0	82.9	65.1	71.5	63.3	100.0

Jaccard index

Sample	Tissue 1 36 mg no det	Tissue 1 36 mg det	Tissue 2 36 mg det	Tissue 2 12 mg det	Tissue 3 36 mg no det	Tissue 3 36 mg det
Tissue 1 36 mg no det	1.00	0.34	0.34	0.35	0.47	0.36
Tissue 1 36 mg det		1.00	0.49	0.56	0.53	0.63
Tissue 2 36 mg det			1.00	0.60	0.48	0.55
Tissue 2 12 mg det				1.00	0.52	0.59
Tissue 3 36 mg no det					1.00	0.56
Tissue 3 36 mg det						1.00

Table 19: Comparison of ATAC-seq peaks between adipose tissue, mature adipocytes, and SGBS preadipocytes

Percent overlap of bases in peaks

Sample	Tissue 1	Tissue 2	Tissue 3	SGBS 1	SGBS 2	ENCODE	Mature adipocytes
Tissue 1	100.0	63.8	71.8	49.9	51.8	67.8	55.5
Tissue 2	80.3	100.0	77.9	48.0	50.0	63.9	60.2
Tissue 3	80.3	69.2	100.0	47.5	49.5	64.0	56.9
SGBS 1	30.3	23.2	25.8	100.0	82.0	43.9	29.5
SGBS 2	33.6	25.8	28.7	87.6	100.0	47.9	32.4
ENCODE	43.8	32.8	37.0	46.7	47.7	100.0	35.9
Mature adipocytes	57.9	49.8	53.0	50.6	52.1	58.0	100.0

Percent overlap of top 50,000 peaks of union sets

Sample	METSIM	SGBS	Allum	ENCODE
METSIM	100.0	54.0	54.7	60.0
SGBS	30.5	100.0	31.2	41.9
Allum	49.9	50.4	100.0	48.7
ENCODE	44.3	54.7	39.4	100.0

Jaccard index

Sample	Tissue 1	Tissue 2	Tissue 3	SGBS 1	SGBS 2	ENCODE	Mature adipocytes
Tissue 1	1.00	0.55	0.61	0.23	0.26	0.36	0.40
Tissue 2		1.00	0.58	0.19	0.20	0.28	0.37
Tissue 3			1.00	0.20	0.22	0.31	0.38
SGBS 1				1.00	0.73	0.29	0.23
SGBS 2					1.00	0.31	0.25
ENCODE						1.00	0.29
Mature adipocytes							1.00

Jaccard index top 50,000 peaks of union sets

Sample	METSIM	SGBS	Allum	ENCODE
METSIM	1.00	0.24	0.35	0.34
SGBS		1.00	0.24	0.31
Allum			1.00	0.28
ENCODE				1.00

Percent overlap and jaccard index was determined using the top 25,000 peaks in each sample, ranked by *p*- value.
ENCODE= subcutaneous adipose tissue ATAC-seq ENCODE ID: ENCSR540BML
Mature adipocyte ATAC-seq was previously published (33)

Table 20: Overlap of ATAC peaks with adipose nuclei and rank among 98 chromatin states

Adipose Tissue

All peaks			Top 50,000 peaks			Top 50,000 peaks, center 200 bp		
Chromatin State	Percent covered bases in adipose nuclei	Adipose rank among all 98 tissues	Chromatin State	Percent covered bases in adipose nuclei	Adipose rank among all 98 tissues	Chromatin State	Percent covered bases in adipose nuclei	Adipose rank among all 98 tissues
Promoter+Enhancer	85.7	1	Promoter+Enhancer	88.9	1	Promoter+Enhancer	84.3	1
Promoter	44.5	4	Promoter	49.3	4	Promoter	38.9	4
Enhancer	41.3	1	Enhancer	39.7	1	Enhancer	45.4	1
Transcribed	3.6	98	Transcribed	2.9	98	Transcribed	4.1	98
Polycomb	2.1	92	Polycomb	1.7	92	Polycomb	2.4	91
Heterochromatin	0.3	91	Heterochromatin	0.2	96	Heterochromatin	0.2	94
ZNF repeat	0.3	51	ZNF repeat	0.2	55	ZNF repeat	0.2	52
Quiescent	8.0	98	Quiescent	6.2	98	Quiescent	8.8	98

SGBS

All peaks			Top 50,000 peaks			Top 50,000 peaks, center 200 bp		
Chromatin State	Percent covered bases in adipose nuclei	Adipose rank among all 98 tissues	Chromatin State	Percent covered bases in adipose nuclei	Adipose rank among all 98 tissues	Chromatin State	Percent covered bases in adipose nuclei	Adipose rank among all 98 tissues
Promoter+Enhancer	49.4	14	Promoter+Enhancer	65.9	17	Promoter+Enhancer	58.0	17
Promoter	24.4	8	Promoter	40.8	11	Promoter	29.9	12
Enhancer	25.0	20	Enhancer	25.2	25	Enhancer	28.1	20
Transcribed	13.3	68	Transcribed	9.1	77	Transcribed	11.1	71
Polycomb	8.6	51	Polycomb	5.3	47	Polycomb	6.0	43
Heterochromatin	1.1	43	Heterochromatin	0.7	49	Heterochromatin	0.9	47
ZNF repeat	0.3	30	ZNF repeat	0.3	28	ZNF repeat	0.3	25
Quiescent	27.2	79	Quiescent	18.8	80	Quiescent	23.7	79

ENCODE

All peaks			Top 50,000 peaks			Top 50,000 peaks, center 200 bp		
Chromatin State	Percent covered bases in adipose nuclei	Adipose rank among all 98 tissues	Chromatin State	Percent covered bases in adipose nuclei	Adipose rank among all 98 tissues	Chromatin State	Percent covered bases in adipose nuclei	Adipose rank among all 98 tissues
Promoter+Enhancer	73.9	4	Promoter+Enhancer	81.9	4	Promoter+Enhancer	74.1	4
Promoter	47.5	9	Promoter	57.0	9	Promoter	42.2	9
Enhancer	26.5	17	Enhancer	24.8	18	Enhancer	31.8	17
Transcribed	7.2	92	Transcribed	5.5	91	Transcribed	7.9	91
Polycomb	5.7	56	Polycomb	4.2	61	Polycomb	5.8	58
Heterochromatin	0.7	56	Heterochromatin	0.3	67	Heterochromatin	0.4	62
ZNF repeat	0.3	45	ZNF repeat	0.1	57	ZNF repeat	0.2	54
Quiescent	12.1	90	Quiescent	8.1	90	Quiescent	11.7	90

Mature adipocytes

All peaks			Top 50,000 peaks			Top 50,000 peaks, center 200 bp		
Chromatin State	Percent covered bases in adipose nuclei	Adipose rank among all 98 tissues	Chromatin State	Percent covered bases in adipose nuclei	Adipose rank among all 98 tissues	Chromatin State	Percent covered bases in adipose nuclei	Adipose rank among all 98 tissues
Promoter+Enhancer	33.5	1	Promoter+Enhancer	73.8	1	Promoter+Enhancer	73.1	1
Promoter	15.4	5	Promoter	40.0	7	Promoter	34.4	7
Enhancer	18.1	8	Enhancer	33.8	1	Enhancer	38.7	1
Transcribed	12.4	65	Transcribed	5.1	98	Transcribed	5.9	98
Polycomb	16.1	31	Polycomb	4.8	67	Polycomb	4.9	69
Heterochromatin	6.1	35	Heterochromatin	3.0	43	Heterochromatin	2.2	50
ZNF repeat	3.3	35	ZNF repeat	1.8	41	ZNF repeat	1.6	40
Quiescent	28.6	83	Quiescent	11.5	96	Quiescent	12.4	96

Percent coverage was determined by the percent of ATAC-seq peak bases overlapped by each chromatin state.
Roadmap 18-state chromatin states included in the above categories:
Promoter: 1_TssA, 2_TssFlnk, 3_TssFlnkU, 4_TssFlnkD, 14_TssBiv
Enhancer: 7_EnhG1, 8_EnhG2, 9_EnhA1, 10_EnhA2, 11_EnhWk, 15_EnhBiv
Transcribed: 5_Tx, 6_TxWk
Polycomb repressed: 16_ReprPC, 17_ReprPCWk
Heterochromatin: 13_Het
ZNF repeat: 12_ZNF/Rpts
Quiescent: 18_Quies

Table 21: Roadmap Epigenome IDs

Epigenome ID	Standardized Epigenome name
E001	ES-I3 Cells
E002	ES-WA7 Cells
E003	H1 Cells
E004	H1 BMP4 Derived Mesendoderm Cultured Cells
E005	H1 BMP4 Derived Trophoblast Cultured Cells
E006	H1 Derived Mesenchymal Stem Cells
E007	H1 Derived Neuronal Progenitor Cultured Cells
E008	H9 Cells
E009	H9 Derived Neuronal Progenitor Cultured Cells
E010	H9 Derived Neuron Cultured Cells
E011	hESC Derived CD184+ Endoderm Cultured Cells
E012	hESC Derived CD56+ Ectoderm Cultured Cells
E013	hESC Derived CD56+ Mesoderm Cultured Cells
E014	HUES48 Cells
E015	HUES6 Cells
E016	HUES64 Cells
E017	IMR90 fetal lung fibroblasts Cell Line
E018	iPS-15b Cells
E019	iPS-18 Cells
E020	iPS-20b Cells
E021	iPS DF 6.9 Cells
E022	iPS DF 19.11 Cells
E023	Mesenchymal Stem Cell Derived Adipocyte Cultured Cells
E024	ES-UCSF4 Cells
E025	Adipose Derived Mesenchymal Stem Cell Cultured Cells
E026	Bone Marrow Derived Cultured Mesenchymal Stem Cells
E027	Breast Myoepithelial Primary Cells
E028	Breast variant Human Mammary Epithelial Cells (vHMEC)
E029	Primary monocytes from peripheral blood
E030	Primary neutrophils from peripheral blood
E031	Primary B cells from cord blood
E032	Primary B cells from peripheral blood
E033	Primary T cells from cord blood
E034	Primary T cells from peripheral blood
E035	Primary hematopoietic stem cells
E036	Primary hematopoietic stem cells short term culture
E037	Primary T helper memory cells from peripheral blood 2
E038	Primary T helper naive cells from peripheral blood
E039	Primary T helper naive cells from peripheral blood
E040	Primary T helper memory cells from peripheral blood 1
E041	Primary T helper cells PMA-I stimulated
E042	Primary T helper 17 cells PMA-I stimulated
E043	Primary T helper cells from peripheral blood
E044	Primary T regulatory cells from peripheral blood
E045	Primary T cells effector/memory enriched from peripheral blood
E046	Primary Natural Killer cells from peripheral blood
E047	Primary T CD8+ naive cells from peripheral blood
E048	Primary T CD8+ memory cells from peripheral blood
E049	Mesenchymal Stem Cell Derived Chondrocyte Cultured Cells
E050	Primary hematopoietic stem cells G-CSF-mobilized Female
E051	Primary hematopoietic stem cells G-CSF-mobilized Male
E052	Muscle Satellite Cultured Cells
E053	Cortex derived primary cultured neurospheres
E054	Ganglion Eminence derived primary cultured neurospheres
E055	Foreskin Fibroblast Primary Cells skin01
E056	Foreskin Fibroblast Primary Cells skin02
E057	Foreskin Keratinocyte Primary Cells skin02
E058	Foreskin Keratinocyte Primary Cells skin03
E059	Foreskin Melanocyte Primary Cells skin01
E061	Foreskin Melanocyte Primary Cells skin03
E062	Primary mononuclear cells from peripheral blood

E063	Adipose Nuclei
E065	Aorta
E066	Liver
E067	Brain Angular Gyrus
E068	Brain Anterior Caudate
E069	Brain Cingulate Gyrus
E070	Brain Germinal Matrix
E071	Brain Hippocampus Middle
E072	Brain Inferior Temporal Lobe
E073	Brain_Dorsolateral_Prefrontal_Cortex
E074	Brain Substantia Nigra
E075	Colonic Mucosa
E076	Colon Smooth Muscle
E077	Duodenum Mucosa
E078	Duodenum Smooth Muscle
E079	Esophagus
E080	Fetal Adrenal Gland
E081	Fetal Brain Male
E082	Fetal Brain Female
E083	Fetal Heart
E084	Fetal Intestine Large
E085	Fetal Intestine Small
E086	Fetal Kidney
E087	Pancreatic Islets
E088	Fetal Lung
E089	Fetal Muscle Trunk
E090	Fetal Muscle Leg
E091	Placenta
E092	Fetal Stomach
E093	Fetal Thymus
E094	Gastric
E095	Left Ventricle
E096	Lung
E097	Ovary
E098	Pancreas
E099	Placenta Amnion
E100	Psoas Muscle
E101	Rectal Mucosa Donor 29
E102	Rectal Mucosa Donor 31
E103	Rectal Smooth Muscle
E104	Right Atrium
E105	Right Ventricle
E106	Sigmoid Colon
E107	Skeletal Muscle Male
E108	Skeletal Muscle Female
E109	Small Intestine
E110	Stomach Mucosa
E111	Stomach Smooth Muscle
E112	Thymus
E113	Spleen
E114	A549 EtOH 0.02pct Lung Carcinoma Cell Line
E115	Dnd41 TCell Leukemia Cell Line
E116	GM12878 Lymphoblastoid Cells
E117	HeLa-S3 Cervical Carcinoma Cell Line
E118	HepG2 Hepatocellular Carcinoma Cell Line
E119	HMEC Mammary Epithelial Primary Cells
E120	HSMM Skeletal Muscle Myoblasts Cells
E121	HSMM cell derived Skeletal Muscle Myotubes Cells
E122	HUVEC Umbilical Vein Endothelial Primary Cells
E123	K562 Leukemia Cells
E124	Monocytes-CD14+ RO01746 Primary Cells
E125	NH-A Astrocytes Primary Cells
E126	NHDF-Ad Adult Dermal Fibroblast Primary Cells
E127	NHEK-Epidermal Keratinocyte Primary Cells
E128	NHLF Lung Fibroblast Primary Cells
E129	Osteoblast Primary Cells

Table 22: Transcription factor motifs enriched in adipose tissue- and SGBS-specific ATAC peaks

SGBS-specific

Motif Name	Consensus Sequence	P-value	Log P-value	q-value (Benjamini)	Target sequences with motif (of 10000)	Percent target sequences with motif	Background sequences with motif (of 44753)	Percent background sequences with motif
Atf3	DATGASTCATHN	1.00E-923	-2127.0	0	4993	49.9%	8505	20.5%
Fra1	NNATGASTCATH	1.00E-904	-2083.0	0	4671	46.7%	7624	18.4%
BATF	DATGASTCAT	1.00E-887	-2043.0	0	4887	48.9%	8372	20.2%
AP-1	VTGACTCATC	1.00E-881	-2030.0	0	5074	50.7%	8989	21.7%
Fosl2	NATGASTCABNN	1.00E-603	-1389.0	0	3276	32.8%	5179	12.5%
Jun-AP1	GATGASTCATCN	1.00E-455	-1050.0	0	2610	26.1%	4146	10.0%
Bach2	TGCTGAGTCA	1.00E-206	-475.0	0	1580	15.8%	2834	6.8%
MafK	GCTGASTCAGCA	1.00E-95	-220.4	0	932	9.3%	1836	4.4%
TEAD4	CCWGGAATGY	1.00E-92	-213.7	0	2266	22.7%	6173	14.9%
TEAD	YCWGGAATGY	1.00E-87	-202.2	0	2053	20.5%	5513	13.3%
TEAD2	CCWGGAATGY	1.00E-62	-144.5	0	1488	14.9%	3968	9.6%
RUNX2	NWAACCACADNN	1.00E-54	-125.6	0	1590	15.9%	4455	10.8%
RUNX	SAAACCACAG	1.00E-54	-124.5	0	1416	14.2%	3863	9.3%
NF-E2	GATGACTCAGCA	1.00E-53	-122.6	0	496	5.0%	951	2.3%
RUNX1	AAACCACARM	1.00E-52	-121.8	0	1847	18.5%	5390	13.0%
RUNX-AML	GCTGTGGTTW	1.00E-48	-111.0	0	1402	14.0%	3914	9.5%
MafA	TGCTGACTCA	1.00E-43	-101.0	0	1581	15.8%	4626	11.2%
Nrf2	HTGCTGAGTCAT	1.00E-43	-100.6	0	412	4.1%	795	1.9%
Bach1	AWWNTGCTGAGTCAT	1.00E-42	-97.0	0	478	4.8%	996	2.4%
Pdx1	YCATYAATCA	1.00E-31	-71.8	0	1663	16.6%	5210	12.6%
GATA3	AGATAASR	1.00E-28	-65.7	0	2162	21.6%	7149	17.3%
Gata1	SAGATAAGRV	1.00E-22	-52.1	0	906	9.1%	2685	6.5%
Gata2	BBCTTATCTS	1.00E-22	-51.8	0	998	10.0%	3016	7.3%
Gata4	NBWGATAAGR	1.00E-19	-45.6	0	1428	14.3%	4662	11.3%
Pit1	ATGCATAATTCA	1.00E-10	-23.3	0	514	5.1%	1593	3.9%
Arnt:Ahr	TBGCACGCAA	1.00E-09	-21.7	0	612	6.1%	1968	4.8%
PAX3:FKHR	ACCRTGACTAATTNN	1.00E-08	-20.6	0	337	3.4%	994	2.4%
Hoxb4	TGATTRATGGCY	1.00E-05	-12.6	0	319	3.2%	1018	2.5%
NFAT:AP1	SARTGGAAAAWRTGAGTCAB	1.00E-05	-12.2	0.0001	386	3.9%	1269	3.1%
Pit1	ATGMATATDC	1.00E-05	-12.1	0.0001	1302	13.0%	4797	11.6%
GATA,IR3	NNNNNBAGATAWYATCTVHN	1.00E-05	-12.1	0.0001	309	3.1%	989	2.4%
PAX5	GTCACGCTCSCTGM	1.00E-04	-10.2	0.0004	130	1.3%	373	0.9%
OCT:OCT	ATGAATWATTCATGA	1.00E-04	-9.7	0.0006	36	0.4%	73	0.2%
Brn2	ATGAATATTC	1.00E-04	-9.7	0.0006	167	1.7%	506	1.2%
PAX5	GCAGCCAAGCRTGACH	1.00E-04	-9.5	0.0007	364	3.6%	1230	3.0%

Adipose tissue-specific

Motif Name	Consensus Sequence	P-value	Log P-value	q-value (Benjamini)	Target sequences with motif (of 10000)	Percent target sequences with motif	Background sequences with motif (of 44753)	Percent background sequences with motif
ARE	RGRACASNSTGTYCYB	1.00E-136	-313.2	0	911	9.1%	1616	3.6%
GRE,IR3	VAGRACAKWCTGTYC	1.00E-133	-306.4	0	888	8.9%	1571	3.5%
GRE,IR3	NRGVACABNVTGTYCY	1.00E-114	-264.0	0	617	6.2%	955	2.1%
PGR	AAGAACATWHTGTTC	1.00E-98	-227.5	0	834	8.3%	1660	3.7%
PPARE	TGACCTTTGCCCCA	1.00E-92	-214.0	0	1833	18.3%	5070	11.3%
NF1	YTGCCAAG	1.00E-76	-176.3	0	3888	38.9%	13491	30.2%
RXR,DR1	TAGGGCAAAGGTCA	1.00E-75	-174.2	0	1996	20.0%	5952	13.3%
PR	VAGRACAKNCTGTBC	1.00E-69	-159.9	0	3474	34.7%	11946	26.7%
AR	CCAGGAACAG	1.00E-58	-134.8	0	5181	51.8%	19569	43.7%
CEBPB	ATTGCGCAAC	1.00E-33	-76.2	0	2110	21.1%	7367	16.5%
Esrrb	KTGACCTTGA	1.00E-23	-55.1	0	907	9.1%	2876	6.4%
ELF5	ACVAGGAAGT	1.00E-21	-50.6	0	1397	14.0%	4843	10.8%
SpiB	AAAGRGGAAGTG	1.00E-21	-49.4	0	548	5.5%	1592	3.6%
PU.1	AGAGGAAGTG	1.00E-21	-48.8	0	1071	10.7%	3577	8.0%
Erra	CAAAGGTCAG	1.00E-19	-45.1	0	3036	30.4%	11750	26.3%
HNF4a,DR1	CARRGKBCAAAGTYCA	1.00E-19	-44.0	0	760	7.6%	2430	5.4%
NPAS2	KCCACGTGAC	1.00E-18	-42.6	0	1381	13.8%	4893	10.9%
TR4,DR1	GAGGTCAAAGGTCA	1.00E-17	-39.2	0	211	2.1%	496	1.1%
ETV1	AACGGAAGT	1.00E-15	-36.3	0	2317	23.2%	8881	19.9%
Etv2	NNAYTTCCTGHN	1.00E-14	-34.3	0	1833	18.3%	6892	15.4%
BMAL1	GNCACGTG	1.00E-14	-33.1	0	2370	23.7%	9175	20.5%
Nur77	TGACCTTTNCNT	1.00E-14	-33.0	0	315	3.2%	885	2.0%
THRa	GGTCANYTGAGGWCA	1.00E-14	-32.8	0	633	6.3%	2067	4.6%
EHF	AVCAGGAAGT	1.00E-12	-29.5	0	2175	21.8%	8430	18.8%
USF1	SGTCACGTGR	1.00E-12	-28.9	0	705	7.1%	2394	5.4%
ETS1	ACAGGAAGTG	1.00E-12	-28.0	0	1906	19.1%	7329	16.4%
Max	RCCACGTGGYYN	1.00E-11	-27.1	0	930	9.3%	3315	7.4%
GABPA	RACCGGAAGT	1.00E-11	-26.9	0	1574	15.7%	5963	13.3%
Fli1	NRYTTCCGGH	1.00E-11	-26.6	0	1886	18.9%	7278	16.3%
Usf2	GTCACGTGGT	1.00E-11	-25.5	0	571	5.7%	1915	4.3%
ERG	ACAGGAAGTG	1.00E-10	-24.6	0	2701	27.0%	10803	24.2%
Nr5a2	BTCAAGGTCA	1.00E-09	-22.6	0	893	8.9%	3240	7.2%
MITF	RTCATGTGAC	1.00E-09	-22.5	0	1407	14.1%	5359	12.0%
EWS:ERG	ATTTCCTGTN	1.00E-09	-22.0	0	1491	14.9%	5722	12.8%
CLOCK	GHCACGTG	1.00E-09	-21.4	0	733	7.3%	2615	5.8%
EBF1	GTCCCCWGGGGA	1.00E-09	-21.2	0	1754	17.5%	6849	15.3%
VDR,DR3	ARAGGTCANWGAGTTCANNN	1.00E-09	-21.0	0	253	2.5%	757	1.7%
Lhx2	TAATTAGN	1.00E-08	-19.1	0	1121	11.2%	4244	9.5%
Sox2	BCCATTGTTC	1.00E-07	-17.9	0	1045	10.5%	3955	8.8%
c-Myc	VVCCACGTGG	1.00E-07	-17.6	0	648	6.5%	2335	5.2%
Ets1	MACAGGAAGT	1.00E-07	-17.1	0	715	7.2%	2615	5.8%
CEBP:AP1	DRTGTTGCAA	1.00E-07	-17.0	0	1719	17.2%	6813	15.2%
n-Myc	VRCCACGTGG	1.00E-07	-16.8	0	843	8.4%	3144	7.0%
Nr5a2	BTCAAGGTCA	1.00E-06	-14.6	0	662	6.6%	2447	5.5%
SPDEF	ASWTCCTGBT	1.00E-06	-14.4	0	1597	16.0%	6370	14.2%
LXRE,DR4	RGGTTACTANAGGTCA	1.00E-05	-13.5	0	101	1.0%	272	0.6%
ELF1	AVCCGGAAGT	1.00E-05	-13.1	0	738	7.4%	2789	6.2%
bHLHE40	KCACGTGMCN	1.00E-05	-12.9	0	324	3.2%	1116	2.5%
EWS:FLI1	VACAGGAAAT	1.00E-05	-12.7	0	1043	10.4%	4073	9.1%
ETS:E-box	AGGAARCAGCTG	1.00E-05	-12.4	0	196	2.0%	627	1.4%
EBF	DGTCCCYRGGGA	1.00E-05	-12.0	0	503	5.0%	1848	4.1%
NF1	CYTGGCABNSTGCCAR	1.00E-05	-11.7	0	1039	10.4%	4080	9.1%
ERE	VAGGTCACNSTGACC	1.00E-05	-11.6	0.0001	308	3.1%	1071	2.4%
Atoh1	VNRVCAGCTGGY	1.00E-04	-10.5	0.0002	1768	17.7%	7238	16.2%
c-Myc	VCCACGTG	1.00E-04	-10.0	0.0003	422	4.2%	1556	3.5%
Foxo1	CTGTTTAC	1.00E-04	-9.9	0.0003	2662	26.6%	11149	24.9%
RORgt	AAYTAGGTCA	1.00E-04	-9.8	0.0003	221	2.2%	755	1.7%

Table 23: Genes with adipose tissue- or SGBS-selective ATAC peaks at TSSs

We used the top 50,000 peaks ranked by p-value for the union peak sets.
The isoform of *ADIPOQ-AS1* with a SGBS-specific peak at the TSS is not present in GENCODE version 24 and was therefore not detected in this analysis, but is clearly specific to SGBS (Figure 2).

Adipose Tissue								SGBS							
1kb				5kb				1kb				5kb			
TSS position	Transcript name(s)	Gene name(s)	Tissue samples	TSS position	Transcript name(s)	Gene name(s)	Tissue samples	TSS position	Transcript name(s)	Gene name(s)	SGBS samples	TSS position	Transcript name(s)	Gene name(s)	SGBS samples
10:119235683	ENST00000412075.6	EMX2OS	Tissue 1, Tissue 3	10:60754385	ENST00000432535.1 ENST00000260315.7, ENST00000526056.5,	LINC00844	Tissue 1, Tissue 2, Tissue 3	10:115468103	ENST00000452490.3	CASP7	SGBS 1, SGBS2	10:115464103	ENST00000452490.3	CASP7	SGBS 1, SGBS2
10:5237797	ENST00000263126.2	AKR1C4	Tissue 1, Tissue 3	11:104888862	ENST00000531367.5	CASP5	Tissue 1	10:127660694	ENST00000445458.1	FANK1-AS1	SGBS 1, SGBS2	10:127656694	ENST00000445458.1	FANK1-AS1	SGBS 1, SGBS2
10:60758385	ENST00000432535.1 ENST00000260315.7,EN ST00000526056.5,ENST	LINC00844	Tissue 1, Tissue 2, Tissue 3	11:104888893	ENST00000444749.6 ENST00000393141.6,	CASP5	Tissue 1	10:128109447	ENST00000456514.1	LINC00601	SGBS 1, SGBS2	10:128105447	ENST00000456514.1	LINC00601	SGBS 1, SGBS2
11:104892862	00000531367.5	CASP5	Tissue 1	11:104888894	ENST00000418434.5	CASP5	Tissue 1	10:35367607	ENST00000584484.1	MIR3611	SGBS 1, SGBS2	10:35363607	ENST00000584484.1	MIR3611	SGBS 1, SGBS2
11:104892893	ENST00000444749.6 ENST00000393141.6,EN	CASP5	Tissue 1	11:1088317	ENST00000333592.6	MUC2	Tissue 2, Tissue 3	10:42969990	ENST00000429940.6	LINC00839	SGBS 1, SGBS2	10:42965990	ENST00000429940.6	LINC00839	SGBS 1, SGBS2
11:104892894	ST00000418434.5	CASP5	Tissue 1	11:111410775	ENST00000271324.5	CD53	Tissue 1, Tissue 3	10:5637080	ENST00000478294.1	RP13-463N16.6	SGBS 1, SGBS2	10:5633080	ENST00000478294.1	RP13-463N16.6	SGBS 1, SGBS2
11:1092317	ENST00000333592.6	MUC2	Tissue 2, Tissue 3	11:112027284	ENST00000414219.5	TMIGD3	Tissue 1, Tissue 3	10:6820141	ENST00000628989.1	LINC00706	SGBS 1, SGBS2	10:6816141	ENST00000628989.1	LINC00706	SGBS 1, SGBS2
11:10967592	ENST00000597455.2	LAMTOR5-AS1	Tissue 1, Tissue 2, Tissue 3	11:131235372	ENST00000374791.7	NTM	Tissue 1, Tissue 2, Tissue 3	10:70991245	ENST00000450995.1	RP11-227H15.4	SGBS 1, SGBS2	10:70987245	ENST00000450995.1	RP11-227H15.4	SGBS 1, SGBS2
11:110000728	ENST00000528673.5	ZC3H12C	Tissue 1, Tissue 2, Tissue 3	11:1869199	ENST00000311604.3	LSP1	Tissue 1	10:81141083	ENST00000438554.2	RP11-342M3.5	SGBS 1, SGBS2	10:95863231	ENST00000433038.1	PLCE1-AS2	SGBS 1, SGBS2
11:111414775	ENST00000271324.5	CD53	Tissue 1	11:118722845	ENST00000336338.9	SPAG17	Tissue 1, Tissue 3	10:95867231	ENST00000433038.1	PLCE1-AS2	SGBS 1, SGBS2	10:95863292	ENST00000613585.4	PLCE1-AS2	SGBS 1, SGBS2
11:112031284	ENST00000414219.5	TMIGD3	Tissue 1, Tissue 3	11:45371921	ENST00000524410.1	RP11-430H10.1	Tissue 1, Tissue 2, Tissue 3	10:95867292	ENST00000613585.4	PLCE1-AS2	SGBS 1, SGBS2	10:95863505	ENST00000453183.5	PLCE1-AS2	SGBS 1, SGBS2
11:130734455	ENST00000538312.6	RP11-890B15.2	Tissue 1, Tissue 2, Tissue 3	11:45372051	ENST00000524565.1	RP11-430H10.1	Tissue 1, Tissue 2, Tissue 3	10:95867505	ENST00000453183.5	PLCE1-AS2	SGBS 1, SGBS2	10:95863536	ENST00000438899.5	PLCE1-AS2	SGBS 1, SGBS2
11:131239372	ENST00000374791.7 ENST00000339772.9,EN	NTM	Tissue 1, Tissue 2, Tissue 3	11:4659649	ENST00000396952.5	ORS1E1	Tissue 1, Tissue 2, Tissue 3	10:95867536	ENST00000438899.5	PLCE1-AS2	SGBS 1, SGBS2	10:107724286	ENST00000375682.8	SLC35F2	SGBS 1, SGBS2
11:134200767	ST00000535456.6	GLB1L2	Tissue 1, Tissue 2, Tissue 3	11:49224923	ENST00000533034.1	FOLH1	Tissue 2	11:10714098	ENST00000558540.5	MRVI1	SGBS 1, SGBS2	11:107724913	ENST00000525815.5	SLC35F2	SGBS 1, SGBS2
11:18286720	ENST00000356524.8	SAA1	Tissue 2	11:49225083	ENST00000356696.7	FOLH1	Tissue 2	11:10714122	ENST00000527509.6	MRVI1	SGBS 1, SGBS2	11:118281691	ENST00000602598.1	RP11-770J1.7	SGBS 1, SGBS2
11:118726845	ENST00000336338.9	SPAG17	Tissue 1, Tissue 3	11:52590578	ENST00000335674.1	LCE3A	Tissue 1, Tissue 2, Tissue 3	11:10714168	ENST00000541483.5	MRVI1	SGBS 1, SGBS2	11:126857347	ENST00000531585.1	RP11-688I9.2	SGBS 1, SGBS2
11:18727083	ENST00000527285.1	RP11-1081L13.4	Tissue 1, Tissue 3	11:58796106	ENST00000368141.4	MNDA	Tissue 1	11:10714298	ENST00000423302.6	MRVI1	SGBS 1, SGBS2	11:128045321	ENST00000609911.1	RP11-702B10.2	SGBS 1, SGBS2
11:18731199	ENST00000311604.3	LSP1	Tissue 1	11:67214885	ENST00000438189.6	CABP4	Tissue 1, Tissue 3	11:10714534	ENST00000547195.5	MRVI1	SGBS 1, SGBS2	11:17447152	ENST00000610171.1	RP4-753F5.1	SGBS 1, SGBS2
11:3858096	ENST00000396979.1	RHOG	Tissue 1	11:93966315	ENST00000506309.6	RP11-680H20.2	Tissue 1, Tissue 2, Tissue 3	11:107728286	ENST00000375682.8	SLC35F2	SGBS 1, SGBS2	11:17447678	ENST00000393203.2	PTGFRN	SGBS 1, SGBS2
11:45375921	ENST00000524410.1	RP11-430H10.1	Tissue 1, Tissue 2, Tissue 3	11:130549869	ENST00000535487.1	RP11-474D1.2	Tissue 1, Tissue 2, Tissue 3	11:107728913	ENST00000525815.5	SLC35F2	SGBS 1, SGBS2	11:1766820	ENST00000340134.4	IFITM10	SGBS 1, SGBS2
11:45376051	ENST00000524565.1	RP11-430H10.1	Tissue 1, Tissue 2, Tissue 3	12:15860926	ENST00000543612.5	EPS8	Tissue 1, Tissue 2, Tissue 3	11:118285691	ENST00000602598.1	RP11-770J1.7	SGBS 1, SGBS2	11:19397030	ENST00000528204.1	NAV2-IT1	SGBS 1, SGBS2
11:4663649	ENST00000396952.5	ORS1E1	Tissue 2, Tissue 3	12:41081243	ENST00000547702.5	CNTN1	Tissue 1	11:126861347	ENST00000531585.1	RP11-688I9.2	SGBS 1, SGBS2	11:33769944	ENST00000532057.5	FBXO3	SGBS 1, SGBS2
11:49228923	ENST00000533034.1	FOLH1	Tissue 2	12:48095492	ENST00000366478.2	OR2L13	Tissue 1, Tissue 3	11:128049321	ENST00000609911.1	RP11-702B10.2	SGBS 1, SGBS2	11:33769945	ENST00000531080.5	FBXO3	SGBS 1, SGBS2
11:49229083	ENST00000356696.7	FOLH1	Tissue 2	12:6228840	ENST00000261405.9	VWF	Tissue 1, Tissue 2, Tissue 3	11:17451152	ENST00000610171.1	RP4-753F5.1	SGBS 1, SGBS2	11:4714071	ENST00000396950.3	ORS1E2	SGBS 1, SGBS2
11:52594578	ENST00000335674.1	LCE3A	Tissue 1, Tissue 2, Tissue 3	12:69737120	ENST00000261267.6	LYZ	Tissue 1, Tissue 3	11:174516778	ENST00000393203.2	PTGFRN	SGBS 1, SGBS2	11:147747854	ENST00000428035.1	RP11-495P10.6	SGBS 1, SGBS2
1:158800106	ENST00000368141.4	MNDA	Tissue 1	12:69737163	ENST00000549690.1	LYZ	Tissue 1, Tissue 3	11:1770820	ENST00000340134.4	IFITM10	SGBS 1, SGBS2	11:148243103	ENST00000427732.1	RP11-89F3.2	SGBS 1, SGBS2
11:61015804	ENST00000541528.1	PGAS	Tissue 1, Tissue 2, Tissue 3	12:69737165	ENST00000548839.1	LYZ	Tissue 1, Tissue 3	11:19401030	ENST00000528204.1	NAV2-IT1	SGBS 1, SGBS2	11:14898048	ENST00000380378.1	ORS1T1	SGBS 1, SGBS2
1:161038734	ENST00000368013.7	ARHGAP30	Tissue 1, Tissue 3	12:7843371	ENST00000329913.3	GDF3	Tissue 1, Tissue 2, Tissue 3	11:33773944	ENST00000532057.5	FBXO3	SGBS 1, SGBS2	11:14898129	ENST00000322049.1	ORS1T1	SGBS 1, SGBS2
1:161038759	ENST00000368015.1 ENST00000271450.10,E	ARHGAP30	Tissue 1, Tissue 3	12:9263752	ENST00000318602.11	A2M	Tissue 1, Tissue 2, Tissue 3	11:33773945	ENST00000531080.5	FBXO3	SGBS 1, SGBS2	1:153343124	ENST00000368737.4	S100A12	SGBS 1, SGBS2
1:161474219	NST00000367972.8 ENST00000544868.2,EN ST00000617791.1,ENST	FCGR2A	Tissue 1, Tissue 3	12:94671619	ENST00000547927.1	RP11-1105G2.3	Tissue 1, Tissue 3	11:35235531	ENST00000528869.1	RP1-68D18.4	SGBS 1, SGBS2	1:159919009	ENST00000368092.7	SLAMF9	SGBS 1, SGBS2
11:65265512	00000620902.1	MALAT1	Tissue 1	13:23750090	ENST00000218867.3	SGCG	Tissue 1, Tissue 2, Tissue 3	1:145074881	ENST00000585156.5	PDE4DIP	SGBS 1, SGBS2	1:159919043	ENST00000368093.3	SLAMF9	SGBS 1, SGBS2
11:67218885	ENST00000438189.6	CABP4	Tissue 1, Tissue 3	13:24823576	ENST00000430733.1	SPATA13-AS1	Tissue 1	1:145074888	ENST00000620605.4	PDE4DIP	SGBS 1, SGBS2	11:60686947	ENST00000453848.6	TMEM132A	SGBS 1, SGBS2
1:177319721	ENST00000451341.1	RP1-35C21.1	Tissue 1, Tissue 2, Tissue 3	14:47807320	ENST00000357362.7	MDGA2	Tissue 1, Tissue 2, Tissue 3	1:145074900	ENST00000524974.5	PDE4DIP	SGBS 1, SGBS2	11:60686952	ENST00000005286.8 ENST00000236938.10 ,ENST00000350710.3, ENST00000367949.6, ENST00000367959.6, ENST00000540521.5,	TMEM132A	SGBS 1, SGBS2
11:77773906	ENST00000281030.2	THRSP	Tissue 1, Tissue 2, Tissue 3	14:65932819	ENST00000385047.1 ENST00000271153.8,	MIR625	Tissue 2	1:145075185	ENST00000530472.5	PDE4DIP	SGBS 1, SGBS2	1:161671761	ENST00000546024.5	FCRLA	SGBS 1, SGBS2
1:192126586	ENST00000367460.3	RGS18	Tissue 1	1:47259717	ENST00000371923.8 ENST00000371919.8,	CYP4B1	Tissue 1, Tissue 2, Tissue 3	11:4718071	ENST00000396950.3	OR51E2	SGBS 1, SGBS2	11:66098872	ENST00000627248.1	RIN1	SGBS 1, SGBS2
11:93270080	ENST00000527149.1	SMCO4	Tissue 1, Tissue 2, Tissue 3	1:47259753	ENST00000614163.4	CYP4B1	Tissue 1, Tissue 2, Tissue 3	1:1474729	ENST00000624426.1	TMEM240	SGBS 1, SGBS2	11:66098999	ENST00000311320.8	RIN1	SGBS 1, SGBS2
11:93970315	ENST00000506309.6	RP11-680H20.2	Tissue 1, Tissue 2, Tissue 3	1:49237589	ENST00000371833.3	BEND5	Tissue 1, Tissue 3	1:1474736	ENST00000378733.8	TMEM240	SGBS 1, SGBS2	11:66099310	ENST00000530056.1	RIN1	SGBS 1, SGBS2
1:198607169	ENST00000442510.6	PTPRC	Tissue 1	14:94979180	ENST00000341228.2	SERPINA12	Tissue 3	1:1474832	ENST00000425828.1	TMEM240	SGBS 1, SGBS2	1:17511277	ENST00000412427.1	RP11-380I14.1	SGBS 1, SGBS2
1:198607224	ENST00000348564.10 ENST00000367364.5,EN	PTPRC	Tissue 1	15:67346590	ENST00000558071.2	RP11-798K3.2	Tissue 1, Tissue 2, Tissue 3	1:147751854	ENST00000428035.1	RP11-495P10.6	SGBS 1, SGBS2	1:17511283	ENST00000634972.1	RP11-380I14.1	SGBS 1, SGBS2
1:198607226	ST00000413409.6	PTPRC	Tissue 1	15:69217972	ENST00000455873.7 ENST00000331849.8,	NOXS	Tissue 1	1:148247103	ENST00000427732.1	RP11-89F3.2	SGBS 1, SGBS2	11:75922068	ENST00000531538.1 ENST00000315938.4,	RP11-619A14.3	SGBS 1, SGBS2
1:209928376	ENST00000400959.7	TRAF3IP3	Tissue 1	16:20415855	ENST00000575584.5	ACSM5	Tissue 1, Tissue 2, Tissue 3	11:4902048	ENST00000380378.1	OR51T1	SGBS 1, SGBS2	11:76993462	ENST00000376217.6 ENST00000367019.5,	GDPD4	SGBS 1, SGBS2
1:209928393	ENST00000367025.7	TRAF3IP3	Tissue 1	16:20415867	ENST00000575070.1	ACSM5	Tissue 1, Tissue 2, Tissue 3	11:4902129	ENST00000322049.1	OR51T1	SGBS 1, SGBS2	1:210106518	ENST00000537238.5 ENST00000472886.5,	SYT14	SGBS 1, SGBS2
1:209928523	ENST00000367026.7	TRAF3IP3	Tissue 1	16:29752326	ENST00000329410.3	C16orf54	Tissue 1, Tissue 3	11:5225270	ENST00000418080.1	AC104389.16	SGBS 1, SGBS2	1:210106595	ENST00000629778.2	SYT14	SGBS 1, SGBS2
12:105626227	ENST00000539978.6	APPL2	Tissue 2	16:31266310	ENST00000287497.12	ITGAM	Tissue 1	1:153347124	ENST00000368737.4	S100A12	SGBS 1, SGBS2	12:104669499	ENST00000549807.1	RP11-818F20.5	SGBS 1, SGBS2
12:109026734	ENST00000550948.1	SELPLG	Tissue 1, Tissue 3	16:31266314	ENST00000544665.7	ITGAM	Tissue 1	11:58671896	ENST00000454512.1	AP001652.1	SGBS 1, SGBS2	12:210607926	ENST00000367009.2	HHAT	SGBS 1, SGBS2
1:123186889	ENST00000328880.5	HCAR2	Tissue 1, Tissue 3	16:55352671	ENST00000290552.7	IRX6	Tissue 1, Tissue 2, Tissue 3	1:159923009	ENST00000368092.7	SLAMF9	SGBS 1, SGBS2	12:121194394	ENST00000542620.1	RP11-173P15.7	SGBS 1, SGBS2
12:130553869	ENST00000535487.1	RP11-474D1.2	Tissue 1, Tissue 2, Tissue 3	16:55418611	ENST00000570308.5	MMP2	Tissue 1, Tissue 2, Tissue 3	1:159923043	ENST00000368093.3	SLAMF9	SGBS 1, SGBS2	12:121623670	ENST00000619282.1	RP11-340F14.6	SGBS 1, SGBS2
12:15864926	ENST00000543612.5	EPS8	Tissue 1, Tissue 2, Tissue 3	16:89038611	ENST00000268679.8	CBFA2T3	Tissue 1, Tissue 3	11:60690947	ENST00000453848.6	TMEM132A	SGBS 1, SGBS2	12:131256090	ENST00000624414.1	RP11-989F5.4	SGBS 1, SGBS2
1:226861767	ENST00000412918.1	ITPKB-IT1	Tissue 1, Tissue 2, Tissue 3	17:18275975	ENST00000399134.4	EVPLL	Tissue 1, Tissue 2	11:60690952	ENST00000005286.8 ENST00000236938.10 ,ENST00000350710.3, ENST00000367949.6,E NST00000367959.6,E NST00000540521.5,E	TMEM132A	SGBS 1, SGBS2	12:133194558	ENST00000623606.1	RP13-554M15.5	SGBS 1, SGBS2
1:22961998	ENST00000374642.7	C1QA	Tissue 1	17:2694731	ENST00000366401.8	RAP1GAP2	Tissue 1	1:161675761	NST00000546024.5	FCRLA	SGBS 1, SGBS2	12:15810625	ENST00000540613.5	EPS8	SGBS 1, SGBS2
12:32637905	ENST00000531134.5	FGD4	Tissue 1, Tissue 3	17:34308793	ENST00000618404.4	CCL14	Tissue 1	11:66102872	ENST00000627248.1	RIN1	SGBS 1, SGBS2	1:222658041	ENST00000621440.1	RP11-378J18.9	SGBS 1, SGBS2
12:41085243	ENST00000547702.5	CNTN1	Tissue 1	17:34309039	ENST00000614009.1	CCL14	Tissue 1	11:66102999	ENST00000311320.8	RIN1	SGBS 1, SGBS2	12:22847874	ENST00000536744.5	RP11-114G22.1	SGBS 1, SGBS2
12:45685465	ENST00000441606.2	ANO6	Tissue 1, Tissue 2, Tissue 3	17:49512013	ENST00000588341.1	RP11-1018N14.5	Tissue 1	11:66103310	ENST00000530056.1	RIN1	SGBS 1, SGBS2	12:22847961	ENST00000628326.1	RP11-114G22.1	SGBS 1, SGBS2
1:248099492	ENST00000366478.2	OR2L13	Tissue 1, Tissue 3	17:53346013	ENST00000574716.1	RP11-515D17.2	Tissue 1, Tissue 2, Tissue 3	1:17515277	ENST00000412427.1	RP11-380I14.1	SGBS 1, SGBS2	1:225883305	ENST00000448264.1	RP11-145A3.1	SGBS 1, SGBS2
12:56105194	ENST00000452168.6	ITGA7	Tissue 1, Tissue 2, Tissue 3	18:13422478	ENST00000588672.1	LDLRAD4-AS1	Tissue 1, Tissue 2, Tissue 3	1:17515283	ENST00000634972.1	RP11-380I14.1	SGBS 1, SGBS2	1:229571766	ENST00000434311.1	RP5-1068B5.5	SGBS 1, SGBS2
12:6232840	ENST00000261405.9	VWF	Tissue 1, Tissue 2, Tissue 3	18:22562645	ENST00000580984.1	RP11-958F21.1	Tissue 1, Tissue 2, Tissue 3	11:75926068	ENST00000531538.1	RP11-619A14.3	SGBS 1, SGBS2	1:240770448	ENST00000318160.4	GREM2	SGBS 1, SGBS2
12:69741120	ENST00000261267.6	LYZ	Tissue 1, Tissue 3	18:28676550	ENST00000581836.1	DSCAS	Tissue 1, Tissue 2	11:7625968	ENST00000530181.5 ENST00000315938.4,E	PPFIBP2	SGBS 1, SGBS2	1:244075703	ENST00000440494.1	RP11-278H7.1	SGBS 1, SGBS2
12:69741163	ENST00000549690.1	LYZ	Tissue 1, Tissue 3	18:29557861	ENST00000624000.1	RP11-326K13.3	Tissue 1, Tissue 2, Tissue 3	11:76997462	NST00000376217.6	GDPD4	SGBS 1, SGBS2	1:246848348	ENST00000442712.1 ENST00000266579.8,	RP11-439E19.7	SGBS 1, SGBS2
12:69741165	ENST00000548839.1	LYZ	Tissue 1, Tissue 3	18:30711040	ENST00000581852.5	CCDC178	Tissue 1, Tissue 2, Tissue 3	1:179786325	ENST00000415218.1	RP11-12M5.3	SGBS 1, SGBS2	12:47214779	ENST00000447411.5	SLC38A4	SGBS 1, SGBS2
12:739056	ENST00000537514.1	RP11-218M22.1	Tissue 1, Tissue 2, Tissue 3	19:12300831	ENST00000426044.1	CTD-2666L21.1	Tissue 1	1:209740017	ENST00000424696.6	RP1-272L16.1	SGBS 1, SGBS2	12:70299153	ENST00000549710.1	RP11-498M15.1	SGBS 1, SGBS2
12:7847371	ENST00000329913.3	GDF3	Tissue 1, Tissue 2, Tissue 3	19:12300902	ENST00000451691.2	CTD-2666L21.1	Tissue 1	1:209823678	ENST00000391911.5 ENST00000367019.5,E	LAMB3	SGBS 1, SGBS2	12:91500607	ENST00000266718.4	LUM	SGBS 1, SGBS2
1:27951750	ENST00000399173.5	FGR	Tissue 1	19:22188720	ENST00000597040.1 ENST00000397126.8, ENST00000601773.5,	ZNF208	Tissue 1	1:210110518	NST00000537238.5 ENST00000472886.5,E	SYT14	SGBS 1, SGBS2	1:30177393	ENST00000422638.1	RP4-656G21.1	SGBS 1, SGBS2
12:9267752	ENST00000318602.11	A2M	Tissue 1, Tissue 2, Tissue 3	19:22188744	ENST00000609966.5	ZNF208	Tissue 1	1:210110595	NST00000629778.2	SYT14	SGBS 1, SGBS2	13:100100722	ENST00000582220.1	AL583784.1	SGBS 1, SGBS2
12:94675619	ENST00000547927.1	RP11-1105G2.3	Tissue 1, Tissue 3	19:22188750	ENST00000599916.5	ZNF208	Tissue 1	12:104673499	ENST00000549807.1	RP11-818F20.5	SGBS 1, SGBS2	13:31546759	ENST00000433788.1	RP11-252M21.6	SGBS 1, SGBS2
12:96389121	ENST00000538703.5 ENST00000261208.7,EN	HAL	Tissue 1, Tissue 2, Tissue 3	19:22600147	ENST00000357774.9	ZNF98	Tissue 1	1:210611926	ENST00000367009.2	HHAT	SGBS 1, SGBS2	1:33603830	ENST00000622988.1	RP11-131M11.2	SGBS 1, SGBS2
12:96389142	ST00000541929.5	HAL	Tissue 1, Tissue 2, Tissue 3	19:22600155	ENST00000601553.1	ZNF98	Tissue 1	12:109021462	ENST00000550306.1	RP11-689B22.2	SGBS 1, SGBS2	13:96362031	ENST00000516092.1	snR65	SGBS 1, SGBS2
12:96389298	ENST00000551849.1	RP11-256L6.3	Tissue 1, Tissue 2, Tissue 3	19:22600314	ENST00000599129.1	AC011516.1	Tissue 1	12:109089177	ENST00000421578.6	CORO1C	SGBS 1, SGBS2	14:105326616	ENST00000556508.5 ENST00000414716.7,	CEP170B	SGBS 1, SGBS2
13:113632625	ENST00000421756.5	MCF2L	Tissue 1, Tissue 2, Tissue 3	19:22710427	ENST00000598832.1	LINC01233	Tissue 1	12:121198394	ENST00000542620.1	RP11-173P15.7	SGBS 1, SGBS2	14:105326649	ENST00000453495.2	CEP170B	SGBS 1, SGBS2
1:32041085	ENST00000457433.6	TINAGL1	Tissue 3	19:22710472	ENST00000601708.1	LINC01233	Tissue 1	12:121627670	ENST00000619282.1	RP11-340F14.6	SGBS 1, SGBS2	1:44509048	ENST00000431800.1	RP5-1198O20.4	SGBS 1, SGBS2
1:32041135	ENST00000271064.11	TINAGL1	Tissue 3	19:22961908	ENST00000596209.3	ZNF99	Tissue 1, Tissue 2, Tissue 3	12:131260090	ENST00000624414.1	RP11-989F5.4	SGBS 1, SGBS2	1:44509063	ENST00000437643.1	RP5-1198O20.4	SGBS 1, SGBS2
13:23754090	ENST00000218867.3	SGCG	Tissue 1, Tissue 2, Tissue 3	19:39821644	ENST00000602185.5	GMFG	Tissue 1, Tissue 2, Tissue 3	12:133198558	ENST00000623606.1	RP13-554M15.5	SGBS 1, SGBS2	1:44851393	ENST00000362515.1	RNU6-369P	SGBS 1, SGBS2
13:24827576	ENST00000430733.1	SPATA13-AS1	Tissue 1	19:39821645	ENST00000598034.5	GMFG	Tissue 1, Tissue 3	12:15814625	ENST00000540613.5	EPS8	SGBS 1, SGBS2	14:85978277	ENST00000557547.1	CTD-2128A3.2	SGBS 1, SGBS2
13:44731357	ENST00000415082.1 ENST00000329797.7,EN	SMIM2-IT1	Tissue 1, Tissue 2, Tissue 3	19:39821656	ENST00000601387.5	GMFG	Tissue 1, Tissue 3	1:21911964	ENST00000457706.1	RP11-63N8.3	SGBS 1, SGBS2	15:101414580	ENST00000329841.9	ALDH1A3	SGBS 1, SGBS2
14:105530766	ST00000539291.6	GPR132	Tissue 1	19:39821663	ENST00000595636.1	GMFG	Tissue 1, Tissue 3	1:222662041	ENST00000621440.1	RP11-378J18.9	SGBS 1, SGBS2	15:101415043	ENST00000557963.1	ALDH1A3	SGBS 1, SGBS2
14:105530781	ENST00000392585.2	GPR132	Tissue 1	19:39821667	ENST00000253054.12	GMFG	Tissue 1, Tissue 3	12:22851874	ENST00000536744.5	RP11-114G22.1	SGBS 1, SGBS2	15:101415079	ENST00000346623.6	ALDH1A3	SGBS 1, SGBS2
14:24041209	ENST00000544177.1	JPH4	Tissue 1, Tissue 3	19:39821678	ENST00000594700.5	GMFG	Tissue 1, Tissue 3	12:22851961	ENST00000628326.1	RP11-114G22.1	SGBS 1, SGBS2	1:52451435	ENST00000371655.3	RAB3B	SGBS 1, SGBS2
14:47811220	ENST00000357362.7	MDGA2	Tissue 1, Tissue 2, Tissue 3	19:39821824	ENST00000597595.6	GMFG	Tissue 1, Tissue 3	1:225887305	ENST00000448264.1	RP11-145A3.1	SGBS 1, SGBS2	15:29264666	ENST00000560531.1	RP13-126C7.1	SGBS 1, SGBS2
14:65936819	ENST00000385047.1 ENST00000271153.8,EN	MIR625	Tissue 2	19:42050885	ENST00000407170.6 ENST00000595395.1,	CEACAM21	Tissue 1, Tissue 2, Tissue 3	1:229575766	ENST00000434311.1	RP5-1068B5.5	SGBS 1, SGBS2	15:41311696	ENST00000560178.1	RP11-540O11.4	SGBS 1, SGBS2
1:47263717	ST00000371923.8 ENST00000371919.8,EN	CYP4B1	Tissue 1, Tissue 2, Tissue 3	19:42050887	ENST00000601116.5	AC006129.2	Tissue 1, Tissue 2, Tissue 3	1:236106104	ENST00000438371.1	RP5-940F7.2	SGBS 1, SGBS2	15:42714894	ENST00000364207.1	RNU6-188P	SGBS 1, SGBS2
1:47263753	ST00000614163.4	CYP4B1	Tissue 1, Tissue 2, Tissue 3	19:48748027	ENST00000522431.5	CARD8	Tissue 1, Tissue 2, Tissue 3	1:240774448	ENST00000318160.4	GREM2	SGBS 1, SGBS2	15:64158217	ENST00000362286.1	MIR422A	SGBS 1, SGBS2
1:49241589	ENST00000371833.3	BEND5	Tissue 1, Tissue 3	19:48748082	ENST00000520153.5	CARD8	Tissue 1, Tissue 2, Tissue 3	1:244079703	ENST00000440494.1	RP11-278H7.1	SGBS 1, SGBS2	15:78852861	ENST00000299565.9	CHRNA5	SGBS 1, SGBS2
14:94983180	ENST00000341228.2	SERPINA12	Tissue 3	19:48748089	ENST00000520015.5	CARD8	Tissue 1, Tissue 2, Tissue 3	1:246852348	ENST00000442712.1 ENST00000266579.8,E	RP11-439E19.7	SGBS 1, SGBS2	15:78852912	ENST00000559554.5	CHRNA5	SGBS 1, SGBS2
15:100665310	ENST00000560128.1	RP11-90E5.1	Tissue 1, Tissue 2, Tissue 3	19:48748103	ENST00000447740.6 ENST00000376970.6, ENST00000425460.5,	CARD8	Tissue 1, Tissue 2, Tissue 3	12:47218779	NST00000447411.5	SLC38A4	SGBS 1, SGBS2	15:79193468	ENST00000624440.1	RP11-16K12.2	SGBS 1, SGBS2
15:50405700	ENST00000559829.5	ATP8B4	Tissue 1	19:50701884	ENST00000440075.6	MYH14	Tissue 1, Tissue 2, Tissue 3	12:72103153	ENST00000549710.1	RP11-498M15.1	SGBS 1, SGBS2	16:15591122	ENST00000452191.6	C16orf45	SGBS 1, SGBS2

15:67350590	ENST00000558071.2	RP11-798K3.2	Tissue 1, Tissue 2, Tissue 3	19:50701901	ENST00000601313.5	MYH14	Tissue 1, Tissue 2, Tissue 3	12:90340469	ENST00000547094.1	RP11-654D12.3	SGBS 1, SGBS2	16:18990255	ENST00000304381.9, ENST00000569532.5	TMC7	SGBS 1, SGBS2
15:69221838	ENST00000310673.3	SPESP1	Tissue 1	19:54871413	ENST00000474878.5	LAIR1	Tissue 1, Tissue 3	12:91504607	ENST00000266718.4	LUM	SGBS 1, SGBS2	16:18990608	ENST00000421369.3	TMC7	SGBS 1, SGBS2
15:69221842	ENST00000557966.1 ENST00000260364.9,EN	RP11-809H16.2	Tissue 1	19:54871471	ENST00000348231.8	LAIR1	Tissue 1, Tissue 3	1:30181393	ENST00000422638.1	RP4-656G21.1	SGBS 1, SGBS2	16:23676331	ENST00000566996.2	CTD-2196E14.9	SGBS 1, SGBS2
15:69221863	ST00000448182.7	NOX5	Tissue 1	19:54871563	ENST00000391742.6	LAIR1	Tissue 1, Tissue 3	13:100104722	ENST00000582220.1	AL583784.1	SGBS 1, SGBS2	1:64572887	ENST00000412349.6	ROR1-AS1	SGBS 1, SGBS2
15:69221972	ENST00000455873.7	NOX5	Tissue 1	19:54871601	ENST00000434277.6	LAIR1	Tissue 1, Tissue 3	13:31550759	ENST00000438788.1	RP11-252M21.6	SGBS 1, SGBS2	16:50295426	ENST00000394697.6	ADCY7	SGBS 1, SGBS2
15:94405337	ENST00000554530.5 ENST00000331849.8,EN	RP11-76E17.3	Tissue 1, Tissue 3	19:57178127	ENST00000537055.3	ZNF835	Tissue 1, Tissue 2, Tissue 3	1:33607830	ENST00000622988.1	RP11-131M11.2	SGBS 1, SGBS2	16:50295461	ENST00000566433.6	ADCY7	SGBS 1, SGBS2
16:20419855	ST00000575584.5	ACSM5	Tissue 1, Tissue 2, Tissue 3	20:30593252	ENST00000262659.12 ENST00000534862.5,	CCM2L	Tissue 1, Tissue 2, Tissue 3	13:96366031	ENST00000516092.1	snR65	SGBS 1, SGBS2	16:580315	ENST00000581803.1	MIR5587	SGBS 1, SGBS2
16:20419867	ENST00000575070.1	ACSM5	Tissue 1, Tissue 2, Tissue 3	20:30634990	ENST00000538448.5	HCK	Tissue 1, Tissue 3	14:105330616	ENST00000556508.5 ENST00000414716.7,E	CEP170B	SGBS 1, SGBS2	16:71455366	ENST00000567469.1	RP11-510M2.1	SGBS 1, SGBS2
16:29756326	ENST00000329410.3	C16orf54	Tissue 1, Tissue 3	20:30634999	ENST00000375862.6 ENST00000520553.5,	HCK	Tissue 1, Tissue 3	14:105330649	NST00000453495.2	CEP170B	SGBS 1, SGBS2	16:71555153	ENST00000539698.3	CHST4	SGBS 1, SGBS2
16:31270310	ENST00000287497.12	ITGAM	Tissue 1	20:30635044	ENST00000629881.2 ENST00000375852.2,	HCK	Tissue 1, Tissue 3	1:44513048	ENST00000431800.1	RP5-1198O20.4	SGBS 1, SGBS2	16:72816671	ENST00000584072.1	AC004943.1	SGBS 1, SGBS2
16:31270314	ENST00000544665.7	ITGAM	Tissue 1	20:30635063	ENST00000518730.5	HCK	Tissue 1, Tissue 3	1:44513063	ENST00000437643.1	RP5-1198O20.4	SGBS 1, SGBS2	1:6779696	ENST00000432429.5	RP11-242F24.1	SGBS 1, SGBS2
16:55356671	ENST00000290552.7	IRX6	Tissue 1, Tissue 2, Tissue 3	20:31664332	ENST00000375483.3	BPIFB4	Tissue 1, Tissue 2	1:44853393	ENST00000362515.1	RNU6-369P	SGBS 1, SGBS2	1:6779732	ENST00000635092.1	RP11-242F24.1	SGBS 1, SGBS2
16:55422611	ENST00000570308.5	MMP2	Tissue 1, Tissue 2, Tissue 3	20:34073807	ENST00000453914.1	RP3-477O4.14	Tissue 1	14:71564414	ENST00000553682.1	LINC01269	SGBS 1, SGBS2	16:81673962	ENST00000398040.8	CMIP	SGBS 1, SGBS2
16:89042611	ENST00000268679.8	CBFA2T3	Tissue 1, Tissue 2, Tissue 3	20:51709190	ENST00000362348.1	RN7SKP184	Tissue 1, Tissue 3	14:85982277	ENST00000557547.1	CTD-2128A3.2	SGBS 1, SGBS2	16:82221394	ENST00000516079.1	RN7SKP190	SGBS 1, SGBS2
17:18279975	ENST00000399134.4 ENST00000254695.12,E	EVPLL	Tissue 1, Tissue 2	20:9961735	ENST00000449270.1	RP5-839B4.8	Tissue 1, Tissue 2, Tissue 3	15:101418580	ENST00000329841.9	ALDH1A3	SGBS 1, SGBS2	16:83018902	ENST00000565238.1	CTD-3253I12.1	SGBS 1, SGBS2
17:2698731	NST00000366401.8	RAP1GAP2	Tissue 1	2:108140942	ENST00000443205.1	AC096669.3	Tissue 1, Tissue 2, Tissue 3	15:101419043	ENST00000557963.1	ALDH1A3	SGBS 1, SGBS2	16:87835081	ENST00000561567.1	RP4-536B24.3	SGBS 1, SGBS2
17:34312793	ENST00000618404.4	CCL14	Tissue 1	21:15913661	ENST00000400564.5	SAMSN1	Tissue 1	15:101419079	ENST00000346623.6	ALDH1A3	SGBS 1, SGBS2	16:9157410	ENST00000562893.1	RP11-473I1.6	SGBS 1, SGBS2
17:34313039	ENST00000614009.1 ENST00000308423.6,EN	CCL14	Tissue 1	21:15913663	ENST00000400566.5	SAMSN1	Tissue 1	1-52455435	ENST00000371655.3	RAB3B	SGBS 1, SGBS2	17:15177793	ENST00000453339.2	AC005703.2	SGBS 1, SGBS2
17:41002200	ST00000613571.1	AOC3	Tissue 1, Tissue 2	21:15913680	ENST00000619120.4	SAMSN1	Tissue 1	15:29268666	ENST00000560531.1	RP13-126C7.1	SGBS 1, SGBS2	17:39533654	ENST00000394001.1	KRT34	SGBS 1, SGBS2
17:49516013	ENST00000588341.1	RP11-1018N14.5	Tissue 1	21:29322239	ENST00000433344.1	AJ006995.3	Tissue 1, Tissue 2, Tissue 3	15:41315696	ENST00000560178.1	RP11-540O11.4	SGBS 1, SGBS2	17:70021794	ENST00000472655.3	LINC01152	SGBS 1, SGBS2
17:53350013	ENST00000574716.1	RP11-515O17.2	Tissue 1, Tissue 2, Tissue 3	2:143881882	ENST00000295095.10 ENST00000291700.8,	ARHGAP15	Tissue 1	15:42718894	ENST00000364207.1	RNU6-188P	SGBS 1, SGBS2	17:70583942	ENST00000453722.6	LINC00511	SGBS 1, SGBS2
17:6982625	ENST00000254868.8	CLEC10A	Tissue 1, Tissue 3	21:48020120	ENST00000367071.4	S100B	Tissue 1, Tissue 2, Tissue 3	15:64162217	ENST00000362286.1	MIR422A	SGBS 1, SGBS2	17:79604438	ENST00000611590.1	TSPAN10	SGBS 1, SGBS2
17:72965798	ENST00000577295.1	HID1-AS1	Tissue 1, Tissue 3	2:154723425	ENST00000392825.7	GALNT13	Tissue 1, Tissue 2, Tissue 3	15:78856861	ENST00000299565.9	CHRNA5	SGBS 1, SGBS2	17:80535380	ENST00000459591.1	snoU13	SGBS 1, SGBS2
17:73504968	ENST00000433559.6	CASKIN2	Tissue 1, Tissue 2, Tissue 3	2:231085444	ENST00000420434.7	SP140	Tissue 1	15:78856912	ENST00000559554.5	CHRNA5	SGBS 1, SGBS2	18:32342626	ENST00000596954.1	RP11-138H11.1	SGBS 1, SGBS2
18:13426478	ENST00000588672.1	LDLRAD4-AS1	Tissue 1, Tissue 3	2:231085445	ENST00000392045.7	SP140	Tissue 1	15:79197468	ENST00000624440.1	RP11-16K12.2	SGBS 1, SGBS2	18:61437608	ENST00000540675.5	SERPINB7	SGBS 1, SGBS2
18:22566645	ENST00000580984.1	RP11-958F21.1	Tissue 1, Tissue 2, Tissue 3	2:231085456	ENST00000417495.7	SP140	Tissue 1	16:15595122	ENST00000452191.6 ENST00000304381.9,E	C16orf45	SGBS 1, SGBS2	18:61437643	ENST00000398019.6	SERPINB7	SGBS 1, SGBS2
18:28680550	ENST00000581836.1	DSCAS	Tissue 1, Tissue 2	2:231085459	ENST00000343805.10	SP140	Tissue 1	16:18994255	NST00000569532.5	TMC7	SGBS 1, SGBS2	18:68046826	ENST00000582251.1	RP11-41O4.2	SGBS 1, SGBS2
18:29561861	ENST00000624000.1	RP11-326K13.3	Tissue 1, Tissue 2, Tissue 3	2:231085478	ENST00000373645.3	SP140	Tissue 1	16:18994608	ENST00000421369.3	TMC7	SGBS 1, SGBS2	18:73114992	ENST00000579386.2	RP11-321M21.3	SGBS 1, SGBS2
18:30715040	ENST00000581852.5	CCDC178	Tissue 1, Tissue 2, Tissue 3	22:50046077	ENST00000405854.5	C22orf34	Tissue 1, Tissue 2, Tissue 3	16:23680331	ENST00000566996.2	CTD-2196E14.9	SGBS 1, SGBS2	19:14043877	ENST00000254320.7	PODNL1	SGBS 1, SGBS2
18:3450590	ENST00000551541.5	TGIF1	Tissue 1	22:50046151	ENST00000414287.5	C22orf34	Tissue 1, Tissue 2, Tissue 3	1:64576887	ENST00000412349.6	ROR1-AS1	SGBS 1, SGBS2	19:14044288	ENST00000339560.9	PODNL1	SGBS 1, SGBS2
18:59414408	ENST00000567801.2	LINC01544	Tissue 1, Tissue 2, Tissue 3	2:50196292	ENST00000635519.1	NRXN1	Tissue 1, Tissue 2, Tissue 3	16:50299426	ENST00000394697.6	ADCY7	SGBS 1, SGBS2	19:33550793	ENST00000254260.7	RHPN2	SGBS 1, SGBS2
19:1173254	ENST00000587024.5	SBNO2	Tissue 1, Tissue 3	2:50196336	ENST00000412315.5	NRXN1	Tissue 1, Tissue 2, Tissue 3	16:50299461	ENST00000566433.6	ADCY7	SGBS 1, SGBS2	19:35891508	ENST00000536898.5	LINC01531	SGBS 1, SGBS2
19:1173281	ENST00000361757.7	SBNO2	Tissue 1, Tissue 3	2:50196340	ENST00000378262.7 ENST00000415520.5,	NRXN1	Tissue 1, Tissue 2, Tissue 3	16:584315	ENST00000581803.1	MIR5587	SGBS 1, SGBS2	19:36159192	ENST00000443196.1	UPK1A-AS1	SGBS 1, SGBS2
19:12304831	ENST00000426044.1	CTD-2666L21.1	Tissue 1	2:7213010	ENST00000428379.1	AC019048.1	Tissue 1, Tissue 2, Tissue 3	16:58532950	ENST00000566192.5	NDRG4	SGBS 1, SGBS2	19:44759032	ENST00000592581.5	ZNF233	SGBS 1, SGBS2
19:12304902	ENST00000451691.2	CTD-2666L21.1	Tissue 1	2:9241721	ENST00000565044.1	RP11-734K21.5	Tissue 1, Tissue 2, Tissue 3	16:58533010	ENST00000568640.5 ENST00000356752.8,E NST00000562999.5,E NST00000563799.5,E NST00000569923.5,E	NDRG4	SGBS 1, SGBS2	19:44759075	ENST00000391958.6	ZNF233	SGBS 1, SGBS2
19:14550071	ENST00000342216.8	PKN1	Tissue 1, Tissue 3	3:178979758	ENST00000349697.2	KCNMB3	Tissue 1, Tissue 2, Tissue 3	16:58533046	NST00000570248.5	NDRG4	SGBS 1, SGBS2	19:55876740	ENST00000264563.6 ENST00000585513.1,	IL11	SGBS 1, SGBS2
19:17374532	ENST00000431146.6	USHBP1	Tissue 3	3:178979789	ENST00000497599.5	KCNMB3	Tissue 1, Tissue 2, Tissue 3	16:67310412	ENST00000360461.9	PLEKHG4	SGBS 1, SGBS2	19:55876830	ENST00000590625.5	IL11	SGBS 1, SGBS2
19:17374604	ENST00000252597.7	USHBP1	Tissue 3	3:183268476	ENST00000341319.7	KLHL6	Tissue 1, Tissue 3	1:66819064	ENST00000480109.2	PDE4B	SGBS 1, SGBS2	1:96203651	ENST00000603401.1 ENST00000306708.10	RP11-286B14.2	SGBS 1, SGBS2
19:18507420	ENST00000595840.1	LRRC25	Tissue 1	3:18782235	ENST00000453361.1	SATB1-AS1	Tissue 1	16:71459366	ENST00000567469.1	RP11-510M2.1	SGBS 1, SGBS2	19:7948389	,ENST00000618098.4	LRRC8E	SGBS 1, SGBS2
19:18507426	ENST00000339007.3	LRRC25	Tissue 1	3:40802697	ENST00000412811.1	RP11-761N21.1	Tissue 3	16:71559153	ENST00000539698.3	CHST4	SGBS 1, SGBS2	19:9512730	ENST00000634951.1	CTC-325H20.8	SGBS 1, SGBS2
19:22192720	ENST00000597040.1 ENST00000397126.8,EN ST00000601773.5,ENST	ZNF208	Tissue 1	3:56945498	ENST00000496106.5	ARHGEF3	Tissue 1, Tissue 2, Tissue 3	16:72820671	ENST00000584072.1	AC004943.1	SGBS 1, SGBS2	20:18290754	ENST00000436848.1	RP4-568F9.3	SGBS 1, SGBS2
19:22192744	00000609966.5	ZNF208	Tissue 1	4:15424537	ENST00000502344.5	RP11-665G4.1	Tissue 1, Tissue 2, Tissue 3	1:6783696	ENST00000432429.5	RP11-242F24.1	SGBS 1, SGBS2	20:24477980	ENST00000458461.1	RP11-526K17.2	SGBS 1, SGBS2
19:22192750	ENST00000599916.5	ZNF208	Tissue 1	5:160360632	ENST00000518580.5	RP11-109J4.1	Tissue 1, Tissue 2, Tissue 3	1:6783732	ENST00000635092.1	RP11-242F24.1	SGBS 1, SGBS2	20:32600554	ENST00000451556.2	RP5-1125A11.4	SGBS 1, SGBS2
19:22604147	ENST00000577774.9	ZNF98	Tissue 1	5:23502723	ENST00000296682.1	PRDM9	Tissue 1, Tissue 2, Tissue 3	16:81642911	ENST00000665513.5	CMIP	SGBS 1, SGBS2	20:49634665	ENST00000371571.4	CNCNG1	SGBS 1, SGBS2
19:22604155	ENST00000601553.1	ZNF98	Tissue 1	5:42543144	ENST00000618088.4	GHR	Tissue 1, Tissue 3	16:81677962	ENST00000398040.8	CMIP	SGBS 1, SGBS2	20:5887075	ENST00000378961.8	CHGB	SGBS 1, SGBS2
19:22604314	ENST00000599129.1	AC011516.1	Tissue 1	5:42543314	ENST00000612382.4	GHR	Tissue 1, Tissue 3	16:82225394	ENST00000516079.1	RN7SKP190	SGBS 1, SGBS2	20:60007602	ENST00000616819.1	RP5-827E24.1	SGBS 1, SGBS2
19:22714427	ENST00000598832.1	LINC01233	Tissue 1	5:76430073	ENST00000515356.2	ZBED3-AS1	Tissue 1, Tissue 2, Tissue 3	16:83022902	ENST00000565238.1	CTD-3253I12.1	SGBS 1, SGBS2	2:113964098	ENST00000456685.5	PAX8-AS1	SGBS 1, SGBS2
19:22714472	ENST00000601708.1	LINC01233	Tissue 1	6:15990174	ENST00000451285.1	RP11-401E14.2	Tissue 1, Tissue 2, Tissue 3	16:87839081	ENST00000561567.1	RP4-536B24.3	SGBS 1, SGBS2	2:114595932	ENST00000446401.1	AC104653.1	SGBS 1, SGBS2
19:22965908	ENST00000596209.3	ZNF99	Tissue 1, Tissue 2, Tissue 3	6:166183673	ENST00000431017.1	RP11-252P19.2	Tissue 1, Tissue 2, Tissue 3	16:9161410	ENST00000562893.1	RP11-473I1.6	SGBS 1, SGBS2	2:121619977	ENST00000603720.1	RP11-297J22.1	SGBS 1, SGBS2
19:29458803	ENST00000587998.5	LINC00906	Tissue 1, Tissue 2, Tissue 3	6:32186843	ENST00000375023.3	NOTCH4	Tissue 1, Tissue 3	17:15181793	ENST00000453339.2	AC005703.2	SGBS 1, SGBS2	2:122543019	ENST00000438722.5	AC018737.3	SGBS 1, SGBS2
19:3177765	ENST00000246115.4	S1PR4	Tissue 1, Tissue 2, Tissue 3	6:32552581	ENST00000611060.4	HLA-DRB1	Tissue 3	17:39537654	ENST00000394001.1	KRT34	SGBS 1, SGBS2	2:122543040	ENST00000435895.5	AC018737.3	SGBS 1, SGBS2
1:9337449	ENST00000517032.1	Z98044.1	Tissue 3	6:32552624	ENST00000360004.5	HLA-DRB1	Tissue 3	17:52976873	ENST00000575882.5	TOM1L1	SGBS 1, SGBS2	21:27757454	ENST00000421771.1	AP001596.6	SGBS 1, SGBS2
19:39825644	ENST00000602185.5	GMFG	Tissue 1, Tissue 3	6:32629428	ENST00000399082.7	HLA-DQB1	Tissue 1, Tissue 2, Tissue 3	17:52977112	ENST00000445275.6	TOM1L1	SGBS 1, SGBS2	21:30718954	ENST00000504298.1	BACH1-IT1	SGBS 1, SGBS2
19:39825645	ENST00000598034.5	GMFG	Tissue 1, Tissue 3	6:32629440	ENST00000399079.7	HLA-DQB1	Tissue 1, Tissue 2, Tissue 3	17:52977166	ENST00000348161.8	TOM1L1	SGBS 1, SGBS2	21:35823534	ENST00000408471.1	SNORA11	SGBS 1, SGBS2
19:39825656	ENST00000601387.5	GMFG	Tissue 1, Tissue 3	6:32629460	ENST00000374943.8	HLA-DQB1	Tissue 1, Tissue 2, Tissue 3	17:52977194	ENST00000572405.5	TOM1L1	SGBS 1, SGBS2	2:167992464	ENST00000525330.1	XIRP2-AS1	SGBS 1, SGBS2
19:39825663	ENST00000595636.1	GMFG	Tissue 1, Tissue 3	6:32629470	ENST00000434651.6	HLA-DQB1	Tissue 1, Tissue 2, Tissue 3	17:52977196	ENST00000572158.5	TOM1L1	SGBS 1, SGBS2	2:175647655	ENST00000416004.1	AC018890.4	SGBS 1, SGBS2
19:39825667	ENST00000253054.12	GMFG	Tissue 1, Tissue 3	6:32631159	ENST00000399084.5	HLA-DQB1	Tissue 1, Tissue 2, Tissue 3	17:52977213	ENST00000536554.5	TOM1L1	SGBS 1, SGBS2	2:207073351	ENST00000621141.4	GPR1	SGBS 1, SGBS2
19:39825678	ENST00000594700.5	GMFG	Tissue 1, Tissue 3	6:46866207	ENST00000451135.1	RP3-365O12.2	Tissue 3	17:52977217	ENST00000575333.5	TOM1L1	SGBS 1, SGBS2	22:19863686	ENST00000485358.5	TXNRD2	SGBS 1, SGBS2
19:39825824	ENST00000597595.5	GMFG	Tissue 1, Tissue 3	6:62991131	ENST00000281156.4	KHDRBS2	Tissue 1, Tissue 2, Tissue 3	17:70025794	ENST00000472655.3	LINC01152	SGBS 1, SGBS2	22:19863891	ENST00000462330.5	TXNRD2	SGBS 1, SGBS2
19:42054885	ENST00000407170.6 ENST00000595395.1,EN	CEACAM21	Tissue 1, Tissue 2, Tissue 3	6:6583340	ENST00000379953.6	LY86	Tissue 1, Tissue 3	17:70334707	ENST00000581183.1	LINC00511	SGBS 1, SGBS2	2:223158464	ENST00000409828.7	PAX3	SGBS 1, SGBS2
19:42054887	ST00000601116.5	AC006129.2	Tissue 1, Tissue 2, Tissue 3	7:150259364	ENST00000255945.3	GIMAP4	Tissue 1, Tissue 3	17:70587942	ENST00000453722.6	LINC00511	SGBS 1, SGBS2	2:223158470	ENST00000350526.8	PAX3	SGBS 1, SGBS2
19:4539078	ENST00000586020.1	CTB-50L17.14	Tissue 1	7:150259456	ENST00000461940.5	GIMAP4	Tissue 1, Tissue 3	17:79608438	ENST00000611590.1	TSPAN10	SGBS 1, SGBS2	2:223158583	ENST00000258387.5	PAX3	SGBS 1, SGBS2
19:48752027	ENST00000522431.5	CARD8	Tissue 1, Tissue 2, Tissue 3	7:150408644	ENST00000307194.5	GIMAP1	Tissue 1, Tissue 2, Tissue 3	17:80539380	ENST00000459591.1	snoU13	SGBS 1, SGBS2	2:223158613	ENST00000409551.7	PAX3	SGBS 1, SGBS2
19:48752082	ENST00000520153.5	CARD8	Tissue 1, Tissue 2, Tissue 3	7:150408755	ENST00000611999.4	GIMAP1-GIMAP5	Tissue 1, Tissue 2, Tissue 3	18:12418931	ENST00000592149.5	PRELID3A	SGBS 1, SGBS2	2:223158699	ENST00000392069.6 ENST00000336840.10 ,ENST00000344493.8,	PAX3	SGBS 1, SGBS2
19:48752089	ENST00000520015.5	CARD8	Tissue 1, Tissue 2, Tissue 3	7:150414340	ENST00000498181.6	GIMAP5	Tissue 1, Tissue 2, Tissue 3	18:32346626	ENST00000596954.1	RP11-138H11.1	SGBS 1, SGBS2	2:223158714	ENST00000392070.6	PAX3	SGBS 1, SGBS2
19:48752103	ENST00000447740.6 ENST00000376970.6,EN ST00000425460.5,ENST	CARD8	Tissue 1, Tissue 2, Tissue 3	7:153104318	ENST00000454441.2	LINC01287	Tissue 2	18:47375215	ENST00000592688.1	MYO5B	SGBS 1, SGBS2	2:227045086	ENST00000423838.1	AC068138.1	SGBS 1, SGBS2
19:50705884	00000440075.6	MYH14	Tissue 1, Tissue 2, Tissue 3	7:157642276	ENST00000438047.1	AC011899.9	Tissue 1, Tissue 3	18:61441608	ENST00000540675.5	SERPINB7	SGBS 1, SGBS2	2:228730769	ENST00000373666.6	DAW1	SGBS 1, SGBS2
19:50705901	ENST00000601313.5	MYH14	Tissue 1, Tissue 2, Tissue 3	7:37388271	ENST00000442504.5	ELMO1	Tissue 1, Tissue 2, Tissue 3	18:61441643	ENST00000398019.6	SERPINB7	SGBS 1, SGBS2	2:228731322	ENST00000309931.2	DAW1	SGBS 1, SGBS2
19:54326570	ENST00000391775.7	NLRP12	Tissue 1	7:57467422	ENST00000624231.1	RP11-114G11.4	Tissue 1, Tissue 2, Tissue 3	18:68050826	ENST00000582251.1	RP11-41O4.2	SGBS 1, SGBS2	2:29422463	ENST00000325660.3	ZNRF3-AS1	SGBS 1, SGBS2
19:54326596	ENST00000324134.10 ENST00000345770.9,EN ST00000391772.1,ENST	NLRP12	Tissue 1	7:63356200	ENST00000450544.1	RP11-340I6.8	Tissue 1, Tissue 2, Tissue 3	18:73118992	ENST00000579386.2	RP11-321M21.3	SGBS 1, SGBS2	2:231768566	ENST00000454890.1	AC012507.4	SGBS 1, SGBS2
19:54326647	00000391773.5	NLRP12	Tissue 1	7:63356445	ENST00000585312.1	AC092634.1	Tissue 1, Tissue 2, Tissue 3	19:14047877	ENST00000254320.7	PODNL1	SGBS 1, SGBS2	2:239031030	ENST00000409506.1	ESPNL	SGBS 1, SGBS2
19:54875413	ENST00000474878.5	LAIR1	Tissue 1, Tissue 3	7:63382213	ENST00000582948.1 ENST00000456806.2,	AC092634.2	Tissue 1, Tissue 2, Tissue 3	19:14048288	ENST00000339560.9	PODNL1	SGBS 1, SGBS2	22:40919136	ENST00000516315.1	AL031594.1	SGBS 1, SGBS2
19:54875471	ENST00000348231.8	LAIR1	Tissue 1, Tissue 3	7:63500820	ENST00000550760.3	ZNF727	Tissue 1, Tissue 2, Tissue 3	19:33554793	ENST00000254260.7	RHPN2	SGBS 1, SGBS2	2:70909273	ENST00000457851.1	AC007395.3	SGBS 1, SGBS2
19:54875563	ENST00000391742.6	LAIR1	Tissue 1, Tissue 3	7:75236019	ENST00000616821.4	HIP1	Tissue 1, Tissue 2, Tissue 3	19:35895508	ENST00000536898.5	LINC01531	SGBS 1, SGBS2	3:11173778	ENST00000438284.2	HRH1	SGBS 1, SGBS2
19:54875601	ENST00000434277.6	LAIR1	Tissue 1, Tissue 3	8:101310486	ENST00000519449.5	RNF19A	Tissue 1, Tissue 2, Tissue 3	19:36163192	ENST00000443196.1	UPK1A-AS1	SGBS 1, SGBS2	3:113928159	ENST00000481773.1	RP11-553L6.2	SGBS 1, SGBS2
19:57182127	ENST00000537055.3	ZNF835	Tissue 1, Tissue 2, Tissue 3	8:112383867	ENST00000521904.5	RP11-1101K5.1	Tissue 1, Tissue 2, Tissue 3	19:44763032	ENST00000592581.5	ZNF233	SGBS 1, SGBS2	3:113928237	ENST00000493033.1	RP11-553L6.2	SGBS 1, SGBS2
19:6771724	ENST00000304076.6	VAV1	Tissue 1	8:112383891	ENST00000521753.5	RP11-1101K5.1	Tissue 1, Tissue 2, Tissue 3	19:44763075	ENST00000391958.6	ZNF233	SGBS 1, SGBS2	3:115337170	ENST00000305124.10	GAP43	SGBS 1, SGBS2
19:6771736	ENST00000602142.5	VAV1	Tissue 1	8:40005973	ENST00000315792.4	C8orf4	Tissue 1, Tissue 2, Tissue 3	19:52075580	ENST00000436511.2	ZNF175	SGBS 1, SGBS2	3:115337356	ENST00000393780.3	GAP43	SGBS 1, SGBS2
19:6771738	ENST00000596764.5	VAV1	Tissue 1	8:56428535	ENST00000522918.1	RP11-628E19.2	Tissue 2, Tissue 3	19:55880740	ENST00000264563.6 ENST00000585513.1,E	IL11	SGBS 1, SGBS2	3:120621918	ENST00000273666.10 ENST00000340059.11	STXBP5L	SGBS 1, SGBS2
20:1471041	ENST00000359801.7	SIRPB2	Tissue 1, Tissue 2	8:61875333	ENST00000525556.1	AC022182.3	Tissue 1, Tissue 3	19:55880830	NST00000590625.5	IL11	SGBS 1, SGBS2	3:120622051	ENST00000471454.5	STXBP5L	SGBS 1, SGBS2
20:1471126	ENST00000444444.2	SIRPB2	Tissue 1, Tissue 2	8:93676115	ENST00000521307.1 ENST00000343817.9,	RP11-100L22.1	Tissue 1, Tissue 3	1:96207651	ENST00000603401.1 ENST00000306708.10	RP11-286B14.2	SGBS 1, SGBS2	3:120622135	ENST00000472879.5 ENST00000492541.5,	STXBP5L	SGBS 1, SGBS2
20:17942588	ENST00000390930.1	SNORD17	Tissue 1, Tissue 2	9:116258706	ENST00000394646.7	RGS3	Tissue 1, Tissue 2, Tissue 3	19:7952389	,ENST00000618098.4	LRRC8E	SGBS 1, SGBS2	3:120622160	ENST00000497029.5 ENST00000340059.11	STXBP5L	SGBS 1, SGBS2
20:30597252	ENST00000262659.12 ENST00000534862.5,EN	CCM2L	Tissue 1, Tissue 3	9:130519696	ENST00000429553.5	SH2D3C	Tissue 1, Tissue 2, Tissue 3	19:9516730	ENST00000634951.1	CTC-325H20.8	SGBS 1, SGBS2	3:146666864	ENST00000515991.1	U3	SGBS 1, SGBS2
20:30638990	ST00000538448.5	HCK	Tissue 1, Tissue 3	9:268001	ENST00000429661.1	RP11-59O6.3	Tissue 1	20:18294754	ENST00000436848.1	RP4-568F9.3	SGBS 1, SGBS2	3:155388827	,ENST00000447496.6	PLCH1	SGBS 1, SGBS2
20:30638999	ENST00000375862.6 ENST00000520553.5,EN	HCK	Tissue 1, Tissue 3	9:66531443	ENST00000609749.1 ENST00000375751.8,	RP11-262H14.3	Tissue 1, Tissue 2, Tissue 3	20:24481980	ENST00000458461.1	RP11-526K17.2	SGBS 1, SGBS2	3:155389204	ENST00000494598.5	PLCH1	SGBS 1, SGBS2
20:30639044	ST00000629881.2 ENST00000375852.2,EN	HCK	Tissue 1, Tissue 3	9:93559068	ENST00000375754.8	SYK	Tissue 1	20:2794613	ENST00000380593.4	RP5-860F19.8	SGBS 1, SGBS2	3:178131910	ENST00000614557.1	RP11-385J1.3	SGBS 1, SGBS2
20:30639063	ST00000518730.5	HCK	Tissue 1, Tissue 3	9:93559208	ENST00000375747.5	SYK	Tissue 1	20:2794632	ENST00000603872.1	C20orf141	SGBS 1, SGBS2	3:178131988	ENST00000455307.5	LINC01014	SGBS 1, SGBS2
20:31668332	ENST00000375483.3	BPIFB4	Tissue 1, Tissue 2	X:136643300	ENST00000287538.9	ZIC3	Tissue 1, Tissue 2, Tissue 3	20:2794656	ENST00000380589.4	C20orf141	SGBS 1, SGBS2	3:183261522	ENST00000491676.1	KLHL6-AS1	SGBS 1, SGBS2
20:34077807	ENST00000453914.1	RP3-477O4.14	Tissue 1	X:78617855	ENST00000373298.6	ITM2A	Tissue 1, Tissue 2, Tissue 3	20:32604554	ENST00000451556.2	RP5-1125A11.4	SGBS 1, SGBS2	3:21979057	ENST00000448980.5	ZNF385D-AS2	SGBS 1, SGBS2
20:36973758	ENST00000217407.2	LBP	Tissue 1, Tissue 2, Tissue 3	X:78618048	ENST00000434584.2	ITM2A	Tissue 1, Tissue 2, Tissue 3	20:49638665	ENST00000371571.4	KCNG1	SGBS 1, SGBS2	3:52877568	ENST00000624222.1	RP11-894J14.2	SGBS 1, SGBS2
20:51713190	ENST00000362348.1	RN7SKP184	Tissue 1, Tissue 3					20:52198635	ENST00000302342.3	ZNF217	SGBS 1, SGBS2	4:100060426	ENST00000505590.5	ADH4	SGBS 1, SGBS2
20:589909	ENST00000621005.1	TCF15	Tissue 1, Tissue 2, Tissue 3					20:5891075	ENST00000378961.8	CHGB	SGBS 1, SGBS2	4:100060431	ENST00000629236.2	ADH4	SGBS 1, SGBS2

20:590041	ENST00000246080.3	TCF15	Tissue 1, Tissue 2, Tissue 3
20:9965735	ENST00000449270.1	RP5-839B4.8	Tissue 2, Tissue 3
2:102607305	ENST00000332549.7	IL1R2	Tissue 1
2:108144942	ENST00000443205.1	AC096669.3	Tissue 1, Tissue 2, Tissue 3
2:113930547	ENST00000245796.10	PSD4	Tissue 1
2:113930573	ENST00000441564.7	PSD4	Tissue 1
21:15917661	ENST00000400564.5	SAMSN1	Tissue 1
21:15917663	ENST00000400566.5	SAMSN1	Tissue 1
21:15917680	ENST00000619120.4	SAMSN1	Tissue 1
21:29326239	ENST00000433344.1	AJ006995.3	Tissue 1, Tissue 3
21:36420594	ENST00000437180.5	RUNX1	Tissue 1
21:36420640	ENST00000303085.7	RUNX1	Tissue 1
2:138858453	ENST00000446520.1	AC069394.1	Tissue 1, Tissue 2, Tissue 3
21:39869399	ENST00000453032.6	ERG	Tissue 1, Tissue 2, Tissue 3
21:39869406	ENST00000288319.11	ERG	Tissue 1, Tissue 2, Tissue 3
2:143885882	ENST00000295095.10	ARHGAP15	Tissue 1
21:44898413	ENST00000441283.1	LINC00313	Tissue 1, Tissue 2, Tissue 3
21:48024120	ST00000367071.4	S100B	Tissue 1, Tissue 2, Tissue 3
2:154727425	ENST00000392825.7	GALNT13	Tissue 1, Tissue 2, Tissue 3
2:158299653	ENST00000264192.7	CYTIP	Tissue 1
2:204570197	ENST00000374481.7	CD28	Tissue 1
2:204570246	ENST00000458610.6	CD28	Tissue 1
2:204570270	ENST00000324106.8	CD28	Tissue 1
2:209223437	ENST00000617735.4	PTH2R	Tissue 1, Tissue 2
22:18246024	ENST00000583102.1	MIR3198-1	Tissue 1
2:219124737	ENST00000519574.1,EN ST00000522678.5	GPBAR1	Tissue 1, Tissue 2, Tissue 3
2:219124756	ENST00000521462.1	GPBAR1	Tissue 1, Tissue 2, Tissue 3
2:219245751	ENST00000233202.10	SLC11A1	Tissue 1
2:231089444	ENST00000420434.7	SP140	Tissue 1
2:231089445	ENST00000392045.7	SP140	Tissue 1
2:231089456	ENST00000417495.7	SP140	Tissue 1
2:231089459	ENST00000343805.10	SP140	Tissue 1
2:231089478	ENST00000373645.3 ENST00000248899.10,E	SP140	Tissue 1
22:37256029	NST00000397147.6	NCF4	Tissue 1
22:37308669	ENST00000403662.7	CSF2RB	Tissue 1
22:37308674	ENST00000262825.9	CSF2RB	Tissue 1
22:50050077	ENST00000405854.5	C22orf34	Tissue 1, Tissue 2, Tissue 3
22:50050151	ENST00000414287.5	C22orf34	Tissue 1, Tissue 2, Tissue 3
2:32488948	ENST00000402280.5	NLRC4	Tissue 1
2:43328828	ENST00000434020.1	AC093609.1	Tissue 1, Tissue 2, Tissue 3
2:50200292	ENST00000635519.1	NRXN1	Tissue 1, Tissue 2, Tissue 3
2:50200336	ENST00000412315.5	NRXN1	Tissue 1, Tissue 2, Tissue 3
2:50200340	ENST00000378262.7	NRXN1	Tissue 1, Tissue 2, Tissue 3
2:68591304	ENST00000234313.7 ENST00000409030.7,EN	PLEK	Tissue 1
2:69000932	ST00000409220.5	ARHGAP25	Tissue 1, Tissue 2, Tissue 3
2:69001071	ENST00000497079.5 ENST00000415520.5,EN	ARHGAP25	Tissue 1, Tissue 2, Tissue 3
2:7217010	ST00000428379.1	AC019048.1	Tissue 1, Tissue 2, Tissue 3
2:9245721	ENST00000565044.1	RP11-734K21.5	Tissue 1, Tissue 2, Tissue 3
3:111851097	ENST00000460387.6	GCSAM	Tissue 1
3:111851100	ENST00000484193.5	GCSAM	Tissue 1
3:111851151	ENST00000308910.8	GCSAM	Tissue 1
3:128293928	ENST00000356020.2	LINC01565	Tissue 3
3:150995254	ENST00000309170.7	P2RY14	Tissue 1
3:167587764	ENST00000470523.2	RP11-298O21.3	Tissue 1, Tissue 2, Tissue 3
3:178983758	ENST00000349697.2	KCNMB3	Tissue 1, Tissue 2, Tissue 3
3:178983789	ENST00000497599.5	KCNMB3	Tissue 1, Tissue 2, Tissue 3
3:183272476	ENST00000341319.7	KLHL6	Tissue 1, Tissue 2, Tissue 3
3:186559478	ENST00000320741.6	ADIPOQ	Tissue 1, Tissue 2, Tissue 3
3:186559492	ENST00000444204.2	ADIPOQ	Tissue 1, Tissue 2, Tissue 3
3:18786235	ENST00000453361.1	SATB1-AS1	Tissue 1
3:40806697	ENST00000412811.1	RP11-761N21.1	Tissue 3
3:56949498	ENST00000496106.5	ARHGEF3	Tissue 1, Tissue 2, Tissue 3
3:9920934	ENST00000423850.5	CIDEC	Tissue 1, Tissue 2, Tissue 3
3:9920937	ENST00000618572.4 ENST00000305046.12,E	CIDEC	Tissue 1, Tissue 2, Tissue 3
4:100241555	NST00000625860.2	ADH1B	Tissue 2
4:100241557	ENST00000506651.5	ADH1B	Tissue 2
4:123804866	ENST00000609680.1	RP11-170N16.3	Tissue 1, Tissue 2, Tissue 3
4:15428537	ENST00000502344.5	RP11-665G4.1	Tissue 1, Tissue 2, Tissue 3
4:185725941	ENST00000513317.5	ACSL1	Tissue 1, Tissue 2, Tissue 3
4:8320902	ENST00000509453.1	RP11-774O3.2	Tissue 1, Tissue 2, Tissue 3
4:88243037	ENST00000302219.10	HSD17B13	Tissue 1, Tissue 3
4:88243057	ENST00000328546.4	HSD17B13	Tissue 1, Tissue 3
5:160364632	ENST00000518580.5 ENST00000327705.13,E	RP11-109J4.1	Tissue 1, Tissue 2, Tissue 3
5:180466224	NST00000376841.6	BTNL9	Tissue 1, Tissue 2, Tissue 3
5:180466265	ENST00000515271.1	BTNL9	Tissue 1, Tissue 2, Tissue 3
5:23506723	ENST00000296682.3	PRDM9	Tissue 1, Tissue 3
5:36724274	ENST00000510740.1	CTD-2353F22.1	Tissue 1, Tissue 3
5:36724277	ENST00000512329.1	CTD-2353F22.1	Tissue 1, Tissue 3
5:36724296	ENST00000508745.1	CTD-2353F22.1	Tissue 1, Tissue 3
5:38447357	ENST00000514476.1	EGFLAM	Tissue 1, Tissue 2, Tissue 3
5:42547144	ENST00000618088.4	GHR	Tissue 1, Tissue 3
5:42547314	ENST00000612382.4 ENST00000507920.5,EN	GHR	Tissue 1, Tissue 3
5:42810993	ST00000511224.5	SEPP1	Tissue 1, Tissue 2, Tissue 3
5:42811180	ENST00000514985.5	SEPP1	Tissue 1, Tissue 2, Tissue 3
5:55411777	ENST00000443982.2	ANKRD55	Tissue 1
5:59316075	ENST00000502993.1	CTD-2254N19.1	Tissue 1
5:60932534	ENST00000507461.1	C5orf64	Tissue 1, Tissue 2, Tissue 3
5:60932635	ENST00000505642.5	C5orf64	Tissue 1, Tissue 2, Tissue 3
5:7325176	ENST00000512746.1	RP11-404K5.4	Tissue 1, Tissue 2, Tissue 3
5:76434073	ENST00000515356.2	ZBED3-AS1	Tissue 1, Tissue 2, Tissue 3
6:107831639	ENST00000441532.5	RP1-67A8.3	Tissue 1, Tissue 2, Tissue 3
6:107831924	ENST00000437040.1	RP1-67A8.3	Tissue 1, Tissue 2, Tissue 3
6:11381580	ENST00000504387.5	NEDD9	Tissue 1, Tissue 3
6:134860129	ENST00000606039.1	RP11-557H15.4	Tissue 1, Tissue 2, Tissue 3
6:154676868	ENST00000367220.8	IPCEF1	Tissue 1
6:154676899	ENST00000265198.8	IPCEF1	Tissue 1
6:15994174	ENST00000451285.1	RP11-401E14.2	Tissue 1, Tissue 2, Tissue 3
6:166187673	ENST00000431017.1	RP11-252P19.2	Tissue 1, Tissue 2, Tissue 3
6:166800623	ENST00000623222.1	RP1-168L15.6	Tissue 1, Tissue 2, Tissue 3
6:31147507	ENST00000441888.7	POU5F1	Tissue 1, Tissue 2, Tissue 3
6:32190843	ENST00000375023.3	NOTCH4	Tissue 1, Tissue 3
6:32556581	ENST00000611060.4	HLA-DRB1	Tissue 3
6:32556624	ENST00000360004.5	HLA-DRB1	Tissue 3
6:32633428	ENST00000399082.7	HLA-DQB1	Tissue 2, Tissue 3
6:32633440	ENST00000399079.7	HLA-DQB1	Tissue 2, Tissue 3
20:60011602	ENST00000616819.1	RP5-827E24.1	SGBS 1, SGBS2
2:113968098	ENST00000456685.5	PAX8-AS1	SGBS 1, SGBS2
2:114599932	ENST00000446401.1	AC104653.1	SGBS 1, SGBS2
2:121623977	ENST00000603720.1	RP11-297J22.1	SGBS 1, SGBS2
2:122547019	ENST00000438722.5	AC018737.3	SGBS 1, SGBS2
2:122547040	ENST00000435895.5	AC018737.3	SGBS 1, SGBS2
21:27761454	ENST00000421771.1	AP001596.6	SGBS 1, SGBS2
21:30722954	ENST00000504298.1	BACH1-IT1	SGBS 1, SGBS2
21:35827534	ENST00000408471.1	SNORA11	SGBS 1, SGBS2
21:40297103	ENST00000623098.3	AF064858.6	SGBS 1, SGBS2
21:40297200	ENST00000613045.1	AF064858.6	SGBS 1, SGBS2
2:167996464	ENST00000525330.1	XIRP2-AS1	SGBS 1, SGBS2
2:175651655	ENST00000416004.1	AC018890.4	SGBS 1, SGBS2
2:197123747	ENST00000430904.1	AC020571.3	SGBS 1, SGBS2
2:197123837	ENST00000433933.1	AC020571.3	SGBS 1, SGBS2
2:207077351	ENST00000621141.4	GPR1	SGBS 1, SGBS2
2:208117975	ENST00000432413.2	AC007879.5	SGBS 1, SGBS2
22:19867686	ENST00000485358.5	TXNRD2	SGBS 1, SGBS2
22:19867891	ENST00000462330.5	TXNRD2	SGBS 1, SGBS2
2:220324976	ENST00000396686.1	SPEG	SGBS 1, SGBS2
2:220325106	ENST00000396689.2	SPEG	SGBS 1, SGBS2
2:223162464	ENST00000409828.7	PAX3	SGBS 1, SGBS2
2:223162470	ENST00000350526.8	PAX3	SGBS 1, SGBS2
2:223162583	ENST00000258387.5	PAX3	SGBS 1, SGBS2
2:223162613	ENST00000409551.7	PAX3	SGBS 1, SGBS2
2:223162699	ENST00000392069.6 ENST00000336840.10 ,ENST00000344493.8,	PAX3	SGBS 1, SGBS2
2:223162714	ENST00000392070.6	PAX3	SGBS 1, SGBS2
22:24911475	ENST00000432032.1	AP000355.2	SGBS 1, SGBS2
2:227049086	ENST00000423838.1	AC068138.1	SGBS 1, SGBS2
2:228734769	ENST00000373666.6	DAW1	SGBS 1, SGBS2
2:228735322	ENST00000309931.2	DAW1	SGBS 1, SGBS2
22:29426463	ENST00000325660.3	ZNRF3-AS1	SGBS 1, SGBS2
2:231772566	ENST00000454890.1	AC012507.4	SGBS 1, SGBS2
2:239035030	ENST00000409506.1	ESPNL	SGBS 1, SGBS2
22:40923136	ENST00000516315.1	AL031594.1	SGBS 1, SGBS2
2:38373556	ENST00000362733.1	RNU6-951P	SGBS 1, SGBS2
2:42251441	ENST00000451514.1	AC013480.2	SGBS 1, SGBS2
2:70913273	ENST00000457851.1	AC007395.3	SGBS 1, SGBS2
2:75144274	ENST00000435984.1	LINC01291	SGBS 1, SGBS2
3:11177778	ENST00000438284.2	HRH1	SGBS 1, SGBS2
3:113932159	ENST00000481773.1	RP11-553L6.2	SGBS 1, SGBS2
3:113932237	ENST00000493033.1	RP11-553L6.2	SGBS 1, SGBS2
3:115341170	ENST00000305124.10	GAP43	SGBS 1, SGBS2
3:115341356	ENST00000393780.3	GAP43	SGBS 1, SGBS2
3:11609397	ENST00000424529.6	VGLL4	SGBS 1, SGBS2
3:120625918	ENST00000273666.10	STXBP5L	SGBS 1, SGBS2
3:120626051	ENST00000471454.5	STXBP5L	SGBS 1, SGBS2
3:120626135	ENST00000472879.5 ENST00000492541.5,E	STXBP5L	SGBS 1, SGBS2
3:120626160	NST00000497029.5	STXBP5L	SGBS 1, SGBS2
3:141084978	ENST00000507698.1	RP11-438D8.2	SGBS 1, SGBS2
3:146670864	ENST00000515991.1 ENST00000340059.11	U3	SGBS 1, SGBS2
3:155392827	,ENST00000447496.6	PLCH1	SGBS 1, SGBS2
3:155393204	ENST00000494598.5	PLCH1	SGBS 1, SGBS2
3:178135910	ENST00000614557.1	RP11-385J1.3	SGBS 1, SGBS2
3:178135988	ENST00000455307.5	LINC01014	SGBS 1, SGBS2
3:183265525	ENST00000491676.1	KLHL6-AS1	SGBS 1, SGBS2
3:187663008	ENST00000413056.1	RP11-44H4.1	SGBS 1, SGBS2
3:21983057	ENST00000448980.5	ZNF385D-AS2	SGBS 1, SGBS2
3:52881568	ENST00000624222.1	RP11-894J14.2	SGBS 1, SGBS2
3:66118284	ENST00000354883.10	SLC25A26	SGBS 1, SGBS2
4:100064426	ENST00000505590.5	ADH4	SGBS 1, SGBS2
4:100064431	ENST00000629236.2 ENST00000265512.11	ADH4	SGBS 1, SGBS2
4:100064447	,ENST00000508393.5	ADH4	SGBS 1, SGBS2
4:113193471	ENST00000610220.1	RP11-73K9.3	SGBS 1, SGBS2
4:119273157	ENST00000296498.3	PRSS12	SGBS 1, SGBS2
4:129521554	ENST00000487657.2	RP11-184M15.2	SGBS 1, SGBS2
4:145955283	ENST00000410662.1	RN7SKP23S	SGBS 1, SGBS2
4:154177554	ENST00000482578.2	TRIM2	SGBS 1, SGBS2
4:157606815	ENST00000512609.1	RP11-171N4.4	SGBS 1, SGBS2
4:3370605	ENST00000338806.4	RGS12	SGBS 1, SGBS2
4:74863495	ENST00000296027.4	CXCL5	SGBS 1, SGBS2
4:81186752	ENST00000312465.11	FGF5	SGBS 1, SGBS2
4:81186792	ENST00000456523.3	FGF5	SGBS 1, SGBS2
5:10478482	ENST00000416930.2	LINC01513	SGBS 1, SGBS2
5:111332160	ENST00000560761.6	NREP	SGBS 1, SGBS2
5:112823526	ENST00000408903.6	MCC	SGBS 1, SGBS2
5:121464207	ENST00000296600.4	ZNF474	SGBS 1, SGBS2
5:127872295	ENST00000619499.4	FBN2	SGBS 1, SGBS2
5:127872513	ENST00000508989.5	FBN2	SGBS 1, SGBS2
5:127872734	ENST00000262464.8	FBN2	SGBS 1, SGBS2
5:127872915	ENST00000620257.1	FBN2	SGBS 1, SGBS2
5:135393537	ENST00000508076.5	TGFBI	SGBS 1, SGBS2
5:140704776	ENST00000606901.1	AC005618.6	SGBS 1, SGBS2
5:140704806	ENST00000606674.1	AC005618.6	SGBS 1, SGBS2
5:140867687	ENST00000252087.2	PCDHGC5	SGBS 1, SGBS2
5:140867733	ENST00000610789.1	PCDHGC5	SGBS 1, SGBS2
5:141999906	ENST00000359370.10	FGF1	SGBS 1, SGBS2
5:167717655	ENST00000265293.8	WWC1	SGBS 1, SGBS2
5:167718154	ENST00000521089.5	WWC1	SGBS 1, SGBS2
5:171985016	ENST00000517916.1	CTB-54I1.1	SGBS 1, SGBS2
5:19885380	ENST00000506372.5	CDH18	SGBS 1, SGBS2
5:19987293	ENST00000502796.5	CDH18	SGBS 1, SGBS2
5:19987315	ENST00000511273.1	CDH18	SGBS 1, SGBS2
5:19987338	ENST00000382275.5 ENST00000509140.5,E NST00000514088.5,E NST00000542609.5,E NST00000619676.4,E	CDH18	SGBS 1, SGBS2
5:35117223	NST00000620785.4	PRLR	SGBS 1, SGBS2
5:41793684	ENST00000510634.5	OXCT1	SGBS 1, SGBS2
5:72743186	ENST00000615637.1	FOXD1	SGBS 1, SGBS2
5:75918237	ENST00000504899.1	F2RL2	SGBS 1, SGBS2
5:75918258	ENST00000296641.4	F2RL2	SGBS 1, SGBS2
5:95544930	ENST00000384270.1	RNU6-524P	SGBS 1, SGBS2
6:132405223	ENST00000430428.1	RP3-523C21.2	SGBS 1, SGBS2
6:150345606	ENST00000367341.5	RAET1L	SGBS 1, SGBS2
6:150345667	ENST00000286380.2	RAET1L	SGBS 1, SGBS2
6:166400401	ENST00000581850.1	LINC00473	SGBS 1, SGBS2
6:166400409	ENST00000444465.1	LINC00473	SGBS 1, SGBS2
4:100060447	ENST00000265512.11 ,ENST00000508393.5	ADH4	SGBS 1, SGBS2
4:113189471	ENST00000610220.1	RP11-73K9.3	SGBS 1, SGBS2
4:119269157	ENST00000296498.3	PRSS12	SGBS 1, SGBS2
4:129517554	ENST00000487657.2	RP11-184M15.2	SGBS 1, SGBS2
4:145951283	ENST00000410662.1	RN7SKP23S	SGBS 1, SGBS2
4:154173554	ENST00000482578.2	TRIM2	SGBS 1, SGBS2
4:157602815	ENST00000512609.1	RP11-171N4.4	SGBS 1, SGBS2
4:3366605	ENST00000338806.4	RGS12	SGBS 1, SGBS2
4:74859495	ENST00000296027.4	CXCL5	SGBS 1, SGBS2
4:81182752	ENST00000312465.11	FGF5	SGBS 1, SGBS2
4:81182792	ENST00000456523.3	FGF5	SGBS 1, SGBS2
5:10474482	ENST00000416930.2	LINC01513	SGBS 1, SGBS2
5:111328160	ENST00000450761.6	NREP	SGBS 1, SGBS2
5:111328733	ENST00000610184.1	RP11-542K23.10	SGBS 1, SGBS2
5:112819526	ENST00000408903.6	MCC	SGBS 1, SGBS2
5:121460207	ENST00000296600.4	ZNF474	SGBS 1, SGBS2
5:127868295	ENST00000619499.4	FBN2	SGBS 1, SGBS2
5:127868513	ENST00000508989.5	FBN2	SGBS 1, SGBS2
5:127868734	ENST00000262464.8	FBN2	SGBS 1, SGBS2
5:127868915	ENST00000620257.1	FBN2	SGBS 1, SGBS2
5:140863687	ENST00000252087.2	PCDHGC5	SGBS 1, SGBS2
5:140863733	ENST00000610789.1	PCDHGC5	SGBS 1, SGBS2
5:171981016	ENST00000517916.1	CTB-54I1.1	SGBS 1, SGBS2
5:19881380	ENST00000506372.5	CDH18	SGBS 1, SGBS2
5:19983293	ENST00000502796.5	CDH18	SGBS 1, SGBS2
5:19983315	ENST00000511273.1	CDH18	SGBS 1, SGBS2
5:19983338	ENST00000382275.5 ENST00000509140.5, ENST00000514088.5, ENST00000542609.5, ENST00000619676.4,	CDH18	SGBS 1, SGBS2
5:35113223	ENST00000620785.4	PRLR	SGBS 1, SGBS2
5:41789684	ENST00000510634.5	OXCT1	SGBS 1, SGBS2
5:72739186	ENST00000615637.1	FOXD1	SGBS 1, SGBS2
5:75914237	ENST00000504899.1	F2RL2	SGBS 1, SGBS2
5:75914258	ENST00000296641.4	F2RL2	SGBS 1, SGBS2
5:95540930	ENST00000384270.1	RNU6-524P	SGBS 1, SGBS2
6:132401223	ENST00000430428.1	RP3-523C21.2	SGBS 1, SGBS2
6:150341606	ENST00000367341.5	RAET1L	SGBS 1, SGBS2
6:150341667	ENST00000286380.2	RAET1L	SGBS 1, SGBS2
6:166396401	ENST00000581850.1	LINC00473	SGBS 1, SGBS2
6:166396409	ENST00000444465.1	LINC00473	SGBS 1, SGBS2
6:166396478	ENST00000584179.1	LINC00473	SGBS 1, SGBS2
6:166396535	ENST00000584911.1	LINC00473	SGBS 1, SGBS2
6:28758811	ENST00000577570.1	AL662890.3	SGBS 1, SGBS2
6:5049656	ENST00000606227.1	RP11-428J1.4	SGBS 1, SGBS2
6:56502657	ENST00000518935.5	DST	SGBS 1, SGBS2
6:56502693	ENST00000370765.10	DST	SGBS 1, SGBS2
6:56502793	ENST00000244364.10	DST	SGBS 1, SGBS2
7:114866408	ENST00000419883.1	LINC01392	SGBS 1, SGBS2
7:137023519	ENST00000393083.2	PTN	SGBS 1, SGBS2
7:137023610	ENST00000348225.6	PTN	SGBS 1, SGBS2
7:143627258	ENST00000408955.2	OR2F2	SGBS 1, SGBS2
7:224556	ENST00000514988.1	AC145676.2	SGBS 1, SGBS2
7:23431064	ENST00000390962.1	SNORD65	SGBS 1, SGBS2
7:30003682	ENST00000631335.1	SCRN1	SGBS 1, SGBS2
7:30003886	ENST00000409497.5	SCRN1	SGBS 1, SGBS2
7:50855665	ENST00000403097.5	GRB10	SGBS 1, SGBS2
7:50856132	ENST00000402497.5	GRB10	SGBS 1, SGBS2
7:50856158	ENST00000335866.7	GRB10	SGBS 1, SGBS2
7:75396574	ENST00000505180.8	CCL26	SGBS 1, SGBS2
7:75711761	ENST00000515989.1	RNU6-863P	SGBS 1, SGBS2
8:102133385	ENST00000524369.1	KB-1460A1.1	SGBS 1, SGBS2
8:125278923	ENST00000517482.1	RP11-383J24.1	SGBS 1, SGBS2
8:16983689	ENST00000513892.2	RP11-468H14.2	SGBS 1, SGBS2
8:18739543	ENST00000523619.5	PSD3	SGBS 1, SGBS2
8:32618642	ENST00000607314.1	RP11-1002K11.1	SGBS 1, SGBS2
8:37346430	ENST00000522718.5	LINC01605	SGBS 1, SGBS2
8:40750343	ENST00000297737.10	ZMAT4	SGBS 1, SGBS2
8:40750351	ENST00000315769.11 ENST00000289734.11	ZMAT4	SGBS 1, SGBS2
8:41650139	,ENST00000347528.8	ANK1	SGBS 1, SGBS2
8:68653577	ENST00000518549.1	CPA6	SGBS 1, SGBS2
8:68653579	ENST00000297770.8	CPA6	SGBS 1, SGBS2
8:69532693	ENST00000325233.3 ENST00000260128.8,	C8orf34	SGBS 1, SGBS2
8:70373858	ENST00000458141.6	SULF1	SGBS 1, SGBS2
8:74263695	ENST00000514599.5	RDH10-AS1	SGBS 1, SGBS2
9:102126195	ENST00000605377.5	NAMA	SGBS 1, SGBS2
9:116322233	ENST00000342620.9	RGS3	SGBS 1, SGBS2
9:119046047	ENST00000438048.1	PAPPA-AS2	SGBS 1, SGBS2
9:125160184	ENST00000439471.2	RP11-542K23.10	SGBS 1, SGBS2
9:132510325	ENST00000340607.4	PTGES	SGBS 1, SGBS2
9:132929856	ENST00000372398.5	NCS1	SGBS 1, SGBS2
9:21554667	ENST00000304425.3	MIR31HG	SGBS 1, SGBS2
9:32640023	ENST00000413291.1	RP11-555J4.3	SGBS 1, SGBS2
9:71814926	ENST00000535702.5	TJP2	SGBS 1, SGBS2
9:71815077	ENST00000539225.1	TJP2	SGBS 1, SGBS2
X:40117169	ENST00000452690.1	RP11-320G24.1	SGBS 1, SGBS2
X:51234447	ENST00000375992.3	NUDT11	SGBS 1, SGBS2
X:77909824	ENST00000321110.1	ZCCHC5	SGBS 1, SGBS2
X:96867257	ENST00000445414.1	DIAPH2-AS1	SGBS 1, SGBS2

6:32633460	ENST00000374943.8	HLA-DQB1	Tissue 2, Tissue 3
6:32633470	ENST00000434651.6	HLA-DQB1	Tissue 2, Tissue 3
6:32635159	ENST00000399084.5	HLA-DQB1	Tissue 2, Tissue 3
6:32783824	ENST00000438763.6	HLA-DOB	Tissue 1
6:46870207	ENST00000451135.1	RP3-365O12.2	Tissue 3
6:62995131	ENST00000281156.4	KHDRBS2	Tissue 1, Tissue 2, Tissue 3
6:6587340	ENST00000379953.6	LY86	Tissue 1, Tissue 3
6:74103470	ENST00000370336.4	DDX43	Tissue 1, Tissue 2, Tissue 3
6:74103815	ENST00000370363.5	OOEP	Tissue 1, Tissue 2, Tissue 3
7:139527951	ENST00000414508.6	TBXAS1	Tissue 1
7:139528022	ENST00000448866.5	TBXAS1	Tissue 1
	ENST00000455353.5,EN		
7:139528052	ST00000458722.5	TBXAS1	Tissue 1
7:139528062	ENST00000411653.5	TBXAS1	Tissue 1
7:150037748	ENST00000482669.1	RARRES2	Tissue 1, Tissue 2, Tissue 3
7:150037762	ENST00000223271.7	RARRES2	Tissue 1, Tissue 2, Tissue 3
7:150037859	ENST00000563946.1	RP4-584D14.7	Tissue 1, Tissue 2, Tissue 3
7:150146717	ENST00000307271.3	GIMAP8	Tissue 1, Tissue 2, Tissue 3
7:150263364	ENST00000255945.3	GIMAP4	Tissue 1, Tissue 3
7:150263456	ENST00000461940.5	GIMAP4	Tissue 1, Tissue 3
7:150412644	ENST00000307194.5	GIMAP1	Tissue 1, Tissue 2, Tissue 3
7:150412755	ENST00000611999.4	GIMAP1-GIMAP5	Tissue 1, Tissue 2, Tissue 3
7:150418340	ENST00000498181.6	GIMAP5	Tissue 1, Tissue 2, Tissue 3
7:153108318	ENST00000454441.2	LINC01287	Tissue 2
7:157646276	ENST00000438047.1	AC011899.9	Tissue 1, Tissue 3
7:30600131	ENST00000578293.6	AC005154.6	Tissue 1
7:73792271	ENST00000442504.5	ELMO1	Tissue 1, Tissue 2, Tissue 3
7:50343323	ENST00000413698.5	IKZF1	Tissue 1
	ENST00000331340.7,EN ST00000343574.9,ENST 00000346667.8,ENST00 000349824.8,ENST0000 0357364.8,ENST000003 59197.9,ENST00000440 768.6,ENST0000061265		
7:50343377	8.4,ENST00000615491.4	IKZF1	Tissue 1
7:57471422	ENST00000624231.1	RP11-114G11.4	Tissue 1, Tissue 2, Tissue 3
7:63360200	ENST00000450544.1	RP11-340I6.8	Tissue 1, Tissue 2, Tissue 3
7:63360445	ENST00000585312.1	AC092634.1	Tissue 1, Tissue 2, Tissue 3
7:63386213	ENST00000582948.1	AC092634.2	Tissue 1, Tissue 2, Tissue 3
	ENST00000456806.2,EN		
7:63504820	ST00000550760.3	ZNF727	Tissue 1, Tissue 2, Tissue 3
7:75240019	ENST00000616821.4	HIP1	Tissue 1, Tissue 2, Tissue 3
	ENST00000433696.6,EN ST00000538969.5,ENST		
7:80266831	00000544133.5	CD36	Tissue 1, Tissue 2, Tissue 3
	ENST00000394788.7,EN		
7:80266991	ST00000447544.6	CD36	Tissue 1, Tissue 2, Tissue 3
8:101314486	ENST00000519449.5	RNF19A	Tissue 1, Tissue 2
8:112387867	ENST00000521904.5	RP11-1101K5.1	Tissue 1, Tissue 2, Tissue 3
8:112387891	ENST00000521753.5	RP11-1101K5.1	Tissue 1, Tissue 2, Tissue 3
8:144294067	ENST00000622500.1	GPIHBP1	Tissue 1, Tissue 2, Tissue 3
8:22784420	ENST00000256404.7	PEBP4	Tissue 1, Tissue 3
8:27467873	ENST00000523500.5	CLU	Tissue 1, Tissue 2, Tissue 3
8:27467947	ENST00000405140.7	CLU	Tissue 1, Tissue 2, Tissue 3
8:40009973	ENST00000315792.4	C8orf4	Tissue 1, Tissue 2, Tissue 3
8:56432535	ENST00000522918.1	RP11-628E19.2	Tissue 2, Tissue 3
8:61879333	ENST00000525556.1	AC022182.3	Tissue 1, Tissue 3
8:74963574	ENST00000519323.1	RP11-6I2.3	Tissue 1
8:93680115	ENST00000521307.1	RP11-100L22.1	Tissue 1, Tissue 2, Tissue 3
	ENST00000343817.9,EN		
9:116262706	ST00000394646.7	RGS3	Tissue 1, Tissue 3
9:130523696	ENST00000429553.5	SH2D3C	Tissue 1, Tissue 2, Tissue 3
9:133768224	ENST00000343079.1	QRFP	Tissue 1
9:134144879	ENST00000372269.7	FAM78A	Tissue 1, Tissue 2, Tissue 3
9:139556378	ENST00000308874.11	EGFL7	Tissue 1, Tissue 2, Tissue 3
9:139556381	ENST00000406555.7	EGFL7	Tissue 1, Tissue 2, Tissue 3
9:139926461	ENST00000314412.6	FUT7	Tissue 1
9:272001	ENST00000429661.1	RP11-59O6.3	Tissue 1
9:66535443	ENST00000669749.1	RP11-262H14.3	Tissue 2, Tissue 3
9:92253728	ENST00000617436.1	UNQ6494	Tissue 1, Tissue 2, Tissue 3
	ENST00000375751.8,EN		
9:93563068	ST00000375754.8	SYK	Tissue 1
9:93563208	ENST00000375747.5	SYK	Tissue 1
X:136647300	ENST00000287538.9	ZIC3	Tissue 1, Tissue 2, Tissue 3
	ENST00000377934.4,EN		
X:45059145	ST00000398000.6	CXorf36	Tissue 1, Tissue 2, Tissue 3
X:50385682	ENST00000460112.3	SHROOM4	Tissue 1, Tissue 2, Tissue 3
X:78621855	ENST00000373298.6	ITM2A	Tissue 1, Tissue 2, Tissue 3
X:78622048	ENST00000434584.2	ITM2A	Tissue 1, Tissue 2, Tissue 3

6:166400478	ENST00000584179.1	LINC00473	SGBS 1, SGBS2
6:166400535	ENST00000584911.1	LINC00473	SGBS 1, SGBS2
6:28762811	ENST00000577570.1	AL662890.3	SGBS 1, SGBS2
6:31081607	ENST00000259881.9	PSORS1C1	SGBS 1, SGBS2
6:41884263	ENST00000384641.1	Y_RNA	SGBS 1, SGBS2
6:5053656	ENST00000606227.1	RP11-428J1.4	SGBS 1, SGBS2
6:56506657	ENST00000518935.5	DST	SGBS 1, SGBS2
6:56506693	ENST00000370765.10	DST	SGBS 1, SGBS2
6:56506793	ENST00000244364.10	DST	SGBS 1, SGBS2
6:71947300	ENST00000612512.1	RP11-154D6.1	SGBS 1, SGBS2
7:114870408	ENST00000419883.1	LINC01392	SGBS 1, SGBS2
7:137027519	ENST00000393083.2	PTN	SGBS 1, SGBS2
7:137027610	ENST00000348225.6	PTN	SGBS 1, SGBS2
7:142510083	ENST00000417977.2	TRBV30	SGBS 1, SGBS2
7:143631258	ENST00000408955.2	OR2F2	SGBS 1, SGBS2
7:228556	ENST00000514988.1	AC145676.2	SGBS 1, SGBS2
7:23435064	ENST00000390962.1	SNORD65	SGBS 1, SGBS2
7:30007682	ENST00000631335.1	SCRN1	SGBS 1, SGBS2
7:30007886	ENST00000409497.5	SCRN1	SGBS 1, SGBS2
7:30959914	ENST00000413284.6	AQP1	SGBS 1, SGBS2
7:30960048	ENST00000409899.5	AQP1	SGBS 1, SGBS2
7:30960104	ENST00000409611.1	AQP1	SGBS 1, SGBS2
7:50859665	ENST00000403097.5	GRB10	SGBS 1, SGBS2
7:50860132	ENST00000402497.5	GRB10	SGBS 1, SGBS2
7:50860158	ENST00000335866.7	GRB10	SGBS 1, SGBS2
7:75400574	ENST00000005180.8	CCL26	SGBS 1, SGBS2
7:75715761	ENST00000515989.1	RNU6-863P	SGBS 1, SGBS2
8:102137385	ENST00000524369.1	KB-1460A1.1	SGBS 1, SGBS2
8:125282923	ENST00000517482.1	RP11-383J24.1	SGBS 1, SGBS2
8:129439161	ENST00000520206.5	LINC00824	SGBS 1, SGBS2
8:16987689	ENST00000513892.2	RP11-468H14.2	SGBS 1, SGBS2
8:18743543	ENST00000523619.5	PSD3	SGBS 1, SGBS2
8:32622642	ENST00000607314.1	RP11-1002K11.1	SGBS 1, SGBS2
8:37350430	ENST00000522718.5	LINC01605	SGBS 1, SGBS2
8:40754343	ENST00000297737.10	ZMAT4	SGBS 1, SGBS2
8:40754351	ENST00000315769.11 ENST00000289734.11	ZMAT4	SGBS 1, SGBS2
8:41654139	,ENST00000347528.8	ANK1	SGBS 1, SGBS2
8:68657577	ENST00000518549.1	CPA6	SGBS 1, SGBS2
8:68657579	ENST00000297770.8	CPA6	SGBS 1, SGBS2
8:69536693	ENST00000325233.3 ENST00000260128.8,E	C8orf34	SGBS 1, SGBS2
8:70377858	NST00000458141.6	SULF1	SGBS 1, SGBS2
8:74267695	ENST00000514599.5	RDH10-AS1	SGBS 1, SGBS2
9:102130195	ENST00000605377.5	NAMA	SGBS 1, SGBS2
9:116326233	ENST00000342620.9	RGS3	SGBS 1, SGBS2
9:119050047	ENST00000438048.1	PAPPA-AS2	SGBS 1, SGBS2
9:125164184	ENST00000439471.2	RP11-542K23.10	SGBS 1, SGBS2
9:132514325	ENST00000340607.4	PTGES	SGBS 1, SGBS2
9:132933856	ENST00000372398.5	NCS1	SGBS 1, SGBS2
9:21558667	ENST00000304425.3	MIR31HG	SGBS 1, SGBS2
9:32644023	ENST00000413291.1	RP11-55SJ4.3	SGBS 1, SGBS2
9:6565424	ENST00000516301.1	snoU13	SGBS 1, SGBS2
9:71818926	ENST00000535702.5	TJP2	SGBS 1, SGBS2
9:71819077	ENST00000539225.1	TJP2	SGBS 1, SGBS2
X:40121169	ENST00000452690.1	RP11-320G24.1	SGBS 1, SGBS2
X:51238447	ENST00000375992.3	NUDT11	SGBS 1, SGBS2
X:77913824	ENST00000321110.1	ZCCHC5	SGBS 1, SGBS2
X:96871257	ENST00000445414.1	DIAPH2-AS1	SGBS 1, SGBS2

Table 24: GWAS loci used in GREGOR enrichment analyses

Trait category	Number of loci	Number of variants	Traits included
Cardiovascular outcomes	76	2049	Coronary artery disease, coronary heart disease, large artery atherosclerosis, myocardial infarction, stroke
HDL cholesterol	250	5084	HDL cholesterol, HDL cholesterol:triglycerides
Blood pressure traits	57	2184	Diastolic blood pressure, hypertension, mean arterial pressure, pulse pressure, systolic blood pressure
Body mass index	91	3188	Body mass index
Type 2 diabetes	154	3046	Type 2 diabetes
LDL cholesterol	291	3941	LDL cholesterol
Waist to hip ratio adjusted for BMI	64	1385	Waist to hip ratio adjusted for BMI
Insulin	31	735	Fasting insulin, fasting insulin adjusted for BMI
Triglycerides	159	2753	Triglycerides
Adiponectin	21	424	Adiponectin
Total cholesterol	231	3801	Total cholesterol
Glucose	63	1135	1 hour glucose, 2 hour glucose, 2 hour glucose adjusted for BMI, fasting glucose, fasting glucose adjusted for BMI

GWAS loci are from the GWAS catalog (www.ebi.ac.uk/gwas, accessed December 2016)

Table 25: Enrichment of GWAS loci in ATAC peaks using GREGOR

Trait	Sample	Overlaps	Expected overlaps	Standard Deviation	Fold change	Z score	P value	Total loci in trait category	Percent loci in ATAC peaks	Total variants	Number of variants in ATAC peaks	Percent of variants in ATAC peaks
Adiponectin	Adipose tissue all peaks	8	2.46	1.39	3.25	3.99	8.03E-04	21	38.1	424	12	2.8
Adiponectin	Adipose tissue enhancer peaks	5	0.82	0.88	6.07	4.76	8.66E-04	21	23.8	424	8	1.9
Adiponectin	Adipose tissue promoter peaks	2	1.05	0.96	1.91	0.99	2.82E-01	21	9.5	424	3	0.7
Adiponectin	Roadmap adipose enhancers	20	5.66	1.92	3.53	7.47	4.94E-12	21	95.2	424	70	16.5
Adiponectin	Roadmap adipose promoters	7	2.91	1.5	2.41	2.72	1.41E-02	21	33.3	424	18	4.2
Adiponectin	SGBS all peaks	13	5.99	1.86	2.17	3.78	4.37E-04	21	61.9	424	21	5.0
Adiponectin	SGBS enhancer peaks	9	1.28	1.07	7.06	7.22	5.78E-07	21	42.9	424	11	2.6
Adiponectin	SGBS promoter peaks	4	1.42	1.1	2.82	2.35	4.11E-02	21	19.0	424	4	0.9
Blood pressure traits	Adipose tissue all peaks	25	10.45	2.69	2.39	5.41	9.96E-07	57	43.9	2184	36	1.6
Blood pressure traits	Adipose tissue enhancer peaks	13	3.75	1.79	3.47	5.16	2.82E-05	57	22.8	2184	11	0.5
Blood pressure traits	Adipose tissue promoter peaks	16	5.16	2.03	3.1	5.34	6.45E-06	57	28.1	2184	17	0.8
Blood pressure traits	Roadmap adipose enhancers	38	19.84	3.33	1.92	5.45	1.23E-07	57	66.7	2184	162	7.4
Blood pressure traits	Roadmap adipose promoters	21	11.45	2.77	1.83	3.44	1.00E-03	57	36.8	2184	51	2.3
Blood pressure traits	SGBS all peaks	32	22.86	3.28	1.4	2.78	4.57E-03	57	56.1	2184	110	5.0
Blood pressure traits	SGBS enhancer peaks	13	5.02	2.08	2.59	3.84	7.70E-04	57	22.8	2184	23	1.1
Blood pressure traits	SGBS promoter peaks	16	6.57	2.25	2.44	4.2	1.77E-04	57	28.1	2184	19	0.9
Body mass index	Adipose tissue all peaks	19	15.31	3.16	1.24	1.17	1.56E-01	91	20.9	3188	39	1.2
Body mass index	Adipose tissue enhancer peaks	3	5.28	2.15	0.57	-1.06	9.12E-01	91	3.3	3188	4	0.1
Body mass index	Adipose tissue promoter peaks	15	7.66	2.4	1.96	3.05	4.08E-03	91	16.5	3188	27	0.8
Body mass index	Roadmap adipose enhancers	33	29.62	4.11	1.11	0.82	2.40E-01	91	36.3	3188	181	5.7
Body mass index	Roadmap adipose promoters	27	16.6	3.3	1.63	3.16	2.05E-03	91	29.7	3188	80	2.5
Body mass index	SGBS all peaks	38	32.5	4.03	1.17	1.37	1.08E-01	91	41.8	3188	97	3.0
Body mass index	SGBS enhancer peaks	9	7.58	2.51	1.19	0.56	3.42E-01	91	9.9	3188	11	0.3
Body mass index	SGBS promoter peaks	15	9.7	2.67	1.55	1.98	4.13E-02	91	16.5	3188	35	1.1
Cardiometabolic outcomes	Adipose tissue all peaks	26	12.91	2.92	2.01	4.47	3.49E-05	76	34.2	2049	40	2.0
Cardiometabolic outcomes	Adipose tissue enhancer peaks	16	4.37	1.95	3.66	5.97	1.76E-06	76	21.1	2049	19	0.9
Cardiometabolic outcomes	Adipose tissue promoter peaks	10	6.84	2.28	1.46	1.38	1.24E-01	76	13.2	2049	16	0.8
Cardiometabolic outcomes	Roadmap adipose enhancers	49	25.91	3.83	1.89	6.03	5.60E-09	76	64.5	2049	267	13.0
Cardiometabolic outcomes	Roadmap adipose promoters	21	15.41	3.09	1.36	1.81	5.34E-02	76	27.6	2049	67	3.3
Cardiometabolic outcomes	SGBS all peaks	39	27.49	3.65	1.42	3.15	1.54E-03	76	51.3	2049	129	6.3
Cardiometabolic outcomes	SGBS enhancer peaks	20	7	2.39	2.86	5.43	3.45E-06	76	26.3	2049	41	2.0
Cardiometabolic outcomes	SGBS promoter peaks	12	8.45	2.51	1.42	1.42	1.14E-01	76	15.8	2049	28	1.4
Glucose	Adipose tissue all peaks	15	8.98	2.6	1.67	2.31	2.14E-02	63	23.8	1135	21	1.9
Glucose	Adipose tissue enhancer peaks	5	2.97	1.65	1.68	1.23	1.71E-01	63	7.9	1135	5	0.4
Glucose	Adipose tissue promoter peaks	10	4.15	1.87	2.41	3.14	5.08E-03	63	15.9	1135	13	1.1
Glucose	Roadmap adipose enhancers	36	20.22	3.52	1.78	4.48	1.43E-05	250	14.0	5084	48	0.9
Glucose	Roadmap adipose promoters	22	10.38	2.78	2.12	4.17	1.23E-04	63	34.9	1135	40	3.5
Glucose	SGBS all peaks	29	21.33	3.43	1.36	2.24	1.97E-02	63	46.0	1135	57	5.0
Glucose	SGBS enhancer peaks	5	5.08	2.11	0.98	-0.04	5.85E-01	63	7.9	1135	7	0.6
Glucose	SGBS promoter peaks	12	5.27	2.09	2.28	3.23	3.48E-03	63	19.0	1135	19	1.7
HDL cholesterol	Adipose tissue all peaks	78	36.77	5.18	2.12	7.96	4.92E-13	250	31.2	5084	90	1.8
HDL cholesterol	Adipose tissue enhancer peaks	42	12.34	3.33	3.4	8.92	6.17E-13	250	16.8	5084	35	0.7
HDL cholesterol	Adipose tissue promoter peaks	24	17.82	3.82	1.35	1.62	7.26E-02	250	9.6	5084	36	0.7
HDL cholesterol	Roadmap adipose enhancers	167	79.42	6.88	2.1	12.72	3.62E-34	250	66.8	5084	720	14.2
HDL cholesterol	Roadmap adipose promoters	80	43.23	5.52	1.85	6.67	5.04E-10	250	32.0	5084	136	2.7
HDL cholesterol	SGBS all peaks	122	86.54	6.66	1.41	5.32	1.42E-07	250	48.8	5084	218	4.3
HDL cholesterol	SGBS enhancer peaks	35	19.21	4.07	1.82	3.88	2.73E-04	63	55.6	1135	144	12.7
HDL cholesterol	SGBS promoter peaks	38	23.45	4.32	1.62	3.37	1.09E-03	250	15.2	5084	59	1.2
Insulin	Adipose tissue all peaks	14	4.97	1.9	2.82	4.75	3.21E-05	31	45.2	735	19	2.6
Insulin	Adipose tissue enhancer peaks	9	1.65	1.23	5.47	5.99	1.32E-05	31	29.0	735	14	1.9
Insulin	Adipose tissue promoter peaks	0	NA	1.39805	NA	NA	NA	31	NA	735	0	NA
Insulin	Roadmap adipose enhancers	23	10.56	2.47	2.18	5.03	1.33E-06	31	74.2	735	142	19.3
Insulin	Roadmap adipose promoters	5	6.3	2.05	0.79	-0.63	8.08E-01	31	16.1	735	12	1.6
Insulin	SGBS all peaks	20	11.44	2.44	1.75	3.51	5.87E-04	31	64.5	735	33	4.5
Insulin	SGBS enhancer peaks	7	2.76	1.54	2.54	2.76	1.41E-02	31	22.6	735	11	1.5
Insulin	SGBS promoter peaks	1	3.2	1.59	0.31	-1.39	9.73E-01	31	3.2	735	1	0.1
LDL cholesterol	Adipose tissue all peaks	54	34.04	5.05	1.59	3.95	1.42E-04	291	18.6	3941	58	1.5
LDL cholesterol	Adipose tissue enhancer peaks	11	11.53	3.23	0.95	-0.16	6.10E-01	291	3.8	3941	12	0.3
LDL cholesterol	Adipose tissue promoter peaks	35	15.99	3.68	2.19	5.17	3.83E-06	291	12.0	3941	33	0.8
LDL cholesterol	Roadmap adipose enhancers	118	81.97	7.2	1.44	5	8.85E-07	291	40.2	3941	327	8.3
LDL cholesterol	Roadmap adipose promoters	76	42.41	5.58	1.79	6.02	1.93E-08	291	26.1	3941	138	3.5
LDL cholesterol	SGBS all peaks	98	82.31	6.85	1.19	2.29	1.42E-02	291	33.7	3941	180	4.6
LDL cholesterol	SGBS enhancer peaks	25	17.84	3.96	1.4	1.81	5.14E-02	291	8.6	3941	29	0.7
LDL cholesterol	SGBS promoter peaks	41	21.99	4.22	1.86	4.5	2.93E-05	291	14.1	3941	52	1.3
Total cholesterol	Adipose tissue all peaks	55	32.63	4.75	1.69	4.71	8.06E-06	231	23.8	3801	56	1.5
Total cholesterol	Adipose tissue enhancer peaks	10	11.07	3.14	0.9	-0.34	6.80E-01	231	4.3	3801	8	0.2
Total cholesterol	Adipose tissue promoter peaks	33	15.39	3.53	2.14	4.98	6.96E-06	231	14.3	3801	34	0.9
Total cholesterol	Roadmap adipose enhancers	101	69.46	6.51	1.45	4.84	1.84E-06	231	43.3	3801	346	9.1
Total cholesterol	Roadmap adipose promoters	68	38.71	5.17	1.76	5.67	9.71E-08	231	29.4	3801	154	4.1
Total cholesterol	SGBS all peaks	96	73.05	6.2	1.31	3.7	1.95E-04	231	41.6	3801	194	5.1
Total cholesterol	SGBS enhancer peaks	25	17.04	3.81	1.47	2.09	2.98E-02	231	10.8	3801	31	0.8
Total cholesterol	SGBS promoter peaks	40	20.65	4.03	1.94	4.81	9.12E-06	231	17.3	3801	59	1.6
Triglycerides	Adipose tissue all peaks	64	27.45	4.3	2.33	8.51	1.81E-14	159	40.3	2753	57	2.1
Triglycerides	Adipose tissue enhancer peaks	28	9.53	2.88	2.94	6.4	7.70E-08	159	17.6	2753	22	0.8
Triglycerides	Adipose tissue promoter peaks	22	13.49	3.26	1.63	2.61	9.85E-03	159	13.8	2753	18	0.7
Triglycerides	Roadmap adipose enhancers	115	54.22	5.58	2.12	10.9	3.70E-26	159	72.3	2753	390	14.2
Triglycerides	Roadmap adipose promoters	58	31.27	4.63	1.86	5.78	5.52E-08	159	36.5	2753	88	3.2
Triglycerides	SGBS all peaks	77	59.35	5.41	1.3	3.26	8.49E-04	159	48.4	2753	127	4.6
Triglycerides	SGBS enhancer peaks	28	13.82	3.44	2.03	4.12	1.55E-04	159	17.6	2753	35	1.3

Triglycerides	SGBS promoter peaks	25	17.4	3.66	1.44	2.08	3.01E-02	159	15.7	2753	26	0.9
Type 2 diabetes	Adipose tissue all peaks	39	22.17	3.92	1.76	4.3	5.09E-05	154	25.3	3046	48	1.6
Type 2 diabetes	Adipose tissue enhancer peaks	15	7.13	2.53	2.1	3.1	4.30E-03	154	9.7	3046	13	0.4
Type 2 diabetes	Adipose tissue promoter peaks	12	11.25	2.9	1.07	0.26	4.51E-01	154	7.8	3046	19	0.6
Type 2 diabetes	Roadmap adipose enhancers	72	48.74	5.37	1.48	4.33	1.93E-05	154	46.8	3046	403	13.2
Type 2 diabetes	Roadmap adipose promoters	33	26.75	4.2	1.23	1.49	8.74E-02	154	21.4	3046	86	2.8
Type 2 diabetes	SGBS all peaks	73	50.29	5.11	1.45	4.44	1.22E-05	154	47.4	3046	170	5.6
Type 2 diabetes	SGBS enhancer peaks	24	11.8	3.15	2.03	3.88	3.76E-04	154	15.6	3046	44	1.4
Type 2 diabetes	SGBS promoter peaks	18	14.9	3.28	1.21	0.95	2.10E-01	154	11.7	3046	34	1.1
Waist to hip ratio adjusted for BMI	Adipose tissue all peaks	26	7.37	2.39	3.53	7.79	1.16E-10	64	40.6	1385	45	3.2
Waist to hip ratio adjusted for BMI	Adipose tissue enhancer peaks	19	2.59	1.54	7.33	10.63	6.43E-13	64	29.7	1385	32	2.3
Waist to hip ratio adjusted for BMI	Adipose tissue promoter peaks	4	3.25	1.67	1.23	0.45	4.11E-01	64	6.3	1385	6	0.4
Waist to hip ratio adjusted for BMI	Roadmap adipose enhancers	44	18.03	3.38	2.44	7.69	9.42E-13	64	68.8	1385	294	21.2
Waist to hip ratio adjusted for BMI	Roadmap adipose promoters	22	9.2	2.58	2.39	4.96	9.21E-06	64	34.4	1385	65	4.7
Waist to hip ratio adjusted for BMI	SGBS all peaks	36	19.38	3.28	1.86	5.07	1.43E-06	64	56.3	1385	124	9.0
Waist to hip ratio adjusted for BMI	SGBS enhancer peaks	25	4.09	1.89	6.12	11.09	1.15E-15	64	39.1	1385	52	3.8
Waist to hip ratio adjusted for BMI	SGBS promoter peaks	9	4.49	1.93	2	2.33	2.64E-02	64	14.1	1385	23	1.7

Table 26: GWAS loci and variants overlapping transcription factor motifs

Sample	Trait	Total loci in trait category	Number of loci in ATAC peaks	Loci with overlapping TF motifs	Percent of loci overlapping TF motifs	Variants in ATAC peaks	Variants in ATAC peaks overlapping TF motif	Percent of variants in ATAC peaks overlapping TF motif
Adipose tissue	Waist to hip ratio adjusted for BMI	64	26	23	88.5	45	34	75.6
Adipose tissue	Body mass index	91	19	13	68.4	39	23	59.0
Adipose tissue	Adiponectin	21	8	6	75.0	12	6	50.0
Adipose tissue	Type 2 diabetes	154	39	28	71.8	48	28	58.3
Adipose tissue	Insulin	31	14	14	100.0	19	14	73.7
Adipose tissue	Glucose	63	15	13	86.7	21	15	71.4
Adipose tissue	HDL cholesterol	250	78	61	78.2	90	62	68.9
Adipose tissue	LDL cholesterol	291	54	43	79.6	58	38	65.5
Adipose tissue	Total cholesterol	231	55	46	83.6	56	38	67.9
Adipose tissue	Triglycerides	159	64	33	51.6	57	33	57.9
Adipose tissue	Cardiovascular outcomes	76	26	21	80.8	40	25	62.5
Adipose tissue	Blood pressure traits	57	25	16	64.0	36	22	61.1
Adipose tissue	All traits	1245	345	257	74.5	371	242	65.2
SGBS	Waist to hip ratio adjusted for BMI	64	36	33	91.7	124	71	57.3
SGBS	Body mass index	91	38	30	78.9	97	51	52.6
SGBS	Adiponectin	21	13	10	76.9	21	15	71.4
SGBS	Type 2 diabetes	154	73	63	86.3	170	90	52.9
SGBS	Insulin	31	20	12	60.0	33	16	48.5
SGBS	Glucose	63	29	23	79.3	57	33	57.9
SGBS	HDL cholesterol	250	122	101	82.8	218	132	60.6
SGBS	LDL cholesterol	291	98	81	82.7	180	99	55.0
SGBS	Total cholesterol	231	96	77	80.2	194	108	55.7
SGBS	Triglycerides	159	77	58	75.3	127	73	57.5
SGBS	Cardiovascular outcomes	76	39	39	100.0	129	74	57.4
SGBS	Blood pressure traits	57	32	26	81.3	110	58	52.7
SGBS	All traits	1245	575	476	82.8	1123	631	56.2

Table 27: Variants at GWAS-eQTL colocalized loci that overlap ATAC peaks

GWAS Trait	GWAS locus	GWAS index	eQTL Gene(s)	eSNP(s)	ATAC variant	MAF	r2 with GWAS	Total LD proxies at locus	ATAC tissue samples	Transcription factor motifs overlapping tissue peaks	Tissue Footprints	SGBS samples	Transcription factor motifs overlapping SGBS peaks	SGBS Footprints
Abdominal aortic aneurysm	*DAB2IP*	rs7025486	*DAB2IP*	rs885150	rs10818578	0.254191	0.989032	16				SGBS 1, SGBS 2		
Adiponectin	*ARL15*	rs4311394	*FST*	rs59061738	rs255757	0.234005	0.983153	85				SGBS 1, SGBS 2	STAT1,Stat4,ZNF410	
Adiponectin	*ARL15*	rs4311394	*FST*	rs59061738	rs6876198	0.233905	0.983699	85				SGBS 1	Stat6	
Adiponectin	*CDH13*	rs12051272	*CDH13*	rs12051272	rs12051272	0.106983	N/A	9				SGBS 1, SGBS 2	NFYA,NFYB,RREB1	RREB1
Adiponectin	*GNL3*	rs2590838	*GNL3,NEK4*	rs35212380,rs7612511	rs11108842	0.472473	1	21	Tissue 1, Tissue 2, Tissue 3			SGBS 1, SGBS 2		
Adiponectin,WHRadjBMI,HDL cholesterol	*CMIP*	rs2925979	*CMIP*	rs56823429	rs2925979	0.314007	N/A	2				SGBS 1, SGBS 2	DUX4	
Body mass index	*DMXL2*	rs3736485	*CYP19A1*	rs930920	rs11070860	0.470241	0.81589	194				SGBS 1		
Body mass index	*DMXL2*	rs3736485	*CYP19A1*	rs930920	rs2278989	0.47034	0.815551	194	Tissue 1			SGBS 1, SGBS 2		
Body mass index	*DMXL2*	rs3736485	*CYP19A1*	rs930920	rs2278990	0.47034	0.815551	194	Tissue 1	NFKB1,NFKB2		SGBS 1, SGBS 2	NFKB1,NFKB2	
Body mass index	*DMXL2*	rs3736485	*CYP19A1*	rs930920	rs7183479	0.463694	0.82174	194	Tissue 1			SGBS 1, SGBS 2		
Body mass index	*FOXO3*	rs9400239	*FOXO3*	rs3800228	rs4946936	0.386122	0.948583	48				SGBS 1, SGBS 2	Tcf7	
Body mass index	*FOXO3*	rs9400239	*FOXO3*	rs3800228	rs9400239	0.393562	N/A	48				SGBS 1, SGBS 2		
Body mass index	*INO80E*	rs4787491	*HIRIP3,INO80E,TBX6,*	rs4788211,rs7204797,rs9928448,	rs10871451	0.420792	0.891016	87	Tissue 1, Tissue 3	EOMES,Nr1h3::Rxra,Rarg(var.2)		SGBS 1	EOMES,Nr1h3::Rxra,Rarg(var.2)	
Body mass index	*INO80E*	rs4787491	*HIRIP3,INO80E,TBX6,*	rs4788211,rs7204797,rs9928448,	rs4318227	0.41975	0.894621	87	Tissue 1, Tissue 2, Tissue 3	NR2C2,NRF1		SGBS 1, SGBS 2	NR2C2	
Body mass index	*INO80E*	rs4787491	*TMEM219,YPEL3*	rs9972866	rs4788204	0.418659	0.899638	87				SGBS 1, SGBS 2		
Body mass index	*INO80E*	rs4787491	*HIRIP3,INO80E,TBX6,*	rs4788211,rs7204797,rs9928448,	rs6565173	0.436961	0.827771	87				SGBS 1, SGBS 2	IRF1	IRF1
Body mass index	*MTCH2*	rs10838738	*MTCH2*	rs11604825	rs2293578	0.294862	0.823474	99				SGBS 1, SGBS 2	EWSR1-FLI1,ZNF263	
Body mass index	*USP37*	rs492400	*PLCD4,RQCD1,ZNF142*	rs4674320,rs62182125,rs832810	2:219446813	0.425156	0.995336	262				SGBS 1, SGBS 2	FOXP1,IRF1	
Body mass index	*USP37*	rs492400	*PLCD4,RQCD1,ZNF142*	rs4674320,rs62182125,rs832810	rs2241527	0.398968	0.893145	262	Tissue 1			SGBS 1, SGBS 2		
Body mass index	*USP37*	rs492400	*PLCD4,RQCD1,ZNF142*	rs4674320,rs62182125,rs832810	rs2556391	0.398076	0.889009	262				SGBS 1, SGBS 2		
Body mass index	*USP37*	rs492400	*PLCD4,RQCD1,ZNF142*	rs4674320,rs62182125,rs832810	rs3731867	0.398919	0.893358	262				SGBS 1, SGBS 2		
Body mass index	*USP37*	rs492400	*PLCD4,RQCD1,ZNF142*	rs4674320,rs62182125,rs832810	rs4674319	0.427438	0.985675	262				SGBS 1	Vdr	
Body mass index	*USP37*	rs492400	*PLCD4,RQCD1,ZNF142*	rs4674320,rs62182125,rs832810	rs4674324	0.425305	0.994729	262				SGBS 1, SGBS 2		
Body mass index	*USP37*	rs492400	*PLCD4,RQCD1,ZNF142*	rs4674320,rs62182125,rs832810	rs500317	0.424958	0.994524	262	Tissue 1, Tissue 2, Tissue 3	GUS3,ZBTB33		SGBS 1, SGBS 2	ZBTB33	
Body mass index	*USP37*	rs492400	*PLCD4,RQCD1,ZNF142*	rs4674320,rs62182125,rs832810	rs500422	0.432546	0.966107	262	Tissue 1, Tissue 2, Tissue 3	IRF8,Spz1		SGBS 1, SGBS 2	IRF8	
Body mass index	*USP37*	rs492400	*PLCD4,RQCD1,ZNF142*	rs4674320,rs62182125,rs832810	rs524012	0.427487	0.984262	262	Tissue 1, Tissue 2, Tissue 3	E2F4,E2F6,E2F7,E2F8,EGR1,EGR2,EGR3,EGR4,SP1,SP2,ZNF263		SGBS 1, SGBS 2	E2F6,EGR1,EGR2,EGR3,EGR4,ZNF263	
Body mass index	*USP37*	rs492400	*PLCD4,RQCD1,ZNF142*	rs4674320,rs62182125,rs832810	rs524902	0.424958	0.990472	262	Tissue 1, Tissue 2, Tissue 3	HNF4G,Hnf4a,KLF14,NR1H2::RXRA,NR2C2,Nr2f6,Nr2f6(var.2),Pparg::Rxra,RXRB,RXRG,Rxra,SP3,SP4,SP8		SGBS 1, SGBS 2	HNF4G,Hnf4a,KLF16,NR1H2::RXRA,NR2C2,Nr2f6,Nr2f6(var.2),Pparg::Rxra,RXRB,RXRG,Rxra	
Body mass index	*USP37*	rs492400	*PLCD4,RQCD1,ZNF142*	rs4674320,rs62182125,rs832810	rs526897	0.398919	0.893358	262	Tissue 1, Tissue 2, Tissue 3	FOXP1,NFATC1,NFATC3,POU6F2,SRY,Sox6		SGBS 1, SGBS 2	FOXP1,SRY,Sox6	
Body mass index	*USP37*	rs492400	*PLCD4,RQCD1,ZNF142*	rs4674320,rs62182125,rs832810	rs611203	0.425007	0.993916	262				SGBS 1, SGBS 2		
Body mass index	*USP37*	rs492400	*PLCD4,RQCD1,ZNF142*	rs4674320,rs62182125,rs832810	rs689116	0.425206	0.995539	262				SGBS 1, SGBS 2		
Cardiac hypertrophy	*COL17A1*	rs1320448	*OBFC1,RP11-541N10.3*	rs79342925	rs111447985	0.05803	0.880233	57	Tissue 1, Tissue 2, Tissue 3	ELK1,ELK3,ELK4,ERF,ETS1,ETV1,ETV4,ETV5,FEV,FLI1,Gabpa,TFAP2A(var.3),TFAP2B(var.3),TFAP2C(var.3)	ETS1	SGBS 1, SGBS 2	ELK1,ELK3,ELK4,ERF,ETS1,ETV1,ETV4,ETV5,FEV,FLI1,Gabpa,TFAP2B(var.3),TFAP2C(var.3)	
Cardiac hypertrophy	*COL17A1*	rs1320448	*OBFC1,RP11-541N10.3*	rs79342925	rs73329737	0.0609067	0.996543	57				SGBS 1, SGBS 2		
Cardiac hypertrophy	*COL17A1*	rs1320448	*OBFC1,RP11-541N10.3*	rs79342925	rs805702	0.0611546	0.999136	57				SGBS 1		
Cardiac hypertrophy	*COL17A1*	rs1320448	*OBFC1,RP11-541N10.3*	rs79342925	rs805703	0.061105	1	57				SGBS 1, SGBS 2		
Coronary heart disease	*ATP5G1-GIP-SNF8-UBE2Z*	rs46522	*CALCOCO2,UBE2Z*	rs62075846,rs9904645	17:46970081	0.458089	0.994809	116	Tissue 1, Tissue 2, Tissue 3	Sox1		SGBS 1, SGBS 2	Sox1	
Coronary heart disease	*ATP5G1-GIP-SNF8-UBE2Z*	rs46522	*CALCOCO2,UBE2Z*	rs62075846,rs9904645	17:46970083	0.45804	0.993412	116	Tissue 1, Tissue 2, Tissue 3	Sox1		SGBS 1, SGBS 2	Sox1	
Coronary heart disease	*ATP5G1-GIP-SNF8-UBE2Z*	rs46522	*CALCOCO2,UBE2Z*	rs62075846,rs9904645	rs12601872	0.457693	0.999201	116				SGBS 1	Zfx	
Coronary heart disease	*ATP5G1-GIP-SNF8-UBE2Z*	rs46522	*CALCOCO2,UBE2Z*	rs62075846,rs9904645	rs12602746	0.493155	0.866593	116				SGBS 1		
Coronary heart disease	*ATP5G1-GIP-SNF8-UBE2Z*	rs46522	*CALCOCO2,UBE2Z*	rs62075846,rs9904645	rs12603612	0.457693	0.999201	116				SGBS 1		
Coronary heart disease	*ATP5G1-GIP-SNF8-UBE2Z*	rs46522	*CALCOCO2,UBE2Z*	rs62075846,rs9904645	rs2270574	0.457693	0.999201	116	Tissue 1, Tissue 2, Tissue 3			SGBS 1, SGBS 2		
Coronary heart disease	*ATP5G1-GIP-SNF8-UBE2Z*	rs46522	*CALCOCO2,UBE2Z*	rs62075846,rs9904645	rs2291725	0.493056	0.866553	116				SGBS 1, SGBS 2		
Coronary heart disease	*ATP5G1-GIP-SNF8-UBE2Z*	rs46522	*CALCOCO2,UBE2Z*	rs62075846,rs9904645	rs2411375	0.457693	0.999201	116				SGBS 1	SP2	
Coronary heart disease	*ATP5G1-GIP-SNF8-UBE2Z*	rs46522	*CALCOCO2,UBE2Z*	rs62075846,rs9904645	rs3744608	0.456602	0.994816	116				SGBS 1, SGBS 2	GLIS2,KLF14,KLF16,NR2C2,SP3,SP4,SP8,ZIC3,ZNF740	
Coronary heart disease	*ATP5G1-GIP-SNF8-UBE2Z*	rs46522	*CALCOCO2,UBE2Z*	rs62075846,rs9904645	rs4378658	0.457693	0.999201	116				SGBS 1, SGBS 2		
Coronary heart disease	*ATP5G1-GIP-SNF8-UBE2Z*	rs46522	*CALCOCO2,UBE2Z*	rs62075846,rs9904645	rs61576918	0.458189	0.997205	116	Tissue 1, Tissue 2, Tissue 3	EHF,ELK3,ELK4,ETV4,ETV6,Gabpa		SGBS 1, SGBS 2	EHF,ELK3,ELK4,ETV4,ETV6,Gabpa	
Coronary heart disease	*ATP5G1-GIP-SNF8-UBE2Z*	rs46522	*CALCOCO2,UBE2Z*	rs62075846,rs9904645	rs62075824	0.457544	0.989622	116				SGBS 1, SGBS 2	ZNF263	
Coronary heart disease	*ATP5G1-GIP-SNF8-UBE2Z*	rs46522	*CALCOCO2,UBE2Z*	rs62075846,rs9904645	rs9904645	0.456998	0.992018	116	Tissue 1, Tissue 2, Tissue 3	ZNF143		SGBS 1, SGBS 2		
Coronary heart disease	*CDKN2A-CDKN2B*	rs1333048	*CDKN2B*	rs1333046	rs10757279	0.404821	0.989115	53				SGBS 2		
Coronary heart disease	*CDKN2A-CDKN2B*	rs1333048	*CDKN2B*	rs1333046	rs1333042	0.410178	0.948515	53	Tissue 1, Tissue 2, Tissue 3			SGBS 1		
Coronary heart disease	*CDKN2A-CDKN2B*	rs1333048	*CDKN2B*	rs1333046	rs1537372	0.371987	0.8594	53				SGBS 1, SGBS 2	RREB1,ZNF263	
Coronary heart disease	*CDKN2A-CDKN2B*	rs1333048	*CDKN2B*	rs1333046	rs1537373	0.410078	0.948102	53	Tissue 1, Tissue 2, Tissue 3	GCM2		SGBS 1, SGBS 2		
Coronary heart disease	*CDKN2A-CDKN2B*	rs1333048	*CDKN2B*	rs1333046	rs2383206	0.420792	0.929841	53				SGBS 1		
Coronary heart disease	*CDKN2A-CDKN2B*	rs1333048	*CDKN2B*	rs1333046	rs2891168	0.422875	0.89792	53	Tissue 1					
Coronary heart disease	*CDKN2A-CDKN2B*	rs1333048	*CDKN2B*	rs1333046	rs4977575	0.416328	0.943139	53				SGBS 1, SGBS 2		
Coronary heart disease	*CDKN2A-CDKN2B*	rs1333048	*CDKN2B*	rs1333046	rs7341786	0.42223	0.925536	53				SGBS 1, SGBS 2	Gfi1b,MYBL2	
Coronary heart disease	*LIPA*	rs1412444	*LIPA*	rs1412445	rs1332328	0.402837	0.948118	8	Tissue 3	EGR2,EGR4,GCM1,GCM2,ZEB1		SGBS 1, SGBS 2	EGR2,EGR4,GCM1,GCM2,ZEB1	
Coronary heart disease	*MRAS*	rs2306374	*MRAS*	rs34905952	rs9848655	0.100833	0.998906	15	Tissue 1, Tissue 2, Tissue 3	HNF4G,MZF1(var.2),NR2C2,Nr2f6,PRDM1,Pparg::Rxra,RXRB,RXRG,Rxra,SP1,SP2,Sox3,ZNF263		SGBS 1, SGBS 2	HNF4G,MZF1(var.2),NR2C2,Nr2f6,PRDM1,Pparg::Rxra,RXRB,RXRG,Rxra,SP1,SP2,ZNF263	
Coronary heart disease	*MRAS*	rs2306374	*MRAS*	rs34905952	rs9872754	0.100784	0.999453	15				SGBS 1, SGBS 2	HIC2,Vdr	
Coronary heart disease	*WDR12*	rs6725887	*ICA1L*	rs72928613	2:203736202	0.117697	0.910861	286	Tissue 1, Tissue 2	NRF1,Tcfl5		SGBS 1, SGBS 2	NRF1,Tcfl5	
Coronary heart disease	*WDR12*	rs6725887	*ICA1L*	rs72928613	rs143035655	0.111348	0.989983	286				SGBS 1, SGBS 2	Nr2f6(var.2),Rarb,SP2,THAP1	
Coronary heart disease	*WDR12*	rs6725887	*ICA1L*	rs72928613	rs144505847	0.1111	0.990467	286	Tissue 1, Tissue 2, Tissue 3			SGBS 1, SGBS 2		
Coronary heart disease	*WDR12*	rs6725887	*ICA1L*	rs72928613	rs149846585	0.11358	0.966107	286	Tissue 1, Tissue 2	SP2,SP8,TFAP2A(var.2),TFAP2B		SGBS 1, SGBS 2	SP2,SP8,TFAP2A(var.2),TFAP2B	
Coronary heart disease	*WDR12*	rs6725887	*ICA1L*	rs72928613	rs72932575	0.110555	0.997986	286				SGBS 1, SGBS 2		
Coronary heart disease	*WDR12*	rs6725887	*ICA1L*	rs72928613	rs72932770	0.11115	0.990971	286	Tissue 1	MEF2A,TBP		SGBS 1, SGBS 2	MEF2A,TBP	
Coronary heart disease	*WDR12*	rs6725887	*ICA1L*	rs72928613	rs72932777	0.111001	0.992467	286				SGBS 1, SGBS 2	NRF1,TP53,Zfx	
Coronary heart disease	*WDR12*	rs6725887	*ICA1L*	rs72928613	rs72934512	0.110654	0.996979	286				SGBS 1, SGBS 2		
Coronary heart disease	*WDR12*	rs6725887	*ICA1L*	rs72928613	rs72936309	0.110207	0.991444	286				SGBS 1, SGBS 2		
Coronary heart disease	*WDR12*	rs6725887	*ICA1L*	rs72928613	rs75869289	0.110555	0.997986	286				SGBS 1, SGBS 2		
Coronary heart disease	*WDR12*	rs6725887	*ICA1L*	rs72928613	rs79642273	0.110455	0.996979	286				SGBS 1		
Fasting insulin-related traits	*IRS1*	rs2943634	*IRS1*	rs1515098	2:227052396	0.350461	0.989772	107				SGBS 1	FOXA1,FOXB1,FOXC1,FOXC2,RARA	
Fasting insulin-related traits	*IRS1*	rs2943634	*IRS1*	rs1515098	rs2943646	0.375608	0.882662	107	Tissue 2	Dux				
Fasting insulin-related traits	*IRS1*	rs2943634	*IRS1*	rs1515098	rs2943654	0.37526	0.879829	107				SGBS 1	HMBOX1,IRF2,IRF8	
Fasting insulin-related traits	*IRS1*	rs2943634	*IRS1*	rs1515098	rs2943656	0.393066	0.81567	107				SGBS 1, SGBS 2	FOXP1,Sox3,Sox6	
Fasting insulin-related traits	*PDGFC*	rs4691380	*PDGFC*	rs4568281	4:157656410	0.267583	0.930021	94				SGBS 1, SGBS 2		
Fasting insulin-related traits	*PDGFC*	rs4691380	*PDGFC*	rs4568281	rs10857295	0.280925	0.869658	94				SGBS 1		
Fasting insulin-related traits	*PDGFC*	rs4691380	*PDGFC*	rs4568281	rs6819797	0.26049	0.998456	94				SGBS 1		
Glycated hemoglobins	*ANK1*	rs4737009	*IKBKB*	rs4737010	rs552904	0.288215	0.897496	15				SGBS 1		
Glycated hemoglobins	*ANK1*	rs4737009	*IKBKB*	rs4737010	rs7461534	0.284347	0.922446	15				SGBS 1, SGBS 2	RFX5	
Glycated hemoglobins	*C1orf85-CCT3-TMEM79*	rs6684514	*CCT3,GLMP (C1orf85)*	rs12022657,rs2273832	1:156308031	0.282412	0.984018	111	Tissue 1, Tissue 2, Tissue 3	HSF1,HSF2,HSF4		SGBS 1, SGBS 2	HSF1,HSF2,HSF4	
Glycated hemoglobins	*C1orf85-CCT3-TMEM79*	rs6684514	*CCT3,GLMP (C1orf85)*	rs12022657,rs2273832	1:156329883	0.284446	0.990044	111				SGBS 1, SGBS 2		
Glycated hemoglobins	*C1orf85-CCT3-TMEM79*	rs6684514	*CCT3,GLMP (C1orf85)*	rs12022657,rs2273832	rs10908498	0.285537	0.999757	111				SGBS 1, SGBS 2	ZNF143	
Glycated hemoglobins	*C1orf85-CCT3-TMEM79*	rs6684514	*CCT3,GLMP (C1orf85)*	rs12022657,rs2273832	rs12132794	0.285537	0.999757	111	Tissue 1, Tissue 3	JUN(var.2),MAF::NFE2,Nr2e3				
Glycated hemoglobins	*C1orf85-CCT3-TMEM79*	rs6684514	*CCT3,GLMP (C1orf85)*	rs12022657,rs2273832	rs2075166	0.281619	0.97968	111				SGBS 1	NRF1	
Glycated hemoglobins	*C1orf85-CCT3-TMEM79*	rs6684514	*CCT3,GLMP (C1orf85)*	rs12022657,rs2273832	rs2273832	0.282908	0.985945	111				SGBS 1, SGBS 2		
Glycated hemoglobins	*C1orf85-CCT3-TMEM79*	rs6684514	*CCT3,GLMP (C1orf85)*	rs12022657,rs2273832	rs2273833	0.282908	0.985945	111				SGBS 1, SGBS 2	ID4,NHLH1,NRF1,SNAI2,TCF3,TCF4	
Glycated hemoglobins	*C1orf85-CCT3-TMEM79*	rs6684514	*CCT3,GLMP (C1orf85)*	rs12022657,rs2273832	rs28372828	0.285339	0.998784	111				SGBS 1, SGBS 2	Pou5f1::Sox2,RFX5,Rfx1,Stat6	
Glycated hemoglobins	*C1orf85-CCT3-TMEM79*	rs6684514	*CCT3,GLMP (C1orf85)*	rs12022657,rs2273832	rs3806408	0.285339	0.997812	111				SGBS 2	Nr2e3	
Glycated hemoglobins	*C1orf85-CCT3-TMEM79*	rs6684514	*CCT3,GLMP (C1orf85)*	rs12022657,rs2273832	rs3806409	0.285339	0.998784	111	Tissue 1, Tissue 2, Tissue 3	EGR4,EWSR1-FLI1,GCM2,RARA::RXRA,RREB1,THAP1,ZNF263		SGBS 1, SGBS 2	EWSR1-FLI1,GCM2,RARA::RXRA,RREB1,ZNF263	RREB1
HDL cholesterol	*ANGPTL8*	rs4804154	*ANGPTL8*	rs12463177	rs12463177	0.117895	0.971939	11	Tissue 3					
HDL cholesterol	*CETP*	rs1532624	*CETP*	rs4784741	rs12720926	0.411864	0.989602	10	Tissue 2, Tissue 3			SGBS 1, SGBS 2		
HDL cholesterol	*CETP*	rs1532624	*CETP*	rs4784741	rs1532624	0.413798	N/A	10				SGBS 1, SGBS 2		

HDL cholesterol	*CETP*	rs1532624	*CETP*	rs4784741	rs1532625	0.413848	0.999796	10				SGBS 1, SGBS 2		
HDL cholesterol	*GALNT2*	rs4846914	*GALNT2*	rs4631704	rs4846922	0.416229	0.815339	23				SGBS 1, SGBS 2	EMX2,EVX1,EVX2,GBX1,HOXA2 ,HOXA5,HOXB2,Hoxb5,Hoxd3,	
HDL cholesterol	*GSK3B*	rs6805251	*GSK3B*	rs334533	rs334558	0.368366	0.899155	61	Tissue 1, Tissue 2, Tissue 3	INSM1,NR2C2		SGBS 1, SGBS 2	MEOX1,MEOX2	
HDL cholesterol	*LCAT*	rs2271293	*GFOD2,NUTF2*	rs6499143,rs7199443	rs1124324	0.169775	1	92	Tissue 1, Tissue 3	SOX21,ZBTB7A,ZBTB7B,ZBTB7C		SGBS 1, SGBS 2	ZBTB7A,ZBTB7B,ZBTB7C	
HDL cholesterol	*LCAT*	rs2271293	*GFOD2,NUTF2*	rs6499143,rs7199443	rs20549	0.200278	0.814092	92	Tissue 1, Tissue 2, Tissue 3			SGBS 1, SGBS 2		
HDL cholesterol	*LCAT*	rs2271293	*GFOD2,NUTF2*	rs6499143,rs7199443	rs56047901	0.162335	0.947686	92				SGBS 1		
HDL cholesterol	*LCAT*	rs2271293	*GFOD2,NUTF2*	rs6499143,rs7199443	rs58588228	0.164319	0.961546	92				SGBS 1, SGBS 2	BHLHE22,BHLHE23,NEUROD2, Neurog1,OLIG2,OLIG3	
HDL cholesterol	*LCAT*	rs2271293	*GFOD2,NUTF2*	rs6499143,rs7199443	rs7196789	0.193384	0.852952	92	Tissue 1, Tissue 2, Tissue 3	TFAP2A(var.3),TFAP2B(var.3),T FAP2C(var.3) EWSR1-FLI1,NR2C2,Pparg::Rxra,ZNF26 3		SGBS 1, SGBS 2	TFAP2A(var.3),TFAP2B(var.3),T FAP2C(var.3)	
HDL cholesterol	*LCAT*	rs2271293	*GFOD2,NUTF2*	rs6499143,rs7199443	rs7199443	0.193532	0.852139	92	Tissue 1, Tissue 2, Tissue 3			SGBS 1, SGBS 2	NR2C2,Pparg::Rxra,ZNF263	
HDL cholesterol	*MMAB-MVK*	rs7134594	*MMAB*	rs3782894	rs10850358	0.460817	0.999202	57				SGBS 1	ESR2	
HDL cholesterol	*MMAB-MVK*	rs7134594	*MMAB*	rs3782894	rs7308864	0.491072	0.831764	57				SGBS 2	JUND(var.2),Sox17	
HDL cholesterol	*NR0B2-PIGV*	rs12748152	*ARID1A,PIGV*	rs34217609,rs6656815	rs11535549	0.0864498	0.88805	158	Tissue 1, Tissue 2, Tissue 3	NR2C2,REST,ZIC3		SGBS 1, SGBS 2	REST	
HDL cholesterol	*NR0B2-PIGV*	rs12748152	*ARID1A,PIGV*	rs34217609,rs6656815	rs11555809	0.0870945	0.904682	158	Tissue 1, Tissue 2, Tissue 3			SGBS 1, SGBS 2		
HDL cholesterol	*NR0B2-PIGV*	rs12748152	*ARID1A,PIGV*	rs34217609,rs6656815	rs12737946	0.0856562	1	158	Tissue 1, Tissue 2, Tissue 3	PLAG1,RUNX1		SGBS 1, SGBS 2	PLAG1	
HDL cholesterol	*NR0B2-PIGV*	rs12748152	*ARID1A,PIGV*	rs34217609,rs6656815	rs12738345	0.0856562	1	158	Tissue 1, Tissue 2, Tissue 3	REST,ZBTB33		SGBS 1, SGBS 2	REST,ZBTB33	
HDL cholesterol	*NR0B2-PIGV*	rs12748152	*ARID1A,PIGV*	rs34217609,rs6656815	rs12742115	0.0856562	1	158	Tissue 1, Tissue 2, Tissue 3	SOX8		SGBS 1, SGBS 2	SOX8	
HDL cholesterol	*NR0B2-PIGV*	rs12748152	*ARID1A,PIGV*	rs34217609,rs6656815	rs12760759	0.085805	0.995576	158	Tissue 1, Tissue 2					
HDL cholesterol	*NR0B2-PIGV*	rs12748152	*ARID1A,PIGV*	rs34217609,rs6656815	rs17162387	0.0864498	0.88805	158	Tissue 1	ZNF410				
HDL cholesterol	*NR0B2-PIGV*	rs12748152	*ARID1A,PIGV*	rs34217609,rs6656815	rs34618114	0.0856562	1	158				SGBS 1, SGBS 2	PLAG1,ZNF263,Zfx	
HDL cholesterol	*NR0B2-PIGV*	rs12748152	*ARID1A,PIGV*	rs34217609,rs6656815	rs35108146	0.0867473	0.884653	158	Tissue 1			SGBS 1, SGBS 2		
HDL cholesterol	*NR0B2-PIGV*	rs12748152	*ARID1A,PIGV*	rs34217609,rs6656815	rs57217461	0.0856562	1	158				SGBS 1, SGBS 2	SOX21,Sox1	
HDL cholesterol	*NR0B2-PIGV*	rs12748152	*ARID1A,PIGV*	rs34217609,rs6656815	rs58421016	0.0856562	1	158				SGBS 1, SGBS 2		
HDL cholesterol	*NR0B2-PIGV*	rs12748152	*ARID1A,PIGV*	rs34217609,rs6656815	rs71636779	0.0856562	1	158				SGBS 1, SGBS 2	Foxd3	
HDL cholesterol	*NR0B2-PIGV*	rs12748152	*ARID1A,PIGV*	rs34217609,rs6656815	rs71636780	0.0859538	0.996213	158	Tissue 1, Tissue 2, Tissue 3	HES5,HES7,NRF1,Zfx		SGBS 1, SGBS 2	NRF1,Zfx	
HDL cholesterol	*NR0B2-PIGV*	rs12748152	*ARID1A,PIGV*	rs34217609,rs6656815	rs71640337	0.102024	0.824543	158	Tissue 1			SGBS 1, SGBS 2		
HDL cholesterol	*PDE3A*	rs7134375	*PDE3A*	rs7134375	rs7134375	0.397332	N/A	3	Tissue 1	ELK1,ELK3,ERF,ERG,ETS1,ETV1, ETV3,ETV4,ETV5,FEV,FLI1		SGBS 2	ELF1,ELK1,ELK3,ERF,ERG,ETS1,E TV1,ETV2,ETV3,ETV4,ETV5,FEV, FLI1,SPDEF	
HDL cholesterol	*PMVK-HDGF*	rs12145743	*RRNAD1*	rs3806415	rs11264533	0.323777	0.99932	18				SGBS 1, SGBS 2		
HDL cholesterol	*PMVK-HDGF*	rs12145743	*RRNAD1*	rs3806415	rs12030654	0.323579	0.998414	18				SGBS 1	EGR1,INSM1,KLF16,NR2C2,Ppa rgc::Rxra,SP1,SP2,SP3,THAP1	
HDL cholesterol	*PMVK-HDGF*	rs12145743	*RRNAD1*	rs3806415	rs3806415	0.325017	0.969816	18	Tissue 1, Tissue 2, Tissue 3			SGBS 1, SGBS 2		
HDL cholesterol	*PMVK-HDGF*	rs12145743	*RRNAD1*	rs3806415	rs4501833	0.323777	0.99932	18	Tissue 1, Tissue 3	RFX2,RFX4		SGBS 1, SGBS 2	RFX2,RFX3,RFX4,RFX5,Rfx1	
HDL cholesterol	*RBMS*	rs2013208	*RBM6*	rs11130233	3:50138580	0.452187	0.977931	96				SGBS 1, SGBS 2	PHOX2A,PROP1,Phox2b	
HDL cholesterol	*RBMS*	rs2013208	*RBM6*	rs11130233	3:50176259	0.449757	0.93049	96				SGBS 1	EGR1,EWSR1-FLI1,KLF16,KLF5,MZF1(var.2),N R2C2,PLAG1,SP1,SP2,SP3,SP8,Z NF263,ZNF740	
HDL cholesterol	*RBMS*	rs2013208	*RBM6*	rs11130233	rs2252833	0.450749	0.972485	96				SGBS 1	Six3,Spz1	
HDL cholesterol	*RBMS*	rs2013208	*RBM6*	rs11130233	rs2526754	0.447178	0.995193	96	Tissue 1	GLIS3		SGBS 2		
HDL cholesterol	*RBMS*	rs2013208	*RBM6*	rs11130233	rs2883057	0.450749	0.972485	96				SGBS 1, SGBS 2		
HDL cholesterol	*RBMS*	rs2013208	*RBM6*	rs11130233	rs35065728	0.450749	0.972485	96				SGBS 1		
HDL cholesterol	*RBMS*	rs2013208	*RBM6*	rs11130233	rs6772095	0.450699	0.972284	96	Tissue 1			SGBS 1, SGBS 2		
HDL cholesterol	*STAB1*	rs13326165	*PBRM1*	rs13326165	rs13326146	0.172453	N/A	1				SGBS 1		
HDL cholesterol	*STARD3*	rs11869286	*PGAP3,STARD3*	rs11869286,rs2517951	rs1053651	0.273981	0.82125	62				SGBS 1		
HDL cholesterol	*STARD3*	rs11869286	*PGAP3,STARD3*	rs11869286,rs2517951	rs2271308	0.27408	0.82166	62	Tissue 3				NR1H2::RXRA,Pparg::Rxra,RXR	
HDL cholesterol	*STARD3*	rs11869286	*PGAP3,STARD3*	rs11869286,rs2517951	rs2952151	0.287124	0.831459	62				SGBS 1, SGBS 2	G,Rxra	
HDL cholesterol	*STARD3*	rs11869286	*PGAP3,STARD3*	rs11869286,rs2517951	rs881844	0.314751	0.991042	62				SGBS 1	PPARG,Znf423	
HDL cholesterol	*UBE2L3*	rs181362	*UBE2L3,YDJC*	rs11089620,rs12158299	rs11089620	0.324769	0.884579	91				SGBS 1, SGBS 2	NRF1,TFAP2A(var.2)	
HDL cholesterol	*UBE2L3*	rs181362	*UBE2L3,YDJC*	rs11089620,rs12158299	rs140491	0.352892	0.995228	91				SGBS 1, SGBS 2	Klf12,SP4	
HDL cholesterol	*UBE2L3*	rs181362	*UBE2L3,YDJC*	rs11089620,rs12158299	rs140492	0.3519	0.999565	91				SGBS 1, SGBS 2		
HDL cholesterol	*UBE2L3*	rs181362	*UBE2L3,YDJC*	rs11089620,rs12158299	rs181360	0.3519	0.999565	91				SGBS 1, SGBS 2		
HDL cholesterol	*UBE2L3*	rs181362	*UBE2L3,YDJC*	rs11089620,rs12158299	rs2266959	0.324769	0.885437	91				SGBS 1, SGBS 2		
HDL cholesterol	*UBE2L3*	rs181362	*UBE2L3,YDJC*	rs11089620,rs12158299	rs3747093	0.341335	0.831789	91	Tissue 1, Tissue 2, Tissue 3	SP4,Zfx		SGBS 1, SGBS 2	SP4	
HiPadjBMI	*KLHL31*	rs7739232	*KLHL31*	rs113505085	6:53538019	0.0757861	0.997168	52				SGBS 2		
HiPadjBMI	*KLHL31*	rs7739232	*KLHL31*	rs113505085	rs75582203	0.0757861	0.997168	52				SGBS 2	CEBPA	
HiPadjBMI	*KLHL31*	rs7739232	*KLHL31*	rs113505085	rs76936629	0.0750918	0.991498	52	Tissue 1, Tissue 3			SGBS 1, SGBS 2		
HiPadjBMI	*KLHL31*	rs7739232	*KLHL31*	rs113505085	rs77162760	0.0751414	0.992206	52				SGBS 1, SGBS 2	Stat6	
HiPadjBMI	*KLHL31*	rs7739232	*KLHL31*	rs113505085	rs77597023	0.0757861	0.997168	52				SGBS 2		
HiPadjBMI	*KLHL31*	rs7739232	*KLHL31*	rs113505085	rs77937059	0.0757861	0.997168	52				SGBS 2	EWSR1-FLI1	
HiPadjBMI	*KLHL31*	rs7739232	*KLHL31*	rs113505085	rs79823138	0.0750918	0.991498	52	Tissue 1			SGBS 1, SGBS 2		
HiPadjBMI	*KLHL31*	rs7739232	*KLHL31*	rs113505085	rs80087193	0.0757861	0.997168	52				SGBS 2	ZNF263	
HiPadjBMI	*KLHL31*	rs7739232	*KLHL31*	rs113505085	rs80271463	0.0756373	0.999291	52	Tissue 1			SGBS 2		
HiPadjBMI	*VPS53*	rs2034088	*FAM101B,VPS53*	rs10153248,rs2289620	rs2289620	0.369755	0.832991	22				SGBS 1		
Intracranial aneurysm	*PLCL1-BOLL*	rs700651	*SF3B1*	rs2564383	2:198540701	0.410178	0.949014	296				SGBS 1, SGBS 2	IRF1	IRF1
Intracranial aneurysm	*PLCL1-BOLL*	rs700651	*SF3B1*	rs2564383	2:198692686	0.39138	0.962402	296				SGBS 1, SGBS 2		
Intracranial aneurysm	*PLCL1-BOLL*	rs700651	*SF3B1*	rs2564383	rs1116734	0.427686	0.859539	296	Tissue 1, Tissue 2, Tissue 3			SGBS 1, SGBS 2		
Intracranial aneurysm	*PLCL1-BOLL*	rs700651	*SF3B1*	rs2564383	rs4850812	0.3889	0.957302	296	Tissue 1	Hic1,NFIA,NFIX,Sox1				
Intracranial aneurysm	*PLCL1-BOLL*	rs700651	*SF3B1*	rs2564383	rs700674	0.39133	0.962612	296				SGBS 1, SGBS 2		
Intracranial aneurysm	*PLCL1-BOLL*	rs700651	*SF3B1*	rs2564383	rs770659	0.39133	0.962201	296				SGBS 1, SGBS 2	PAX5	
Intracranial aneurysm	*PLCL1-BOLL*	rs700651	*SF3B1*	rs2564383	rs787883	0.428033	0.858292	296				SGBS 1, SGBS 2		
Intracranial aneurysm	*STARD13*	rs9315204	*KL,STARD13*	rs1998728,rs614691	rs1980781	0.343914	0.999121	22	Tissue 1, Tissue 2, Tissue 3	HNF4G,Nr2f6(var.2),RORA,ROR A(var.2),Rarb,Rarg		SGBS 1, SGBS 2	HNF4G,Nr2f6(var.2),RORA,ROR A(var.2),Rarb,Rarg	
Intracranial aneurysm	*STARD13*	rs9315204	*KL,STARD13*	rs1998728,rs614691	rs671937	0.36792	0.882484	22				SGBS 2		
Intracranial aneurysm	*STARD13*	rs9315204	*KL,STARD13*	rs1998728,rs614691	rs9315204	0.344013	N/A	22				SGBS 1, SGBS 2	Pou5f1::Sox2,TEAD1,TEAD3,TE AD4	
Intracranial aneurysm	*STARD13*	rs9315204	*KL,STARD13*	rs1998728,rs614691	rs9527098	0.355967	0.948381	22				SGBS 1		
LDL cholesterol	*HMGCR*	rs12916	*HMGCR*	rs34341	rs1423527	0.462206	0.919682	54				SGBS 1	CEBPA	
LDL cholesterol	*HMGCR*	rs12916	*HMGCR*	rs34341	rs2878417	0.457048	0.940258	54				SGBS 1	EWSR1-FLI1,ZNF263	
LDL cholesterol	*HMGCR*	rs12916	*HMGCR*	rs34341	rs7703282	0.451394	0.860682	54				SGBS 1, SGBS 2		
LDL cholesterol	*PLEC*	rs11136341	*PLEC*	rs10107388	8:145030900	0.361075	0.918912	26				SGBS 1, SGBS 2	INSM1,RUNX1	
LDL cholesterol	*PLEC*	rs11136341	*PLEC*	rs10107388	8:145030903	0.350164	0.874653	26				SGBS 1, SGBS 2	INSM1	
LDL cholesterol	*PLEC*	rs11136341	*PLEC*	rs10107388	rs62523994	0.394554	0.801675	26				SGBS 1, SGBS 2	GLI2,SP4	
LDL cholesterol	*PLEC*	rs11136341	*PLEC*	rs10107388	rs7832643	0.394455	0.802827	26				SGBS 1, SGBS 2	EGR1	
LDL cholesterol,CHD	*PSRC1*	rs599839	*CELSR2*	rs1277930	rs12740374	0.216794	0.980703	11				SGBS 1, SGBS 2	TFAP2A(var.2),TFAP2B,TFAP2C	
LDL cholesterol,CHD	*PSRC1*	rs599839	*CELSR2*	rs1277930	rs646776	0.216794	0.980703	11				SGBS 1, SGBS 2		
LDL cholesterol,CHD	*PSRC1*	rs599839	*CELSR2*	rs1277930	rs660240	0.215901	0.975547	11				SGBS 1, SGBS 2	RREB1	
Lipid metabolism phenotypes (VLDL-D)	*MLXIPL*	rs13247874	*MLXIPL*	rs13493868	rs55747707	0.157772	0.959112	33	Tissue 1, Tissue 2, Tissue 3					
Lipoprotein-associated phospholipase A2 activity and mass	*PLA2G7*	rs7756935	*PLA2G7*	rs3757241	rs7767605	0.246503	0.976309	25	Tissue 3					
Metabolic traits	*ACADM*	rs211718	*ACADM*	rs12094842	rs1251075	0.294862	0.885642	142	Tissue 1, Tissue 2, Tissue 3	KLF14,NFIC::TLX1,PAX5,SP8,TF AP2A(var.2),TFAP2B,TFAP2C NR2F1,NR4A2,Nr2e1,Pparg::Rx ra,Sox3		SGBS 1, SGBS 2	NFIC::TLX1,TFAP2A(var.2),TFAP 2B,TFAP2C	
Metabolic traits	*ACADM*	rs211718	*ACADM*	rs12094842	rs61797339	0.27165	0.990489	142	Tissue 1, Tissue 2, Tissue 3			SGBS 1, SGBS 2		
Metabolic traits	*ENPEP*	rs2087160	*ENPEP*	rs9997379	rs10033071	0.208957	0.991914	33				SGBS 2		
Metabolic traits	*ETFDH*	rs8396	*ETFDH*	rs17843929	rs2070629	0.232963	0.994457	22	Tissue 1, Tissue 2, Tissue 3			SGBS 1, SGBS 2		
Metabolic traits	*ETFDH*	rs8396	*ETFDH*	rs17843929	rs2070630	0.232963	0.994457	22	Tissue 1, Tissue 2, Tissue 3	TFAP2A(var.3),TFAP2C(var.3)		SGBS 1, SGBS 2	TFAP2A(var.3),TFAP2C(var.3)	
Metabolic traits	*LACTB*	rs2652822	*LACTB*	rs2652822	rs2292038	0.478871	0.965617	16	Tissue 1, Tissue 2, Tissue 3	HIC2,MTF1,RARA::RXRA		SGBS 1, SGBS 2	MTF1,RARA::RXRA	
Metabolic traits	*LACTB*	rs2652822	*LACTB*	rs2652822	rs7170846	0.476143	0.976409	16	Tissue 1, Tissue 3	USF1		SGBS 1, SGBS 2	USF1	
Metabolic traits	*NAT8*	rs13391552	*ALMS1*	rs6740766	rs11903916	0.234253	0.880665	180				SGBS 1	CHF,ELF3,ELF4,ELF5	
Metabolic traits	*NAT8*	rs13391552	*ALMS1*	rs6740766	rs4547554	0.180091	0.813533	180	Tissue 1, Tissue 2, Tissue 3	GCM1,Hic1				
Metabolic traits	*NAT8*	rs13391552	*ALMS1*	rs6740766	rs73947808	0.212281	0.99911	180				SGBS 1	Crx,Znf423	
Metabolic traits	*NAT8*	rs13391552	*ALMS1*	rs6740766	rs7564890	0.212082	0.996738	180				SGBS 1, SGBS 2		
Metabolic traits	*OPLAH*	rs6558295	*CYC1*	rs7836836	8:145134151	0.0729094	1	114				SGBS 1, SGBS 2	Nr1h3::Rxra,Sox2,Sox3,Sox6	
Metabolic traits	*OPLAH*	rs6558295	*CYC1*	rs7836836	8:145136999	0.0729094	1	114				SGBS 1	RARA,RHOXF1	
Metabolic traits	*OPLAH*	rs6558295	*CYC1*	rs7836836	8:145159708	0.0862018	0.833673	114	Tissue 1, Tissue 2	Klf12		SGBS 1, SGBS 2	Klf12	
Metabolic traits	*OPLAH*	rs6558295	*CYC1*	rs7836836	8:145181859	0.0707271	0.93719	114				SGBS 1		
Metabolic traits	*OPLAH*	rs6558295	*CYC1*	rs7836836	rs11136254	0.0736314	0.989105	114				SGBS 1		
Metabolic traits	*OPLAH*	rs6558295	*CYC1*	rs7836836	rs11136255	0.0800516	0.903764	114	Tissue 1, Tissue 2, Tissue 3			SGBS 1, SGBS 2	EGR1,SMAD2::SMAD3::SMAD4	

Metabolic traits	*OPLAH*	rs6558295	*CYC1*	rs7836836	rs11541475	0.0787124	0.920478	114	Tissue 1	Tcfl5		SGBS 1, SGBS 2		
Metabolic traits	*OPLAH*	rs6558295	*CYC1*	rs7836836	rs11986647	0.0727606	0.882018	114				SGBS 1, SGBS 2	CTCF,HINFP	CTCF
Metabolic traits	*OPLAH*	rs6558295	*CYC1*	rs7836836	rs12549149	0.0737526	0.987669	114				SGBS 1, SGBS 2	EGR1,EGR3,EGR4,KLF16,KLF5,N R2C2,RREB1,SP1,SP2,SP3,SP8	RREB1
Metabolic traits	*OPLAH*	rs6558295	*CYC1*	rs7836836	rs13255347	0.0805972	0.897114	114	Tissue 1	CTCF,GCM1,GCM2,NR2C2,THA P1		SGBS 1, SGBS 2	CTCF,EGR4,NR2C2	CTCF
Metabolic traits	*OPLAH*	rs6558295	*CYC1*	rs7836836	rs13256214	0.073951	0.984808	114	Tissue 1, Tissue 2, Tissue 3	Bach1::Mafk,GLIS1		SGBS 1, SGBS 2	EGR4,SP2,ZNF740	
Metabolic traits	*OPLAH*	rs6558295	*CYC1*	rs7836836	rs144843136	0.0687432	0.938639	114				SGBS 1, SGBS 2	TFAP2A(var.3),TFAP2B(var.3),T FAP2C(var.3)	
Metabolic traits	*OPLAH*	rs6558295	*CYC1*	rs7836836	rs180768286	0.0724135	0.992666	114				SGBS 1, SGBS 2	Sox2,Sox3,Sox6	
Metabolic traits	*OPLAH*	rs6558295	*CYC1*	rs7836836	rs6984764	0.084813	0.848611	114	Tissue 1, Tissue 2, Tissue 3	NRF1,REST		SGBS 1, SGBS 2		
Metabolic traits	*OPLAH*	rs6558295	*CYC1*	rs7836836	rs73379171	0.0687432	0.938639	114				SGBS 1		
Metabolic traits	*OPLAH*	rs6558295	*CYC1*	rs7836836	rs75361772	0.073703	0.988386	114	Tissue 1, Tissue 2, Tissue 3	GLIS1		SGBS 1, SGBS 2	GLIS1	
Metabolic traits	*PDXDC1*	rs7200543	*PDXDC1*	rs35574015	rs16966953	0.345105	0.999781	37				SGBS 1, SGBS 2		
Metabolic traits,Phospholipid	*SCD*	rs603424	*SCD*	rs603424	rs603424	0.11611	N/A	1	Tissue 1, Tissue 2, Tissue 3	TBX2,TBX20,TBX21		SGBS 1, SGBS 2		
Moyamoya disease	*RNF213*	rs6565681	*RNF213*	rs9901680	rs13096485	0.191449	0.93418	14				SGBS 1, SGBS 2	MAFF,MAFK,NRL	
Obesity	*ATP2A1-SH2B1*	rs7498665	*ATXN2L,SH2B1*	rs11863370,rs62037367	16:28851377	0.431306	0.997174	107	Tissue 1	NFIC::TLX1		SGBS 1, SGBS 2	NFIC::TLX1	
Obesity	*ATP2A1-SH2B1*	rs7498665	*ATXN2L,SH2B1*	rs11863370,rs62037367	rs11864750	0.429471	0.989334	107	Tissue 1, Tissue 2, Tissue 3	EGR1,RREB1,SP2,ZNF263		SGBS 1, SGBS 2	EGR1,SP2,ZNF263	
Obesity	*ATP2A1-SH2B1*	rs7498665	*ATXN2L,SH2B1*	rs11863370,rs62037367	rs62037367	0.432001	1	107				SGBS 1, SGBS 2	MTF1,Nr5a2	
Obesity	*ATP2A1-SH2B1*	rs7498665	*ATXN2L,SH2B1*	rs11863370,rs62037367	rs7187776	0.448715	0.933628	107	Tissue 1, Tissue 2, Tissue 3	SPI1,SPIB	SPI1	SGBS 1, SGBS 2		
Obesity	*ATP2A1-SH2B1*	rs7498665	*ATXN2L,SH2B1*	rs11863370,rs62037367	rs7193733	0.429422	0.988327	107				SGBS 1, SGBS 2		
Obesity	*ATP2A1-SH2B1*	rs7498665	*ATXN2L,SH2B1*	rs11863370,rs62037367	rs7198606	0.432001	1	107	Tissue 1, Tissue 2, Tissue 3	ZNF263		SGBS 1, SGBS 2	ZNF263	
Obesity	*ATP2A1-SH2B1*	rs7498665	*ATXN2L,SH2B1*	rs11863370,rs62037367	rs8055138	0.432001	1	107	Tissue 1	NR2C2,Nr2f6		SGBS 1, SGBS 2	NR2C2,Nr2f6,ZNF263	
Palmitoleic acid (16:1n-7)	*NDUFB8-SCD-SEC31B-WNT8B-HIF1AN*	rs11190604	*HIF1AN,SEC31B*	rs2489041,rs7071271	rs10883504	0.220514	0.896472	240				SGBS 1, SGBS 2	Nr1h3::Rxra,Vdr	
Palmitoleic acid (16:1n-7)	*NDUFB8-SCD-SEC31B-WNT8B-HIF1AN*	rs11190604	*HIF1AN,SEC31B*	rs2489041,rs7071271	rs10883505	0.220514	0.896472	240				SGBS 1, SGBS 2	ZIC3	
Palmitoleic acid (16:1n-7)	*NDUFB8-SCD-SEC31B-WNT8B-HIF1AN*	rs11190604	*HIF1AN,SEC31B*	rs2489041,rs7071271	rs10883506	0.220067	0.894783	240	Tissue 1, Tissue 2, Tissue 3			SGBS 1, SGBS 2		
Palmitoleic acid (16:1n-7)	*NDUFB8-SCD-SEC31B-WNT8B-HIF1AN*	rs11190604	*HIF1AN,SEC31B*	rs2489041,rs7071271	rs11190587	0.220365	0.8961	240				SGBS 1, SGBS 2		
Palmitoleic acid (16:1n-7)	*NDUFB8-SCD-SEC31B-WNT8B-HIF1AN*	rs11190604	*HIF1AN,SEC31B*	rs2489041,rs7071271	rs11190599	0.220514	0.897046	240	Tissue 1, Tissue 2, Tissue 3			SGBS 1, SGBS 2		
Palmitoleic acid (16:1n-7)	*NDUFB8-SCD-SEC31B-WNT8B-HIF1AN*	rs11190604	*HIF1AN,SEC31B*	rs2489041,rs7071271	rs17113137	0.218332	0.862079	240				SGBS 1, SGBS 2	OTX1,OTX2,PITX3,Zfx	
Palmitoleic acid (16:1n-7)	*NDUFB8-SCD-SEC31B-WNT8B-HIF1AN*	rs11190604	*HIF1AN,SEC31B*	rs2489041,rs7071271	rs1800662	0.220514	0.897046	240				SGBS 1, SGBS 2		
Palmitoleic acid (16:1n-7)	*NDUFB8-SCD-SEC31B-WNT8B-HIF1AN*	rs11190604	*HIF1AN,SEC31B*	rs2489041,rs7071271	rs2295776	0.220514	0.897046	240	Tissue 1, Tissue 2, Tissue 3	Klf1,RREB1,SP3		SGBS 1, SGBS 2	Klf1,RREB1,SP3	RREB1
Palmitoleic acid (16:1n-7)	*NDUFB8-SCD-SEC31B-WNT8B-HIF1AN*	rs11190604	*HIF1AN,SEC31B*	rs2489041,rs7071271	rs2495758	0.20246	0.999079	240	Tissue 1, Tissue 3			SGBS 1, SGBS 2		
Palmitoleic acid (16:1n-7)	*NDUFB8-SCD-SEC31B-WNT8B-HIF1AN*	rs11190604	*HIF1AN,SEC31B*	rs2489041,rs7071271	rs28522614	0.203055	0.983788	240				SGBS 1		
Palmitoleic acid (16:1n-7)	*NDUFB8-SCD-SEC31B-WNT8B-HIF1AN*	rs11190604	*HIF1AN,SEC31B*	rs2489041,rs7071271	rs3750626	0.220365	0.8961	240				SGBS 1, SGBS 2		
Palmitoleic acid (16:1n-7)	*NDUFB8-SCD-SEC31B-WNT8B-HIF1AN*	rs11190604	*HIF1AN,SEC31B*	rs2489041,rs7071271	rs4919466	0.221803	0.882931	240				SGBS 1		
Palmitoleic acid (16:1n-7)	*NDUFB8-SCD-SEC31B-WNT8B-HIF1AN*	rs11190604	*HIF1AN,SEC31B*	rs2489041,rs7071271	rs7080356	0.220365	0.8961	240				SGBS 1, SGBS 2		
Palmitoleic acid (16:1n-7)	*NDUFB8-SCD-SEC31B-WNT8B-HIF1AN*	rs11190604	*HIF1AN,SEC31B*	rs2489041,rs7071271	rs7099913	0.20251	0.999386	240				SGBS 1, SGBS 2	Ddit3::Cebpa	
Palmitoleic acid (16:1n-7)	*NDUFB8-SCD-SEC31B-WNT8B-HIF1AN*	rs11190604	*HIF1AN,SEC31B*	rs2489041,rs7071271	rs7900678	0.221803	0.882931	240				SGBS 1, SGBS 2		
Palmitoleic acid (16:1n-7)	*NDUFB8-SCD-SEC31B-WNT8B-HIF1AN*	rs11190604	*HIF1AN,SEC31B*	rs2489041,rs7071271	rs9420796	0.20241	0.999386	240				SGBS 1, SGBS 2	RREB1	
Phospholipid	*ABHD3*	rs11662721	*ESCO1*	rs8086339	rs11662724	0.224978	0.944892	67	Tissue 1, Tissue 2, Tissue 3	EGR3,EGR4,PAX5,PAX9,TFAP2A (var.3),TFAP2B(var.3),TFAP2C(v ar.3)		SGBS 1, SGBS 2	PAX5,PAX9,TFAP2A(var.3),TFAP 2B(var.3),TFAP2C(var.3)	
Phospholipid	*C11orf10*	rs174538	*TMEM258 (C11orf10)*	rs108499	rs174538	0.336723	N/A	6	Tissue 1, Tissue 3	Creb3l2,XBP1	Creb3l2	SGBS 1, SGBS 2	XBP1	
Proinsulin	*LARP6*	rs1549318	*LARP6*	rs4777325	rs7175862	0.477234	0.911488	32				SGBS 1, SGBS 2	NFIA,NFIC::TLX1,NFIX,REST	NFIA
Proinsulin	*MADD*	rs10501320	*ACP2,FNBP4*	rs10501320,rs11039149	rs11039149	0.179546	0.97985	7	Tissue 1, Tissue 2, Tissue 3	PLAG1,Pparg::Rxra		SGBS 1, SGBS 2	PLAG1,Pparg::Rxra	
Stearic acid (18:0)	*RWDD3*	rs6671200	*RWDD3*	rs34266846	rs35917978	0.089624	0.881211	87				SGBS 1	TP53,TP63,TP73	
Total Cholesterol	*DOCK7-ANGPTL3*	rs2131925	*DOCK7*	rs631106	1:63153199	0.241147	0.995938	237				SGBS 1, SGBS 2		
Total Cholesterol	*DOCK7-ANGPTL3*	rs2131925	*DOCK7*	rs631106	rs10889356	0.228499	0.911899	237				SGBS 1, SGBS 2	ELK3,ELK4,ERF,ERG,ETS1,ETV2, ETV6,FEV,FLI1	
Total Cholesterol	*DOCK7-ANGPTL3*	rs2131925	*DOCK7*	rs631106	rs1168042	0.241246	0.999729	237				SGBS 1		
Total Cholesterol	*DOCK7-ANGPTL3*	rs2131925	*DOCK7*	rs631106	rs1168114	0.228846	0.910518	237				SGBS 1, SGBS 2		
Total Cholesterol	*DOCK7-ANGPTL3*	rs2131925	*DOCK7*	rs631106	rs2131925	0.241296	N/A	237				SGBS 1		
Total Cholesterol	*DOCK7-ANGPTL3*	rs2131925	*DOCK7*	rs631106	rs630144	0.241494	0.997834	237				SGBS 1, SGBS 2	Gata1	
Total Cholesterol	*DOCK7-ANGPTL3*	rs2131925	*DOCK7*	rs631106	rs631106	0.241444	0.997022	237	Tissue 1			SGBS 1, SGBS 2		
Total Cholesterol	*GPR146*	rs1997243	*GPR146*	rs11764937	rs10252404	0.115316	0.807176	132	Tissue 1, Tissue 2, Tissue 3	EGR1,INSM1,KLF14,Klf12,NFKB 1,PLAG1,TEAD1,ZNF263		SGBS 1, SGBS 2	EGR1,INSM1,KLF14,Klf12,NFKB 1,PLAG1,ZNF263	
Total Cholesterol	*GPR146*	rs1997243	*GPR146*	rs11764937	rs113119264	0.13213	0.999568	132				SGBS 1, SGBS 2		
Total Cholesterol	*GPR146*	rs1997243	*GPR146*	rs11764937	rs113575110	0.131336	0.905992	132				SGBS 1, SGBS 2		
Total Cholesterol	*GPR146*	rs1997243	*GPR146*	rs11764937	rs113642700	0.132229	0.991799	132	Tissue 1, Tissue 2, Tissue 3			SGBS 1, SGBS 2		
Total Cholesterol	*GPR146*	rs1997243	*GPR146*	rs11764937	rs11763793	0.132923	0.99097	132				SGBS 1	GLIS2,ZIC1,ZIC4	
Total Cholesterol	*GPR146*	rs1997243	*GPR146*	rs11764937	rs118132455	0.132179	0.99913S	132				SGBS 1, SGBS 2		
Total Cholesterol	*GPR146*	rs1997243	*GPR146*	rs11764937	rs140138013	0.132229	0.999568	132				SGBS 2	Stat4,ZIC3	
Total Cholesterol	*GPR146*	rs1997243	*GPR146*	rs11764937	rs1881123	0.132179	1	132				SGBS 1, SGBS 2		
Total Cholesterol	*GPR146*	rs1997243	*GPR146*	rs11764937	rs1997243	0.132179	N/A	132	Tissue 3	EWSR1-FLI1,KLF5,MZF1(var.2),SP1,SP2, ZNF263		SGBS 1, SGBS 2	EWSR1-FLI1,ZNF263	
Total Cholesterol	*GPR146*	rs1997243	*GPR146*	rs11764937	rs28522260	0.135899	0.959189	132	Tissue 1, Tissue 3	Nr1h3::Rxra		SGBS 1, SGBS 2	Nr1h3::Rxra	
Total Cholesterol	*GPR146*	rs1997243	*GPR146*	rs11764937	rs61753396	0.132527	0.994394	132				SGBS 1, SGBS 2	TP53,TP63,TP73,ZIC1,Znf423	
Total Cholesterol	*GPR146*	rs1997243	*GPR146*	rs11764937	rs74360401	0.131882	0.995678	132				SGBS 1, SGBS 2	KLF14,KLF5,Klf1,Klf4,SP1,SP2,S P8	
Total Cholesterol	*GPR146*	rs1997243	*GPR146*	rs11764937	rs77760339	0.132279	0.999136	132	Tissue 1, Tissue 2, Tissue 3			SGBS 1, SGBS 2		
Total Cholesterol	*GPR146*	rs1997243	*GPR146*	rs11764937	rs77943789	0.13585	0.961279	132				SGBS 1, SGBS 2		
Total Cholesterol	*GPR146*	rs1997243	*GPR146*	rs11764937	rs78628466	0.133221	0.990132	132				SGBS 1		
Total Cholesterol	*PXK*	rs13315871	*PXK*	rs35557787	3:58311576	0.097411	0.983296	184				SGBS 1	IRF1,MEF2A,MEF2D	
Total Cholesterol	*PXK*	rs13315871	*PXK*	rs35557787	rs11708590	0.0987005	0.998885	184				SGBS 1, SGBS 2		
Total Cholesterol	*PXK*	rs13315871	*PXK*	rs35557787	rs13066269	0.0977582	0.983849	184				SGBS 1, SGBS 2		
Total Cholesterol	*PXK*	rs13315871	*PXK*	rs35557787	rs13317400	0.0987005	1	184				SGBS 1, SGBS 2		
Total Cholesterol	*PXK*	rs13315871	*PXK*	rs35557787	rs13320620	0.0987005	1	184				SGBS 1, SGBS 2		
Total Cholesterol	*PXK*	rs13315871	*PXK*	rs35557787	rs13327709	0.0987005	1	184				SGBS 1, SGBS 2	Znf423	
Total Cholesterol	*PXK*	rs13315871	*PXK*	rs35557787	rs36037390	0.101379	0.963081	184				SGBS 1, SGBS 2		
Total Cholesterol	*PXK*	rs13315871	*PXK*	rs35557787	rs71311853	0.0978574	0.978297	184				SGBS 1, SGBS 2	Sp1	
Total Cholesterol	*PXK*	rs13315871	*PXK*	rs35557787	rs71311854	0.0976094	0.981069	184				SGBS 1, SGBS 2		
Total Cholesterol	*PXK*	rs13315871	*PXK*	rs35557787	rs77825960	0.102916	0.914103	184	Tissue 1, Tissue 2, Tissue 3			SGBS 1, SGBS 2		
Total Cholesterol	*PXK*	rs13315871	*PXK*	rs35557787	rs78578590	0.103462	0.919208	184	Tissue 1, Tissue 2, Tissue 3	EGR1,EGR4,Klf4,SP1,SP2,SP4,Z NF740		SGBS 1, SGBS 2	EGR1,Klf4,SP2,SP4	
Total Cholesterol	*PXK*	rs13315871	*PXK*	rs35557787	rs78946071	0.0993453	0.96196	184	Tissue 1, Tissue 2, Tissue 3	GLIS2,GLIS3		SGBS 1, SGBS 2	GLIS2,GLIS3	
Total Cholesterol	*PXK*	rs13315871	*PXK*	rs35557787	rs79813245	0.0972622	0.979402	184	Tissue 1, Tissue 2, Tissue 3	CTCF,EGR1,SP1,SP2,TFAP2A(va r.2)		SGBS 1, SGBS 2	CTCF,EGR1,SP1,SP2,TFAP2A(va r.2)	CTCF
Total Cholesterol	*PXK*	rs13315871	*PXK*	rs35557787	rs9818581	0.0973614	0.98274	184				SGBS 1, SGBS 2	ZNF263	
Total Cholesterol	*PXK*	rs13315871	*PXK*	rs35557787	rs9864325	0.0987005	0.998885	184				SGBS 1		
Total Cholesterol	*PXK*	rs13315871	*PXK*	rs35557787	rs9871725	0.0973614	0.98274	184				SGBS 1, SGBS 2		
Total Cholesterol	*TMEM57*	rs10903129	*RHCE,TMEM50A,TMEM57*	rs3093578,rs35343405, rs7880982 9,rs3093578,rs35343405,	rs1053438	0.485765	0.865797	179				SGBS 1, SGBS 2		
Total Cholesterol	*TMEM57*	rs10903129	*RHCE,TMEM50A,TMEM57*	rs3093578,rs35343405, rs7880982 9,rs3093578,rs35343405,	rs10903129	0.46538	N/A	179				SGBS 1, SGBS 2		
Total Cholesterol	*TMEM57*	rs10903129	*RHCE,TMEM50A,TMEM57*	rs3093578,rs35343405, rs7880982 9,rs3093578,rs35343405,	rs35172831	0.458337	0.965725	179				SGBS 1, SGBS 2	CTCF	
Total Cholesterol	*TMEM57*	rs10903129	*RHCE,TMEM50A,TMEM57*	rs3093578,rs35343405, rs7880982 9,rs3093578,rs35343405,	rs35260034	0.458337	0.965725	179				SGBS 2	ZNF263	
Total Cholesterol	*TMEM57*	rs10903129	*RHCE,TMEM50A,TMEM57*	rs3093578,rs35343405, rs7880982 9,rs3093578,rs35343405,	rs35850196	0.465232	0.999003	179				SGBS 1, SGBS 2		
Total Cholesterol	*TMEM57*	rs10903129	*RHCE,TMEM50A,TMEM57*	rs3093578,rs35343405, rs7880982 9,rs3093578,rs35343405,	rs599691	0.485864	0.865437	179				SGBS 1, SGBS 2		
Total Cholesterol	*TMEM57*	rs10903129	*RHCE,TMEM50A,TMEM57*	rs3093578,rs35343405, rs7880982 9,rs3093578,rs35343405,	rs61776810	0.457693	0.963206	179				SGBS 1, SGBS 2	RARA	
Total Cholesterol	*TMEM57*	rs10903129	*RHCE,TMEM50A,TMEM57*	rs3093578,rs35343405, rs7880982 9,rs3093578,rs35343405,	rs631133	0.465628	0.993427	179				SGBS 1, SGBS 2	BATF3,Creb5,JDP2(var.2),JUND (var.2)	
Total Cholesterol	*TMEM57*	rs10903129	*RHCE,TMEM50A,TMEM57*	rs7880982 9	rs9438904	0.462553	0.972424	179	Tissue 1, Tissue 2, Tissue 3	Bach1::Mafk,CREB3L1,XBP1		SGBS 1, SGBS 2		
Total Cholesterol	*TDM1*	rs138777	*HMGXB4*	rs9306298	rs138728	0.343468	0.88858	130	Tissue 1, Tissue 2, Tissue 3	EWSR1-FLI1,ZNF263		SGBS 1, SGBS 2	EWSR1-FLI1,ZNF263	
Total Cholesterol	*TDM1*	rs138777	*HMGXB4*	rs9306298	rs138733	0.336227	0.860215	130	Tissue 1, Tissue 3			SGBS 1, SGBS 2		
Total Cholesterol	*TDM1*	rs138777	*HMGXB4*	rs9306298	rs138774	0.369259	1	130				SGBS 1, SGBS 2	ATF4,RUNX1	
Total Cholesterol	*TDM1*	rs138777	*HMGXB4*	rs9306298	rs138775	0.369457	0.999149	130				SGBS 1, SGBS 2	ZIC1	
Total Cholesterol	*TDM1*	rs138777	*HMGXB4*	rs9306298	rs1534877	0.369557	0.991919	130				SGBS 1, SGBS 2	RXRA::VDR	
Total Cholesterol	*TDM1*	rs138777	*HMGXB4*	rs9306298	rs2235145	0.369507	0.992133	130				SGBS 1	RFX5,TFAP2B(var.2)	
Total Cholesterol	*TDM1*	rs138777	*HMGXB4*	rs9306298	rs3215443	0.369408	0.992555	130	Tissue 1, Tissue 2, Tissue 3	KLF16,SP4,SP8		SGBS 1, SGBS 2	KLF16,SP4,SP8	

Total Cholesterol	*TOM1*	rs138777	*HMGXB4*	rs9306298	rs4461	0.368565	0.991492	130	Tissue 1, Tissue 2, Tissue 3			SGBS 1, SGBS 2		
Total Cholesterol	*TOM1*	rs138777	*HMGXB4*	rs9306298	rs4462	0.36911	0.993829	130				SGBS 1, SGBS 2		
Total Cholesterol	*TOM1*	rs138777	*HMGXB4*	rs9306298	rs4463	0.372384	0.979502	130	Tissue 3	SP1,SP2		SGBS 1, SGBS 2	RREB1,SP1,SP2	RREB1
Total Cholesterol	*TOM1*	rs138777	*HMGXB4*	rs9306298	rs4465	0.369259	1	130				SGBS 1, SGBS 2	Spz1,TFAP4	
Total Cholesterol	*TOM1*	rs138777	*HMGXB4*	rs9306298	rs5755674	0.330771	0.833501	130	Tissue 1, Tissue 2, Tissue 3	Rfx1		SGBS 1, SGBS 2		
Total Cholesterol	*TOM1*	rs138777	*HMGXB4*	rs9306298	rs67921523	0.370747	0.988118	130		E2F6,EGR3,EGR4,EWSR1-FLI1,KLF16,KLF5,MZF1(var.2),NR2C2,RREB1,SP1,SP2,SP3,SP8,ZNF740		SGBS 1, SGBS 2	Stat5a::Stat5b E2F6,EGR3,EGR4,SP1,SP2,ZNF263	Stat5a::Stat5
Triglycerides	*APOE-TOMM40*	rs439401	*APOE*	rs439401	rs439401	0.270558	N/A	3	Tissue 1, Tissue 3			SGBS 1, SGBS 2		
Triglycerides	*FADS1*	rs174548	*FADS1*	rs174555	11:61594920	0.427438	0.887534	48				SGBS 1, SGBS 2		
Triglycerides	*FADS1*	rs174548	*NDST1*	rs174555	11:61596322	0.427587	0.888178	48				SGBS 1, SGBS 2	Pparg::Rxra	
Triglycerides	*FADS1*	rs174548	*FADS1*	rs174555	rs174529	0.426991	0.899905	48				SGBS 1, SGBS 2	REST	
Triglycerides	*FADS1*	rs174548	*FADS1*	rs174555	rs174535	0.425305	0.906955	48				SGBS 1		
Triglycerides	*FADS1*	rs174548	*FADS1*	rs174555	rs174561	0.395496	0.968415	48	Tissue 1, Tissue 2, Tissue 3	EGR3,NFKB1		SGBS 1, SGBS 2	EGR3 ESRRB,NFIC::TLX1,NR4A2,Nr5a2,RUNX1	
Triglycerides	*FADS1*	rs174548	*FADS1*	rs174555	rs97384	0.434977	0.870961	48				SGBS 2		
Triglycerides	*JMJD1C*	rs10761731	*JMJD1C,NRBF2*	rs10761739,rs7896518	rs10761731	0.366531	N/A	16				SGBS 1, SGBS 2		
Triglycerides	*JMJD1C*	rs10761731	*JMJD1C,NRBF2*	rs10761739,rs7896518	rs7073753	0.366333	0.998719	16				SGBS 1	HSF1,HSF2,HSF4,IRF7,IRF8	
Triglycerides	*MLXIPL*	rs3812316	*BCL7B,TBL2*	rs2240466,rs799166	7:73022934	0.117597	0.996183	55				SGBS 1	IRF1,PRDM1,SPIC,STAT1::STAT2	
Triglycerides	*MLXIPL*	rs3812316	*BCL7B,TBL2*	rs2240466,rs799166	rs13231516	0.116209	0.964795	55	Tissue 1, Tissue 3			SGBS 1, SGBS 2		
Triglycerides	*MLXIPL*	rs3812316	*BCL7B,TBL2*	rs2240466,rs799166	rs13233571	0.118143	0.984322	55	Tissue 1	MEIS2		SGBS 1, SGBS 2	MEIS2	
Triglycerides	*MLXIPL*	rs3812316	*BCL7B,TBL2*	rs2240466,rs799166	rs13246490	0.118143	0.984322	55				SGBS 1	EWSR1-FLI1	
Triglycerides	*MLXIPL*	rs3812316	*BCL7B,TBL2*	rs2240466,rs799166	rs34430945	0.116853	0.907456	55				SGBS 1, SGBS 2		
Triglycerides	*MLXIPL*	rs3812316	*BCL7B,TBL2*	rs2240466,rs799166	rs799165	0.117399	0.901648	55	Tissue 1	ELK3,ELK4,ERF,ERG,ETS1,ETV2,ETV6,FEV,FLI1,Gabpa			CTCF,Myod1,ZIC1,ZIC3,ZIC4,ZNF410	
Triglycerides	*MPP3*	rs8077889	*DUSP3,MPP3*	rs2342310,rs55768269	rs15359	0.246801	0.822301	97				SGBS 2		
Triglycerides	*MPP3*	rs8077889	*DUSP3,MPP3*	rs2342310,rs55768269	rs2074143	0.243875	0.847199	97	Tissue 1			SGBS 1, SGBS 2		
Triglycerides	*MPP3*	rs8077889	*DUSP3,MPP3*	rs2342310,rs55768269	rs72836556	0.213917	0.985647	97	Tissue 1, Tissue 3	FOXA1,FOXC2,FOXF2,Foxa2,Foxj3,NFATC1,SOX8		SGBS 1, SGBS 2	FOXA1,FOXC2,FOXF2,Foxa2,Foxj3,NFATC1,SOX8	
Triglycerides	*MPP3*	rs8077889	*DUSP3,MPP3*	rs2342310,rs55768269	rs746742	0.191896	0.808194	97	Tissue 1, Tissue 2, Tissue 3	EWSR1-FLI1,ZNF263				
Triglycerides	*MPP3*	rs8077889	*DUSP3,MPP3*	rs2342310,rs55768269	rs9894873	0.213917	0.985647	97	Tissue 1, Tissue 3	NFATC1,NFATC2		SGBS 1, SGBS 2	NFATC1,NFATC2	
Triglycerides	*MPP3*	rs8077889	*DUSP3,MPP3*	rs2342310,rs55768269	rs901676	0.215455	0.976618	97	Tissue 1, Tissue 3	FOXA1,FOXC2,FOXD1,FOXD2,FOXF2,FOXG1,FOXI1,FOXL1,FOXO3,FOXO4,FOXO6,FOXP3,Foxa2,Foxj3,Foxq1,SOX8		SGBS 1, SGBS 2	FOXA1,FOXC2,FOXD1,FOXD2,FOXF2,FOXG1,FOXI1,FOXL1,FOXO3,FOXO4,FOXO6,FOXP3,Foxa2,Foxj3,Foxq1,SOX8	
Type 2 diabetes	*AP3S2*	rs2028299	*ARPIN (C15orf38)*	rs7174878	15:90371590	0.255729	0.986212	58	Tissue 1			SGBS 1, SGBS 2		
Type 2 diabetes	*AP3S2*	rs2028299	*ARPIN (C15orf38)*	rs7174878	rs12910992	0.256225	0.995061	58				SGBS 1	CTCF	
Type 2 diabetes	*AP3S2*	rs2028299	*ARPIN (C15orf38)*	rs7174878	rs2165069	0.254389	0.989056	58				SGBS 1, SGBS 2		
Type 2 diabetes	*AP3S2*	rs2028299	*ARPIN (C15orf38)*	rs7174878	rs4932143	0.255778	0.996875	58	Tissue 1, Tissue 3			SGBS 1, SGBS 2		
Type 2 diabetes	*AP3S2*	rs2028299	*ARPIN (C15orf38)*	rs7174878	rs4932144	0.254538	0.989316	58				SGBS 1, SGBS 2	Rfx1	
Type 2 diabetes	*AP3S2*	rs2028299	*ARPIN (C15orf38)*	rs7174878	rs6496609	0.255282	1	58				SGBS 1	EHF,ELF4,ELF5	
Type 2 diabetes	*AP3S2*	rs2028299	*ARPIN (C15orf38)*	rs7174878	rs8031576	0.255282	1	58				SGBS 1, SGBS 2		
Type 2 diabetes	*AP3S2*	rs2028299	*ARPIN (C15orf38)*	rs7174878	rs986505	0.256473	0.977689	58	Tissue 1			SGBS 1, SGBS 2		
Type 2 diabetes	*IGF2BP2*	rs4402960	*IGF2BP2*	rs35430985	rs150111048	0.324621	0.999095	80				SGBS 1, SGBS 2	FOXP1,IRF1,PRDM1	IRF1
Type 2 diabetes	*MPHOSPH9*	rs1727313	*C12orf65 (SPG55, COXPD7),CDK2AP1,SBNO1*	rs11057206,rs1616131, rs28583837	rs1106240	0.194028	0.999683	215				SGBS 1	CTCF,HIC2	
Type 2 diabetes	*MPHOSPH9*	rs1727313	*C12orf65 (SPG55, COXPD7),CDK2AP1,SBNO1*	rs11057206,rs1616131, rs28583837	rs1106241	0.19512	0.993376	215				SGBS 1		
Type 2 diabetes	*MPHOSPH9*	rs1727313	*C12orf65 (SPG55, COXPD7),CDK2AP1,SBNO1*	rs11057206,rs1616131, rs28583837	rs7313483	0.217141	0.864152	215				SGBS 2		
Type 2 diabetes	*MPHOSPH9*	rs1727313	*C12orf65 (SPG55, COXPD7),CDK2AP1,SBNO1*	rs11057206,rs1616131, rs28583837	rs7485502	0.223738	0.81857	215	Tissue 1	NRF1		SGBS 1, SGBS 2	NRF1	
Type 2 diabetes	*MPHOSPH9*	rs1727313	*C12orf65 (SPG55, COXPD7),CDK2AP1,SBNO1*	rs11057206,rs1616131, rs28583837	rs884548	0.19512	0.993376	215				SGBS 2		
Type 2 diabetes	*MPHOSPH9*	rs1727313	*C12orf65 (SPG55, COXPD7),CDK2AP1,SBNO1*	rs11057206,rs1616131, rs28583837	rs949142	0.19507	0.99369	215				SGBS 2		
Type 2 diabetes	*SDHAF4 (C6orf57)*	rs1048886	*SDHAF4 (C6orf57)*	rs6901232	6:71224414	0.248041	0.932155	303	Tissue 1, Tissue 2, Tissue 3	PAX7		SGBS 1, SGBS 2		
Type 2 diabetes	*SDHAF4 (C6orf57)*	rs1048886	*SDHAF4 (C6orf57)*	rs6901232	rs112809632	0.224581	0.877106	303	Tissue 2	EGR1,NR2C2,NRF1,SP2,SP8		SGBS 1, SGBS 2	EGR1,NR2C2,NRF1,SP2,SP8	
Type 2 diabetes	*SDHAF4 (C6orf57)*	rs1048886	*SDHAF4 (C6orf57)*	rs6901232	rs12526164	0.224531	0.920498	303				SGBS 1, SGBS 2		
Type 2 diabetes	*SDHAF4 (C6orf57)*	rs1048886	*SDHAF4 (C6orf57)*	rs6901232	rs2293297	0.224581	0.920218	303				SGBS 1, SGBS 2		
Type 2 diabetes	*SDHAF4 (C6orf57)*	rs1048886	*SDHAF4 (C6orf57)*	rs6901232	rs4283855	0.248041	0.932155	303	Tissue 1					
Type 2 diabetes	*SDHAF4 (C6orf57)*	rs1048886	*SDHAF4 (C6orf57)*	rs6901232	rs55968813	0.224531	0.920498	303	Tissue 1					
Type 2 diabetes	*SDHAF4 (C6orf57)*	rs1048886	*SDHAF4 (C6orf57)*	rs6901232	rs6455370	0.248041	0.932682	303				SGBS 2	MEIS2,MEIS3	
Type 2 diabetes	*SDHAF4 (C6orf57)*	rs1048886	*SDHAF4 (C6orf57)*	rs6901232	rs6911983	0.224482	0.920778	303				SGBS 1, SGBS 2		
Type 2 diabetes	*SDHAF4 (C6orf57)*	rs1048886	*SDHAF4 (C6orf57)*	rs6901232	rs75589353	0.221109	0.902394	303				SGBS 1		
Type 2 diabetes	*SDHAF4 (C6orf57)*	rs1048886	*SDHAF4 (C6orf57)*	rs6901232	rs77554252	0.221109	0.902394	303				SGBS 1		
Type 2 diabetes	*SDHAF4 (C6orf57)*	rs1048886	*SDHAF4 (C6orf57)*	rs6901232	rs9455144	0.24809	0.932432	303				SGBS 1		
Venous thromboembolism	*F5*	rs6427196	*F5*	rs2420371	rs6427194	0.0470687	1	16				SGBS 1, SGBS 2	Gfi1b,MYF6	
Venous thromboembolism	*F5*	rs6427196	*F5*	rs2420371	rs6427195	0.0470687	1	16				SGBS 1, SGBS 2	LEF1,T	
Venous thromboembolism	*F5*	rs6427196	*F5*	rs2420371	rs6427196	0.0470687	N/A	16				SGBS 1, SGBS 2	MEF2C	
Venous thromboembolism	*intergenic*	rs10504130	*PCMTD1*	rs80143499	rs10504130	0.253199	N/A	94				SGBS 1, SGBS 2	STAT1::STAT2	
Venous thromboembolism	*intergenic*	rs10504130	*PCMTD1*	rs80143499	rs11781330	0.253199	1	94				SGBS 1, SGBS 2		
Venous thromboembolism	*intergenic*	rs10504130	*PCMTD1*	rs80143499	rs12681203	0.279536	0.872865	94				SGBS 1, SGBS 2	ZNF143	
Venous thromboembolism	*intergenic*	rs6695223	*WDR63*	rs10782538	rs6681402	0.0905168	0.996397	20	Tissue 1, Tissue 2, Tissue 3	Hoxc9,MEF2A,MEF2B		SGBS 1, SGBS 2		
Venous thromboembolism	*intergenic*	rs6695223	*WDR63*	rs10782538	rs6681619	0.0910624	0.989825	20	Tissue 1, Tissue 2, Tissue 3			SGBS 1, SGBS 2		
Venous thromboembolism	*intergenic*	rs10504130	*PCMTD1*	rs80143499	rs76649221	0.2531	0.999475	94				SGBS 1		
Waist-hip ratio	*CPEB4*	rs6861681	*CPEB4*	rs72812818	5:173315416	0.446087	0.893898	79	Tissue 1, Tissue 3	CTCF,EGR1,EGR3,SP2		SGBS 1, SGBS 2	CTCF,EGR1,EGR3,SP2 RARA(var.2),RXRA::VDR,Rarb(var.2)	CTCF
Waist-hip ratio	*CPEB4*	rs6861681	*CPEB4*	rs72812818	rs10516107	0.465331	0.999801	79				SGBS 1		
Waist-hip ratio	*CPEB4*	rs6861681	*CPEB4*	rs72812818	rs17695555	0.465331	0.999801	79				SGBS 1, SGBS 2		
Waist-hip ratio	*CPEB4*	rs6861681	*CPEB4*	rs72812818	rs17763772	0.462851	0.981536	79				SGBS 1		
Waist-hip ratio	*CPEB4*	rs6861681	*CPEB4*	rs72812818	rs55946741	0.465331	0.999801	79	Tissue 1, Tissue 3			SGBS 1, SGBS 2		
Waist-hip ratio	*CPEB4*	rs6861681	*CPEB4*	rs72812818	rs72810974	0.452733	0.947975	79				SGBS 1, SGBS 2	CLOCK,MYF6,SCRT1,TFEB	
Waist-hip ratio	*CPEB4*	rs6861681	*CPEB4*	rs72812818	rs72812811	0.465331	0.999801	79				SGBS 1, SGBS 2	REL,RELA	
Waist-hip ratio	*CPEB4*	rs6861681	*CPEB4*	rs72812818	rs747472	0.465182	0.999203	79				SGBS 1, SGBS 2		
Waist-hip ratio	*NISCH-STAB1*	rs6784615	*NISCH*	rs728408	3:52505756	0.0309493	0.937835	99	Tissue 1, Tissue 2, Tissue 3					
Waist-hip ratio	*NISCH-STAB1*	rs6784615	*NISCH*	rs728408	3:52505758	0.0309493	0.937835	99	Tissue 1, Tissue 2, Tissue 3					
Waist-hip ratio	*NISCH-STAB1*	rs6784615	*NISCH*	rs728408	rs13065677	0.0306517	0.885413	99	Tissue 1, Tissue 2, Tissue 3			SGBS 1, SGBS 2		
Waist-hip ratio	*NISCH-STAB1*	rs6784615	*NISCH*	rs728408	rs13076508	0.0319413	0.848076	99				SGBS 1		
Waist-hip ratio	*NISCH-STAB1*	rs6784615	*NISCH*	rs728408	rs13096571	0.0302053	0.899079	99				SGBS 1	ATF4	
Waist-hip ratio	*NISCH-STAB1*	rs6784615	*NISCH*	rs728408	rs34827613	0.0300069	0.905284	99				SGBS 1, SGBS 2		
Waist-hip ratio	*NISCH-STAB1*	rs6784615	*NISCH*	rs728408	rs3774409	0.0309493	0.937835	99				SGBS 1		
Waist-hip ratio	*NISCH-STAB1*	rs6784615	*NISCH*	rs728408	rs4312651	0.0309493	0.937835	99				SGBS 1		
Waist-hip ratio	*NISCH-STAB1*	rs6784615	*NISCH*	rs728408	rs758800	0.0310485	0.934718	99	Tissue 3	GATA1::TAL1				
Waist-hip ratio	*NISCH-STAB1*	rs6784615	*NISCH*	rs728408	rs929529	0.0309493	0.937835	99	Tissue 3	ZNF263				
Waist-hip ratio	*NISCH-STAB1*	rs6784615	*NISCH*	rs728408	rs9863753	0.0329332	1	99				SGBS 1, SGBS 2		
Weight	*AIF1*	rs115963686	*CSNK2B*	rs114899308	6:31588020	0.324224	0.820463	31	Tissue 1, Tissue 2, Tissue 3	EGR1,EWSR1-FLI1,IRF1,PRDM1,ZNF263	IRF1	SGBS 1, SGBS 2	EGR1,EWSR1-FLI1,IRF1,PRDM1,ZNF263	IRF1
Weight	*AIF1*	rs115963686	*CSNK2B*	rs114899308	rs114167507	0.289803	0.949383	31		ATF4,ATF7,BATF3,Creb5,Crem,DBP,HLF,JDP2(var.2),JUND(var.2),Lhx3,NFIL3,PHOX2A,PROP1,Phox2b,TEF		SGBS 1, SGBS 2	ATF4,ATF7,BATF3,Creb5,Crem,DBP,HLF,JDP2(var.2),JUND(var.2),NFIL3,TEF	
Weight	*AIF1*	rs115963686	*CSNK2B*	rs114899308	rs114195058	0.324323	0.822271	31	Tissue 1, Tissue 2, Tissue 3			SGBS 1, SGBS 2		
Weight	*AIF1*	rs115963686	*CSNK2B*	rs114899308	rs116665625	0.323133	0.815493	31	Tissue 1, Tissue 2, Tissue 3	TBP		SGBS 1, SGBS 2		
WHRadjBMI	*ADAMTS9*	rs2371767	*ADAMTS9*	rs4616635	rs66815886	0.220811	0.995977	9	Tissue 3	FIGLA,TCF3,TCF4				
WHRadjBMI	*ADAMTS9*	rs2371767	*ADAMTS9*	rs4616635	rs7433808	0.189912	0.817086	9	Tissue 3	HSF4,SMAD3		SGBS 1	SMAD3	
WHRadjBMI	*GDF5*	rs224333	*UQCC*	rs2425055	rs143384	0.442863	0.8836	104				SGBS 1, SGBS 2		

WHRadjBMI	*GDF5*	rs224333	*UQCC*	rs2425055	rs224331	0.399564	0.939297	104				SGBS 1		
WHRadjBMI	*GDF5*	rs224333	*UQCC*	rs2425055	rs2425066	0.426247	0.911495	104				SGBS 1, SGBS 2		
WHRadjBMI	*GDF5*	rs224333	*UQCC*	rs2425055	rs3795159	0.458536	0.822012	104				SGBS 1, SGBS 2	ESR1,RREB1	RREB1
WHRadjBMI	*GDF5*	rs224333	*UQCC*	rs2425055	rs6088791	0.458437	0.822343	104				SGBS 1, SGBS 2	Stat4,ZNF263	
WHRadjBMI	*GDF5*	rs224333	*UQCC*	rs2425055	rs6579234	0.430017	0.923435	104				SGBS 1		
WHRadjBMI	*GDF5*	rs224333	*UQCC*	rs2425055	rs78110303	0.41603	0.984763	104				SGBS 1, SGBS 2	EGR1,IRF2,KLF16,PRDM1,SP8	
WHRadjBMI	*GORAB*	rs10919388	*PRRX1*	rs6427242	rs10753805	0.233161	0.941605	27				SGBS 1, SGBS 2		
WHRadjBMI	*GORAB*	rs10919388	*PRRX1*	rs6427242	rs6427242	0.23336	0.934523	27				SGBS 1, SGBS 2	Nkx2-5(var.2)	
WHRadjBMI	*GORAB*	rs10919388	*PRRX1*	rs6427242	rs824840	0.234054	0.946186	27				SGBS 2	GSX2,HNF1A,HNF1B,HOXB13,HOXD13,SRY	
WHRadjBMI	*JUND*	rs12608504	*JUND*	rs7247222	rs10416600	0.341633	0.998457	19				SGBS 1, SGBS 2	KLF5,Klf1,Klf12,Klf4,RREB1,SP3,SP4	
WHRadjBMI	*JUND*	rs12608504	*JUND*	rs7247222	rs7247222	0.347981	0.939994	19	Tissue 1, Tissue 2, Tissue 3	PAX1,PAX5,PAX9		SGBS 1, SGBS 2	PAX1,PAX5,PAX9	
WHRadjBMI	*LEKR1*	rs17451107	*TIPARP*	rs10049090	rs13322435	0.315346	0.936043	16	Tissue 1, Tissue 2, Tissue 3	ZEB1		SGBS 1, SGBS 2	ZEB1	
WHRadjBMI	*LEKR1*	rs17451107	*TIPARP*	rs10049090	rs56406311	0.301359	0.999529	16	Tissue 1, Tissue 2, Tissue 3	NFIC::TLX1,Nr5a2,PLAG1,TFAP2A,TFAP2A(var.2),TFAP2B,TFAP2B(var.2),TFAP2C,TFAP2C(var.2)		SGBS 1, SGBS 2	NFIC::TLX1,Nr5a2,TFAP2A,TFAP2A(var.2),TFAP2B,TFAP2B(var.2),TFAP2C,TFAP2C(var.2)	
WHRadjBMI	*LEKR1*	rs17451107	*TIPARP*	rs10049090	rs9817452	0.301805	0.997412	16	Tissue 1, Tissue 2, Tissue 3	NFIC::TLX1,Nr5a2,TFAP2A,TFAP2A(var.2),TFAP2B,TFAP2B(var.2),TFAP2C,TFAP2C(var.2)		SGBS 1, SGBS 2	NFIC::TLX1,Nr5a2,TFAP2A,TFAP2A(var.2),TFAP2B,TFAP2B(var.2),TFAP2C,TFAP2C(var.2)	
WHRadjBMI	*LEKR1*	rs17451107	*TIPARP*	rs10049090	rs9854955	0.315296	0.936258	16	Tissue 1, Tissue 2, Tissue 3			SGBS 1, SGBS 2		
WHRadjBMI	*MACROD1*	rs11231693	*VEGFB*	rs56271783	rs2701540	0.0958734	0.902817	12				SGBS 1, SGBS 2	ALX3,Dlx4	
WHRadjBMI	*MACROD1*	rs11231693	*VEGFB*	rs56271783	rs2845885	0.0958734	0.902817	12				SGBS 1, SGBS 2	ATF7,JDP2(var.2),JUN,JUND(var.2),NR2F1,Rarg(var.2),Vdr	
WHRadjBMI	*NKX3-6*	rs7830933	*STC1*	rs6983481	rs11986562	0.196607	0.910969	37	Tissue 1, Tissue 3			SGBS 1, SGBS 2		
WHRadjBMI	*NKX3-6*	rs7830933	*STC1*	rs6983481	rs13273073	0.208908	0.99191	37				SGBS 2		
WHRadjBMI	*NKX3-6*	rs7830933	*STC1*	rs6983481	rs57220043	0.209057	1	37				SGBS 1, SGBS 2	Bcl6,STAT1	
WHRadjBMI	*NKX3-6*	rs7830933	*STC1*	rs6983481	rs6983481	0.196607	0.910969	37	Tissue 1, Tissue 3			SGBS 1, SGBS 2	KLF14,KLF16,KLF5,Klf1,Klf12,Klf4,SP1,SP2,SP3,SP4,SP8	
WHRadjBMI	*NTSDC2*	rs12489828	*NTSDC2,SMIM4 (C3orf78)*	rs6778735,rs7639267	3:52552922	0.474556	0.846162	18				SGBS 2		
WHRadjBMI	*NTSDC2*	rs12489828	*NTSDC2,SMIM4 (C3orf78)*	rs6778735,rs7639267	rs12489828	0.456453	N/A	18				SGBS 1		
WHRadjBMI	*NTSDC2*	rs12489828	*NTSDC2,SMIM4 (C3orf78)*	rs6778735,rs7639267	rs66782572	0.452832	0.916749	18				SGBS 1		
WHRadjBMI	*NTSDC2*	rs12489828	*NTSDC2,SMIM4 (C3orf78)*	rs6778735,rs7639267	rs7614981	0.456552	0.999201	18				SGBS 1	Stat5a::Stat5b	
WHRadjBMI	*NTSDC2*	rs12489828	*NTSDC2,SMIM4 (C3orf78)*	rs6778735,rs7639267	rs7639267	0.448071	0.965537	18				SGBS 2		
WHRadjBMI	*RSPO3*	rs72959041	*RSPO3*	rs72961007	6:127440047	0.0682968	0.947962	5				SGBS 1, SGBS 2		
WHRadjBMI	*SNX10*	rs1534696	*CBX3,SNX10*	rs1534696	rs1534696	0.464041	N/A	1	Tissue 1	MEF2C				
WHRadjBMI	*TNFAIP8*	rs1045241	*TNFAIP8*	rs7703744	rs1045241	0.326753	N/A	11				SGBS 1, SGBS 2	POU1F1,POU2F1,POU2F2,POU3F1,POU3F2,POU3F3,POU3F4,POU5F1B,Pou2f3,Pou5f1::Sox2	
WHRadjBMI	*TNFAIP8*	rs1045241	*TNFAIP8*	rs7703744	rs1045242	0.337268	0.953692	11				SGBS 1, SGBS 2		
WHRadjBMI	*TNFAIP8*	rs1045241	*TNFAIP8*	rs7703744	rs11064	0.313213	0.939215	11				SGBS 1, SGBS 2	MYF6,RREB1	RREB1
WHRadjBMI	*TNFAIP8*	rs1045241	*TNFAIP8*	rs7703744	rs3203922	0.326902	0.999324	11				SGBS 1, SGBS 2	LEF1	

Table 28: Allelic imbalance in ATAC peaks

Sample	Number of peaks	Heterozygous Sites	Heterozygous sites overlapping ATAC signala	Imbalanced (*p* <0.05)
Tissue 1	58,279	1,894,559	6,461	381
Tissue 2	29,409	1,890,071	3,631	230
Tissue 3	49,631	1,885,394	3,836	220
Sites in at least 2 samples	N/A	1,451,379	2,267	18
Sites in all 3 samples	N/A	269,858	318	1

Required at least 10 total reads and >= 1 read per allele

p- values were generated using a two-tailed beta-binomial test

Table 29: Variants at GWAS-eQTL colocalized loci with allelic imbalance in ATAC-seq reads

			Tissue 1			Tissue 3							
ATAC variant	MAF	Alleles[a]	Allele 1 count	Allele 2 count	*P* value	Allele 1 count	Allele 2 count	*P* value	GWAS Trait	eQTL Gene(s)	eQTL beta[a]	Adipose Nuclei Chromatin State	Transcription Factor Motifs
rs11662724	0.22	C/T	20	5	0.02	homozygous		N/A	Phospholipids	*ESCO1*	-0.63	Promoter	EGR3, EGR4, PAX5, PAX9, TFAP2A, TFAP2B, TFAP2C
rs7187776	0.45	A/G	25	3	7.05E-04	homozygous		N/A	Obesity	*SH2B1, ATXN2L*	0.40, 0.34	Promoter	SPI1, SPIB
rs9854955	0.32	G/A	15	5	0.04	14	3	0.01	Obesity	*TIPARP*	0.41	Enhancer	
rs13322435	0.32	G/A	18	1	3.13E-04	15	3	3.23E-03	Obesity	*TIPARP*	0.41	Enhancer	ZEB1
rs56406311	0.30	T/C	12	3	0.04	homozygous		N/A	Obesity	*TIPARP*	0.41	Enhancer	NFIC::TLX1, NR5A2, PLAG1, TFAP2A, TFAP2B, TFAP2C
rs9817452	0.30	T/G	14	4	0.04	homozygous		N/A	Obesity	*TIPARP*	0.41	Enhancer	NFIC::TLX1, NR5A2, TFAP2A, TFAP2B, TFAP2C
rs111447985	0.06	C/A	17	5	0.04	11	6	0.33	Cardiac hypertrophy	*RP11-541N10.3, OBFC1*	0.49, 1.04	Promoter	ELK1, ELK3, ELK4, ERF, ETS1, ETV1, ETV4, ETV5, FEV, FLI1, GABPA, TFAP2A, TFAP2B, TFAP2C
rs73597582	0.15	A/G	9	1	0.02	homozygous		N/A	HDL cholesterol	*GFOD, NUTF2*	-0.34, -0.30	Flanking Bivalent TSS/Enhancer	

No significant allelic imbalances were identified in tissue sample 2
Required at least 10 total reads and >= 1 read per allele

[a]Alleles: effect allele/non-effect allele. Effect allele for ATAC-seq is the allele with greater ATAC-seq reads.

CHAPTER 5: DISCUSSION

GWAS have identified hundreds of loci associated with cardiometabolic diseases and traits (www.ebi.ac.uk/gwas/). To date, relatively few mechanisms have been identified at these loci. GWAS are powerful studies that identify many associated loci; however, functionally characterizing these loci presents a multi-faceted challenge and can be time consuming. Challenges include the noncoding nature of many GWAS loci, multiple variants in LD at each locus, the complexity of the underlying mechanisms themselves, and the time and cost of evaluating each locus individually. Statistical fine-mapping, overlap of predicted genomic regulatory elements, and functional assays can be combined to prioritize and identify causal variants. Genome-wide approaches are needed to better characterize the regulatory landscape of cardiometabolic tissues. Identifying causal variants is a critical step in understanding the genetic mechanisms and etiology of complex cardiometabolic diseases.

The noncoding nature of many GWAS loci makes characterizing biological function challenging. The most probable mechanism is that variants act by altering gene regulation (223). Thus, understanding the regulatory landscape of cardiometabolic-relevant tissues is necessary to prioritize candidate variants. The regulatory landscape includes regions of open chromatin, sites of transcription factor binding, and histone marks; these data can be combined to create predicted chromatin states such as promoters, enhancers, and repressors. The ENCODE and Roadmap Epigenomics projects (59,60) have generated a wealth of regulatory datasets in diverse cell types; however, data for some cell types is still limited, particularly for adipose tissue and

pancreatic islets. In Chapter 4, we generated ATAC-seq open chromatin profiles from human subcutaneous adipose tissue and the preadipocyte cell strain, SGBS. By interrogating the open chromatin profiles of adipose tissue and preadipocytes, we can better prioritize candidate variants for functional experiments and identify disease-relevant regulatory regions. As shown in Chapter 4, ATAC-seq open chromatin alone identified a regulatory variant at the *SNX10* WHR GWAS locus, and ATAC-seq peaks were enriched in disease-relevant adipose enhancers. Cellular environment can alter chromatin state, especially for cardiometabolic phenotypes, where both genetic and environmental factors contribute to phenotype; additional experiments are needed to characterize regulatory landscapes across various cellular conditions (e.g. mature adipocytes and insulin or glucose treatment). Together, the ATAC-seq data generated from both adipose tissue and preadipocytes add to the growing understanding of gene regulation and the genetics of cardiometabolic traits and diseases.

Variants at a GWAS locus can be prioritized using statistical fine-mapping analyses such as MANTRA, CAVIAR, and PAINTOR and annotation with evidence of regulatory elements. While statistical fine-mapping analyses are informative and can help the variant prioritization process, several limitations exist. Most statistical fine-mapping analyses return a list of variants, or "credible set", that could contribute to the association signal, which requires all or a subset of variants to be tested in functional assays requiring significant time and resources. Another limitation of statistical fine-mapping are the inconsistent results that can arise by comparing existing statistical methods. Each method makes different assumptions about the contributions of underlying variants and applies different strategies incorporating genetic association, LD, and functional annotation. At the *ANGPTL8* locus, we used three fine-mapping methods (MANTRA, CAVIAR, and PAINTOR) to prioritize candidate variants. MANTRA identified a credible set of

ten variants. CAVIAR identified 24 variants in the Finnish credible set and two in the African American credible set. PAINTOR identified ten, seven, and five variants in Finnish, African American, and the combined studies, respectively. Of the 39 variants identified in at least one fine-mapping analysis, only 12 were identified in at least two analyses, and only 4 in all three analyses. From my experience, current fine-mapping analyses are most informative when combined with regulatory overlap and functional testing of all probable variants.

The genetic mechanisms at GWAS loci can complicate the identification of causal variants. Identifying the target gene(s) at a GWAS locus is a fundamental part of elucidating the molecular mechanism; however, these genes remain largely unknown for most GWAS loci. GWAS loci in noncoding regions can have multiple genes that appear to be good functional candidates based on gene function and expression, rare coding variant effects, chromosome interactions, and/or literature review. To complicate matters further, multiple genes can be targeted at a single locus (224). eQTL data can be used to identify which gene(s) are most likely to be targeted at a given locus. In Chapter 3, we showed a subcutaneous adipose eQTL association for both *ANGPTL8* and *DOCK6.* Biological roles and a stronger eQTL association suggest *ANGPTL8* is the target gene; however, further experiments are needed to fully delineate the role of *ANGPTL8* and/or *DOCK6.* When we published the *CDC123/CAMK1D* work, eQTL associations had been identified for both *CDC123* and *CAMK1D* in whole blood and lung tissue (84,225). Since our publication, in conjunction with collaborators, we have published a large pancreatic islet eQTL study, which identified an eQTL association with GWAS variants, including rs11257655, for *CAMK1D* (P=4.2x10^{-9}) but not *CDC123* (61). The eQTL association in pancreatic islets provides strong evidence that *CAMK1D* is a target gene and that *CAMK1D* plays a role in T2D biology. eQTL studies in additional tissues with larger sample sizes,

especially in liver and other cardiometabolic tissues, are needed to help delineate additional target genes at these and other GWAS loci. Additionally, variants may only affect expression in a specific cellular environment (e.g. when transcriptional regulators are induced); thus, context-specific eQTL studies may be necessary to identify the missing variant-gene connections. Understanding how associated variants act on gene expression is a vital piece to the complex genetics puzzle.

While a single functional variant is the simplest explanation of a GWAS signal, and such mechanisms exist (53,226,227), many loci consist of multiple functional variants (54,138). At a single locus, variants may be acting in concert in promoters, enhancers, and repressors to alter gene expression, which adds to the complexity of determining mechanisms at GWAS loci. At the *CDC123/CAMK1D* T2D GWAS locus, we identified a straightforward molecular mechanism with one functional variant, rs11257655, which displayed allelic differences in both transcriptional activity and binding of FOXA1 and FOXA2 (Chapter 2). At the *ANGPTL8* HDL-C GWAS locus, the mechanism was more complex (Chapter 3). What began as a straightforward locus with only four candidate regulatory variants (r^2>0.8 in European populations), turned into an extremely complex locus with 42 candidate regulatory variants (r^2>0.5 in African and/or European populations), in which I identified seven variants that exhibited allelic differences in regulatory activity in cells or *in vitro*. Evaluating variants at individual GWAS loci allows for the most thorough and accurate identification of potentially functional variants.

The functional assays described in this dissertation (i.e. transcriptional reporter and *in vitro* DNA-protein binding assays) elucidate limitations to and effort involved in experimentally validating candidate variants. Transcriptional reporter luciferase assays are limited in that they do not represent the native chromatin context. The cloned plasmid does not contain nucleosomes

and cannot loop to form the proper chromatin structure present in the cell. We identified functional variants at the *CDC123/CAMK1D* and *ANGPTL8* GWAS loci, but we may have missed potentially functional variants. In addition, regulatory regions tested may show a stronger effect if investigated within the native chromatin context. EMSA experiments determine if proteins are binding to the region of DNA containing a variant; however, they are *in vitro* assays and transcription factors may bind *in vitro* more readily than in cells. Additionally, it is often difficult to determine the identity/identities of the transcription factor(s) responsible for regulatory mechanisms. Transcription factor binding motifs and chromatin immunoprecipitation (ChIP-seq) data can be informative for identifying proteins binding to variant alleles. Moreover, transcription factor footprints identified in sequencing data can be more informative than motifs alone, as demonstrated in Chapter 4 using ATAC-seq data. Even with motifs, ChIP-seq data, and footprints, identifying the correct factor remains challenging. For example, at the *ANGPTL8* locus (Chapter 3), I performed supershift assays using ~40 antibodies for candidate transcription factors, many of which had predicted binding motifs or previous evidence of at least indirect binding from ChIP-seq data, to attempt to determine the identity of the allele-specific protein complex binding to rs737337-C. Only one antibody (RXRα) showed a supershift; however, the supershift was observed in both alleles, suggesting RXRα is not the allele-specific protein. The lack of supershifts may be due to poor antibody specificity, insufficient length of the probe, incorrect assay conditions, or simply not testing for binding of the correct target protein/factor. Once a transcription factor is identified using *in vitro* assays, allelic differences in binding can be confirmed by performing ChIP assays in cells or tissues of differing genotypes. Elucidating the effects of variants on transcriptional activity and protein binding provide strong evidence of

potentially functional variants at GWAS loci and are among the best initial experiments to examine GWAS variant function.

Additional experiments will be necessary to fully characterize the mechanism at the *CDC123/CAMK1D* and *ANGPTL8* GWAS loci. At *ANGPTL8,* the transcription factors binding to variant alleles need to be identified. Additionally, at both loci, physical chromatin interaction and/or genome editing experiments would confirm the target genes, effect on transcription, and downstream phenotype effects. Physical interaction experiments are used to identify chromatin loops between promoters, enhancers, and other regulatory regions. Chromatin conformation assays would solidify the relationship of the identified regulatory regions in Chapters 2 and 3 and target gene promoters. In addition, entire regulatory regions can be deleted or specific variants mutated using CRISPR-Cas9 genome editing to evaluate the effect of altered regulatory elements *in vivo* (228-230). After isolating edited clones, RNA levels of nearby genes are measured and downstream metabolites or other phenotypes can be quantified to determine the effect of the variant or regulatory region. Such experiments were conducted at the *ADCY5* T2D GWAS locus (227), where deletion of the regulatory region containing rs11708067 resulted in decreased *ADCY5* expression levels and decreased insulin secretion in pancreatic islet beta cells. Physical chromatin interaction and genome editing experiments are robust approaches because the chromatin remains in its native context throughout the experiment. Future experiments will further solidify the mechanism at the *CDC123/CAMK1D* and *ANGPTL8* GWAS loci.

Future experiments will increase the power of ATAC-seq; we are currently expanding to include ~400 clinical subcutaneous adipose samples from the METSIM study. As mentioned previously, the METSIM study contains genotypes, expression, and clinical trait data in the same samples. With ~400 ATAC-seq profiles, we will have reasonable statistical power to identify

chromatin quantitative trait loci associations (cQTL). Similar to allelic imbalance in open chromatin sequencing reads, these cQTL associations will delineate where in the genome a variant allele alters chromatin accessibility. Variants that alter chromatin accessibility are more likely to be functional variants (172) and cQTL associations will add to strength our ability to prioritize candidate GWAS variants.

Since the advent of GWAS, we have gained significant insight to the genetics of complex cardiometabolic traits. One model to explain the hundreds of GWAS loci suggests an "omnigenic" mode of inheritance, where gene regulatory networks are so highly interconnected that all genes expressed in relevant tissues contribute to all cardiometabolic traits (231). Regardless of whether this or other hypotheses are true, significant work is needed to understand how the thousands of noncoding variants alter gene expression and downstream phenotypes. Given the small effect sizes of most GWAS loci, significant work is needed to understand the network of genes contributing to cardiometabolic phenotypes and how these genes interact together to increase risk for cardiometabolic diseases. Although some massively-parallel regulatory assays exist (232,233), fully elucidating the mechanism of each locus will require locus-specific analysis. Testing candidate variants locus-by-locus is time-consuming and expensive. However, new methods including CRISPR-Cas9 genome editing can improve the speed of regulatory element identification and allow for testing variants in the native chromatin context (227,228). In this set of studies, I have tested candidate variants at two GWAS loci and generated open chromatin profiles for adipose tissue and preadipocyte cells. Together, these results contribute to a broad understanding of genetic regulation of the human genome and mechanisms of complex cardiometabolic traits.

REFERENCES

1. Claussnitzer, M., Dankel, S.N., Kim, K.H., Quon, G., Meuleman, W., Haugen, C., Glunk, V., Sousa, I.S., Beaudry, J.L., Puviindran, V. *et al.* (2015) FTO Obesity Variant Circuitry and Adipocyte Browning in Humans. *N. Engl. J. Med.*, 373, 895-907.

2. Flannick, J., Thorleifsson, G., Beer, N.L., Jacobs, S.B., Grarup, N., Burtt, N.P., Mahajan, A., Fuchsberger, C., Atzmon, G., Benediktsson, R. *et al.* (2014) Loss-of-function mutations in SLC30A8 protect against type 2 diabetes. *Nat. Genet.*, 46, 357-363.

3. Wu, Y., Waite, L.L., Jackson, A.U., Sheu, W.H., Buyske, S., Absher, D., Arnett, D.K., Boerwinkle, E., Bonnycastle, L.L., Carty, C.L. *et al.* (2013) Trans-ethnic fine-mapping of lipid loci identifies population-specific signals and allelic heterogeneity that increases the trait variance explained. *PLoS genetics*, 9, e1003379.

4. Global_Lipids_Genetics_Consortium. (2013) Discovery and refinement of loci associated with lipid levels. *Nat. Genet.*, 45, 1274-1283.

5. DIAbetes_Genetics_Replication_Meta-analysis_Consortium, Asian_Genetic_Epidemiology_Network_Type_2_Diabetes_(AGEN-T2D)_Consortium, South_Asian_Type_2_Diabetes_(SAT2D)_Consortium, Mexican_American_Type_2_Diabetes_(MAT2D)_Consortium and Type_2_Diabetes_Genetic_Exploration_by_Next-generation_sequencing_in_multi-Ethnic_Samples_(T2D-GENES)_Consortium. (2014) Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. *Nat. Genet.*, 46, 234-244.

6. Shungin, D., Winkler, T.W., Croteau-Chonka, D.C., Ferreira, T., Locke, A.E., Magi, R., Strawbridge, R.J., Pers, T.H., Fischer, K., Justice, A.E. *et al.* (2015) New genetic loci link adipose and insulin biology to body fat distribution. *Nature*, 518, 187-196.

7. Locke, A.E., Kahali, B., Berndt, S.I., Justice, A.E., Pers, T.H., Day, F.R., Powell, C., Vedantam, S., Buchkovich, M.L., Yang, J. *et al.* (2015) Genetic studies of body mass index yield new insights for obesity biology. *Nature*, 518, 197-206.

8. Astle, W.J., Elding, H., Jiang, T., Allen, D., Ruklisa, D., Mann, A.L., Mead, D., Bouman, H., Riveros-Mckay, F., Kostadima, M.A. *et al.* (2016) The Allelic Landscape of Human Blood Cell Trait Variation and Links to Common Complex Disease. *Cell*, 167, 1415-1429 e1419.

9. Balakumar, P., Maung, U.K. and Jagadeesh, G. (2016) Prevalence and prevention of cardiovascular disease and diabetes mellitus. *Pharmacol. Res.*, 113, 600-609.

10. James, P.T., Rigby, N. and Leach, R. (2004) The obesity epidemic, metabolic syndrome and future prevention strategies. *Eur J Cardiovasc Prev Rehabil*, 11, 3-8.

11. Shaw, J.E., Sicree, R.A. and Zimmet, P.Z. (2010) Global estimates of the prevalence of diabetes for 2010 and 2030. *Diabetes Res. Clin. Pract.*, 87, 4-14.

12. Writing Group, M., Mozaffarian, D., Benjamin, E.J., Go, A.S., Arnett, D.K., Blaha, M.J., Cushman, M., Das, S.R., de Ferranti, S., Despres, J.P. *et al.* (2016) Heart Disease and Stroke Statistics-2016 Update: A Report From the American Heart Association. *Circulation*, 133, e38-360.

13. Stumvoll, M., Goldstein, B.J. and van Haeften, T.W. (2005) Type 2 diabetes: principles of pathogenesis and therapy. *Lancet*, 365, 1333-1346.

14. Casella, S., Bielli, A., Mauriello, A. and Orlandi, A. (2015) Molecular Pathways Regulating Macrovascular Pathology and Vascular Smooth Muscle Cells Phenotype in Type 2 Diabetes. *International journal of molecular sciences*, 16, 24353-24368.

15. Buse, J.B., Ginsberg, H.N., Bakris, G.L., Clark, N.G., Costa, F., Eckel, R., Fonseca, V., Gerstein, H.C., Grundy, S., Nesto, R.W. *et al.* (2007) Primary prevention of cardiovascular diseases in people with diabetes mellitus: a scientific statement from the American Heart Association and the American Diabetes Association. *Circulation*, 115, 114-126.

16. Arena, R., Guazzi, M., Lianov, L., Whitsel, L., Berra, K., Lavie, C.J., Kaminsky, L., Williams, M., Hivert, M.F., Franklin, N.C. *et al.* (2015) Healthy Lifestyle Interventions to Combat Noncommunicable Disease-A Novel Nonhierarchical Connectivity Model for Key Stakeholders: A Policy Statement From the American Heart Association, European Society of Cardiology, European Association for Cardiovascular Prevention and Rehabilitation, and American College of Preventive Medicine. *Mayo Clin. Proc.*, 90, 1082-1103.

17. Sharrett, A.R., Ballantyne, C.M., Coady, S.A., Heiss, G., Sorlie, P.D., Catellier, D., Patsch, W. and Atherosclerosis Risk in Communities Study, G. (2001) Coronary heart disease prediction from lipoprotein cholesterol levels, triglycerides, lipoprotein(a), apolipoproteins A-I and B, and HDL density subfractions: The Atherosclerosis Risk in Communities (ARIC) Study. *Circulation*, 104, 1108-1113.

18. Cooney, M.T., Dudina, A., De Bacquer, D., Wilhelmsen, L., Sans, S., Menotti, A., De Backer, G., Jousilahti, P., Keil, U., Thomsen, T. *et al.* (2009) HDL cholesterol protects against cardiovascular disease in both genders, at all ages and at all levels of risk. *Atherosclerosis*, 206, 611-616.

19. Farmer, J.A. and Liao, J. (2011) Evolving concepts of the role of high-density lipoprotein in protection from atherosclerosis. *Current atherosclerosis reports*, 13, 107-114.

20. Chilton, R.J. (2004) Pathophysiology of coronary heart disease: a brief review. *J. Am. Osteopath. Assoc.*, 104, S5-8.

21. Cornier, M.A., Despres, J.P., Davis, N., Grossniklaus, D.A., Klein, S., Lamarche, B., Lopez-Jimenez, F., Rao, G., St-Onge, M.P., Towfighi, A. *et al.* (2011) Assessing adiposity: a scientific statement from the American Heart Association. *Circulation*, 124, 1996-2019.

22. Ogden, C.L., Carroll, M.D., Kit, B.K. and Flegal, K.M. (2012) Prevalence of obesity in the United States, 2009-2010. *NCHS data brief*, 1-8.

23. Hurt, R.T., Frazier, T.H., McClave, S.A. and Kaplan, L.M. (2011) Obesity epidemic: overview, pathophysiology, and the intensive care unit conundrum. *JPEN. Journal of parenteral and enteral nutrition*, 35, 4S-13S.

24. Hurt, R.T., Kulisek, C., Buchanan, L.A. and McClave, S.A. (2010) The obesity epidemic: challenges, health initiatives, and implications for gastroenterologists. *Gastroenterology & hepatology*, 6, 780-792.

25. NCEP_Expert_Panel. (2002) Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) final report. *Circulation*, 106, 3143-3421.

26. Bowes, A., Begley, J. and Kerr, D. (2017) Lifestyle change reduces cardiometabolic risk factors and glucagon-like peptide-1 levels in obese first-degree relatives of people with diabetes. *J. Hum. Nutr. Diet.*

27. Almgren, P., Lehtovirta, M., Isomaa, B., Sarelin, L., Taskinen, M.R., Lyssenko, V., Tuomi, T., Groop, L. and Botnia Study, G. (2011) Heritability and familiality of type 2 diabetes and related quantitative traits in the Botnia Study. *Diabetologia*, 54, 2811-2819.

28. Willemsen, G., Ward, K.J., Bell, C.G., Christensen, K., Bowden, J., Dalgard, C., Harris, J.R., Kaprio, J., Lyle, R., Magnusson, P.K. *et al.* (2015) The Concordance and Heritability of Type 2 Diabetes in 34,166 Twin Pairs From International Twin Registers: The Discordant Twin (DISCOTWIN) Consortium. *Twin research and human genetics : the official journal of the International Society for Twin Studies*, 18, 762-771.

29. Iliadou, A., Snieder, H., Wang, X., Treiber, F.A. and Davis, C.L. (2005) Heritabilities of lipids in young European American and African American twins. *Twin research and human genetics : the official journal of the International Society for Twin Studies*, 8, 492-498.

30. Middelberg, R.P., Martin, N.G. and Whitfield, J.B. (2007) A longitudinal genetic study of plasma lipids in adolescent twins. *Twin research and human genetics : the official journal of the International Society for Twin Studies*, 10, 127-135.

31. Souren, N.Y., Paulussen, A.D., Loos, R.J., Gielen, M., Beunen, G., Fagard, R., Derom, C., Vlietinck, R. and Zeegers, M.P. (2007) Anthropometry, carbohydrate and lipid metabolism in the East Flanders Prospective Twin Survey: heritabilities. *Diabetologia*, 50, 2107-2116.

32. Weissglas-Volkov, D. and Pajukanta, P. (2010) Genetic causes of high and low serum HDL-cholesterol. *J. Lipid Res.*, 51, 2032-2057.

33. Browning, S.R. and Browning, B.L. (2013) Identity-by-descent-based heritability analysis in the Northern Finland Birth Cohort. *Hum. Genet.*, 132, 129-138.

34. van Dongen, J., Willemsen, G., Chen, W.M., de Geus, E.J. and Boomsma, D.I. (2013) Heritability of metabolic syndrome traits in a large population-based sample. *J. Lipid Res.*, 54, 2914-2923.

35. Schleinitz, D., Bottcher, Y., Bluher, M. and Kovacs, P. (2014) The genetics of fat distribution. *Diabetologia*, 57, 1276-1286.

36. Selby, J.V., Newman, B., Quesenberry, C.P., Jr., Fabsitz, R.R., Carmelli, D., Meaney, F.J. and Slemenda, C. (1990) Genetic and behavioral influences on body fat distribution. *Int. J. Obes.*, 14, 593-602.

37. Rose, K.M., Newman, B., Mayer-Davis, E.J. and Selby, J.V. (1998) Genetic and behavioral determinants of waist-hip ratio and waist circumference in women twins. *Obesity research*, 6, 383-392.

38. Willer, C.J., Schmidt, E.M., Sengupta, S., Peloso, G.M., Gustafsson, S., Kanoni, S., Ganna, A., Chen, J., Buchkovich, M.L., Mora, S. *et al.* (2013) Discovery and refinement of loci associated with lipid levels. *Nat. Genet.*, 45, 1274-1283.

39. Kathiresan, S., Willer, C.J., Peloso, G.M., Demissie, S., Musunuru, K., Schadt, E.E., Kaplan, L., Bennett, D., Li, Y., Tanaka, T. *et al.* (2009) Common variants at 30 loci contribute to polygenic dyslipidemia. *Nat. Genet.*, 41, 56-65.

40. Nikpay, M., Goel, A., Won, H.H., Hall, L.M., Willenborg, C., Kanoni, S., Saleheen, D., Kyriakou, T., Nelson, C.P., Hopewell, J.C. *et al.* (2015) A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat. Genet.*, 47, 1121-1130.

41. Schunkert, H., Konig, I.R., Kathiresan, S., Reilly, M.P., Assimes, T.L., Holm, H., Preuss, M., Stewart, A.F., Barbalic, M., Gieger, C. *et al.* (2011) Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nat. Genet.*, 43, 333-338.

42. Consortium, U.K., Walter, K., Min, J.L., Huang, J., Crooks, L., Memari, Y., McCarthy, S., Perry, J.R., Xu, C., Futema, M. *et al.* (2015) The UK10K project identifies rare variants in health and disease. *Nature*, 526, 82-90.

43. Coram, M.A., Duan, Q., Hoffmann, T.J., Thornton, T., Knowles, J.W., Johnson, N.A., Ochs-Balcom, H.M., Donlon, T.A., Martin, L.W., Eaton, C.B. *et al.* (2013) Genome-wide characterization of shared and distinct genetic components that influence blood lipid levels in ethnically diverse human populations. *Am. J. Hum. Genet.*, 92, 904-916.

44. Surakka, I., Horikoshi, M., Magi, R., Sarin, A.P., Mahajan, A., Lagou, V., Marullo, L., Ferreira, T., Miraglio, B., Timonen, S. *et al.* (2015) The impact of low-frequency and rare variants on lipid levels. *Nat. Genet.*, 47, 589-597.

45. Levy, D., Ehret, G.B., Rice, K., Verwoert, G.C., Launer, L.J., Dehghan, A., Glazer, N.L., Morrison, A.C., Johnson, A.D., Aspelund, T. *et al.* (2009) Genome-wide association study of blood pressure and hypertension. *Nat. Genet.*, 41, 677-687.

46. Kato, N., Takeuchi, F., Tabara, Y., Kelly, T.N., Go, M.J., Sim, X., Tay, W.T., Chen, C.H., Zhang, Y., Yamamoto, K. *et al.* (2011) Meta-analysis of genome-wide association studies identifies common variants associated with blood pressure variation in east Asians. *Nat. Genet.*, 43, 531-538.

47. Fuchsberger, C., Flannick, J., Teslovich, T.M., Mahajan, A., Agarwala, V., Gaulton, K.J., Ma, C., Fontanillas, P., Moutsianas, L., McCarthy, D.J. *et al.* (2016) The genetic architecture of type 2 diabetes. *Nature*, 536, 41-47.

48. Gaulton, K.J., Ferreira, T., Lee, Y., Raimondo, A., Magi, R., Reschen, M.E., Mahajan, A., Locke, A., Rayner, N.W., Robertson, N. *et al.* (2015) Genetic fine mapping and genomic annotation defines causal mechanisms at type 2 diabetes susceptibility loci. *Nat. Genet.*, 47, 1415-1425.

49. Huang, Q. (2015) Genetic study of complex diseases in the post-GWAS era. *Journal of genetics and genomics = Yi chuan xue bao*, 42, 87-98.

50. Visscher, P.M., Brown, M.A., McCarthy, M.I. and Yang, J. (2012) Five years of GWAS discovery. *Am. J. Hum. Genet.*, 90, 7-24.

51. Willer, C.J., Sanna, S., Jackson, A.U., Scuteri, A., Bonnycastle, L.L., Clarke, R., Heath, S.C., Timpson, N.J., Najjar, S.S., Stringham, H.M. *et al.* (2008) Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nat. Genet.*, 40, 161-169.

52. Burkhardt, R., Toh, S.A., Lagor, W.R., Birkeland, A., Levin, M., Li, X., Robblee, M., Fedorov, V.D., Yamamoto, M., Satoh, T. *et al.* (2010) Trib1 is a lipid- and myocardial infarction-associated gene that regulates hepatic lipogenesis and VLDL production in mice. *The Journal of clinical investigation*, 120, 4410-4414.

53. Fogarty, M.P., Panhuis, T.M., Vadlamudi, S., Buchkovich, M.L. and Mohlke, K.L. (2013) Allele-specific transcriptional activity at type 2 diabetes-associated single nucleotide polymorphisms in regions of pancreatic islet open chromatin at the JAZF1 locus. *Diabetes*, 62, 1756-1762.

54. Roman, T.S., Marvelle, A.F., Fogarty, M.P., Vadlamudi, S., Gonzalez, A.J., Buchkovich, M.L., Huyghe, J.R., Fuchsberger, C., Jackson, A.U., Wu, Y. *et al.* (2015) Multiple Hepatic Regulatory Variants at the GALNT2 GWAS Locus Associated with High-Density Lipoprotein Cholesterol. *Am. J. Hum. Genet.*, 97, 801-815.

55. Spain, S.L. and Barrett, J.C. (2015) Strategies for fine-mapping complex traits. *Hum. Mol. Genet.*, 24, R111-119.

56. Hormozdiari, F., Kostem, E., Kang, E.Y., Pasaniuc, B. and Eskin, E. (2014) Identifying causal variants at loci with multiple signals of association. *Genetics*, 198, 497-508.

57. Kichaev, G., Yang, W.Y., Lindstrom, S., Hormozdiari, F., Eskin, E., Price, A.L., Kraft, P. and Pasaniuc, B. (2014) Integrating functional data to prioritize causal variants in statistical fine-mapping studies. *PLoS genetics*, 10, e1004722.

58. Morris, A.P. (2011) Transethnic meta-analysis of genomewide association studies. *Genet. Epidemiol.*, 35, 809-822.

59. Encode_Project_Consortium. (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature*, 489, 57-74.

60. Bernstein, B.E., Stamatoyannopoulos, J.A., Costello, J.F., Ren, B., Milosavljevic, A., Meissner, A., Kellis, M., Marra, M.A., Beaudet, A.L., Ecker, J.R. *et al.* (2010) The NIH Roadmap Epigenomics Mapping Consortium. *Nat. Biotechnol.*, 28, 1045-1048.

61. Varshney, A., Scott, L.J., Welch, R.P., Erdos, M.R., Chines, P.S., Narisu, N., Albanus, R.D., Orchard, P., Wolford, B.N., Kursawe, R. *et al.* (2017) Genetic regulatory signatures underlying islet gene expression and type 2 diabetes. *Proc. Natl. Acad. Sci. U. S. A.*, 114, 2301-2306.

62. Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y. and Greenleaf, W.J. (2013) Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods*, 10, 1213-1218.

63. Civelek, M., Wu, Y., Pan, C., Raulerson, C.K., Ko, A., He, A., Tilford, C., Saleem, N.K., Stancakova, A., Scott, L.J. *et al.* (2017) Genetic Regulation of Adipose Gene Expression and Cardio-Metabolic Traits. *Am. J. Hum. Genet.*, 100, 428-443.

64. Ward, L.D. and Kellis, M. (2012) Interpreting noncoding genetic variation in complex traits and human disease. *Nat. Biotechnol.*, 30, 1095-1106.

65. Tak, Y.G. and Farnham, P.J. (2015) Making sense of GWAS: using epigenomics and genome engineering to understand the functional relevance of SNPs in non-coding regions of the human genome. *Epigenetics & chromatin*, 8, 57.

66. Weissglas-Volkov, D., Aguilar-Salinas, C.A., Nikkola, E., Deere, K.A., Cruz-Bautista, I., Arellano-Campos, O., Munoz-Hernandez, L.L., Gomez-Munguia, L., Ordonez-Sanchez, M.L., Reddy, P.M. *et al.* (2013) Genomic study in Mexicans identifies a new locus for triglycerides and refines European lipid loci. *J. Med. Genet.*, 50, 298-308.

67. Poulsen, P., Kyvik, K.O., Vaag, A. and Beck-Nielsen, H. (1999) Heritability of type II (non-insulin-dependent) diabetes mellitus and abnormal glucose tolerance--a population-based twin study. *Diabetologia*, 42, 139-145.

68. Morris, A.P., Voight, B.F., Teslovich, T.M., Ferreira, T., Segre, A.V., Steinthorsdottir, V., Strawbridge, R.J., Khan, H., Grallert, H., Mahajan, A. *et al.* (2012) Large-scale

association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat. Genet.*, 44, 981-990.

69. Zeggini, E., Scott, L.J., Saxena, R., Voight, B.F., Marchini, J.L., Hu, T., de Bakker, P.I., Abecasis, G.R., Almgren, P., Andersen, G. *et al.* (2008) Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat. Genet.*, 40, 638-645.

70. Shu, X.O., Long, J., Cai, Q., Qi, L., Xiang, Y.B., Cho, Y.S., Tai, E.S., Li, X., Lin, X., Chow, W.H. *et al.* (2010) Identification of new genetic risk variants for type 2 diabetes. *PLoS genetics*, 6, e1001127.

71. Kooner, J.S., Saleheen, D., Sim, X., Sehmi, J., Zhang, W., Frossard, P., Been, L.F., Chia, K.S., Dimas, A.S., Hassanali, N. *et al.* (2011) Genome-wide association study in individuals of South Asian ancestry identifies six new type 2 diabetes susceptibility loci. *Nat. Genet.*, 43, 984-989.

72. Stitzel, M.L., Sethupathy, P., Pearson, D.S., Chines, P.S., Song, L., Erdos, M.R., Welch, R., Parker, S.C., Boyle, A.P., Scott, L.J. *et al.* (2010) Global epigenomic analysis of primary human pancreatic islets provides insights into type 2 diabetes susceptibility loci. *Cell Metab.*, 12, 443-455.

73. Paul, D.S., Albers, C.A., Rendon, A., Voss, K., Stephens, J., van der Harst, P., Chambers, J.C., Soranzo, N., Ouwehand, W.H., Deloukas, P. *et al.* (2013) Maps of open chromatin highlight cell type-restricted patterns of regulatory sequence variation at hematological trait loci. *Genome Res.*, 23, 1130-1141.

74. Gaulton, K.J., Nammo, T., Pasquali, L., Simon, J.M., Giresi, P.G., Fogarty, M.P., Panhuis, T.M., Mieczkowski, P., Secchi, A., Bosco, D. *et al.* (2010) A map of open chromatin in human pancreatic islets. *Nat. Genet.*, 42, 255-259.

75. Paul, D.S., Nisbet, J.P., Yang, T.P., Meacham, S., Rendon, A., Hautaviita, K., Tallila, J., White, J., Tijssen, M.R., Sivapalaratnam, S. *et al.* (2011) Maps of open chromatin guide the functional follow-up of genome-wide association signals: application to hematological traits. *PLoS genetics*, 7, e1002139.

76. Kooner, J.S., Saleheen, D., Sim, X., Sehmi, J., Zhang, W., Frossard, P., Been, L.F., Chia, K.S., Dimas, A.S., Hassanali, N. *et al.* (2011) Genome-wide association study in individuals of South Asian ancestry identifies six new type 2 diabetes susceptibility loci. *Nat. Genet.*, 43, 984-989.

77. Dimas, A.S., Lagou, V., Barker, A., Knowles, J.W., Magi, R., Hivert, M.F., Benazzo, A., Rybin, D., Jackson, A.U., Stringham, H.M. *et al.* (2013) Impact of type 2 diabetes susceptibility variants on quantitative glycemic traits reveals mechanistic heterogeneity. *Diabetes*.

78. Soler Artigas, M., Loth, D.W., Wain, L.V., Gharib, S.A., Obeidat, M., Tang, W., Zhai, G., Zhao, J.H., Smith, A.V., Huffman, J.E. *et al.* (2011) Genome-wide association and

large-scale follow up identifies 16 new loci influencing lung function. *Nat. Genet.*, 43, 1082-1090.

79. Bieganowski, P., Shilinski, K., Tsichlis, P.N. and Brenner, C. (2004) Cdc123 and checkpoint forkhead associated with RING proteins control the cell cycle by controlling eIF2gamma abundance. *J. Biol. Chem.*, 279, 44656-44666.

80. Perzlmaier, A.F., Richter, F. and Seufert, W. (2013) Translation initiation requires cell division cycle 123 (Cdc123) to facilitate biogenesis of the eukaryotic initiation factor 2 (eIF2). *J. Biol. Chem.*, 288, 21537-21546.

81. Sakagami, H., Kamata, A., Nishimura, H., Kasahara, J., Owada, Y., Takeuchi, Y., Watanabe, M., Fukunaga, K. and Kondo, H. (2005) Prominent expression and activity-dependent nuclear translocation of Ca2+/calmodulin-dependent protein kinase Idelta in hippocampal neurons. *Eur. J. Neurosci.*, 22, 2697-2707.

82. Verploegen, S., Lammers, J.W., Koenderman, L. and Coffer, P.J. (2000) Identification and characterization of CKLiK, a novel granulocyte Ca(++)/calmodulin-dependent kinase. *Blood*, 96, 3215-3223.

83. Voight, B.F., Scott, L.J., Steinthorsdottir, V., Morris, A.P., Dina, C., Welch, R.P., Zeggini, E., Huth, C., Aulchenko, Y.S., Thorleifsson, G. *et al.* (2010) Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nat. Genet.*, 42, 579-589.

84. Consortium, G.T. (2013) The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.*, 45, 580-585.

85. Consortium, E.P., Bernstein, B.E., Birney, E., Dunham, I., Green, E.D., Gunter, C. and Snyder, M. (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature*, 489, 57-74.

86. Song, L., Zhang, Z., Grasfeder, L.L., Boyle, A.P., Giresi, P.G., Lee, B.K., Sheffield, N.C., Gräf, S., Huss, M., Keefe, D. *et al.* (2011) Open chromatin defined by DNaseI and FAIRE identifies regulatory elements that shape cell-type identity. *Genome Res.*, 21, 1757-1767.

87. ENCODE_Project_Consortium. (2011) A user's guide to the encyclopedia of DNA elements (ENCODE). *PLoS Biol.*, 9, e1001046.

88. Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A. *et al.* (2010) Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci. U. S. A.*, 107, 21931-21936.

89. Heintzman, N.D., Hon, G.C., Hawkins, R.D., Kheradpour, P., Stark, A., Harp, L.F., Ye, Z., Lee, L.K., Stuart, R.K., Ching, C.W. *et al.* (2009) Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature*, 459, 108-112.

90. Zhou, X., Maricque, B., Xie, M., Li, D., Sundaram, V., Martin, E.A., Koebbe, B.C., Nielsen, C., Hirst, M., Farnham, P. *et al.* (2011) The Human Epigenome Browser at Washington University. *Nat Methods*, 8, 989-990.

91. Hohmeier, H.E., Mulder, H., Chen, G., Henkel-Rieger, R., Prentki, M. and Newgard, C.B. (2000) Isolation of INS-1-derived cell lines with robust ATP-sensitive K+ channel-dependent and -independent glucose-stimulated insulin secretion. *Diabetes*, 49, 424-430.

92. Miyazaki, J., Araki, K., Yamato, E., Ikegami, H., Asano, T., Shibasaki, Y., Oka, Y. and Yamamura, K. (1990) Establishment of a pancreatic beta cell line that retains glucose-inducible insulin secretion: special reference to expression of glucose transporter isoforms. *Endocrinology*, 127, 126-132.

93. Fujita, P.A., Rhead, B., Zweig, A.S., Hinrichs, A.S., Karolchik, D., Cline, M.S., Goldman, M., Barber, G.P., Clawson, H., Coelho, A. *et al.* (2011) The UCSC Genome Browser database: update 2011. *Nucleic Acids Res.*, 39, D876-882.

94. Bryne, J.C., Valen, E., Tang, M.H., Marstrand, T., Winther, O., da Piedade, I., Krogh, A., Lenhard, B. and Sandelin, A. (2008) JASPAR, the open access database of transcription factor-binding profiles: new content and tools in the 2008 update. *Nucleic Acids Res.*, 36, D102-106.

95. Parker, S.C., Stitzel, M.L., Taylor, D.L., Orozco, J.M., Erdos, M.R., Akiyama, J.A., van Bueren, K.L., Chines, P.S., Narisu, N., Program, N.C.S. *et al.* (2013) Chromatin stretch enhancer states drive cell-specific gene regulation and harbor human disease risk variants. *Proc. Natl. Acad. Sci. U. S. A.*, 110, 17921-17926.

96. Rees, M.G., Wincovitch, S., Schultz, J., Waterstradt, R., Beer, N.L., Baltrusch, S., Collins, F.S. and Gloyn, A.L. (2012) Cellular characterisation of the GCKR P446L variant associated with type 2 diabetes risk. *Diabetologia*, 55, 114-122.

97. Beer, N.L., Tribble, N.D., McCulloch, L.J., Roos, C., Johnson, P.R., Orho-Melander, M. and Gloyn, A.L. (2009) The P446L variant in GCKR associated with fasting plasma glucose and triglyceride levels exerts its effect through increased glucokinase activity in liver. *Hum. Mol. Genet.*, 18, 4081-4088.

98. Travers, M.E., Mackay, D.J., Dekker Nitert, M., Morris, A.P., Lindgren, C.M., Berry, A., Johnson, P.R., Hanley, N., Groop, L.C., McCarthy, M.I. *et al.* (2013) Insights into the molecular mechanism for type 2 diabetes susceptibility at the KCNQ1 locus from temporal changes in imprinting status in human islets. *Diabetes*, 62, 987-992.

99. Nicolson, T.J., Bellomo, E.A., Wijesekara, N., Loder, M.K., Baldwin, J.M., Gyulkhandanyan, A.V., Koshkin, V., Tarasov, A.I., Carzaniga, R., Kronenberger, K. *et al.* (2009) Insulin storage and glucose homeostasis in mice null for the granule zinc transporter ZnT8 and studies of the type 2 diabetes-associated variants. *Diabetes*, 58, 2070-2083.

100. Lecompte, S., Pasquetti, G., Hermant, X., Grenier-Boley, B., Gonzalez-Gross, M., De Henauw, S., Molnar, D., Stehle, P., Beghin, L., Moreno, L.A. *et al.* (2013) Genetic and molecular insights into the role of PROX1 in glucose metabolism. *Diabetes*, 62, 1738-1745.

101. Ng, H.J. and Gloyn, A.L. (2013) Bridging the gap between genetic associations and molecular mechanisms for type 2 diabetes. *Current diabetes reports*, 13, 778-785.

102. Camp, J.G., Jazwa, A.L., Trent, C.M. and Rawls, J.F. (2012) Intronic cis-regulatory modules mediate tissue-specific and microbial control of angptl4/fiaf transcription. *PLoS genetics*, 8, e1002585.

103. Ishibashi, M., Mechaly, A.S., Becker, T.S. and Rinkwitz, S. (2013) Using zebrafish transgenesis to test human genomic sequences for specific enhancer activity. *Methods*, 62, 216-225.

104. Lee, C.S., Friedman, J.R., Fulmer, J.T. and Kaestner, K.H. (2005) The initiation of liver development is dependent on Foxa transcription factors. *Nature*, 435, 944-947.

105. Gao, N., LeLay, J., Vatamaniuk, M.Z., Rieck, S., Friedman, J.R. and Kaestner, K.H. (2008) Dynamic regulation of Pdx1 enhancers by Foxa1 and Foxa2 is essential for pancreas development. *Genes Dev.*, 22, 3435-3448.

106. Gao, N., Le Lay, J., Qin, W., Doliba, N., Schug, J., Fox, A.J., Smirnova, O., Matschinsky, F.M. and Kaestner, K.H. (2010) Foxa1 and Foxa2 maintain the metabolic and secretory features of the mature beta-cell. *Mol. Endocrinol.*, 24, 1594-1604.

107. Zhang, L., Rubins, N.E., Ahima, R.S., Greenbaum, L.E. and Kaestner, K.H. (2005) Foxa2 integrates the transcriptional response of the hepatocyte to fasting. *Cell Metab.*, 2, 141-148.

108. Serandour, A.A., Avner, S., Percevault, F., Demay, F., Bizot, M., Lucchetti-Miganeh, C., Barloy-Hubler, F., Brown, M., Lupien, M., Metivier, R. *et al.* (2011) Epigenetic switch involved in activation of pioneer factor FOXA1-dependent enhancers. *Genome Res.*, 21, 555-565.

109. Cirillo, L.A., Lin, F.R., Cuesta, I., Friedman, D., Jarnik, M. and Zaret, K.S. (2002) Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. *Mol. Cell*, 9, 279-289.

110. Sekiya, T., Muthurajan, U.M., Luger, K., Tulin, A.V. and Zaret, K.S. (2009) Nucleosome-binding affinity as a primary determinant of the nuclear mobility of the pioneer transcription factor FoxA. *Genes Dev.*, 23, 804-809.

111. Pasquali, L., Gaulton, K.J., Rodriguez-Segui, S.A., Mularoni, L., Miguel-Escalada, I., Akerman, I., Tena, J.J., Moran, I., Gomez-Marin, C., van de Bunt, M. *et al.* (2014) Pancreatic islet enhancer clusters enriched in type 2 diabetes risk-associated variants. *Nat. Genet.*, 46, 136-143.

112. Flutre, T., Wen, X., Pritchard, J. and Stephens, M. (2013) A statistical framework for joint eQTL analysis in multiple tissues. *PLoS genetics*, 9, e1003486.

113. Nica, A.C., Ongen, H., Irminger, J.C., Bosco, D., Berney, T., Antonarakis, S.E., Halban, P.A. and Dermitzakis, E.T. (2013) Cell-type, allelic, and genetic signatures in the human pancreatic beta cell transcriptome. *Genome Res.*, 23, 1554-1562.

114. Verploegen, S., Ulfman, L., van Deutekom, H.W., van Aalst, C., Honing, H., Lammers, J.W., Koenderman, L. and Coffer, P.J. (2005) Characterization of the role of CaMKI-like kinase (CKLiK) in human granulocyte function. *Blood*, 106, 1076-1083.

115. Dalle, S., Quoyer, J., Varin, E. and Costes, S. (2011) Roles and regulation of the transcription factor CREB in pancreatic β -cells. *Curr Mol Pharmacol*, 4, 187-195.

116. Haney, S., Zhao, J., Tiwari, S., Eng, K., Guey, L.T. and Tien, E. (2013) RNAi screening in primary human hepatocytes of genes implicated in genome-wide association studies for roles in type 2 diabetes identifies roles for CAMK1D and CDKAL1, among others, in hepatic glucose regulation. *PloS one*, 8, e64946.

117. Nair, A.K., Piaggi, P., McLean, N.A., Kaur, M., Kobes, S., Knowler, W.C., Bogardus, C., Hanson, R.L. and Baier, L.J. (2016) Assessment of established HDL-C loci for association with HDL-C levels and type 2 diabetes in Pima Indians. *Diabetologia*, 59, 481-491.

118. Spracklen, C.N., Chen, P., Kim, Y.J., Wang, X., Cai, H., Li, S., Long, J., Wu, Y., Wang, Y.X., Takeuchi, F. *et al.* (2017) Association analyses of East Asian individuals and trans-ancestry analyses with European individuals reveal new loci associated with cholesterol and triglyceride levels. *Hum. Mol. Genet.*, 26, 1770-1784.

119. Teslovich, T.M., Musunuru, K., Smith, A.V., Edmondson, A.C., Stylianou, I.M., Koseki, M., Pirruccello, J.P., Ripatti, S., Chasman, D.I., Willer, C.J. *et al.* (2010) Biological, clinical and population relevance of 95 loci for blood lipids. *Nature*, 466, 707-713.

120. Hanson, R.L., Leti, F., Tsinajinnie, D., Kobes, S., Puppala, S., Curran, J.E., Almasy, L., Lehman, D.M., Blangero, J., Duggirala, R. *et al.* (2016) The Arg59Trp variant in ANGPTL8 (betatrophin) is associated with total and HDL-cholesterol in American Indians and Mexican Americans and differentially affects cleavage of ANGPTL3. *Mol. Genet. Metab.*, 118, 128-137.

121. Quagliarini, F., Wang, Y., Kozlitina, J., Grishin, N.V., Hyde, R., Boerwinkle, E., Valenzuela, D.M., Murphy, A.J., Cohen, J.C. and Hobbs, H.H. (2012) Atypical angiopoietin-like protein that regulates ANGPTL3. *Proc. Natl. Acad. Sci. U. S. A.*, 109, 19751-19756.

122. Ren, G., Kim, J.Y. and Smas, C.M. (2012) Identification of RIFL, a novel adipocyte-enriched insulin target gene with a role in lipid metabolism. *Am. J. Physiol. Endocrinol. Metab.*, 303, E334-351.

123. Zhang, R. (2012) Lipasin, a novel nutritionally-regulated liver-enriched factor that regulates serum triglyceride levels. *Biochem. Biophys. Res. Commun.*, 424, 786-792.

124. Espes, D., Martinell, M. and Carlsson, P.O. (2014) Increased circulating betatrophin concentrations in patients with type 2 diabetes. *International journal of endocrinology*, 2014, 323407.

125. Fu, Z., Berhane, F., Fite, A., Seyoum, B., Abou-Samra, A.B. and Zhang, R. (2014) Elevated circulating lipasin/betatrophin in human type 2 diabetes and obesity. *Scientific reports*, 4, 5013.

126. Hu, H., Sun, W., Yu, S., Hong, X., Qian, W., Tang, B., Wang, D., Yang, L., Wang, J., Mao, C. *et al.* (2014) Increased circulating levels of betatrophin in newly diagnosed type 2 diabetic patients. *Diabetes care*, 37, 2718-2722.

127. Espes, D., Lau, J. and Carlsson, P.O. (2014) Increased circulating levels of betatrophin in individuals with long-standing type 1 diabetes. *Diabetologia*, 57, 50-53.

128. Yamada, H., Saito, T., Aoki, A., Asano, T., Yoshida, M., Ikoma, A., Kusaka, I., Toyoshima, H., Kakei, M. and Ishikawa, S.E. (2015) Circulating betatrophin is elevated in patients with type 1 and type 2 diabetes. *Endocr. J.*, 62, 417-421.

129. Lee, Y.H., Lee, S.G., Lee, C.J., Kim, S.H., Song, Y.M., Yoon, M.R., Jeon, B.H., Lee, J.H., Lee, B.W., Kang, E.S. *et al.* (2016) Association between betatrophin/ANGPTL8 and non-alcoholic fatty liver disease: animal and human studies. *Scientific reports*, 6, 24013.

130. GTEx_Consortium. (2013) The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.*, 45, 580-585.

131. Wang, Y., Quagliarini, F., Gusarova, V., Gromada, J., Valenzuela, D.M., Cohen, J.C. and Hobbs, H.H. (2013) Mice lacking ANGPTL8 (Betatrophin) manifest disrupted triglyceride metabolism without impaired glucose homeostasis. *Proc. Natl. Acad. Sci. U. S. A.*, 110, 16109-16114.

132. Zhang, Y., Li, S., Donelan, W., Xie, C., Wang, H., Wu, Q., Purich, D.L., Reeves, W.H., Tang, D. and Yang, L.J. (2015) Angiopoietin-like protein 8 (betatrophin) is a stress-response protein that down-regulates expression of adipocyte triglyceride lipase. *Biochim. Biophys. Acta*, 1861, 130-137.

133. Haller, J.F., Mintah, I.J., Shihanian, L.M., Stevis, P., Buckler, D., Alexa-Braun, C.A., Kleiner, S., Banfi, S., Cohen, J.C., Hobbs, H.H. *et al.* (2017) ANGPTL8 requires ANGPTL3 to inhibit lipoprotein lipase and plasma triglyceride clearance. *J. Lipid Res.*

134. Yi, M., Chen, R.P., Yang, R., Guo, X.F., Zhang, J.C. and Chen, H. (2015) Betatrophin Acts as a Diagnostic Biomarker in Type 2 Diabetes Mellitus and Is Negatively Associated with HDL-Cholesterol. *International journal of endocrinology*, 2015, 479157.

135. Miyamoto, Y., Yamauchi, J., Sanbe, A. and Tanoue, A. (2007) Dock6, a Dock-C subfamily guanine nucleotide exchanger, has the dual specificity for Rac1 and Cdc42 and regulates neurite outgrowth. *Exp. Cell Res.*, 313, 791-804.

136. Shaheen, R., Faqeih, E., Sunker, A., Morsy, H., Al-Sheddi, T., Shamseldin, H.E., Adly, N., Hashem, M. and Alkuraya, F.S. (2011) Recessive mutations in DOCK6, encoding the guanidine nucleotide exchange factor DOCK6, lead to abnormal actin cytoskeleton organization and Adams-Oliver syndrome. *Am. J. Hum. Genet.*, 89, 328-333.

137. Peloso, G.M., Auer, P.L., Bis, J.C., Voorman, A., Morrison, A.C., Stitziel, N.O., Brody, J.A., Khetarpal, S.A., Crosby, J.R., Fornage, M. *et al.* (2014) Association of low-frequency and rare coding-sequence variants with blood lipids and coronary heart disease in 56,000 whites and blacks. *Am. J. Hum. Genet.*, 94, 223-232.

138. Corradin, O., Saiakhova, A., Akhtar-Zaidi, B., Myeroff, L., Willis, J., Cowper-Sal lari, R., Lupien, M., Markowitz, S. and Scacheri, P.C. (2014) Combinatorial effects of multiple enhancer variants in linkage disequilibrium dictate levels of gene expression to confer susceptibility to common traits. *Genome Res.*, 24, 1-13.

139. He, H., Li, W., Liyanarachchi, S., Srinivas, M., Wang, Y., Akagi, K., Wang, Y., Wu, D., Wang, Q., Jin, V. *et al.* (2015) Multiple functional variants in long-range enhancer elements contribute to the risk of SNP rs965513 in thyroid cancer. *Proc. Natl. Acad. Sci. U. S. A.*, 112, 6128-6133.

140. Stancakova, A., Javorsky, M., Kuulasmaa, T., Haffner, S.M., Kuusisto, J. and Laakso, M. (2009) Changes in insulin sensitivity and insulin release in relation to glycemia and glucose tolerance in 6,414 Finnish men. *Diabetes*, 58, 1212-1221.

141. Laakso, M., Kuusisto, J., Stancakova, A., Kuulasmaa, T., Pajukanta, P., Lusis, A.J., Collins, F.S., Mohlke, K.L. and Boehnke, M. (2017) The Metabolic Syndrome in Men study: a resource for studies of metabolic and cardiovascular diseases. *J. Lipid Res.*, 58, 481-493.

142. Inouye, M., Kettunen, J., Soininen, P., Silander, K., Ripatti, S., Kumpula, L.S., Hamalainen, E., Jousilahti, P., Kangas, A.J., Mannisto, S. *et al.* (2010) Metabonomic, transcriptomic, and genomic variation of a population cohort. *Molecular systems biology*, 6, 441.

143. Hays, J., Hunt, J.R., Hubbell, F.A., Anderson, G.L., Limacher, M., Allen, C. and Rossouw, J.E. (2003) The Women's Health Initiative recruitment methods and results. *Ann. Epidemiol.*, 13, S18-77.

144. Reiner, A.P., Beleza, S., Franceschini, N., Auer, P.L., Robinson, J.G., Kooperberg, C., Peters, U. and Tang, H. (2012) Genome-wide association and population genetic analysis of C-reactive protein in African American and Hispanic American women. *Am. J. Hum. Genet.*, 91, 502-512.

145. Davis, J.P., Huyghe, J.R., Jackson, A.U., Stringham, H.M., Teslovich, T.M., Welch, R.P., Fuchsberger, C., Locke, A.E., Narisu, N., Chines, P.S. *et al.* Common, low-frequency and rare variants associated with lipoprotein subclasses and triglyceride measures in Finns. *under review*.

146. Kang, H.M., Sul, J.H., Service, S.K., Zaitlen, N.A., Kong, S.Y., Freimer, N.B., Sabatti, C. and Eskin, E. (2010) Variance component model to account for sample structure in genome-wide association studies. *Nat. Genet.*, 42, 348-354.

147. Li, Y., Willer, C.J., Ding, J., Scheet, P. and Abecasis, G.R. (2010) MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet. Epidemiol.*, 34, 816-834.

148. Schaid, D.J., Rowland, C.M., Tines, D.E., Jacobson, R.M. and Poland, G.A. (2002) Score tests for association between traits and haplotypes when linkage phase is ambiguous. *Am. J. Hum. Genet.*, 70, 425-434.

149. Irizarry, R.A., Hobbs, B., Collin, F., Beazer-Barclay, Y.D., Antonellis, K.J., Scherf, U. and Speed, T.P. (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics*, 4, 249-264.

150. Stegle, O., Parts, L., Durbin, R. and Winn, J. (2010) A Bayesian framework to account for complex non-genetic factors in gene expression levels greatly increases power in eQTL studies. *PLoS computational biology*, 6, e1000770.

151. McCarthy, S., Das, S., Kretzschmar, W., Delaneau, O., Wood, A.R., Teumer, A., Kang, H.M., Fuchsberger, C., Danecek, P., Sharp, K. *et al.* (2016) A reference panel of 64,976 haplotypes for genotype imputation. *Nat. Genet.*, 48, 1279-1283.

152. Parker, S.C., Stitzel, M.L., Taylor, D.L., Orozco, J.M., Erdos, M.R., Akiyama, J.A., van Bueren, K.L., Chines, P.S., Narisu, N., Black, B.L. *et al.* (2013) Chromatin stretch enhancer states drive cell-specific gene regulation and harbor human disease risk variants. *Proc. Natl. Acad. Sci. U. S. A.*, 110, 17921-17926.

153. Sheffield, N.C., Thurman, R.E., Song, L., Safi, A., Stamatoyannopoulos, J.A., Lenhard, B., Crawford, G.E. and Furey, T.S. (2013) Patterns of regulatory activity across diverse human cell types predict tissue identity, transcription factor binding, and long-range interactions. *Genome Res.*, 23, 777-788.

154. Wabitsch, M., Brenner, R.E., Melzner, I., Braun, M., Moller, P., Heinze, E., Debatin, K.M. and Hauner, H. (2001) Characterization of a human preadipocyte cell strain with high capacity for adipose differentiation. *International journal of obesity and related metabolic disorders : journal of the International Association for the Study of Obesity*, 25, 8-15.

155. Kulzer, J.R., Stitzel, M.L., Morken, M.A., Huyghe, J.R., Fuchsberger, C., Kuusisto, J., Laakso, M., Boehnke, M., Collins, F.S. and Mohlke, K.L. (2014) A common functional

regulatory variant at a type 2 diabetes locus upregulates ARAP1 expression in the pancreatic beta cell. *Am. J. Hum. Genet.*, 94, 186-197.

156. Fu, Z., Abou-Samra, A.B. and Zhang, R. (2014) An explanation for recent discrepancies in levels of human circulating betatrophin. *Diabetologia*, 57, 2232-2234.

157. International_HapMap_Consortium. (2003) The International HapMap Project. *Nature*, 426, 789-796.

158. Schmidt, E.M., Zhang, J., Zhou, W., Chen, J., Mohlke, K.L., Chen, Y.E. and Willer, C.J. (2015) GREGOR: evaluating global enrichment of trait-associated variants in epigenomic features using a systematic, data-driven approach. *Bioinformatics*, 31, 2601-2606.

159. Coelho, M., Oliveira, T. and Fernandes, R. (2013) Biochemistry of adipose tissue: an endocrine organ. *Arch Med Sci*, 9, 191-200.

160. Gustafson, B., Hedjazifar, S., Gogg, S., Hammarstedt, A. and Smith, U. (2015) Insulin resistance and impaired adipogenesis. *Trends Endocrinol. Metab.*, 26, 193-200.

161. Fernandez-Veledo, S., Nieto-Vazquez, I., Vila-Bedmar, R., Garcia-Guerra, L., Alonso-Chamorro, M. and Lorenzo, M. (2009) Molecular mechanisms involved in obesity-associated insulin resistance: therapeutical approach. *Arch. Physiol. Biochem.*, 115, 227-239.

162. Porter, S.A., Massaro, J.M., Hoffmann, U., Vasan, R.S., O'Donnel, C.J. and Fox, C.S. (2009) Abdominal subcutaneous adipose tissue: a protective fat depot? *Diabetes care*, 32, 1068-1075.

163. Pischon, T., Boeing, H., Hoffmann, K., Bergmann, M., Schulze, M.B., Overvad, K., van der Schouw, Y.T., Spencer, E., Moons, K.G., Tjonneland, A. *et al.* (2008) General and abdominal adiposity and risk of death in Europe. *N. Engl. J. Med.*, 359, 2105-2120.

164. Abraham, T.M., Pedley, A., Massaro, J.M., Hoffmann, U. and Fox, C.S. (2015) Association between visceral and subcutaneous adipose depots and incident cardiovascular disease risk factors. *Circulation*, 132, 1639-1647.

165. Emilsson, V., Thorleifsson, G., Zhang, B., Leonardson, A.S., Zink, F., Zhu, J., Carlson, S., Helgason, A., Walters, G.B., Gunnarsdottir, S. *et al.* (2008) Genetics of gene expression and its effect on disease. *Nature*, 452, 423-428.

166. Zhong, H., Beaulaurier, J., Lum, P.Y., Molony, C., Yang, X., Macneil, D.J., Weingarth, D.T., Zhang, B., Greenawalt, D., Dobrin, R. *et al.* (2010) Liver and adipose expression associated SNPs are enriched for association to type 2 diabetes. *PLoS genetics*, 6, e1000932.

167. Greenawalt, D.M., Dobrin, R., Chudin, E., Hatoum, I.J., Suver, C., Beaulaurier, J., Zhang, B., Castro, V., Zhu, J., Sieberts, S.K. *et al.* (2011) A survey of the genetics of

stomach, liver, and adipose gene expression from a morbidly obese cohort. *Genome Res.*, 21, 1008-1016.

168. Nica, A.C., Parts, L., Glass, D., Nisbet, J., Barrett, A., Sekowska, M., Travers, M., Potter, S., Grundberg, E., Small, K. *et al.* (2011) The architecture of gene regulatory variation across multiple human tissues: the MuTHER study. *PLoS genetics*, 7, e1002003.

169. Dahlman, I., Ryden, M., Brodin, D., Grallert, H., Strawbridge, R.J. and Arner, P. (2016) Numerous Genes in Loci Associated With Body Fat Distribution Are Linked to Adipose Function. *Diabetes*, 65, 433-437.

170. Lynes, M.D. and Tseng, Y.H. (2017) Deciphering adipose tissue heterogeneity. *Ann. N. Y. Acad. Sci.*

171. Fischer-Posovszky, P., Newell, F.S., Wabitsch, M. and Tornqvist, H.E. (2008) Human SGBS cells - a unique tool for studies of human fat cell biology. *Obesity facts*, 1, 184-189.

172. Degner, J.F., Pai, A.A., Pique-Regi, R., Veyrieras, J.B., Gaffney, D.J., Pickrell, J.K., De Leon, S., Michelini, K., Lewellen, N., Crawford, G.E. *et al.* (2012) DNase I sensitivity QTLs are a major determinant of human expression variation. *Nature*, 482, 390-394.

173. Cheng, C.S., Gate, R.E., Aiden, A.P., Siba, A., Tabaka, M., Lituiev, D., Machol, I., Subramaniam, M., Shamim, M., L., H.K. *et al.* (2016) Genetic determinants of chromatin accessibility and gene regulation in T cell activation across human individuals. *biorxiv preprint*, http://dx.doi.org/10.1101/090241.

174. Chen, L., Ge, B., Casale, F.P., Vasquez, L., Kwan, T., Garrido-Martin, D., Watt, S., Yan, Y., Kundu, K., Ecker, S. *et al.* (2016) Genetic Drivers of Epigenetic and Transcriptional Variation in Human Immune Cells. *Cell*, 167, 1398-1414 e1324.

175. Schmidt, S.F., Larsen, B.D., Loft, A., Nielsen, R., Madsen, J.G. and Mandrup, S. (2015) Acute TNF-induced repression of cell identity genes is mediated by NFkappaB-directed redistribution of cofactors from super-enhancers. *Genome Res.*, 25, 1281-1294.

176. Loft, A., Forss, I., Siersbaek, M.S., Schmidt, S.F., Larsen, A.S., Madsen, J.G., Pisani, D.F., Nielsen, R., Aagaard, M.M., Mathison, A. *et al.* (2015) Browning of human adipocytes requires KLF11 and reprogramming of PPARgamma superenhancers. *Genes Dev.*, 29, 7-22.

177. Allum, F., Shao, X., Guenard, F., Simon, M.M., Busche, S., Caron, M., Lambourne, J., Lessard, J., Tandre, K., Hedman, A.K. *et al.* (2015) Characterization of functional methylomes by next-generation capture sequencing identifies novel disease-associated variants. *Nat Commun*, 6, 7211.

178. Roadmap Epigenomics, C., Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J. *et al.* (2015) Integrative analysis of 111 reference human epigenomes. *Nature*, 518, 317-330.

179. Scott, L.J., Erdos, M.R., Huyghe, J.R., Welch, R.P., Beck, A.T., Wolford, B.N., Chines, P.S., Didion, J.P., Narisu, N., Stringham, H.M. *et al.* (2016) The genetic regulatory signature of type 2 diabetes in human skeletal muscle. *Nat Commun*, 7, 11764.

180. Kasowski, M., Kyriazopoulou-Panagiotopoulou, S., Grubert, F., Zaugg, J.B., Kundaje, A., Liu, Y., Boyle, A.P., Zhang, Q.C., Zakharia, F., Spacek, D.V. *et al.* (2013) Extensive variation in chromatin states across humans. *Science*, 342, 750-752.

181. Kilpinen, H., Waszak, S.M., Gschwind, A.R., Raghav, S.K., Witwicki, R.M., Orioli, A., Migliavacca, E., Wiederkehr, M., Gutierrez-Arcelus, M., Panousis, N.I. *et al.* (2013) Coordinated effects of sequence variation on DNA binding, chromatin structure, and transcription. *Science*, 342, 744-747.

182. McVicker, G., van de Geijn, B., Degner, J.F., Cain, C.E., Banovich, N.E., Raj, A., Lewellen, N., Myrthil, M., Gilad, Y. and Pritchard, J.K. (2013) Identification of genetic variants that affect histone modifications in human cells. *Science*, 342, 747-749.

183. Leung, D., Jung, I., Rajagopal, N., Schmitt, A., Selvaraj, S., Lee, A.Y., Yen, C.A., Lin, S., Lin, Y., Qiu, Y. *et al.* (2015) Integrative analysis of haplotype-resolved epigenomes across human tissues. *Nature*, 518, 350-354.

184. Kumasaka, N., Knights, A.J. and Gaffney, D.J. (2016) Fine-mapping cellular QTLs with RASQUAL and ATAC-seq. *Nat. Genet.*, 48, 206-213.

185. Pique-Regi, R., Degner, J.F., Pai, A.A., Gaffney, D.J., Gilad, Y. and Pritchard, J.K. (2011) Accurate inference of transcription factor binding from DNA sequence and chromatin accessibility data. *Genome Res.*, 21, 447-455.

186. Cannon, M.E., Duan, Q., Wu, Y., Zeynalzadeh, M., Xu, Z., Kangas, A.J., Soininen, P., Ala-Korpela, M., Civelek, M., Lusis, A.J. *et al.* (2017) Trans-ancestry Fine Mapping and Molecular Assays Identify Regulatory Variants at the ANGPTL8 HDL-C GWAS Locus. *G3*.

187. Lassmann, T., Hayashizaki, Y. and Daub, C.O. (2009) TagDust--a program to eliminate artifacts from next generation sequencing data. *Bioinformatics*, 25, 2839-2840.

188. Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods*, 9, 357-359.

189. Quinlan, A.R. (2014) BEDTools: The Swiss-Army Tool for Genome Feature Analysis. *Current protocols in bioinformatics*, 47, 11 12 11-34.

190. Karolchik, D., Hinrichs, A.S., Furey, T.S., Roskin, K.M., Sugnet, C.W., Haussler, D. and Kent, W.J. (2004) The UCSC Table Browser data retrieval tool. *Nucleic Acids Res.*, 32, D493-496.

191. Adey, A., Morrison, H.G., Asan, Xun, X., Kitzman, J.O., Turner, E.H., Stackhouse, B., MacKenzie, A.P., Caruccio, N.C., Zhang, X. *et al.* (2010) Rapid, low-input, low-bias

construction of shotgun fragment libraries by high-density in vitro transposition. *Genome biology*, 11, R119.

192. Jun, G., Flickinger, M., Hetrick, K.N., Romm, J.M., Doheny, K.F., Abecasis, G.R., Boehnke, M. and Kang, H.M. (2012) Detecting and estimating contamination of human DNA samples in sequencing and array-based genotype data. *Am. J. Hum. Genet.*, 91, 839-848.

193. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W. *et al.* (2008) Model-based analysis of ChIP-Seq (MACS). *Genome biology*, 9, R137.

194. Swinton, J. (2009) Vennerable: Venn and Euler area-proportional diagrams. . *R package version 2.0/r75*.

195. R_Core_Team. (2016) R: A languate and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria*.

196. McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M. and Bejerano, G. (2010) GREAT improves functional interpretation of cis-regulatory regions. *Nat. Biotechnol.*, 28, 495-501.

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

198. Mathelier, A., Fornes, O., Arenillas, D.J., Chen, C.Y., Denay, G., Lee, J., Shi, W., Shyr, C., Tan, G., Worsley-Hunt, R. *et al.* (2016) JASPAR 2016: a major expansion and update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.*, 44, D110-115.

199. Grant, C.E., Bailey, T.L. and Noble, W.S. (2011) FIMO: scanning for occurrences of a given motif. *Bioinformatics*, 27, 1017-1018.

200. Saint-Andre, V., Federation, A.J., Lin, C.Y., Abraham, B.J., Reddy, J., Lee, T.I., Bradner, J.E. and Young, R.A. (2016) Models of human core transcriptional regulatory circuitries. *Genome Res.*, 26, 385-396.

201. Warnes, G.R., Bolker, B., Bonebakker, L., Gentelman, R., Liaw, W.H.A., Lumley, T., Maechler, M., Magnusson, A., Moeller, S., Schwartz, M. *et al.* (2016) gplots: Various R Programming Tools for Plotting Data. *R package version 3.0.1*.

202. Buchkovich, M.L., Eklund, K., Duan, Q., Li, Y., Mohlke, K.L. and Furey, T.S. (2015) Removing reference mapping biases using limited or no genotype data identifies allelic differences in protein binding at disease-associated loci. *BMC medical genomics*, 8, 43.

203. Wu, T.D. and Nacu, S. (2010) Fast and SNP-tolerant detection of complex variants and splicing in short reads. *Bioinformatics*, 26, 873-881.

204. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. and Genome Project Data Processing, S. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25, 2078-2079.

205. van de Geijn, B., McVicker, G., Gilad, Y. and Pritchard, J.K. (2015) WASP: allele-specific software for robust molecular quantitative trait locus discovery. *Nat Methods*, 12, 1061-1063.

206. Yee, T.W. (1996) Vector Generalized Linear and Additive Models With an Implementation in R. *Journal of Royal Statistical Society, Series B*, 58, 481-493.

207. Gregoire, F.M. (2001) Adipocyte differentiation: from fibroblast to endocrine cell. *Exp. Biol. Med. (Maywood)*, 226, 997-1002.

208. Zuk, P.A., Zhu, M., Mizuno, H., Huang, J., Futrell, J.W., Katz, A.J., Benhaim, P., Lorenz, H.P. and Hedrick, M.H. (2001) Multilineage cells from human adipose tissue: implications for cell-based therapies. *Tissue Eng.*, 7, 211-228.

209. Ross, S.E., Hemati, N., Longo, K.A., Bennett, C.N., Lucas, P.C., Erickson, R.L. and MacDougald, O.A. (2000) Inhibition of adipogenesis by Wnt signaling. *Science*, 289, 950-953.

210. Sarjeant, K. and Stephens, J.M. (2012) Adipogenesis. *Cold Spring Harbor perspectives in biology*, 4, a008417.

211. Wang, A., Yue, F., Li, Y., Xie, R., Harper, T., Patel, N.A., Muth, K., Palmer, J., Qiu, Y., Wang, J. *et al.* (2015) Epigenetic priming of enhancers predicts developmental competence of hESC-derived endodermal lineage intermediates. *Cell stem cell*, 16, 386-399.

212. An, Y., Kang, Q., Zhao, Y., Hu, X. and Li, N. (2013) Lats2 modulates adipocyte proliferation and differentiation via hippo signaling. *PloS one*, 8, e72042.

213. Segawa, K., Matsuda, M., Fukuhara, A., Morita, K., Okuno, Y., Komuro, R. and Shimomura, I. (2009) Identification of a novel distal enhancer in human adiponectin gene. *J. Endocrinol.*, 200, 107-116.

214. Qiao, L., Maclean, P.S., Schaack, J., Orlicky, D.J., Darimont, C., Pagliassotti, M., Friedman, J.E. and Shao, J. (2005) C/EBPalpha regulates human adiponectin gene transcription through an intronic enhancer. *Diabetes*, 54, 1744-1754.

215. Urs, S., Smith, C., Campbell, B., Saxton, A.M., Taylor, J., Zhang, B., Snoddy, J., Jones Voy, B. and Moustaid-Moussa, N. (2004) Gene expression profiling in human preadipocytes and adipocytes by microarray analysis. *J. Nutr.*, 134, 762-770.

216. Davis, M.R., Arner, E., Duffy, C.R., De Sousa, P.A., Dahlman, I., Arner, P. and Summers, K.M. (2016) Expression of FBN1 during adipogenesis: Relevance to the

lipodystrophy phenotype in Marfan syndrome and related conditions. *Mol. Genet. Metab.*, 119, 174-185.

217. Lapidot, M. and Pilpel, Y. (2006) Genome-wide natural antisense transcription: coupling its regulation to its different regulatory mechanisms. *EMBO Rep.*, 7, 1216-1222.

218. Musri, M.M., Corominola, H., Casamitjana, R., Gomis, R. and Parrizas, M. (2006) Histone H3 lysine 4 dimethylation signals the transcriptional competence of the adiponectin promoter in preadipocytes. *J. Biol. Chem.*, 281, 17180-17188.

219. Schaffler, A., Orso, E., Palitzsch, K.D., Buchler, C., Drobnik, W., Furst, A., Scholmerich, J. and Schmitz, G. (1999) The human apM-1, an adipocyte-specific gene linked to the family of TNF's and to genes expressed in activated T cells, is mapped to chromosome 1q21.3-q23, a susceptibility locus identified for familial combined hyperlipidaemia (FCH). *Biochem. Biophys. Res. Commun.*, 260, 416-425.

220. Yamauchi, T., Kamon, J., Minokoshi, Y., Ito, Y., Waki, H., Uchida, S., Yamashita, S., Noda, M., Kita, S., Ueki, K. *et al.* (2002) Adiponectin stimulates glucose utilization and fatty-acid oxidation by activating AMP-activated protein kinase. *Nat. Med.*, 8, 1288-1295.

221. Yokota, T., Meka, C.S., Medina, K.L., Igarashi, H., Comp, P.C., Takahashi, M., Nishida, M., Oritani, K., Miyagawa, J., Funahashi, T. *et al.* (2002) Paracrine regulation of fat cell formation in bone marrow cultures via adiponectin and prostaglandins. *The Journal of clinical investigation*, 109, 1303-1310.

222. Corces, M.R., Trevino, A.E., Hamilton, E.G., Greenside, P.G., Sinnott-Armstrong, N.A., Vesuna, S., Satpathy, A.T., Rubin, A.J., Montine, K.S., Wu, B. *et al.* (2017) An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat Methods*.

223. Edwards, S.L., Beesley, J., French, J.D. and Dunning, A.M. (2013) Beyond GWASs: illuminating the dark road from association to function. *Am. J. Hum. Genet.*, 93, 779-797.

224. Flister, M.J., Tsaih, S.W., O'Meara, C.C., Endres, B., Hoffman, M.J., Geurts, A.M., Dwinell, M.R., Lazar, J., Jacob, H.J. and Moreno, C. (2013) Identifying multiple causative genes at a single GWAS locus. *Genome Res.*, 23, 1996-2002.

225. Voight, B.F., Kudaravalli, S., Wen, X. and Pritchard, J.K. (2006) A map of recent positive selection in the human genome. *PLoS Biol.*, 4, e72.

226. Musunuru, K., Strong, A., Frank-Kamenetsky, M., Lee, N.E., Ahfeldt, T., Sachs, K.V., Li, X., Li, H., Kuperwasser, N., Ruda, V.M. *et al.* (2010) From noncoding variant to phenotype via SORT1 at the 1p13 cholesterol locus. *Nature*, 466, 714-719.

227. Roman, T.S., Cannon, M.E., Vadlamudi, S., Buchkovich, M.L., Wolford, B.N., Welch, R.P., Morken, M.A., Kwon, G.J., Varshney, A., Kursawe, R. *et al.* (2017) A Type 2

Diabetes-Associated Functional Regulatory Variant in a Pancreatic Islet Enhancer at the Adcy5 Locus. *Diabetes*.

228. Klann, T.S., Black, J.B., Chellappan, M., Safi, A., Song, L., Hilton, I.B., Crawford, G.E., Reddy, T.E. and Gersbach, C.A. (2017) CRISPR-Cas9 epigenome editing enables high-throughput screening for functional regulatory elements in the human genome. *Nat. Biotechnol.*, 35, 561-568

.

229. Canver, M.C., Smith, E.C., Sher, F., Pinello, L., Sanjana, N.E., Shalem, O., Chen, D.D., Schupp, P.G., Vinjamur, D.S., Garcia, S.P. *et al.* (2015) BCL11A enhancer dissection by Cas9-mediated in situ saturating mutagenesis. *Nature*, 527, 192-197.

230. Findlay, G.M., Boyle, E.A., Hause, R.J., Klein, J.C. and Shendure, J. (2014) Saturation editing of genomic regions by multiplex homology-directed repair. *Nature*, 513, 120-123.

231. Boyle, E.A., Li, Y.I. and Pritchard, J.K. (2017) An Expanded View of Complex Traits: From Polygenic to Omnigenic. *Cell*, 169, 1177-1186.

232. Arnold, C.D., Gerlach, D., Stelzer, C., Boryn, L.M., Rath, M. and Stark, A. (2013) Genome-wide quantitative enhancer activity maps identified by STARR-seq. *Science*, 339, 1074-1077.

233. Melnikov, A., Murugan, A., Zhang, X., Tesileanu, T., Wang, L., Rogov, P., Feizi, S., Gnirke, A., Callan, C.G., Jr., Kinney, J.B. *et al.* (2012) Systematic dissection and optimization of inducible enhancers in human cells using a massively parallel reporter assay. *Nat. Biotechnol.*, 30, 271-277.