2021

Role of GPS2 in regulating transcription in adipose tissue and skeletal muscle differentiation

BOSTON UNIVERSITY

SCHOOL OF MEDICINE

Dissertation

ROLE OF GPS2 IN REGULATING TRANSCRIPTION IN ADIPOSE TISSUE AND SKELETAL MUSCLE DIFFERENTIATION

by

JOSEPH OROFINO

M.Sc., Boston University, 2015

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

2021

Approved by

First Reader

Valentina Perissi, Ph.D.
Associate Professor of Biochemistry

Second Reader

Stephen R Farmer, Ph.D.
Professor of Biochemistry

ACKNOWLEDGEMENTS

I would like to acknowledge all members of my lab past and present for being great friends and colleagues. To begin, I am forever indebted to Dr. Valentina Perissi for giving me the opportunity to work with her and to learn from her experience and wisdom as a principal investigator. I always joke that its fate that I, an adopted kid with an Italian name, ended up in a laboratory consisting of multiple Italians. Specifically, I will always be grateful to Dr. Carly Cederquist and Dr. Claudia Lentucci for being so welcoming and kind when I first joined the lab. I would be remiss to not mention Dr. Dafne Cardamone, who was integral in maintaining my sanity through all my 6 years. As well as Dr. Stefanie Chan, with whom I went on many food and coffee adventures. And of course, my thanks to soon-to-be Dr. Justin English for all of his shared efforts in our joint project and for his camaraderie.

I'm also thankful for many members of the Varelas lab as we've shared both laboratory space, protocols, and great conversation over the years. I am especially grateful to Dr. Julia Hicks-Berthet, Dr. Elena Stampouloglou, and Dr. Aleks Szymaniak for all of their advice and support they offered me throughout my time. As well as Nate Kingston and Debbie Chan, my classmates who started their dissertation at the same time and joined the same program.

And finally, I must thank my wonderful mother, sister, and friends who put up with my eccentricities all these 6 years.

ROLE OF GPS2 IN REGULATING TRANSCRIPTION

IN ADIPOSE TISSUE AND SKELETAL MUSCLE DIFFERENTIATION

JOSEPH OROFINO

Boston University School of Medicine, 2021

Major Professor: Valentina Perissi, Ph.D., Associate Professor of Biochemistry

ABSTRACT

Metabolic syndrome is an umbrella term for a large network of interdependent biochemical, metabolic and physiological factors that are associated with a higher risk of cardiovascular disease, and type 2 diabetes mellitus. Physiological risk factors such as insulin resistance, elevated blood pressure, visceral adiposity, and chronic inflammation are often rooted in alterations and dysfunction of the underlying metabolic tissues of the body, such as the adipose tissue and skeletal muscle. This work focuses on elucidating molecular mechanisms of gene regulation that impact the differentiation and function of both adipose tissue and skeletal muscle using well established *in vitro* models as well as genetic mouse models.

The first two chapters describe the transcriptional profiling of GPS2-Adiponectin-Cre specific knockout (GPS2-AKO) mice upon high fat diet feeding. We used various next generation sequencing (NGS) technologies to explore how the lack of GPS2 impact upon adipose tissue adaptation to the dietary stress on a molecular level. In the first Chapter, I discuss how RNAseq experiments performed in the epididymal white adipose tissue (eWAT), and subcutaneous

white adipose tissue (scWAT) reveal a complex role for GPS2 in the regulation of mitochondrial genes, inflammation and metabolism in mature adipocytes. The second chapter utilizes single cell RNA sequencing to examine how loss of GPS2 in mature adipocytes alters the underlying stromal vascular fraction tissue environment. This work has demonstrated key roles for GPS2 in directly regulating the response of adipocytes to high fat diet stress and indirectly affecting the crosstalk between adipocytes and the underlying SVF.

The third chapter explores the role of GPS2 as a key mediator of skeletal muscle cell differentiation. Given the known role of GPS2 as a transcriptional cofactor, we profiled the genomic occupancy of GPS2 across C2C12 differentiation via ChIPseq and found binding to known enhancer regions linked to important muscle differentiation terms. This led to the identification of an unexpected role for GPS2 as a required factor for the differentiation of C2C12 skeletal muscle cells *in vitro*.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

BAT ..... Brown adipose tissue

BCL ..... Base call files

bHLH ..... Basic Helix-Loop-Helix

CEBPα..... CCAAT/enhancer-binding protein alpha

ChIPseq..... Chromatin Immunoprecipitation Sequencing

CVD..... Cardiovascular Disease

DNA..... Deoxyribonucleic acid

DTT ..... Dithiothreitol

ECM ..... Extracellular Matrix

EDTA..... Ethylenediaminetetraacetic acid

EtOH..... Ethanol

eWAT ..... Epididymal white adipose tissue

FBS ..... Fetal Bovine Serum

FDR ..... False Discovery Rate

FGSEA ..... Fast gene set enrichment analysis

gDNA..... Genomic DNA

GPS2..... G protein pathway suppressor 2

GPS2-AKO ..... Gps2 Adiponectin-Cre Specific Knockout

GRO-Seq..... Global Run-On Sequencing

GSEA ..... Gene Set Enrichment Analysis

H3k9 ..... Histone 3, Lysine 9

HAT ..... Histone Acetyltransferase

HFD ..... High Fat Diet

HS ..... Horse Serum

IF ..... Immunofluorescence

IR ..... Insulin resistance

KO ..... Knockout

Log2FC ..... Log2 Fold Change

MHC ..... Myosin Heavy Chain

MRFs ..... Myogenic Regulatory Factors

Mrf4 ..... Myogenic Regulatory Factor 4

Myf5 ..... Myogenic Factor 5

MyoG ..... Myogenin

NGS ..... Next Generation Sequencing

PPARγ ..... Peroxisome proliferator-activated receptor gamma

PTM ..... Post-translational Modifications

RNA ..... Ribonucleic acid

RNAseq ..... RNA sequencing

scRNAseq ..... Single cell RNA sequencing

scWAT ..... Subcutaneous white adipose tissue

SVF ..... Stromal Vascular Fraction

SWI/SNF ..... SWItch/Sucrose NonFermentable

T2DM ..... Type II Diabetes Mellitus

TFs ..... Transcription factors

WB..... Western blot

WCL ..... Whole cell lysate

WT..... Wild-type

INTRODUCTION

Transcription and Transcriptional Regulation

Transcription: Overview

Gene expression, the means by which DNA is ultimately converted to a functionally active gene product, is universally important to all living organisms[1]. One of the earliest steps in this process is transcription, or the synthesis of RNA molecules from DNA via the RNA polymerase complexes (RNAP).

Transcription represents one of the first regulatory checkpoints in the control and coordination of gene expression. The primary regulators of this process are protein transcription factors (TFs) that bind DNA in a sequence-specific manner to directly regulate transcription[2]. Although TFs can vary in how they achieve this regulation, the majority are thought to function primarily by the recruitment of other proteins known as coactivators and corepressors. These mediators of transcriptional activity are often large, multi-subunit protein complexes that enhance or repress transcriptional activity through such means as remodeling of the local chromatin environment[3] or altering the affinity for the recruitment of the core components of RNAPII[4].

Chromatin Environment: Overview

The chromatin environment refers to the molecular scaffold that serves to condense, protect and regulate access to DNA. Specifically, chromatin is defined

as the higher-order structure formed by the aggregation and condensing of many nucleosomes, or stretches of around 147 bp of DNA wrapped around a specialized complex of proteins termed histones. Broadly speaking, chromatin is thought to regulate gene expression in two main ways: First, by physically restricting the ability of factors to bind to the DNA itself. Second, by post-translational modifications (PTMs) to the histone proteins which creates a pattern of recognition sites for regulatory proteins such as transcription factors or chromatin interacting proteins. These chromatin interacting proteins can be generally referred to as writers, readers, or erasers depending on how they chemically modify or interact with histones and chromatin[5].

PTMs of histone tails can occur on different residues and each may have different regulatory effects. One of the well-characterized modifications is lysine methylation. This modification is typically referred to by the name of the modified subunit followed by the amino acid residue position. For example, the canonical methylation sites for lysine methylation include histone H3 (K4, K9, K26, K27, K36, K79), histone H4 (K20), and Histone H1 (K26)[6]. Each lysine may be mono-, di- or trimethylated and methylation alters both the size and hydrophobicity of the residue. The effects of lysine methylation are highly context dependent and have been associated with both gene activation and gene repression. For example, H3K4me1 is considered a mark of activated genes in certain contexts while H3K27me1 and H3k9me3 are typically associated with regions of constitutive heterochromatin and gene silencing[7,8].

Another common modification made to the lysine residues on histone tails is the addition of an acetyl group or acetylation. In contrast to methylation, acetylation is largely correlated with transcriptionally active regions and open chromatin. It is thought that the electrostatic properties of the acetyl group partially neutralize the charge of the lysine, and weakens its interaction with negatively charged DNA allowing nuclear factors to access the region[9].

This regulation of chromatin structure through the pattern of histone modifications introduced the general framework of histone writers, readers, and erasers[10]. These are the broad classes of proteins responsible for recognizing, adding, removing and recognizing specific histone modifications. Histone writers are the proteins that catalyze the addition of various modifications onto histone tails. They can be separated into two main groups: histone acetyltransferases (HATs)[11] and histone methyltransferases (HMTs)[12], each named for the reaction they catalyze. Histone erasers function to catalyze the removal of histone PTMs and include members of the HDAC family (histone deacetylase)[13] and Jumonji C domain-containing family of HDMs (histone demethylases)[14].

The histone readers are made up proteins that recognize various histone modifications and mediate their regulatory effects. They do this through the presence of specialized protein domains that recognize and bind the different PTMs made to histones. This binding may serve to recruit other factors involved in further chromatin remodeling or directly impact transcription through the recruitment of RNAPII and associated factors. Some prominent examples include

the readers of lysine methylation such as ankyrin, Tudor, PHD, and WD40. These families of proteins typically have specific affinities for one or several specific lysine methylation sites on histones. Other well-characterized histone readers are the bromodomain family, which recognize and bind to acetylated histone residues[15].

Transcription during development and differentiation

Complex multicellular organisms, such as mammals, function through a vast variety of specialized cells compartmentalized into higher-order tissues and organs. This process begins during embryonic development as totipotent stem cells specialize and differentiate into a multitude of different cell types and lineages. This patterning is largely orchestrated through precise spatial and temporal regulation of gene expression through large-scale remodeling of the chromatin environment and activation and repression of specific gene programs. Nearly all differentiated and specialized cells are directed through development primarily by regulation of gene expression including blood cells[16] (hematopoiesis), fat cells[17] (adipogenesis), and muscle[18] (myogenesis).

One recurring theme in lineage specification is the presence of master regulatory transcription factors that remodel the chromatin environment to activate lineage-specific gene programs and repress extraneous genes. A classic example of a master regulatory factor is the muscle-specific TF MyoD. MyoD is one of the earliest factors expressed during myogenesis and functions primarily

by interacting with various cofactors to remodel the chromatin environment. It was shown to directly promote histone acetylation on a genome-wide scale by binding the histone acetyltransferases P300 and PCAF to activate transcription[19]. Conversely, MyoD has also been implicated in the repression of gene activity by its presence and association with histone deacetylase HDAC1 on many genes prior to differentiation[20]. Many of the other master transcription factors such as PPARγ, AP-1, SCL all interact with various cofactors to remodel chromatin and influence gene expression through the course of lineage specification and differentiation[21].

GPS2: Overview

G protein suppressor 2 (GPS2) was originally identified during a screen for suppressors of G protein activated MAPK signaling in yeast and mammalian cells[22,23]. It is a 37 kDa, relatively ubiquitous protein that has been demonstrated to have a variety of genomic and non-genomic functions. With regards to transcription and its nuclear functions, early work showed that it was a core component of the N-CoR/SMRT co-repressor complex[24]. In addition to this role, GPS2 has also been shown to be a coactivator and has multiple functional interactions with a variety of different transcriptional regulators including p300, FXR, HNF4, p53, MSH4/MSH5, LXR and PPARγ[25,26].

CHAPTER ONE

Transcriptomic profiling of the subcutaneous and epididymal adipose tissue depots of GPS2-AKO mice

INTRODUCTION

Some of the early work into the functions of GPS2 revealed its key role as an anti-inflammatory agent in both adipose tissue and macrophages. It was demonstrated that overexpression of GPS2 impaired activation of pro-inflammatory pathways *in vitro* and *in vivo* through inhibition of JNK activation. In accordance with these observations, the downregulation of GPS2 and associated corepressor SMRT in the adipose tissue of obese human individuals was found to be inversely correlated with inflammatory status[27]. In a macrophage-specific GPS2-knockout mouse model, isolated GPS2-deficient macrophages were shown to have a pro-inflammatory gene signature and were sensitized to various stress signals relative to WT macrophages[28].

There have been two parallel adipose tissue specific GPS2-KO mouse models described in the literature. In a study by Drareni et al. 2018[29], it was observed that GPS2-AKO mice displayed maladaptive expansion of the WAT when exposed to high fat diet feeding. Specifically, loss of GPS2 was linked to unrestrained HIF1α activation and subsequent adipocyte hypertrophy. After 4 weeks of HFD, GPS2-AKO mice rapidly became insulin resistant compared to WT. Separately, work by Cederquist *et al* 2018[30] also described the generation

and characterization of a GPS2-AKO mouse line utilizing the same Cre driver but differing in the locations of the LoxP sites. They found that GPS2-AKO mice became obese under a standard laboratory chow diet while maintaining insulin sensitivity. This apparent uncoupling of obesity and insulin resistance was not observed in the GPS2-AKO mouse model described by Drareni *et al.* and further studies will be needed to investigate the differences between the two models.

GPS2:Inhibition of K63 ubiquitination

Ubiquitination is another post-translational modification made on the lysine residues of proteins. Ubiquitin (Ub) is a 76 amino acid protein that can be attached to substrate proteins as a single molecule or part of a larger chain of multiple Ub molecules connected through isopeptide bonds[31]. Specific Ub modifications are referred to by the lysine residue present on Ub that serves to link the poly-ubiquitin chains. Some of the commonly studied polyubiquitin chains include K11, K48 and K63[32]. Similar to the aforementioned histone code, the pattern of ubiquitination and deubiquitination acts as a regulatory signal that regulates the protein's activity and fate with different Ub chains promoting specific functions. As an example, K48 Ub chains are most often associated with signaling a protein for degradation through the proteasome[33]. By contrast, K63 ubiquitination has been shown to have regulatory roles in vesicle trafficking, DNA damage repair and inflammatory processes[34].

Previous work from our lab revealed that many of the recognized genomic

and non-genomic functions of GPS2 are mediated through the regulation of K63 ubiquitin. Relevant to the regulation of adipose tissue metabolism, GPS2-mediated regulation of K63 ubiquitination was demonstrated to regulate insulin signaling through the PI3-AKT pathway. Work by Cederquist *et al* demonstrated that K63 non-proteolytic ubiquitination is a key component of AKT activation in the adipocyte insulin signaling pathway and that GPS2 regulates this process through enzymatic inhibition of the Ubc13 E3 ligase. Genetic knockout or acute downregulation of GPS2 in primary adipocytes resulted in constitutive activation of insulin signaling both *in vitro* and *in vivo*. This genetic mouse model displayed an intriguing metabolic phenotype wherein KO mice became obese under a standard laboratory chow diet while remaining metabolically healthy. In short, the adipose tissue of GPS2-AKO mice show no inflammation, and improved systemic insulin sensitivity relative to WT mice.

Adipose Tissue: Overview

The primary purpose of adipose tissue is the storage and metabolism of lipids in response to nutritional and energy balance needs. At a high level, there exist two main types of adipose: white adipose tissue (WAT) and brown adipose tissue (BAT)[35]. In addition to the differences played by the two main types of adipose tissue, there is a wealth of literature that has also demonstrated unique roles for the different adipose tissue depots spread throughout the body[36,37]. WAT houses the majority of lipids found in mammals and white adipocytes are

easily identifiable as cells containing a large, unilocular lipid droplet. WAT can be found organized into depots spread across the body and is canonically associated with energy balance, and lipid storage. BAT, by contrast, is primarily characterized by multilocular lipid droplet containing adipocytes that also express high numbers of mitochondria. The primary recognized function of BAT is the dissipation of stored lipids and energy in the form of heat, a process known as thermogenesis[38]. While it was once thought that BAT was restricted in expression to only human infants and other species, a plethora of recent work has demonstrated that BAT is present in adult humans and that populations of adipocytes exist which display mixed characteristics of both white and brown adipose[39,40].

Adipose tissue: more than just lipid storage

While adipose tissue plays an indispensable role in lipid storage and fatty acid metabolism, it has recently come to be appreciated as a critical secretory organ that has a profound impact on energy balance, inflammation, and other essential systemic responses[41,42]. The term adipokine has been created to describe the wide variety of products secreted by adipocytes that act in a paracrine and endocrine manner to influence physiological responses. One of the most well-studied adipokines is leptin, a hormone almost exclusively expressed and secreted by the adipose tissue[43]. Its primary function is acting as a direct signal to the central nervous system to regulate food intake and energy

expenditure. Other well-known adipokines include adiponectin, resistin, IL-6, and TNFα. Importantly, in addition to regulating the underlying tissue environment, many of these adipokines are known to have long range interactions with other organs such as the muscle, pancreas, and liver. These secreted factors are not only implicated in the regulation of the functions and homeostasis of immune cells but also control diverse metabolic processes[44].

Adipogenesis: Role of transcription

The differentiation of adipocytes is a tightly coordinated series of events directed by a transcriptional cascade regulated through two major factors: PPARγ and C/EBPα. Both of these proteins have been shown to be essential for adipogenesis *in vitro* and *in vivo* and binding experiments demonstrate their presence at a large majority of upregulated genes during differentiation[45]. Upon initiation of adipogenesis, C/EBPβ and C/EBPδ begin to be expressed and induce the expression of C/EBPα and PPARγ. C/EBPβ, in particular, was demonstrated to bind to a large number of sites that were found in closed chromatin region and was required for the subsequent binding of early adipogenic transcription factors such as STAT5A, RXR and Glucocorticoid receptor (GR)[46]. C/EBPα and PPARγ serve to regulate a majority of adipogenic genes as well as mutually inducing the expression of each other. PPARγ is a member of the nuclear hormone receptor family and must heterodimerize with another nuclear receptor in order to be transcriptionally active. PPARγ has been

shown to interact with a number of canonical corepressors and coactivators to remodel chromatin prior to and during differentiation including the NCoR/SMRT complex, CBP/300 and the mediator complex[47].

Of important note to the following work, GPS2 was shown to prime the local chromatin environment for recruitment of the master adipogenic regulatory factor PPARγ at selected gene promoters in 3T3-L1 preadipocytes[48]. The molecular basis for this regulation was shown to be through the demethylation of the promoter region via stabilization of lysine demethylase KDM4A. A similar mechanism was reported for GPS2-mediated regulation of a significant proportion of nuclear encoded mitochondrial genes (neMITO) and confirmed *in vivo* to function through inhibition of UBC13[49]. Chromatin Immunoprecipitation Sequencing (ChIPseq) and Global Run-On Sequencing (GROseq) experiments confirmed that upon loss of GPS2 in the brown adipose tissue of GPS2-AKO mice, mitochondrial gene promoters were aberrantly enriched with H3K9 methylation and transcription was impaired due to RNAPII stalling. At a tissue level, this loss of GPS2 in the brown adipose tissue (BAT) resulted in reduced mitochondrial content compared to WT mice. This regulation of mitochondrial genes was separately observed in the study by Drareni *et al* 2018 where they noted the downregulation of 216 mitochondrial-related genes in their GPS2-AKO mouse model.

In this chapter, we have performed RNAseq to profile gene expression in isolated adipocytes from our GPS2-AKO mouse model. Specifically, we sought to

determine potential differences in the response to high fat diet stress that may explain their altered metabolic phenotype.

Results

Following the development and metabolic characterization of our GPS2-AKO genetic mouse model, we isolated primary adipocytes from two main depots: epididymal WAT and subcutaneous WAT. We profiled these depots under conditions of chow or HFD and performed RNAseq experiments to understand the transcriptomic differences occurring in wild-type or GPS2-deficient mature adipocytes. We compared KO and WT adipocytes from the Epi WAT under chow and HFD conditions, and the SC WAT under HFD conditions. Each experiment included 4 WT and 4 KO mice that were matched with littermate controls.

To begin, we performed RNA sequencing on isolated adipocytes from the Epi WAT of our GPS2-AKO mice under chow diet conditions. At this time point, the GPS2-AKO mice are obese relative to their WT counterparts but seemingly protected from the associated metabolic complications. We performed RNAseq on the isolated adipocytes to determine what transcriptomic changes were occurring in absence of GPS2 and how they might contribute to this protective phenotype.

Using a false discovery rate (FDR) cutoff of <.1, we identified a total of 768 differentially expressed genes with 263 upregulated, and 505 downregulated genes between KO and WT adipocytes (**Figure 1**).

Overrepresentation Analysis: Epididymal Chow Adipocytes

Overrepresentation analysis (ORA) of upregulated genes showed a significant enrichment of Gene Ontology: Biological Process (GO:BP) terms related to the respiratory electron transport chain (GO:0042775, GO:0022904), mitochondrial organization (GO:0007005), and mitochondrial transport (GO:0006839). Downregulated genes are enriched for terms such as platelet degranulation (GO:0002576), extracellular matrix and focal adhesions (GO:0001953, GO:1903053) and angiogenesis (GO:0045765). (**Figure 1b).**

Gene Set Enrichment Analysis: Epididymal Chow Adipocytes

Gene set enrichment analyses based on a Log2FoldChange (log2FC) ranked list of all the genes discovered in the experiment yielded similar results to the ORA. Gene sets were considered significant with a cutoff of FDR <.25 and the top 10 sets with the highest and lowest normalized enrichment score are shown in table 1 and 2. Gene sets with a positive NES include those related to the electron transport chain, oxidative phosphorylation, and mitochondria. Those with a negative NES include pathways related to focal adhesions, and complement cascade and coagulation (**Tables 1, 2).**

Loss of GPS2 in the epididymal depot leads to specific upregulation of mitochondrial genes

Amongst the 263 upregulated genes, we note that 68 of them are identified as genes encoding for mitochondrial proteins as found in the MitoCarta 3.0 database. Consistent with our previous work, the mitochondrial genes regulated by GPS2 appear to be enriched for specific functional terms such as “TCA cycle” and “Electron Transport Chain”. This was observed for enrichment of the upregulated DEGs as well as in gene sets with a positive normalized enrichment score as determined by GSEA performed with a ranked list by log2 fold change. This was unexpected as previous work had demonstrated the downregulation of these genes upon loss of GPS2 in various models. Specifically, *in vitro* knockdown of GPS2 in 3T3-L1, 293T and Hela cells all revealed a similar downregulation of mitochondrial genes.

In the total downregulated genes, we see a concentration of genes related to inflammatory pathways (*Ifitm3, Itih4, Sp100, Il1r1, Src, Stat3, Irak3, Ptpn14, Lgals9, Il6t, Trim56*) (**Figure 1c**). Similarly, the GSEA results also show various gene sets such as Interferon Signaling and Antigen Processing and Presentation enriched amongst the downregulated genes. These findings are again in contrast with what we have previously seen in other settings. GPS2 has been shown to be a key anti-inflammatory mediator in the adipose tissue and macrophages in both mouse and human models.

GPS2-KO adipocytes display a sensitized response to high fat diet as compared to C57BL/6J mice

We subjected our GPS2-AKO mice and WT controls to a 16-week high fat diet and we discovered 946 differentially expressed genes with 482 upregulated and 464 downregulated when comparing between WT and KO adipocytes (**Figure 2a)**.

Overrepresentation Analysis: Epididymal HFD Adipocytes

Overrepresentation analysis (ORA) of upregulated genes showed a significant enrichment of Gene Ontology: Biological Process (GO:BP) terms related to neutrophils (GO:0043312, GO:0002283, GO:0002446), extracellular matrix (GO:0030198, GO:0022617) and nucleosomes/chromatin (GO:0006334, GO:0031497). Downregulated genes were enriched for terms such as muscle contraction (GO:0006936), apoptotic cell clearance (GO:2000427, GO:2000425), and endothelial cell migration (GO:0010596, GO:0043535). (**Figure 2b)**

Gene Set Enrichment Analysis: Epididymal HFD Adipocytes

Gene set enrichment analyses were based on a descending Log2FoldChange (log2FC) ranked list of all the genes discovered in the experiment. Gene sets were considered significant with a cutoff of FDR <.25 and the top 10 sets with positive and negative normalized enrichment scores are shown in tables 3 and 4,

respectively. We note the presence of several gene sets related to potassium channels and the nitric oxide pathway (**Tables 3, 4**).

Overall, we observe a strong upregulation of inflammatory related pathways and downregulation of ECM and remodeling-related genes. These results can be seen in both overrepresentation analysis of DEGs as well as in the GSEA analyses performed using gene lists ranked by log2foldchange. Specifically, upregulated genes under this condition show a strong enrichment for multiple terms related to neutrophil degranulation, and neutrophil immunity. The downregulated genes show a specific enrichment for pathways such as muscle contraction, extracellular matrix organization, and cell migration. GSEA analysis demonstrates pathways related to DNA damage and cell cycle amongst those with a positive NES. Pathways with a negative NES include those related to amino acid metabolism and potassium channels (**Figure 2c)**.

Subcutaneous and Epididymal WAT GPS2-KO adipocytes under HFD are metabolically worse compared to their WT counterparts

In order to study any depot-specific effects in our GPS2-AKO mice, we also profiled the SC WAT after a 16-week high fat diet. It has been shown in the literature that the different adipose tissue depots play distinct roles in the response to a high fat diet stress. The subcutaneous depot, in particular, has been thought to act as a metabolic sink that sequesters excess lipids and protects from systemic complications related to lipid spillover. After a 16-week

HFD stress, we can see that the SC tissue appears to be highly inflamed in our KO mice relative to their WT counterparts. Using a false discovery rate of FDR < .1, we identified 1170 DEGs in the SC WAT under HFD (**Figure 3a)**.

Overrepresentation Analysis: Subcutaneous HFD Adipocytes

Overrepresentation analysis (ORA) of upregulated genes showed a significant enrichment of Gene Ontology: Biological Process (GO:BP) terms related neutrophils (GO:0043312, GO:0002283, GO:0002446), inflammatory responses, and ECM. Downregulated genes were strongly enriched for terms related to various metabolic pathways such as branched-chain amino acid metabolic process (GO:0009081) or cellular amino acid catabolic process (GO:0009063). (**Table 5, 6)**

Gene Set Enrichment Analysis: Subcutaneous *HFD Adipocytes*

As seen in table 5 and table 6, Gene set enrichment analyses revealed an enrichment of gene sets amongst the downregulated genes related to mitochondrial function including pathways for respiratory electron transport, TCA Cycle, and amino acid degradation pathways. General inflammatory pathways were noted to be amongst the upregulated gene sets including KEGG pathways for natural killer cell mediated cytotoxicity, systemic lupus erythematosus, and phagocytosis.

Isolated adipocytes from subcutaneous depot under 16-week HFD show high levels of inflammation and metabolic dysfunction

On a global level, we see a robust enrichment for inflammation and inflammatory pathway terms in both the GSEA and functional enrichment analyses (**Figure 3b, 3c).** Among these are pathways related to neutrophil degranulation, cytokine signaling, and general inflammatory pathways. The downregulated genes are enriched for a large variety of amino acid catabolic pathways as well as mitochondrial gene sets. These changes are consistent with previous reports characterizing wild-type adipocytes subjected to a high fat diet stress. However, it is important to note that our experimental design was designed to compare between WT and KO adipocytes both subjected to a high fat diet stress. GPS2-KO adipocytes appear to be more sensitized to the effects of the high fat diet stress relative to the WT adipocytes.

Though any direct comparison is complicated by the fact that the expression profiling was not performed at the same time, we would like to comment on potential differences we observed between the epididymal and subcutaneous depots. In contrast with the Epi WAT, the SC adipocytes appear to be more perturbed in their metabolic capacity. The Epi WAT under HFD shows a downregulation of genes related to muscle contraction, and extracellular matrix. The SC WAT, however, suggest downregulation of many metabolic pathways related to amino acid and fatty acid metabolism. Similarly, upregulated genes in the Epi WAT appear more related to pathways specifically related to DNA

damage while those in the SC show an enrichment for a wide variety of inflammatory, phagocytic and cytokine signaling pathways.

Overlap between RNAseq data and adipokine secretion study suggests potential paracrine and long-range signaling mechanism for GPS2-deficient adipocytes

Given that we observe whole-body effects on metabolism in our GPS2-AKO mice, even though GPS2 is deleted only in mature adipocytes, we theorized that some of these changes could be mediated by alterations in adipocyte signaling. The genetic deletion in our GPS2-AKO mice is driven by the expression of Adiponectin, which has been demonstrated in many other sources to be a unique marker for mature adipocytes only.

Recently, it has come to be appreciated that adipose tissue is a major secretory organ that can alter the activity of other metabolic organs such as the muscle or liver as well as mediate whole-body physiological changes. We utilized our RNAseq data coupled with recently published data identifying major secreted adipokines to find potential secreted factors that are altered in response to loss of GPS2. Although we acknowledge the imperfect relationship between mRNA expression and secretion, we intend for this analysis to serve as a preliminary proof of principle to provide potential candidates while we actively set up mass spectrometry and exosome isolation experiments to better and more directly characterize secreted factors.

Khan *et. al* 2018[50] used a combination of AHA and pSILAC labeling to identify 1,337 total unique proteins to serve as a candidate list for adipose tissue specific secreted factors. Although their study was focused on uncovering differentially secreted factors between WAT and BAT upon norepinephrine stimulation, we used the total list of discovered adipokines as our background for secreted factors from adipocytes. We filtered the 1,337 proteins to 461 proteins that also displayed evidence of being a secreted protein in external literature. To begin identifying candidate secreted proteins that are altered upon GPS2-deficiency, we overlapped the 461 candidate proteins with our list of differentially expressed genes from each depot and condition. We can see that there are 109, 155, 165 overlapping candidates (data not shown) in the Epi WAT under chow diet, Epi WAT under HFD and SC WAT under HFD, respectively. We performed enrichment analysis on these overlapping candidates using the total list of DEG from each experiment as the background. We can see a strong enrichment for extracellular matrix and remodeling genes in all three conditions. We also note the presence of several key adipokines found in all three conditions that could be important for mediating systemic complications.

Discussion

Previous work has identified GPS2 as an important regulator of nuclear encoded mitochondrial gene expression in adipose tissue. Similarly, the role of GPS2 as an anti-inflammatory agent has also been well-established in both

murine and human models. Given this, we were interested to observe the upregulation of mitochondrial genes and a downregulation of inflammatory genes in the epididymal WAT from our GPS2-AKO mice under chow diet.

A study by Jones *et al* 2019[51] demonstrated that the visceral WAT of C57BL/6J mice show a characteristic downregulation of mitochondrial genes and upregulation of ECM-related genes in response to a high fat diet stress. Taking this study as a model for a generalized response to HFD, we observe the opposite transcriptomic changes while under a chow diet but in a setting of obesity. Intriguingly, we can see that when subjected to even a short high fat diet stress, obesity induced by GPS2 deletion is no longer uncoupled from insulin resistance and inflammation. This can be seen directly from our experiments demonstrating increased body weight and insulin resistance in our GPS2-AKO mice compared to wild-type controls after just 5 weeks of high fat diet (data not shown). Further studies will address this interesting switch between phenotypes.

Addressing our phenotype under a chow diet, on a whole-body level, our GPS2-AKO mice are more insulin sensitive and display no signs of ectopic lipid deposition despite their increased body weight. We observed downregulation of innate immunity and interferon signaling pathways in the KO adipocytes in our RNAseq studies. Previous work also demonstrated no significant differences in inflammatory gene expression or infiltration of macrophages between WT and KO mice at this time. In general, at both the tissue and adipocyte level, we

confirm no signs of inflammation present in our KO despite their increased body weight.

More surprising was the revelation that we observed a strong upregulation of mitochondrial genes. We have demonstrated in a number of *in vitro* and *in vivo* settings that GPS2 is required for the activation of nuclear encoded mitochondrial genes. Given this, we hypothesize that this upregulation of mitochondrial genes in our KO adipocytes is a potential compensatory mechanism to maintain proper mitochondrial function given the lack of transcriptional activation. RNAseq is a steady-state measure of mRNA levels and it is possible that while nascent transcription is impaired, there are compensatory mechanisms that stabilize mRNA transcripts. To test this hypothesis, we plan on performing more comprehensive profiling of the metabolic function of these isolated adipocytes. We will assay the respiratory function of isolated mitochondria from these cells as well as examine the adipocytes at a higher resolution than bulk RNAseq by employing single nuclei sequencing to examine potential adipocyte heterogeneity.

Under chow conditions, our GPS2-AKO mice display an interesting uncoupling of obesity and metabolic complications. Under a high fat diet, we observe the same phenotype as reported in the literature for a separately developed GPS2-AKO conditional mouse model: loss of GPS2 in the setting of energy surplus exacerbates the deleterious effects on the adipose tissue. To better determine when in time this switch occurs, we decided to investigate the

response of the adipose tissue after a shorter exposure to high fat diet in order to capture early changes that may explain the rapid switch between phenotypes.

In order to determine the potential differences in our GPS2-AKO mice and the parallel model reported in the literature by Treuter et al[29], we subjected our mice to a 16-week HFD. Our results are concordant with what was previously reported and we see a rapid decline in metabolic health at both the whole-body level as well as characteristic shifts in the transcriptome of the adipocytes. We observe in the Epi WAT a strong enrichment of pathways related to DNA damage, inflammation and ECM in our KO samples. This enrichment relative to the wild-type adipocytes under HFD strongly agrees with other published data suggesting loss of GPS2 results in worse metabolic outcomes in the adipose tissue and reflected in whole-body metabolic issues. We believe our future studies will better explain the differences we observer under a chow diet where obesity is uncoupled from inflammation and IR.

Given the known importance of adipose tissue as a secretory organ, we theorized that the phenotype we observed in our GPS2-AKO mice might be mediated through the alteration of their secretome. Changes in the profile of secreted adipokines could be responsible for both mediating changes in the local tissue environment and also systemic physiology. As a preliminary test of this hypothesis, we have performed a simple overlap of the differentially expressed genes found in our isolated adipocytes with a list of known and verified adipokines identified in the literature. We identify a significant enrichment for

genes related to fibrosis, inflammation and extracellular matrix. It is important to note that the correlation between expression of mRNA and protein is not always positively correlated due to other layers of post-transcriptional or post-translational regulation. To remedy the shortcomings in this initial analysis, we are planning on performing mass spectrometry experiments and measuring circulating levels of adipokine markers directly in order to identify possible secreted factors that are altered after GPS2 loss in adipose tissue.

To begin exploring whether altered signaling and cross-talk from GPS2-deficient adipocytes is contributing to our whole-body phenotype, we began with single cell RNA sequencing of the local tissue microenvironment of the adipose tissue after a 5-week high fat diet exposure. At this early timepoint, our GPS2-AKO mice already display reduced insulin sensitivity and show a trend towards increased body weight compared to their wild-type counterparts. We hypothesized that early changes in the adipose tissue may be reflected in the local tissue environment and feedback to reinforce the fundamental changes mediated by the adipocytes. We again chose two representative adipose tissue depots, epididymal and subcutaneous, and isolated the stromal vascular fraction after 5 weeks of high fat diet feeding. Our goal was to identify potential changes in populations between our KO and WT mice as well as examine the early response of the adipose tissue as a whole to high fat diet stress. These results will be discussed in chapter 2.

Open questions and Future Directions

One of the unresolved questions from our work is the mechanism behind the upregulation of mitochondrial genes seen in the Epi WAT of chow-fed GPS2-AKO mice under chow diet. Our previous work had demonstrated an essential role for GPS2 in the mediation of mitochondrial retrograde signaling and as a key transcriptional activator of nuclear-encoded mitochondrial genes. These findings were extended *in vivo* as the brown adipose tissue of GPS2-AKO mice shows reduced mitochondrial content relative to their wild-type counterparts. However, the actual respiratory capacity of isolated mitochondria from GPS2-KO adipocytes was equal to those isolated from WT adipocytes in the brown adipose tissue. These results indicated that overall mitochondrial function is not impaired in GPS2-deficient BAT, but instead results in reduced mitochondrial mass.

In context with what we see in the Epi WAT, these results suggest a potential post-transcriptional mechanism that maintains mitochondrial function even in absence of GPS2. Given our observation of upregulation of mitochondrial mRNAs, we posit that this may be a result of either stabilization of these transcripts and/or more efficient translation that may compensate for lack of activation at a transcriptional level. Ongoing work in the laboratory has lent credence to this hypothesis by demonstrating that GPS2 regulates mitochondria-associated mRNA local translation via inhibition of K63 ubiquitination on PABPC1. This interaction was identified through mass spectrometry experiments seeking to identify target proteins regulated by UBC13 and GPS2.

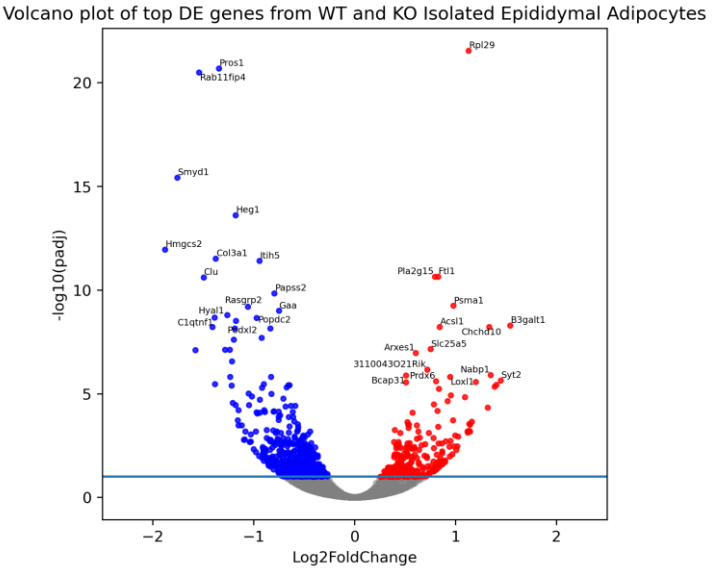
To address this question further, we would like to more comprehensively profile the mitochondria found in our KO adipocytes under both chow and high fat diet in both depots. As an exploratory experiment, we are interested in looking at the relative levels of expression of oxidative phosphorylation complexes found in both WT and KO adipose tissue mitochondria. To more accurately characterize their respiratory capacity, we are planning on utilizing Seahorse XF Cell Mito Stress Test assays to specifically probe the oxygen consumption rate and extracellular acidification rate of isolated mitochondria from the WAT depots. These results should indicate to us whether we observe any potential impacts on mitochondrial function. If our hypothesis regarding the existence of a compensatory mechanism to maintain normal mitochondrial function is correct, then we would expect to not see any major differences in the actual respiratory capacity of isolated mitochondria from our KO mice. To complement these studies of mitochondrial function, we are also considering the use of techniques such as RIBOseq and mRNA stability assays to more directly test whether the upregulation of mitochondrial genes in our GPS2-AKO mice is a result of a compensatory mechanism that allows for the proper functionality throughout development in spite of loss of nuclear encoded mitochondrial gene activation mediated by GPS2.

We also would like to profile the subcutaneous depot of our GPS2-AKO mice under chow diet. Several studies have suggested that the subcutaneous depot acts as a “metabolic sink” to preferentially sequester excess lipids and

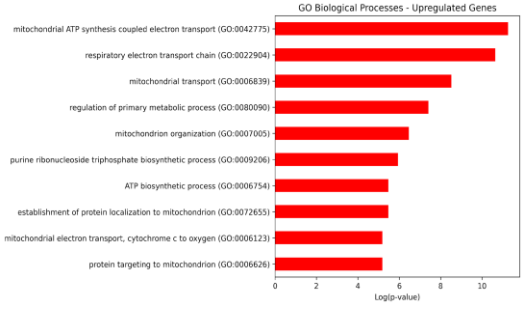
prevent lipotoxicity and other metabolic complications[52]. It is also known that the subcutaneous depot responds differently to high fat diet stress and has also been shown to have a higher capacity for “browning” than visceral WAT in mice[53]. Coupled with our previous studies on the epididymal depot, these studies would provide a more complete picture as to how loss of GPS2 in two major adipose tissue depots contributes to protection from IR and inflammation in a setting of obesity.

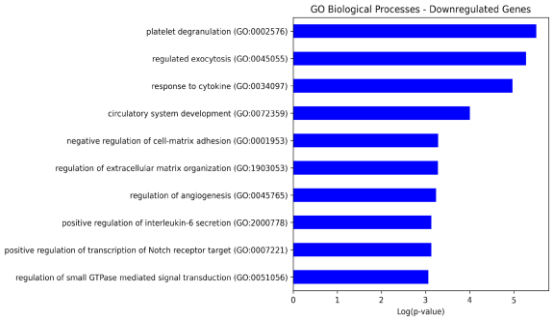
Figures and Figure legends

A

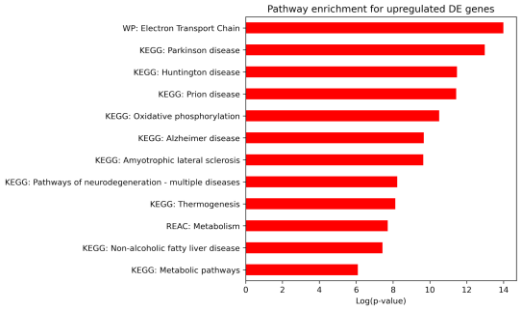


B





C



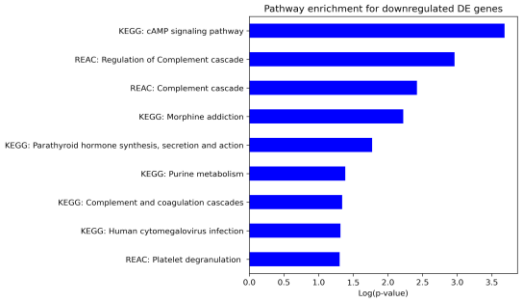
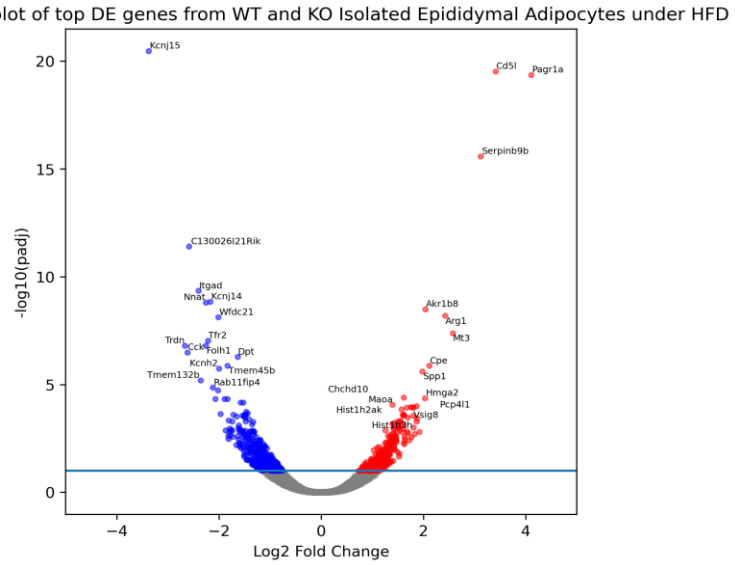


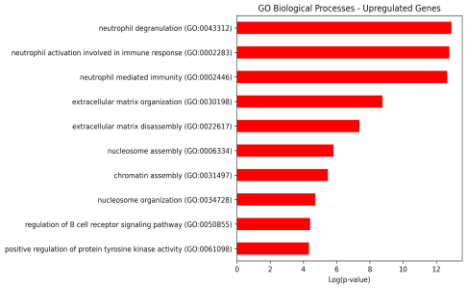
Figure 1.
Loss of GPS2 leads to upregulation of mitochondrial genes and downregulation of inflammatory genes in isolated adipocytes from mice under a chow diet

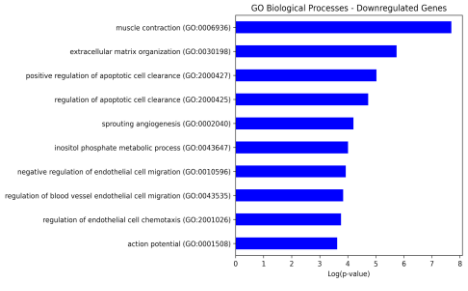
(a) Volcano plot of DEG results with selected up- and down-regulated genes with top ranked genes by p-value annotated
(b) Overrepresentation analysis for Gene Ontology Biological Process terms performed for up- and downregulated genes
(c) Overrepresentation analysis for Pathway enrichment terms from KEGG, REAC, and WP performed for up- and downregulated genes

A

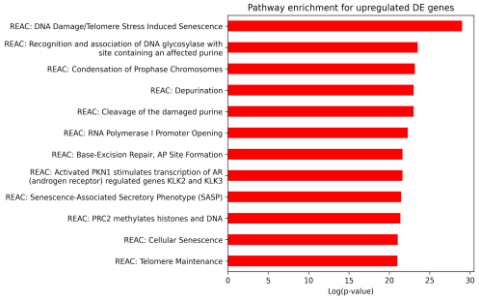

Volcano plot of top DE genes from WT and KO Isolated Epididymal Adipocytes under HFD
-log10(padj)
Log2 Fold Change
Kcnj15
Cd5l
Pagr1a
Serpinb9b
C130026I21Rik
Itgad
Nnat
Kcnj14
Wfdc21
Tfr2
Trdn
Cck
Folh1
Dpt
Kcnh2
Tmem45b
Tmem132b
Rab11fip4
Akr1b8
Arg1
Mt3
Cpe
Spp1
Chchd10
Maoa
Hmga2
Pcp4l1
Hist1h2ak
Ysig8
Hist1h1b

B


GO Biological Processes - Upregulated Genes
neutrophil degranulation (GO:0043312)
neutrophil activation involved in immune response (GO:0002283)
neutrophil mediated immunity (GO:0002446)
extracellular matrix organization (GO:0030198)
extracellular matrix disassembly (GO:0022617)
nucleosome assembly (GO:0006334)
chromatin assembly (GO:0031497)
nucleosome organization (GO:0034728)
regulation of B cell receptor signaling pathway (GO:0050855)
positive regulation of protein tyrosine kinase activity (GO:0061098)
Log(p-value)


GO Biological Processes - Downregulated Genes
muscle contraction (GO:0006936)
extracellular matrix organization (GO:0030198)
positive regulation of apoptotic cell clearance (GO:2000427)
regulation of apoptotic cell clearance (GO:2000425)
sprouting angiogenesis (GO:0002040)
inositol phosphate metabolic process (GO:0043647)
negative regulation of endothelial cell migration (GO:0010596)
regulation of blood vessel endothelial cell migration (GO:0043535)
regulation of endothelial cell chemotaxis (GO:2001026)
action potential (GO:0001508)
Log(p-value)

C


Pathway enrichment for upregulated DE genes
REAC: DNA Damage/Telomere Stress Induced Senescence
REAC: Recognition and association of DNA glycosylase with site containing an affected purine
REAC: Condensation of Prophase Chromosomes
REAC: Depurination
REAC: Cleavage of the damaged purine
REAC: RNA Polymerase I Promoter Opening
REAC: Base-Excision Repair, AP Site Formation
REAC: Activated PKN1 stimulates transcription of AR (androgen receptor) regulated genes KLK2 and KLK3
REAC: Senescence-Associated Secretory Phenotype (SASP)
REAC: PRC2 methylates histones and DNA
REAC: Cellular Senescence
REAC: Telomere Maintenance
Log(p-value)

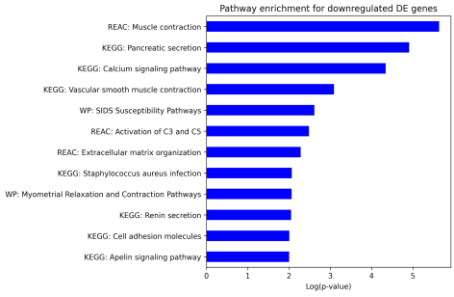
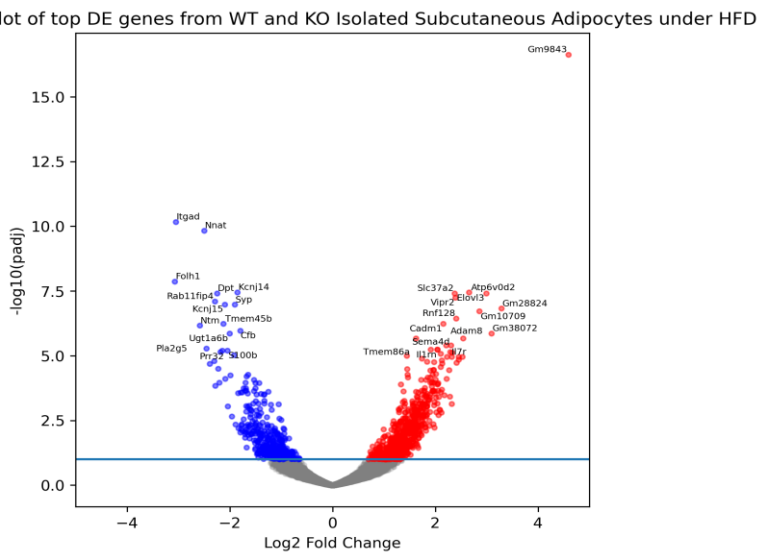

Pathway enrichment for downregulated DE genes
REAC: Muscle contraction
KEGG: Pancreatic secretion
KEGG: Calcium signaling pathway
KEGG: Vascular smooth muscle contraction
WP: SIDS Susceptibility Pathways
REAC: Activation of C3 and C5
REAC: Extracellular matrix organization
KEGG: Staphylococcus aureus infection
WP: Myometrial Relaxation and Contraction Pathways
KEGG: Renin secretion
KEGG: Cell adhesion molecules
KEGG: Apelin signaling pathway
Log(p-value)

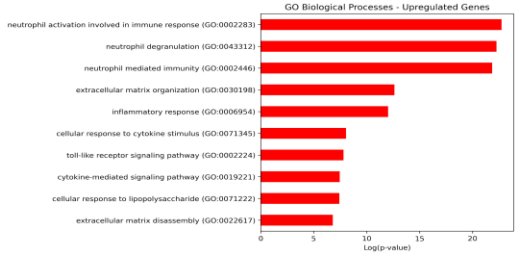
Figure 2.
Loss of GPS2 leads to upregulation of DNA damage genes and downregulation of ECM genes in isolated adipocytes from the epididymal WAT from mice under HFD

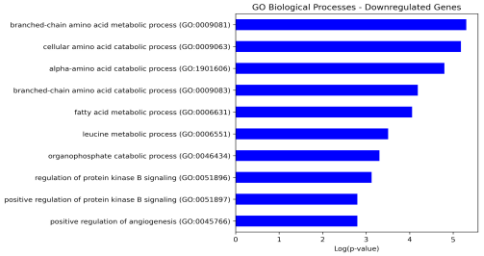
(a) Volcano plot of DEG results with selected up- and down-regulated genes with top ranked genes by p-value annotated
(b) Overrepresentation analysis for Gene Ontology Biological Process terms performed for up- and downregulated genes
(c) Overrepresentation analysis for Pathway enrichment terms from KEGG, REAC, and WP performed for up- and downregulated genes

A

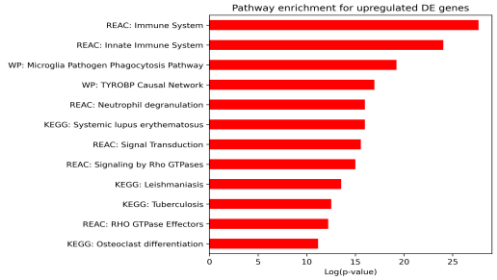

Volcano plot of top DE genes from WT and KO Isolated Subcutaneous Adipocytes under HFD
-log10(padj)
Log2 Fold Change
Gm9843
Itgad
Nnat
Folh1
Dpt
Kcnj14
Rab11fip4
Syp
Kcnj15
Ntm
Tmem45b
Ugt1a6b
Cfb
Pla2g5
Prr32
S100b
Slc37a2
Atp6v0d2
Elovl3
Vipr2
Gm28824
Rnf128
Gm10709
Cadm1
Adam8
Gm38072
Sema4d
Tmem86a
Il1rl1
Il7r

B


GO Biological Processes - Upregulated Genes
neutrophil activation involved in immune response (GO:0002283)
neutrophil degranulation (GO:0043312)
neutrophil mediated immunity (GO:0002446)
extracellular matrix organization (GO:0030198)
inflammatory response (GO:0006954)
cellular response to cytokine stimulus (GO:0071345)
toll-like receptor signaling pathway (GO:0002224)
cytokine-mediated signaling pathway (GO:0019221)
cellular response to lipopolysaccharide (GO:0071222)
extracellular matrix disassembly (GO:0022617)
Log(p-value)


GO Biological Processes - Downregulated Genes
branched-chain amino acid metabolic process (GO:0009081)
cellular amino acid catabolic process (GO:0009063)
alpha-amino acid catabolic process (GO:1901606)
branched-chain amino acid catabolic process (GO:0009083)
fatty acid metabolic process (GO:0006631)
leucine metabolic process (GO:0006551)
organophosphate catabolic process (GO:0046434)
regulation of protein kinase B signaling (GO:0051896)
positive regulation of protein kinase B signaling (GO:0051897)
positive regulation of angiogenesis (GO:0045766)
Log(p-value)

C


Pathway enrichment for upregulated DE genes
REAC: Immune System
REAC: Innate Immune System
WP: Microglia Pathogen Phagocytosis Pathway
WP: TYROBP Causal Network
REAC: Neutrophil degranulation
KEGG: Systemic lupus erythematosus
REAC: Signal Transduction
REAC: Signaling by Rho GTPases
KEGG: Leishmaniasis
KEGG: Tuberculosis
REAC: RHO GTPase Effectors
KEGG: Osteoclast differentiation
Log(p-value)

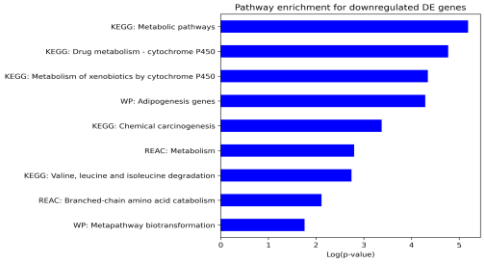

Pathway enrichment for downregulated DE genes
KEGG: Metabolic pathways
KEGG: Drug metabolism - cytochrome P450
KEGG: Metabolism of xenobiotics by cytochrome P450
WP: Adipogenesis genes
KEGG: Chemical carcinogenesis
REAC: Metabolism
KEGG: Valine, leucine and isoleucine degradation
REAC: Branched-chain amino acid catabolism
WP: Metapathway biotransformation
Log(p-value)

Figure 3
Loss of GPS2 leads to upregulation of inflammatory related pathways and downregulation of metabolic processes in isolated adipocytes from the subcutaneous WAT from mice under HFD

(a) Volcano plot of DEG results with selected up- and down-regulated genes with top ranked genes by p-value annotated
(b) Overrepresentation analysis for Gene Ontology Biological Process terms performed for up- and downregulated genes
(c) Overrepresentation analysis for Pathway enrichment terms from KEGG, REAC, and WP performed for up- and downregulated genes

Table 1. Significantly enriched pathways with a positive NES from GSEA performed using gene lists sorted by log2 fold change

Pathway	FDR	NES
REACTOME RESPIRATORY ELECTRON TRANSPORT ATP SYNTHESIS BY CHEMIOSMOTIC COUPLING AND HEAT PRODUCTION BY UNCOUPLING PROTEINS	0.00729	3.42
KEGG OXIDATIVE PHOSPHORYLATION	0.00729	3.40
REACTOME RESPIRATORY ELECTRON TRANSPORT	0.00729	3.36
KEGG PARKINSONS DISEASE	0.00729	3.23
REACTOME TCA CYCLE AND RESPIRATORY ELECTRON TRANSPORT	0.00729	3.07
REACTOME INFLUENZA VIRAL RNA TRANSCRIPTION AND REPLICATION	0.00729	2.93
REACTOME PEPTIDE CHAIN ELONGATION	0.00729	2.92
KEGG RIBOSOME	0.00729	2.88
REACTOME SRP DEPENDENT COTRANSLATIONAL PROTEIN TARGETING TO MEMBRANE	0.00729	2.88
REACTOME TRANSLATION	0.00734	2.83
REACTOME MITOCHONDRIAL PROTEIN IMPORT	0.00729	2.82
REACTOME 3 UTR MEDIATED TRANSLATIONAL REGULATION	0.00729	2.82
REACTOME INFLUENZA LIFE CYCLE	0.00729	2.65
KEGG HUNTINGTONS DISEASE	0.00764	2.61
REACTOME FORMATION OF THE TERNARY COMPLEX AND SUBSEQUENTLY THE 43S COMPLEX	0.00729	2.58

NES represents the normalized enrichment score calculated by GSEA, where positive or negative values correspond to genes in the pathway that are increased or decreased in the knockout condition

Table 2. Signicantly enriched pathways with a negative NES from GSEA performed using gene lists sorted by log2fold change

Pathway	FDR	NES
REACTOME INTERFERON ALPHA BETA SIGNALING	0.00729	-2.15
BIOCARTA INTEGRIN PATHWAY	0.00729	-2.08
PID AR TF PATHWAY	0.00729	-1.98
KEGG ANTIGEN PROCESSING AND PRESENTATION	0.00729	-1.91
KEGG ECM RECEPTOR INTERACTION	0.00729	-1.90
PID REG GR PATHWAY	0.00729	-1.90
BIOCARTA MET PATHWAY	0.00729	-1.88
PID TCPTP PATHWAY	0.00729	-1.88
ST GRANULE CELL SURVIVAL PATHWAY	0.00784	-1.88
REACTOME INTERFERON GAMMA SIGNALING	0.00729	-1.87
PID IL6 7 PATHWAY	0.00729	-1.85
PID FAK PATHWAY	0.00729	-1.84
REACTOME GABA RECEPTOR ACTIVATION	0.01179	-1.83
BIOCARTA IL6 PATHWAY	0.01293	-1.82
KEGG DILATED CARDIOMYOPATHY	0.00962	-1.82

Table 3. Significantly enriched pathways with a positive NES from GSEA performed using gene lists sorted by log2 fold change

Pathway	FDR	NES
REACTOME CHROMOSOME MAINTENANCE	0.002948	2.67
REACTOME PEPTIDE CHAIN ELONGATION	0.002948	2.63
KEGG RIBOSOME	0.002948	2.61
REACTOME INFLUENZA VIRAL RNA TRANSCRIPTION AND REPLICATION	0.002948	2.58
REACTOME TELOMERE MAINTENANCE	0.002948	2.58
REACTOME INFLUENZA LIFE CYCLE	0.002948	2.57
PID PLK1 PATHWAY	0.002948	2.56
REACTOME DEPOSITION OF NEW CENPA CONTAINING NUCLEOSOMES AT THE CENTROMERE	0.002948	2.56
REACTOME DNA REPLICATION	0.002948	2.53
REACTOME SRP DEPENDENT COTRANSLATIONAL PROTEIN TARGETING TO MEMBRANE	0.002948	2.53
REACTOME 3 UTR MEDIATED TRANSLATIONAL REGULATION	0.002948	2.51
REACTOME CELL CYCLE	0.002948	2.49
REACTOME MITOTIC M M G1 PHASES	0.002948	2.45
REACTOME RNA POL I PROMOTER OPENING	0.002948	2.42
REACTOME RNA POL I TRANSCRIPTION	0.002948	2.41

Table 4. Significantly enriched pathways with a negative NES from GSEA performed using gene lists sorted by log2 fold change

Pathway	FDR	NES
REACTOME POTASSIUM CHANNELS	0.00503	-1.96
REACTOME NITRIC OXIDE STIMULATES GUANYLATE CYCLASE	0.01028	-1.95
REACTOME BRANCHED CHAIN AMINO ACID CATABOLISM	0.01028	-1.92
KEGG VALINE LEUCINE AND ISOLEUCINE DEGRADATION	0.00503	-1.92
KEGG PROPANOATE METABOLISM	0.00503	-1.91
PID SHP2 PATHWAY	0.00295	-1.91
NABA BASEMENT MEMBRANES	0.00503	-1.89
KEGG DRUG METABOLISM CYTOCHROME P450	0.00704	-1.87
REACTOME SIGNALING BY BMP	0.02374	-1.84
REACTOME INWARDLY RECTIFYING K CHANNELS	0.01982	-1.84
BIOCARTA NO1 PATHWAY	0.02287	-1.82
KEGG BASAL CELL CARCINOMA	0.01280	-1.82
REACTOME NUCLEAR RECEPTOR TRANSCRIPTION PATHWAY	0.01862	-1.80
REACTOME SIGNALING BY FGFR1 MUTANTS	0.03552	-1.78
KEGG PEROXISOME	0.00503	-1.76

Table 5. Significantly enriched pathways with a positive NES from GSEA performed using gene lists sorted by log2 fold change

Pathway	FDR	NES
KEGG LEISHMANIA INFECTION	0.00207	2.84
KEGG SYSTEMIC LUPUS ERYTHEMATOSUS	0.00207	2.66
PID PLK1 PATHWAY	0.00207	2.54
KEGG NATURAL KILLER CELL MEDIATED CYTOTOXICITY	0.00207	2.50
KEGG B CELL RECEPTOR SIGNALING PATHWAY	0.00207	2.50
REACTOME EXTRACELLULAR MATRIX ORGANIZATION	0.00207	2.49
KEGG CELL ADHESION MOLECULES CAMS	0.00207	2.49
KEGG FC GAMMA R MEDIATED PHAGOCYTOSIS	0.00207	2.44
REACTOME GPVI MEDIATED ACTIVATION CASCADE	0.00207	2.43
REACTOME INTEGRIN CELL SURFACE INTERACTIONS	0.00207	2.41
PID FOXM1 PATHWAY	0.00207	2.39
REACTOME RNA POL I PROMOTER OPENING	0.00207	2.36
REACTOME IL 3 5 AND GM CSF SIGNALING	0.00207	2.36
KEGG TYPE I DIABETES MELLITUS	0.00207	2.35
NABA ECM AFFILIATED	0.00207	2.33

Table 6. Significantly enriched pathways with a negative NES from GSEA performed using gene lists sorted by log2 fold change

Pathway	FDR	NES
REACTOME RESPIRATORY ELECTRON TRANSPORT	0.00207	-2.93
REACTOME TCA CYCLE AND RESPIRATORY ELECTRON TRANSPORT	0.00215	-2.86
REACTOME RESPIRATORY ELECTRON TRANSPORT ATP SYNTHESIS BY CHEMIOSMOTIC COUPLING AND HEAT PRODUCTION BY UNCOUPLING PROTEINS	0.00213	-2.84
KEGG VALINE LEUCINE AND ISOLEUCINE DEGRADATION	0.00207	-2.63
KEGG PROPANOATE METABOLISM	0.00207	-2.50
KEGG OXIDATIVE PHOSPHORYLATION	0.00213	-2.50
REACTOME BRANCHED CHAIN AMINO ACID CATABOLISM	0.00207	-2.48
KEGG PARKINSONS DISEASE	0.00213	-2.45
KEGG PEROXISOME	0.00210	-2.40
REACTOME ER PHAGOSOME PATHWAY	0.00207	-2.37
REACTOME VIF MEDIATED DEGRADATION OF APOBEC3G	0.00207	-2.35
REACTOME SIGNALING BY WNT	0.00207	-2.33
KEGG HUNTINGTONS DISEASE	0.00223	-2.31
REACTOME REGULATION OF ORNITHINE DECARBOXYLASE ODC	0.00207	-2.27
REACTOME AUTODEGRADATION OF CDH1 BY CDH1 APC C	0.00207	-2.26

CHAPTER TWO

Profiling the heterogeneity of the SVF of GPS2-AKO adipose tissue and its impact on whole-body metabolism

Introduction

Adipose Tissue: classical model of dysfunction

The classical model of adipose tissue dysfunction posits that prolonged lipid stress due to diet or other factors leads to unhealthy expansion of the adipose tissue in an attempt to accommodate the excess lipids. Under typical conditions, the adipose tissue adapts to store excess lipids through an expansion in the size of existing adipocytes (hypertrophy) or generation of new adipocytes via adipogenesis (hyperplasia)[54]. When the storage capacity of the tissue is reached through prolonged periods of positive energy balance, the adipose tissue becomes stressed and initiates an inflammatory response[55]. This program has both local and systemic effects that serve to promote inflammation, hypoxia and fibrosis. All of these processes together ultimately impair the function of the adipose tissue itself and eventually lead to ectopic lipid deposition in other tissues of the body such as the muscle, heart and liver. In addition, dysfunctional adipose tissue may have a drastically altered secretome, which has far-reaching effects on systemic inflammation and the cross-talk between adipose tissue and important metabolic organs including the liver and skeletal muscle[56]. It is thought that this myriad of complications together contributes to the development of

various physiological complications such as non-alcoholic fatty liver disease (NAFLD)[57], T2DM and insulin resistance (IR)[58] and cardiovascular disease[59]. Although multifactorial in nature, all of these complications have been demonstrated to result from alterations in gene expression that cascade into wider reaching tissue and systemic effects.

Cellular composition of the adipose tissue

Recent research has highlighted the complexity of the diverse cell populations that support and underlie the adipocytes, which are collectively known as the stromal vascular fraction (SVF). The SVF is a diverse population of both progenitor cells of various origins (e.g. adipocyte precursors, smooth muscle, endothelial) and fully differentiated cells (e.g. immune cells, fibroblasts, pericytes)[60]. Collectively, these cells have been shown to be critical for adipose tissue function and serve essential roles in structural organization[61], vascularization and angiogenesis[62], adipogenesis[63] and the regulation of inflammatory response[64].

Although the stromal vascular fraction was known to be heterogeneous, single-cell technologies have deepened our appreciation for its complexity by revealing even more specialized lineages and subgroups. Burl et al[65] characterized a large cohort of cells derived from beige adipocytes stimulated by CL 316,243. They discovered populations of macrophages with mixed M1 and

M2 characteristics with specific enrichment for migration, adhesion and ECM remodeling.

Other single cell studies have expanded our knowledge of both the underlying SVF as well as heterogeneity in the adipocytes themselves. Recent single cell studies have uncovered a number of different specialized subsets of adipocyte precursor cells (APCs). One such study by Schwalie et al[66] delineated a new subgroup of APCs termed Areg cells due to their ability to negatively regulate the adipogenic potential of surrounding APCs through paracrine signaling. Knockdown of certain dependent genes or removal of these Aregs led to marked increase in adipogenesis in the depot. Another group separately demonstrated the existence of fibro-inflammatory progenitors, which share functional similarities with the Aregs but also display unique fibrotic and inflammatory gene profiles. Hepler et al[67] showed that this novel population was also capable of inhibiting adipogenesis and the proportions of these cells were differentially regulated by high fat diet feeding. These studies as a whole have demonstrated the vital importance of APCs in regulating the remodeling of the adipose tissue and its response to stress related to metabolic imbalance.

In addition to these APCs, immune cells are also critically important for the response of the adipose tissue to various stressors. Single cell studies have enabled the discovery of a number of specialized populations not previously recognized in bulk assays. Jaitin et al[68] mapped the SVF of mouse epididymal WAT across a high fat diet time course and discovered a distinct Cd9+/Trem2+

macrophage cluster with a gene signature enriched for lipid metabolism and classical macrophage phagocytosis genes. This cell population was validated and demonstrated to be present in human visceral adipose tissue and tyheir proportions positively correlated with body mass index (BMI). Rajhbandari et al[69] used scRNAseq to uncover a specific program of adaptive immune cells that expanded in response to thermogenic conditions and upregulated expression of IL10, a known antagonist for thermogenesis. scRNAseq has enabled more precise identification of important immune cell populations that contribute to remodeling, inflammation and thermogenesis in adipose tissue[70].

Remodeling of adipose tissue

Under conditions of stress, adipocytes may release inflammatory cytokines and other paracrine factors that will lead to remodeling and adaptation of the SVF through recruitment of inflammatory cells or altered expression of ECM components[71,72]. Prolonged periods of stress caused by chronic inflammation or nutrient excess in the adipose tissue may impair the formation of new adipocytes or result in deleterious levels of fibrosis from the SVF. These changes may serve to further exacerbate the inflammatory response and lead to systemic complications[73,74].

SVF in GPS2-AKO mice

Given our knowledge of the critical interplay between adipocytes and the SVF, we hypothesized that knockout of GPS2 in adipocytes may also be altering the paracrine signaling to the local tissue environment and/or its endocrine signaling of factors important for systemic metabolic and inflammatory pathways. Our previous experiments were focused on profiling gene expression directly in mature, GPS2-deleted adipocytes. While we found interesting differences between WT and KO in their response to high fat diet stress, none of the changes seemingly explain the unique phenotype of our GPS2-AKO mice. While obesity induced by GPS2 deletion in adipocytes is uncoupled from insulin resistance and inflammation under a chow diet, our other studies demonstrated that GPS2-AKO mice show much higher levels of inflammation, insulin resistance and adipose tissue dysfunction after being subjected to 16 weeks of high fat diet stress. In order to determine when this switch between phenotypes occurs, we decided to examine our GPS2-AKO mice after a shorter, 5-week high fat diet stress and profile how the underlying tissue environment is affected. After 5 weeks of high fat diet, our KO mice show significant weight gain and the development of insulin resistance relative to their WT counterparts. We chose this time point to capture the early changes seen under a HFD stress that may be mediating the switch from being metabolically healthy to metabolically unhealthy in a setting of obesity. We were interested in examining changes in the underlying SVF of the adipose tissue as it is known have extensive cross-talk

with mature adipocytes and to be an important determinant of overall tissue and metabolic health.

Results

We chose to employ single cell RNA sequencing to profile the SVF of the epididymal and subcutaneous adipose depots of our GPS2-AKO mice after a 5-week HFD stress. We chose this time point after observing that our KO mice under HFD for this shorter period of time show a rapid decline in metabolic fitness, and demonstrate signs of insulin resistance without reaching the same body weight as observed in chow conditions. Importantly, this time point is also prior to our previous studies that subjected GPS2-AKO mice to a 16-week HFD and where we observed a strong shift towards inflammation and metabolic dysfunction in the KO white adipose tissue by bulk RNAseq.

Due to timing limitations, SVF fractions were frozen in 10% DMSO, FBS prior to sequencing. Cells were shown to have a >90% live dead ratio after using a trypan blue stain to test for viability. Suspensions were prepared and sequenced using the 10X genomics platform.

For each experiment, we sequenced the SVF from four mice per tissue depot and processed the initial data using the 10X Cell Ranger platform[75]. After alignment to the MM10 reference genome, the resulting matrices from Cell Ranger were combined using the Cellranger Aggr tool to generate a single matrix encompassing all the samples for downstream analysis. Post-processing

including clustering, annotation, and other analyses were all performed using Scanpy[76].

scRNAseq of Epididymal SVF under HFD reveals diverse populations of immune cell populations

To examine any potential differences in the underlying SVF tissue mediated by GPS2-deficient adipocytes under a high fat stress, we performed scRNA sequencing of the SVF of the Epi WAT. We did not sort the cells prior to sequencing in order to obtain a comprehensive view of the entire tissue. We identified 8,797 cells that after unbiased clustering through k-nearest neighbors (KNN) and the Leiden graph-clustering method[77], partitioned into 14 distinct clusters (**Figure 4a, 4b)**.

In order to understand the different cell populations present, we manually curated a list of marker genes from other scRNAseq studies performed in the murine adipose tissue and SVF. We used these markers as a first pass strategy to annotate our clusters to a potential representative population. We can see that several cell populations robustly express previously validated markers (**Figure 4c)**. For example, cluster 8 demonstrates specific expression of known B-cell markers such as *Cd79a, Cd79b*, and *Ms4a1*. This annotation strategy also clearly delineated late stage differentiated APCs from earlier precursors due to their strong and distinct expression of mature markers *Retn* and *Adipoq*. We also clearly identified a population of mesothelial cells expressing *Msln, Lrrn4, and*

Up3kb and these cells have been shown in other scRNAseq studies to be present specifically in the visceral WAT[78].

The remaining cell populations were identified by examining their marker genes (i.e. the top differentially expressed genes in a cluster vs. all the rest of the clusters). Different visualization techniques of the expression of these genes across clusters enabled us to provide general biological identities to all of our clusters. This strategy also allowed us to identify subpopulations by identifying populations with slightly shifted expression profiles but sharing the expression of core marker genes (**Figures 5, 6).**

Diverse macrophage populations are present in the Epi WAT SVF after a short high fat diet stress

Given that we did not sort our cells by any known lineage markers prior to sequencing, the predominant cell types recovered were immune cell populations. Of these, we observe 3 subpopulations of macrophages (cluster 0, cluster 1, cluster 7) which combined represent nearly 46% of all cells identified in the experiment. Cluster 0 expresses general macrophage markers such as *Fcer1g, Tyrobp, Trem2* and is enriched for genes related to phagocytosis pathways. This cluster likely represents resident adipose tissue macrophages involved in cell clearance. Cluster 1 appears to display a M2-like macrophage signature based on *Mrc1* expression. The balance of M1/M2 macrophages are known to play a critical role in the inflammatory response to high fat diet in adipose tissue. Cluster

7 appears to be a population of proliferating macrophages based on their expression of *Trem2, Lyz2* and cell proliferation markers *Stmn, Tubb5, Birc5, and Mki67*.

Diverse myeloid cell populations present in the Epi WAT SVF

We also identified several clusters that appear to be non-macrophage myeloid cells. We can see 4 main groups that appear to represent natural killer cells (cluster 4), dendritic cells (cluster 5), B-cells (cluster 8), and ILC2/T-regs (cluster 9). The B-cell cluster and the natural killer T cell cluster were both identified by well-characterized gene markers, *Cd79a* and *Nkg7*, respectively. Cluster 9 strongly expresses *Gata3* and the Il33 receptor subunit *Il1rl1*. Previously, *Gata3* has been shown to be a universal marker for ILC2 cells across different tissues and *Il1rl1* was specifically shown to be enriched in adipose-derived ILC2 cells[79].

Pdgfrα/Pdgfrβ populations present in the Epi WAT SVF

The remaining cell populations appear to be related to adipose-derived stromal progenitors (cluster 2, cluster 3) or mesothelial cells (cluster 10). Clusters 2 and 3 express a high proportion of collagen and ECM related markers as well as canonical APC markers such as *Dcn, Gsn, Fn1.* Cluster 2 appears to be both Pdgfrα/Pdgfrβ positive while cluster 3 is mainly marked by Pdgfrβ alone. We can

also see that cluster 3 expresses *Dpp4* and *Pi16*, which have previously been used to mark a distinct mesenchymal ASC cluster found in the adipose tissue[80].

Compositional Analysis reveals a potential decrease in differentiating APCs in KO EPI SVF

One of our main interests was in determining any condition-specific effects between our WT and KO SVF that were occurring in the local adipose tissue environment. We have employed several methods to determine if there are any compositional differences. We can see in **Figure 7a, 7b**. the relative proportions of each cluster split by sample and condition. We note several clusters that are consistently changed when accounting for replicate and litter status such as cluster 6 (committed APC) or cluster 8 (B-cells). In this depot, these cell populations were both increased in the KO SVF samples. This observation is in agreement with previous studies showing that high fat diet stress leads to an increase in the infiltration of B-cells into visceral adipose tissue, while their abrogation was shown to prevent the development of insulin resistance[81].

We also visualized the density of cells embedded into clusters based on their condition and note a higher relative density of the macrophage subpopulations, two ASC populations, and a committed APC cluster in our KO SVF (**Figure 7c)**. It is important to note that this is a visual comparison only and is largely dependent on appropriate sequencing depth.

Finally, we used the recently developed tool scCODA[82], which utilizes a Bayesian model, to examine compositional differences using a statistical algorithm. This method requires a cell population to be set as a reference and implicitly assumes that only a few cell types are changing. We employed this method (**Figure 7d)** and we can see that cluster 6, previously identified as differentiating APC cells is significantly increased in the KO condition. In order to remove any potential biases from selection of the reference population, we applied the scCODA test using every cluster as a reference and noted the amount of times each cluster was returned as being significantly changed. We observe that cluster 6 is found to be consistently changed and decreased in the KO SVF even when considering every other cluster as a reference.

To investigate the presence of any condition-specific differences in these committed APCs, we performed differential expression testing between WT and KO cells assigned to cluster 6 using MAST[83]. We can see that there are only 4 significant DE genes *Rpl29, Cish, Sept4, Retnla* at a permissive FDR cutoff. In general, we do not observe any transcriptional differences that would suggest an impairment in their capacity to differentiate or any large-scale alterations in their function (**Figure 8)**.

scRNAseq of SC SVF reveals diverse populations of macrophages

We performed the same analysis for four samples of SVF derived from the subcutaneous depot of GPS2-AKO and WT mice. We identified 6708 cells that

could be resolved into 15 distinct clusters (**Figure 9a, 9b)**. The largest populations of cells present are immune cells with several clusters expressing a diverse range of macrophage markers: Clusters 1, 3, 7. Unexpectedly, we detect a small population of cells expressing *Mpz*, a core component of the myelin sheath found in peripheral nerves[84]. We believe this likely represents a small source of contamination from remnants of a nerve that escaped filtration during the tissue preparation.

We employed the same strategy for biological identification of clusters as for the epididymal depot. We first relied on manually curated marker genes taken from scRNAseq studies of murine SVF from the literature (**Figure 10)**, followed by visual inspection of marker genes delineating each cluster (**Figure 11)**.

Cluster 1 expresses *Mrc1* and *Clecl10a*, both well validated markers for M2-like macrophages, and also shows expression of MHCII genes. Cluster 3 expresses general macrophage markers (*Fcer1g, Tyrobp, Itgam)*. Functional enrichment of the top 100 marker genes for cluster 3 reveals a significant enrichment for genes involved in phagocytosis and likely represents a population of resident adipose tissue macrophages. Cluster 7 macrophages display a highly interesting gene signature where we observe a strong enrichment for both general macrophage markers (*Lgals3, Lyz2, Cd36, Trem2, Tyrobp, and Fcer1g)* as well as ECM and remodeling-related genes (*Actb, Actg1, Col1a1, Col3a1, Col1a2, Col4a2, Col5a1, Col4a1)*. We also observe expression of pro-inflammatory factors (*Cxcl2, Ccl4, Ccl2)*. A similar macrophage population

expressing both ECM and inflammatory genes has been reported in scRNAseq studies of the adipose tissue after a long term 18 week high fat diet[85].

scRNAseq in the SC HF reveals potential cell populations related to DIO high fat diet stress

Other ASC populations present include cluster 2, which are Pdgfrα/ Pdgfrβ positive progenitors which express high levels of collagen and ECM related genes. We detect a small population of cells, cluster 12, that distinctly express late stage adipocyte markers such as *Retn* and *Adipoq*. We speculate that this cluster 12 may represent a small amount of contamination as both genes are highly specific markers of mature adipocytes, which should be largely removed during the preparation of the SVF.

We also note the presence of the recently reported DPP4+ progenitors as likely represented by our cluster 0 due to their specific expression of Pi16 and Dpp4. This cell population has recently been implicated as multipotent mesenchymal progenitors and was also detected in the epididymal SVF.

Interestingly, we see an ASC population that is enriched for various markers associated with diet induced obesity (DIO) as seen in another study profiling the effects of high fat diet on APCs in visceral adipose tissue[86]. These include proinflammatory factors (*Ccl2, Cxcl1, Ccl7, Cxcl3)* as well as ECM related transcripts (*Postn*, *Col1a1*, and *Bgn).* There are no clusters found in the epididymal depot which display a similar gene expression profile as this one.

Compositional analyses of the SC SVF several potential changes in cell populations between WT and KO

We performed the same set of analyses to look at the potential compositional differences in the WT and KO SVF derived from the SC WAT. We can see the proportion of cells from each replicate are more variable than what was observed in the Epi WAT (**Figure 12a, 12b)**. When examining the proportions of each cluster by replicate, there are four clusters that appear to be changed between WT and KO (cluster 7, cluster 9, cluster 10, and cluster 14). We have chosen to focus on clusters 7 and 10 as cluster 14 consists of a total of 12 cells and cluster 9 is likely a contaminant from remnants of peripheral nerve that were included in the preparation. Clusters 7 and 10 were previously identified in this study as collagen-expressing macrophages and smooth muscle cells, respectively. In both cases, we can see that the populations are largely contributed to by KO samples when accounting for replicate status. We note that cluster 7 is largely made up of cells from one replicate, and both contain relatively few cells total.

Similarly, density analyses reveal similar changes in proportion for cluster 7 and 10 (**Figure 12c).** We can see that cluster 7 appears to be increased in the KO condition, however, as noted before, we can see that most of the cells are attributed to one KO replicate. Cluster 10 also shows increased density in the KO condition and this increase in the KO is consistent even when considering

replicates. Given the interesting nature of these clusters, we hope to further investigate this potential difference via flow cytometry and by increasing the depth of our experiment by adding replicates to the scRNAseq.

Using the scCODA tool for cell composition returns no clusters that appear to be credibly changed in proportion between WT and KO (**Figure 12d)**. Overall, examining the number of cells attributed to each replicate demonstrates a high level of variability and we believe that increased depth might reveal more robust changes in cell populations using this method. We observed potential changes in populations using other visualization techniques and believe the lack of statistical support from this method stems from a lower sequencing depth compared to the epididymal depot.

Discussion

In this study, we have profiled the local adipose tissue environment from the epididymal and subcutaneous WAT depots isolated from our GPS2-AKO mice. Our major finding is a significant increase in the abundance of differentiating APCs found in the Epi WAT. We also observe interesting depot-specific responses to HFD between the Epi SVF and SC SVF that may be reflective of their sensitivity to metabolic stress.

Using a combination of descriptive and computational methods, we found a cluster of cells representing differentiating APCs is increased in the SVF of the visceral WAT of GPS2-AKO mice. This cell population expresses multiple APC

markers (*Pparg*, *Fabp4,* and *Adgrf5)* and based on the simultaneous presence of a separate cluster that robustly expresses late stage differentiation markers (*Retn*, *Adipoq*), we identified this cluster as a subset of APCs that are committed but not terminally differentiated APCs. In addition to examining proportional changes, we also performed differential expression testing between WT and KO cells within this cluster of differentiating APCs. Interestingly, we detect no meaningful differential expression between WT and KO cells.

This increased proportion of cells found in our KO mice suggests two potential hypotheses: 1. There is a block in differentiation that prevents these cells from becoming fully mature adipocytes or 2. There is an increase in the amount of differentiating APCs at this time point in our GPS2-AKO mice.

Addressing the first hypothesis, we posit that this increase in proportion found in the KO is likely due to altered paracrine signaling from the adipocytes that prevents terminal differentiation. In past experiments not shown here, we isolated the SVF from the epididymal WAT of GPS2-AKO mice and performed *in vitro* differentiation experiments. We did not note any defect in differentiation capacity in this setting between KO and wild-type SVF. This along with the lack of significant differential expression between the KO and WT APCs is suggestive of the fact that the cells themselves do not have any intrinsic defect in their differentiation capacity and there is an external signal or feedback mechanism from the adipocytes themselves that mediates the block. Supporting this idea, other studies of APCs and their differentiation potential have suggested that the

adipose tissue microenvironment is a stronger determinant of differentiation capacity than cell intrinsic factors[87].

In context with our knowledge of our knockout mouse's whole-body metabolic phenotype, a block in differentiation potential seems to fit with the rapid decrease in adipose tissue health seen after short exposure to a high fat diet. One of the earliest responses of the adipose tissue to increased metabolic load and high fat diet stress is to promote the differentiation of new adipocytes. Any impairment in this ability would have significant deleterious effects on the underlying SVF and thus contribute to the development of associated metabolic complications[88].

It is important to note that the presence of increased numbers of precursor cells found in KO SVF, on its own, does not necessarily imply a block in differentiation capacity. It is possible that the differentiation of APCs is not a constant, continuous process but occurs in a more cyclical nature, and we captured a specific snapshot in time. Alternatively, this increase in APCs could represent a greater amount of proliferation in the KO SVF of our Epi WAT. Generally speaking, increased early proliferation is associated with protection from adipose tissue dysfunction. Given that the GPS2-AKO mice display worse measures of metabolic health, this increase in proliferation appears inconsistent with their whole-body phenotype.

Generally speaking, the question of how to address compositional changes from scRNAseq data is still widely debated and the methods used here

are either descriptive or newly developed[89]. To address this issue and complement our current studies, we have begun to perform flow cytometry experiments to physically assess the presence and change in proportion of various cell populations from our KO and WT SVF.

Looking specifically at the subcutaneous SVF, we see more subtle changes that we hope to investigate further. Although we do not appear to see any differences between the proportions of differentiating APCs, this result was expected as there are well known depot-specific differences in the response of various fat tissues to high fat diet. The subcutaneous depot accommodates excess lipids primarily by undergoing hypertrophy whereas the epi WAT expands through both cellular hypertrophy and hyperplasia. Scherer et al[90] developed a genetic mouse model with inducible and permanent labeling of mature adipocytes that they termed the AdipoChaser mouse. They investigated the early dynamic changes found in both major adipose tissue depots after a high fat diet and observed that adipogenesis was seen in the Epi WAT after only 4 weeks of high fat diet, while the subcutaneous depot was only shown to undergo hyperplasia. The timing of our study aligns well with the study of this model and suggests that the subcutaneous depot should be largely exhibiting hyperplasia. We are currently preparing histology of the actual adipose tissues at this time point and will be able to assess any morphological changes in the tissue.

We also note the presence of several cell clusters that demonstrate an early response to HFD stress that we do not observe in the Epi WAT. Given the

status of the SC adipose tissue as a “metabolic sink” in response to energy surplus, we find several cell populations that appear to be altered more strongly in the SC versus Epi WAT in terms of response to the short high fat diet stress. We see an ASC population that appears to show upregulation of markers known to be induced under high fat diet stress. We also see a macrophage cluster strongly expressing inflammatory markers as well as ECM and remodeling-related genes. We do not observe any clusters in the Epi WAT SVF that display similar expression profiles to either of these two populations. These results seem to be similar to what we observed in our RNAseq experiments on isolated adipocytes from our GPS2-AKO mice after 16 weeks of HFD, where the subcutaneous depot in the KO appears to be more sensitive to metabolic stress and displays higher inflammation.

We observe a few potential depot-specific differences between our two experiments. Previous reports indicated that mesothelial cells are exclusively found in the visceral adipose tissue depots. Indeed, we detect 54 cells that are clearly defined by known mesothelial markers only in the epididymal depot. Similarly, we observe a population of 140 cells displaying known markers for ILC2 cells also exclusively in the epididymal depot. In the subcutaneous depot, we observe a population of macrophages expressing collagen and ECM related markers as well as a group of smooth muscle cells that have no known correlate with cell populations found in the epididymal depot.

Finally, we have planned a series of experiments that will address whether or not there is some block in differentiation of APCs being mediated by signaling from the adipocytes. We will perform *in vitro* differentiation experiments on KO and WT SVF treated with either conditioned media or exosomes derived from WT or KO isolated adipocytes. Previously, we had performed *in vitro* differentiation experiments with isolated SVF and found no defect in differentiation capacity. However, the SVF was isolated, plated and cultured removed from any potential signaling from the adipocytes. We believe our proposed experiments here will better recapitulate the local microenvironment of the native adipose tissue as well as reveal any potential paracrine effects from the adipocytes. If we observe a difference in the ability of our KO adipocytes to impair the differentiation of the SVF, we plan to perform mass spectrometry experiments to identify potential secreted factors that may be mediating this change. In order to more robustly assess if cell populations are changed in our WT and KO adipose tissue, we have also begun preliminary validation of a flow cytometry panel to capture and more accurately quantify our identified cell populations.

We also believe that potential changes in the secretome of adipocytes upon GPS2 knockout may explain the whole-body phenotype we observe. It has been well documented that secreted factors from the adipocytes can mediate changes in systemic insulin sensitivity and inflammation. An early defect in differentiation capacity in the adipose tissue could feedback to generate more

fibrosis and dysfunction in the SVF and further shift the secretome of the adipocytes to a pro-inflammatory, pro-diabetic profile. In the previous chapter, we highlighted potential candidates for differentially secreted factors found in our KO mice based on our RNAseq and ChIPseq experiments.

Open questions and Future Directions

We have profiled the SVF of the Epi WAT and SC WAT of our GPS2-KO mice after a brief exposure to a high fat diet. Previously, we had profiled the isolated adipocytes of our GPS2-AKO mice after a long term 16-week high fat diet. In concordance with published work with a different GPS2-AKO mouse model[29], we can see that GPS2-deficient adipocytes are highly inflamed relative to their WT counterparts. This is in contrast to what we observe under a chow diet where our GPS2-AKO mice are obese and yet protected from local and systemic metabolic complications related to HFD.

In total, we have examined the SVF of both depots from WT and GPS2-AKO mice under high fat diet for 5 weeks. Our previous work generated RNAseq data from the isolated adipocytes of these same depots after 16 weeks of HFD. To match our early timepoint, we would like to profile the isolated adipocytes at 5 weeks to understand what is happening directly in the adipocytes where GPS2 is deleted. Although our past experiments were performed using bulk RNAseq, we would like to begin incorporating single nuclei RNA sequencing of isolated adipocytes to better characterize the heterogeneity of these cells.

As mentioned, we are actively performing flow cytometry experiments to more robustly identify cell populations changed in proportion between our WT and KO. Once we have these results, we will more closely study any cell populations confirmed to be changed between WT and KO SVF. Given that we will have sorted the cells, we will be able to potentially look at differences in gene expression that may be a result of altered paracrine signaling in the SVF due to loss of GPS2 in the adipocytes.

In order to fully understand this switch, we have also planned on performing the same set of experiments in our adipose tissue depots under chow diet. Profiling the underlying cell environment at this time point may reveal potential cell populations responsible for maintaining metabolic health despite the condition of obesity. We would also like to see if our GPS2-AKO mice exposed to a chow diet for a longer period of time eventually show similar signs of inflammation and IR as they do after even short periods of HFD stress. Together with our characterization of our mice under HFD stress, expanding our profiling of the GPS2-AKO mice under chow will allow us to better understand what drives the switch between phenotypes as well as what changes may mediate the protective phenotype observed under chow.

We have taken preliminary steps to observe how paracrine effects from our GPS2-AKO adipocytes may influence the local tissue environment. However, adipokines are known to have more than just paracrine effects and many studies have demonstrated their systemic effects on whole body metabolism and

crosstalk between other organ systems of the body such as skeletal muscle or the pancreas. Once we have performed our preliminary experiments looking at potential secreted factors using mass spectrometry, we will validate whether any of our candidates are found in circulation. Other studies of the GPS2-AKO mice have suggested that long range signaling from the adipocytes may influence pancreatic beta cell function[91]. Given this precedent, we hope to more thoroughly investigate the potential systemic effects of GPS2-deficiency in the adipose tissue.

Figures and Figure Legends

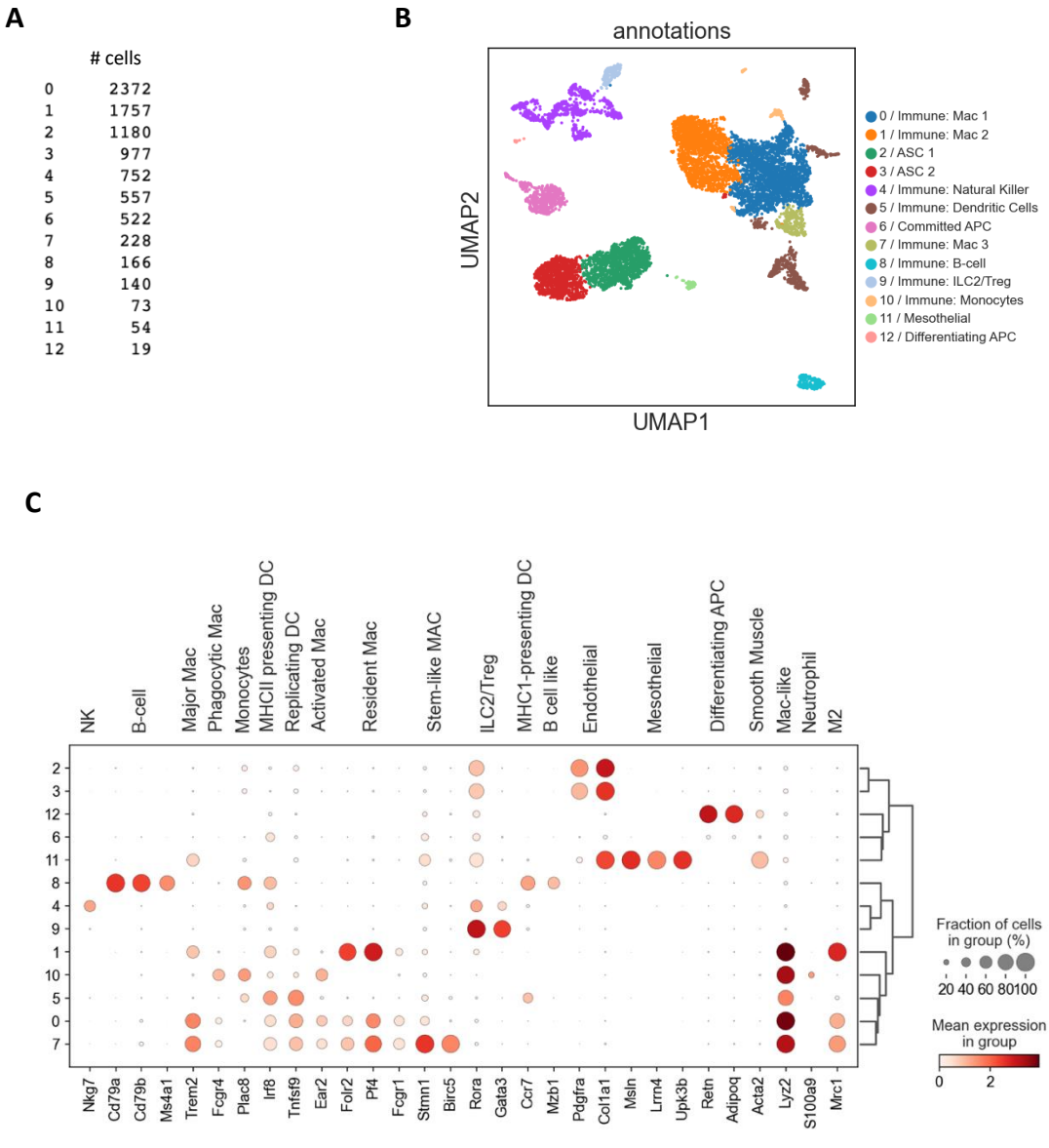


Figure 4.
scRNAseq of the epididymal WAT detects 13 distinct clusters

(a) Total number of cells in each cluster in descending order
(b) UMAP visualization of leiden clustering with fully annotated clusters
(c) Dotplot of gene expression across clusters for curated marker genes from the literature

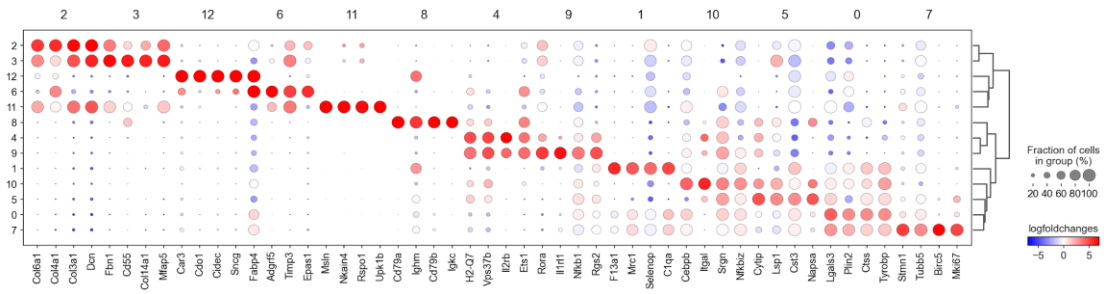


Figure 5.
Top marker genes that defined each cluster in the epididymal white adipose tissue

Dotplot displaying log fold changes top 4 marker genes ranked by p-value used for cluster generation and annotation. Marker genes are determined by performing differential expression of highly variable genes from one cluster vs. all the other clusters using a Wilcoxon rank sum test.

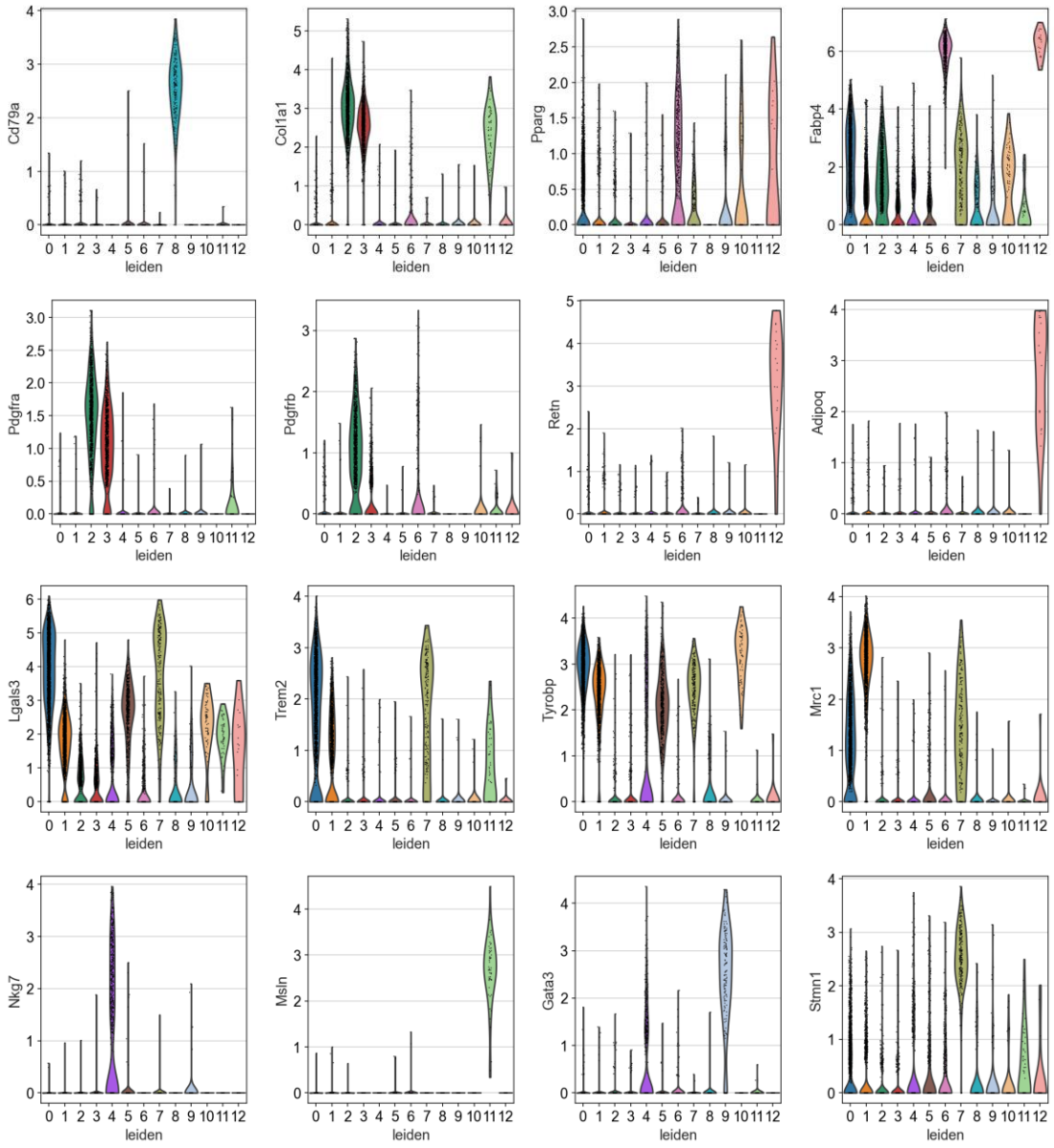
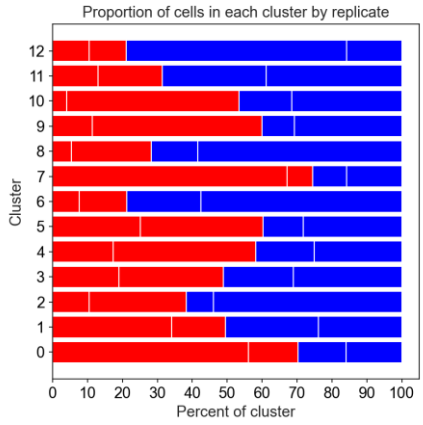


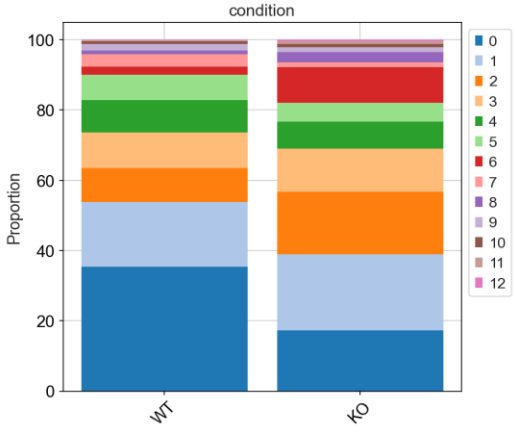
Figure 6.
Visualization of marker genes used for identification of cell populations reveals their specificity and expression across cells from the epididymal white adipose tissue depot

Violin plots for top marker genes used to annotate clusters. A selection of marker genes used for the identification of macrophage, immune cell, and APC populations.

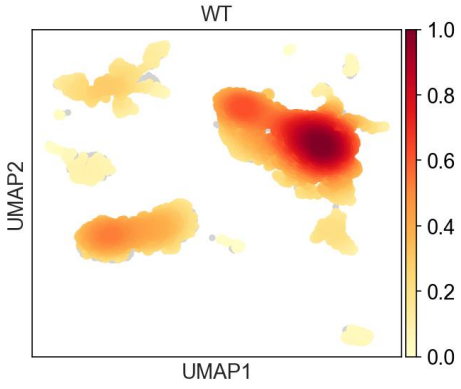
A

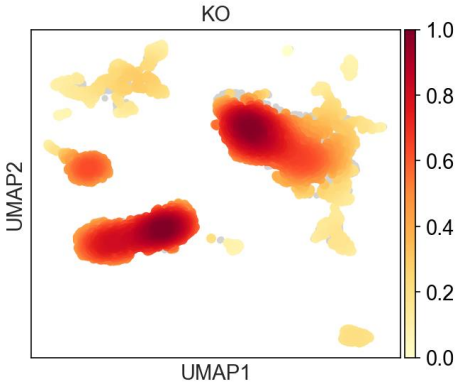


B



C





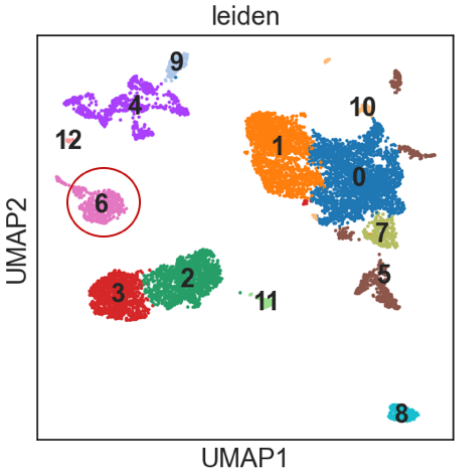
D

Effects:

Covariate	Cell Type	Final Parameter	Expected Sample	log2-fold change
condition[T.WT]	0	0.000000	434.570758	0.043883
	1	0.000000	549.142945	0.043883
	2	0.000000	265.963420	0.043883
	3	0.000000	278.205147	0.043883
	4	0.000000	198.018167	0.043883
	5	0.000000	147.430781	0.043883
	6	-0.592578	83.996998	-0.811026
	7	0.000000	55.832648	0.043883
	8	0.000000	52.214424	0.043883
	9	0.000000	45.302344	0.043883
	10	0.000000	29.322607	0.043883
	11	0.000000	26.638533	0.043883
	12	0.000000	17.696357	0.043883
	13	0.000000	14.914871	0.043883

Figure 7.
Different methods demonstrate potential changes in proportion of APC and macrophage populations between WT and KO samples in epididymal SVF

(a) Proportion of each cluster represented by replicate
(b) Stacked barplot of proportions of each cluster
(c) Heatmaps of cell density split between WT and KO
(d) SCCoda analysis of potential proportional changes between WT and KO



	p_val	avg_logFC	pct.1	pct.2	p_val_adj
Rpl29	6.939244e-09	-0.471029	0.669	0.775	0.000106
Retnla	2.671455e-07	-0.769037	0.073	0.243	0.004067
Hspb1	2.263603e-06	0.355243	0.961	0.928	0.034459
Cish	1.802556e-05	0.313110	0.229	0.054	0.274403
Eif4a2	2.862572e-05	0.313814	0.796	0.676	0.435769
...	...	...	...	...	...
Phlda1	6.941915e-01	-0.274981	0.443	0.423	1.000000
Gm13889	7.081295e-01	-0.251182	0.141	0.153	1.000000
Pgf	7.124586e-01	-0.279975	0.141	0.153	1.000000
Cpe	8.296987e-01	-0.251875	0.144	0.144	1.000000
Loxl2	8.501074e-01	-0.263722	0.105	0.108	1.000000

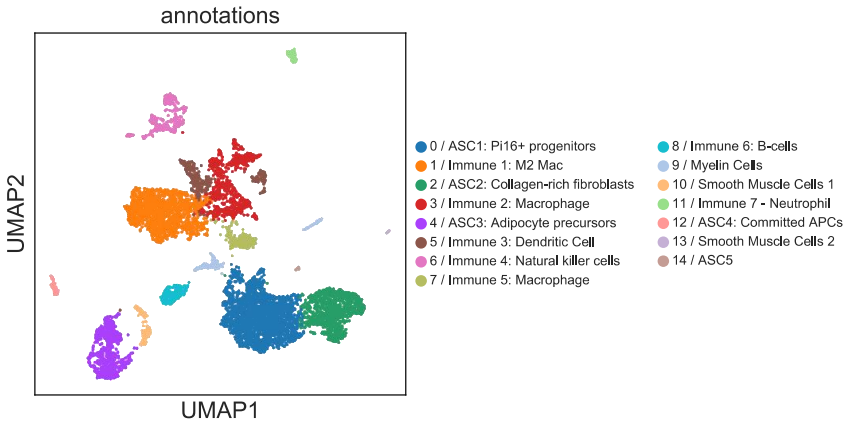
Figure 8.
Differential expression between WT and KO cells in cluster 6 (Committed APCs) reveals no significant changes in gene expression

Differential expression was performed using the MAST methodology developed specifically for DEG analysis in single cell RNAseq

A

	# cells
0	1955
1	1239
2	864
3	644
4	593
5	308
6	277
7	244
8	166
9	153
10	128
11	60
12	53
13	12
14	12

B



C

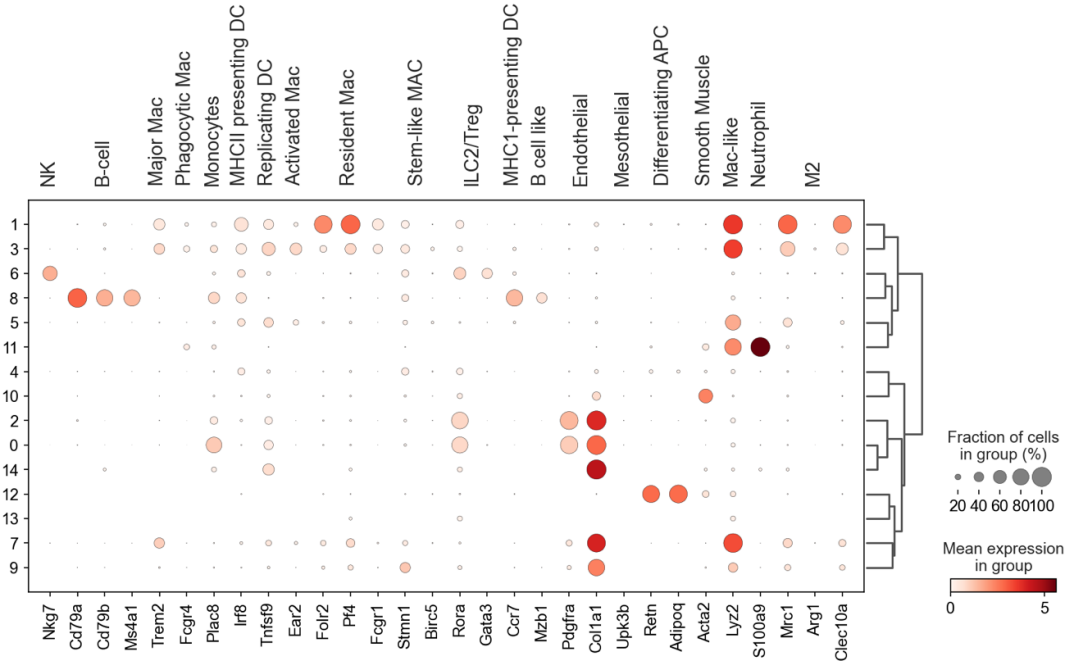


Figure 9.
scRNAseq of the subcutaneous WAT detects 15 distinct clusters

(a) Total number of cells in each cluster in descending order
(b) UMAP visualization of leiden clustering with fully annotated clusters
(c) Dotplot of gene expression across clusters for curated marker genes from the literature

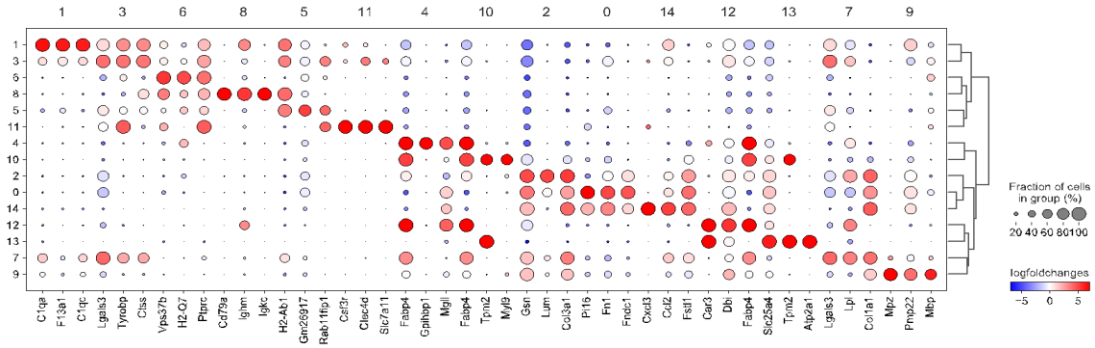


Figure 10.
Top marker genes that defined each cluster in the subcutaneous white adipose tissue

Dotplot displaying log fold changes top 4 marker genes ranked by p-value used for cluster generation and annotation. Marker genes are determined by performing differential expression of highly variable genes from one cluster vs. all the other clusters using a Wilcoxon rank sum test.

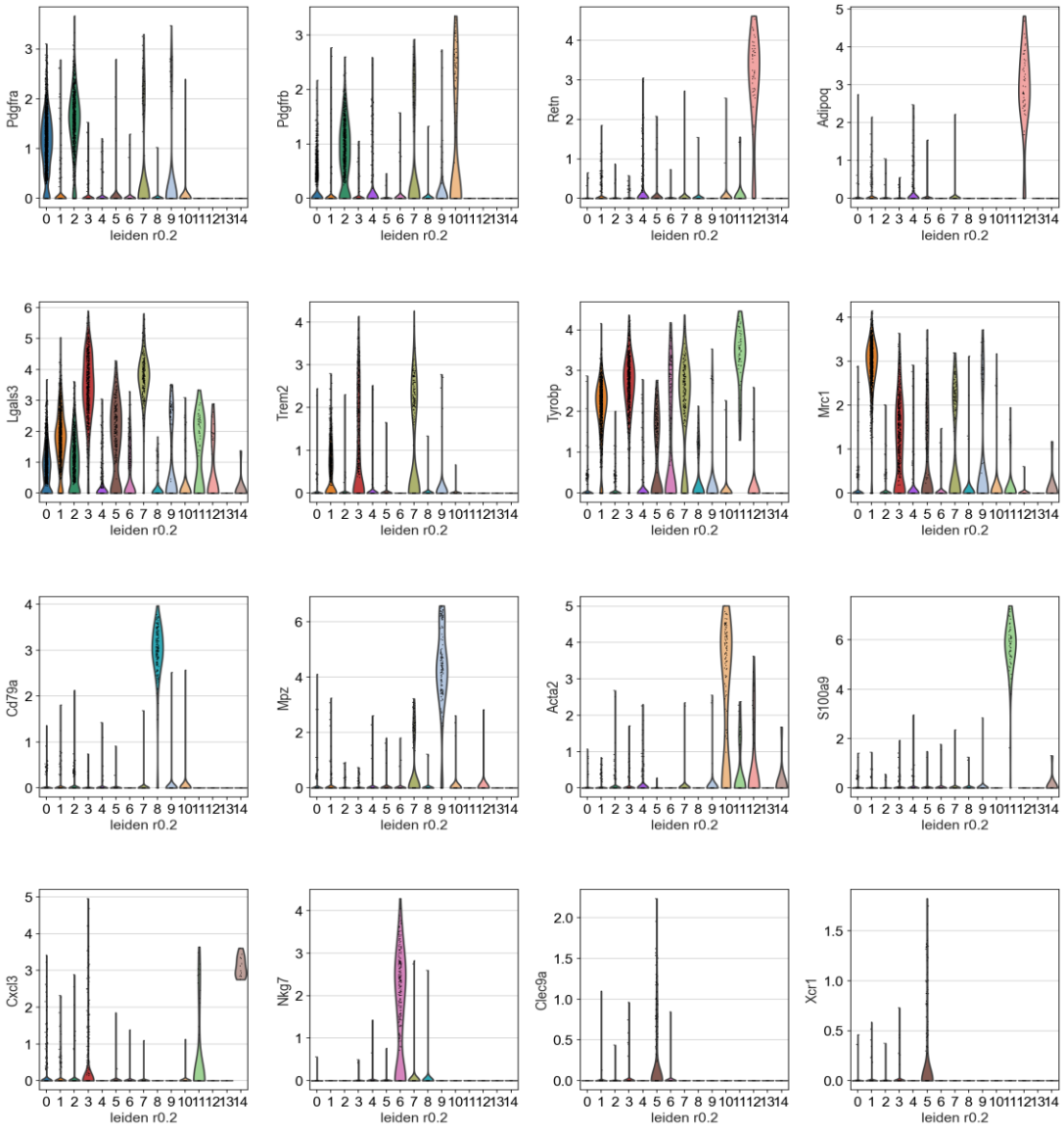
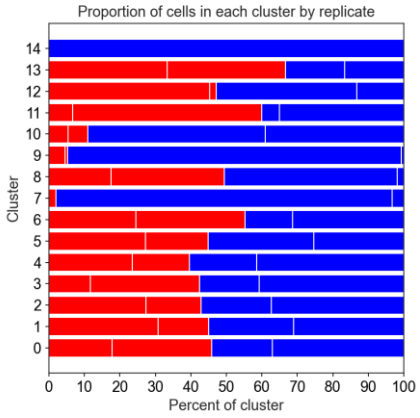


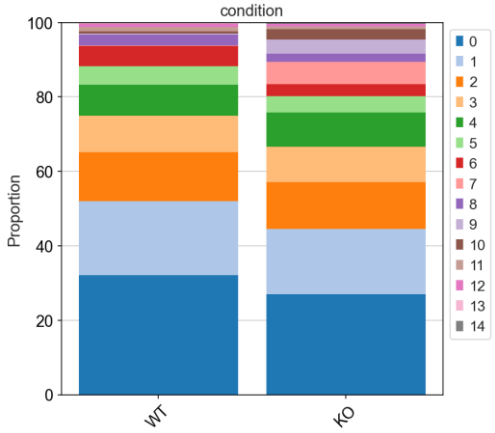
Figure 11.
Visualization of marker genes used for identification of cell populations reveals their specificity and expression across cells from the subcutaneous white adipose tissue depot

Violin plots for top marker genes used to annotate clusters. A selection of marker genes used for the identification of macrophage, immune cell, and APC populations.

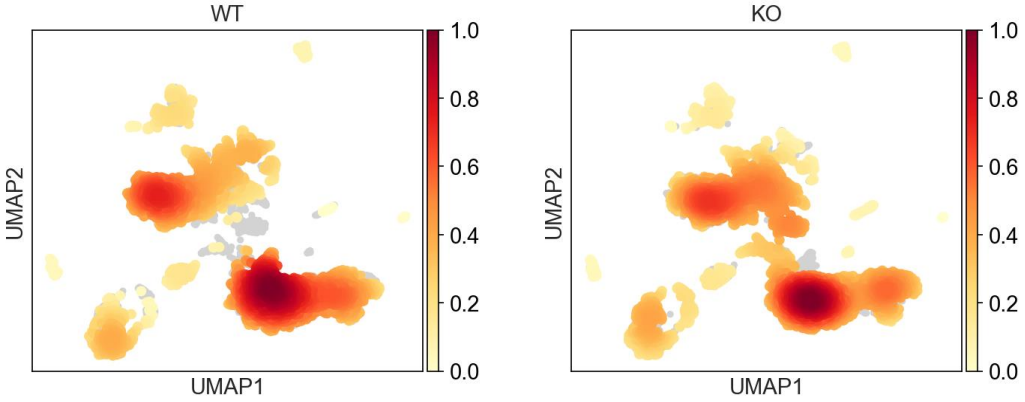
A



B



C



D

Effects:

Covariate	Cell Type	Final Parameter	Expected Sample	log2-fold change
condition[T.WT]	0	0.0	476.095403	0.0
	1	0.0	315.645030	0.0
	2	0.0	222.431293	0.0
	3	0.0	161.034441	0.0
	4	0.0	155.651302	0.0
	5	0.0	89.802989	0.0
	6	0.0	77.759308	0.0
	7	0.0	24.061625	0.0
	8	0.0	38.924282	0.0
	9	0.0	19.660632	0.0
	10	0.0	36.511171	0.0
	11	0.0	22.256147	0.0
	12	0.0	21.319417	0.0
	13	0.0	14.106254	0.0
	14	0.0	9.240706	0.0

Figure 12.

Different methods demonstrate potential changes in proportion of smooth muscle cells and macrophage populations between WT and KO samples in subcutaneous SVF

(a) Proportion of each cluster represented by replicate
(b) Stacked barplot of proportions of each cluster
(c) Heatmaps of cell density split between WT and KO
(d) SCCoda analysis of potential proportional changes between WT and KO

CHAPTER 3

Role of GPS2 in mediating skeletal muscle cell differentiation *in vitro*

INTRODUCTION

Muscle differentiation: a model for transcriptional regulation

Another well-characterized model that has elucidated the importance of transcriptional regulation during differentiation is skeletal muscle. Muscle is a highly complex tissue that in addition to its most well-known function in enabling locomotion and movement also serves to mediate whole-body metabolism. Fully developed muscle tissue is tightly organized into large, bundled myofibers consisting of multiple myofibrils. These bundles are surrounded and supported by a robust extracellular matrix and muscle tissue is also highly innervated and vascularized[92]. Force generation and subsequent locomotion is achieved through the contraction of the sarcomere, a highly organized network of filaments and motor proteins. To enable contraction and movement, muscle contains a specialized network of mitochondria and metabolic processes that adapt and support the energy demands of the tissue[93].

Development of skeletal muscle is a visually striking process that transforms individual cells resembling fibroblasts into highly striated and multinucleated fibers. It was discovered in the early 1970s that this morphological

process is largely governed by the actions of a single transcription factor named MyoD[94]. This transcription factor, often referred to as the master regulator of muscle differentiation, was shown to be sufficient for proper development of muscle tissue. Surprisingly, it was also shown that forced expression of MyoD in completely distinct cell types such as fibroblasts, liver cells, or fat cells was capable of converting these lineages to skeletal muscle cells[95].

The vast majority of skeletal muscle tissue is established and patterned during embryogenesis. A wide variety of morphogens including Sonic Hedgehog, Wnt, and BMP partition the developing embryo into the various tissue layers that will eventually become mature muscle tissue located throughout the body of the organism[96]. Outside of embryogenesis, post-natal skeletal muscle is repaired and regenerated through the proliferation and differentiation of satellite cells. These cells are muscle-resident stem cells that are essential for muscle regeneration as shown by the impairment of muscle regeneration *in vivo* after their specific ablation[97].

Myogenic Regulatory Factors: MRFs

MyoD was the first myogenic regulatory factor discovered and was termed the master regulator of muscle differentiation largely based on its ability to reprogram non-muscle cells to a muscle lineage. However, MyoD belongs to a larger family of transcription factors termed the Myogenic Regulatory Factors

(MRFs) which also include Myogenin (MyoG), Myf5 (Myogenic Factor 5), and Mrf4 (Myogenic Regulatory Factor 4).

Each of these factors contain a basic helix-loop-helix domain (bHLH), which recognizes the E-Box domain motif (CANNTG), which is located throughout the genome in different promoter and enhancer regions of both muscle-related and non-muscle related specific genes. Prior to binding of the E-Box, the MRFs will either heterodimerize with E-proteins or homodimerize to exert their regulatory effects on gene transcription.

Initially, it was suggested that MyoD and Myf5 are primarily responsible for the commitment and early stages of differentiation while Mrf4 and MyoG direct the terminal differentiation of committed cells. However, later work has emphasized the broad functional overlap between the MRFs throughout all stages of differentiation as well as the existence of an extensive auto- and cross-regulatory network between them[98]. For example, MyoD has been shown to remodel the chromatin environment in the MyoG promoter region prior to direct binding and activation through MyoD[99].

MyoD: Mechanism of Action

Lineage determination of cells is accomplished in part by the remodeling of chromatin to silence genes extraneous to the ultimate function of the fully differentiated cell and activate lineage-specific genes. Initial attempts to uncover the mechanism by which MyoD facilitates skeletal muscle differentiation focused

on its role as a pioneering factor capable of inducing remodeling and transcription in previously inaccessible chromatin regions[100].

Mechanistically, MyoD was shown to directly bind to the histone acetyltransferase (HAT) p300 and p300/CBP-associated factor (PCAF) to form complexes with distinct activity. Both of these coactivators were shown to strongly activate MyoD-mediated transcription with p300 responsible for the acetylation of histones H3 and H4 to recruit pCAF, which directly acetylates MyoD and enhances transactivation[19]. MyoD also directly recruits the SWI/SNF chromatin remodeling complex through interactions with the BAF60c subunit prior to gene activation[101]. MyoD has also been shown to play a role in gene repression by binding through HDAC1 to regulate MyoD-dependent gene transcription[20]

C2C12: *In vitro* model of muscle differentiation

One of the most widely used tools in the study of muscle differentiation was developed in 1977 by Yaffe and Saxel[102]. Their work isolated a cloned population of cells derived from the thigh muscle of adult mice post-injury. A subclone of these original cells became known as the C2C12 line and have been widely used for the *in vitro* study of muscle differentiation. They proliferate rapidly under conditions of high serum and are induced to differentiate into multinucleated myotubes upon serum withdrawal[103]. Importantly, the

transcriptional networks that direct muscle differentiation in C2C12 cells closely resemble what is observed *in vivo*[104,105].

Role of GPS2 in skeletal muscle differentiation

Prior work by Treuter et al[28] demonstrated that c-Jun is largely responsible for recruiting GPS2 to enhancer sites in macrophages. They demonstrated a significant genome wide overlap between GPS2 and c-Jun as well as showed a direct physical binding between the two factors. It has been well established that c-Jun preferentially binds to enhancer regions in C2C12s. It was also shown to physically interact with MyoD and loss of c-Jun at enhancer regions reduced levels of enhancer-specific histone marks such as H3K4me1 and H3K27ac in these regions[106]. Given the known association between GPS2 and c-Jun, we wished to determine if c-Jun was mediating the binding and actions of GPS2 in the context of MyoD-mediated regulation of gene expression during muscle differentiation.

We have separately also demonstrated that GPS2 plays an important role as an activator for the transcription and expression of nuclear encoded mitochondrial genes in a model of fat cell differentiation. Both adipose tissue and skeletal muscle are both metabolic organs that

Given the similarities between the underlying transcriptional processes and the importance of mitochondrial gene expression, we sought to determine if we observed similar mechanisms for the regulation of gene expression through

GPS2 in muscle cells. A preliminary experiment in our laboratory had demonstrated that knockdown of GPS2 via siRNA in undifferentiated cells led to a subsequent impairment of muscle differentiation as measured by immunofluorescence staining for myosin heavy chain (data not shown).

Results

GPS2 is expressed in a model of in vitro muscle differentiation

In order to begin exploring the role of GPS2 in muscle differentiation, we first profiled the expression of GPS2 in the C2C12 cell line across a 6-day time course. Western blot experiments that GPS2 levels increase during differentiation, either through upregulated expression or stabilization (**Figure 13a)**. As a positive control, we profiled the expression of known muscle markers over the time course (**Figure 13a)**. MyoD has been typically shown to be an early to mid-stage regulator of the muscle differentiation process[107]. As expected, in our experiment, MyoD is expressed in undifferentiated myoblasts and the expression is lost by d4 of differentiation. MyHC is commonly used as a marker for distinguishing undifferentiated vs. differentiated cells and we can clearly observe its expression only post induction of differentiation[108]. Based on separate siRNA experiments, we determined that the three bands at or below 50 kDa are specific for GPS2 (**Figure 13b)**. Though we speculate that they potentially represent post-translationally modified forms of GPS2, that discussion is beyond the scope of this work. These dynamic changes in the expression of GPS2

across differentiation suggest a potential role in mediating either the development or the functionality of differentiated myoblasts.

GPS2 binds primarily to non-promoter genomic regions throughout muscle differentiation

Based on our knowledge of GPS2 in other models of cellular differentiation and its role as a transcriptional co-factor, we wanted to investigate the binding of GPS2 on a genome-wide level across the course of differentiation in C2C12s. To address this, we performed a ChIPseq experiment in WT C2C12 cells and profiled GPS2 occupancy at four timepoints: i) undifferentiated myoblasts, ii) 3hrs post-differentiation, and iii) 2 days post differentiation to capture early changes in the transcriptional profile of differentiating cells, and iv) 4 days post differentiation to represent terminally differentiated myotubes. We performed two ChIP experiments for GPS2 at each timepoint and submitted the extracted DNA for library preparation and sequencing. We processed the ChIPseq according to the methods described and combined the replicates at each time point.

We find a total of 30071, 15348, 22630, and 13271 GPS2 peaks from undifferentiated myoblasts, myoblasts differentiated for 3hrs, myoblasts differentiated for 2 days, and myoblasts differentiated for 4 days, respectively (**Figure 14)**. We can see by observing the proportion of peaks annotated to various genomic regions that a high percentage of peaks from each time point are located in intronic or intergenic regions. Overall these distal peaks represent

more than 90% of the peaks present in each time point. Given this binding pattern, we sought to determine if these peaks were located in enhancer regions, which are often in intronic / intergenic regions and regulate expression of distant target genes.

To enable the identification of potential enhancer regions in C2C12s, we leveraged the availability of existing published ChIPseq data stored in the GEO repository for known enhancer marks. We analyzed and processed data for H3K27ac[109], H3K4me3[109], H3K4me1[110], MyoD[110], p300[111], and c-Jun[112] taken at several points throughout the differentiation process. These ChIPseq data are publicly available under the following GEO accession numbers: H3K27ac/H3K4me3 – GSE76010, H3K4me1/MyoD – GSE25549, p300 – GSE25308, c-Jun – GSE109636. H3k27ac and H3k4me1 are known, highly conserved histone marks found enriched in enhancer regions and H3K4me3 is often used as a marker of active transcription. MyoD and p300 are important factors shown to bind to a majority of muscle-specific enhancers in both myoblasts and myotubes.

We generated normalized coverage tracks for each mark and examined the signal distribution surrounding our GPS2 peaks at each time point. **Figure 15** shows a heatmap displaying the signal strength of H3K27ac, H3K4me1, H3K4me3 and MyoD centered around each of the 30,071 peaks found for GPS2 in myoblasts and we used k-means clustering to group the peaks into three main clusters.

In myoblasts, Cluster 1 shows H3K27ac and H3K4me1 enriched at the peak center, and enriched regions of H3K4me3 flanking the peak center. Cluster 2 consists of GPS2 peaks with both strong enrichment for H3K27ac and H3K4me1 at their peak center, and MyoD binding shifted up- or downstream of the peak center and likely represent a subset of classical muscle enhancers. Cluster 3 contains a subset of peaks that display H3K27ac and H3K4me1 at their peak center with the vast majority of peaks in this cluster seemingly binding away from any enriched areas for H3K27ac and H3K4me1.

Previous work demonstrated that c-Jun is the main factor responsible for recruiting GPS2 to AP-1 binding sites to repress inflammatory gene expression in macrophages. Given this data, we generated similar coverage tracks for c-Jun binding in undifferentiated myoblasts and observe a pattern similar to MyoD and p300 binding (**Figure 16)**. Namely, we see a shifted binding pattern where c-Jun binding appears centered at roughly 500bp up- or downstream of GPS2 peaks. In regards to its overlap with MyoD binding, it is known in the literature that MyoD and c-Jun are capable of physically interacting. As well, c-Jun was shown to be enriched at the center of enhancers found in myoblasts and this enrichment was also linked to MyoD binding.

We performed the same analyses for the other timepoints found in our ChIPseq and generated similar heatmaps of signal enrichment for various histone marks (**Figures 17, 18, 19).** We observe the same general pattern of GPS2 peak binding in relation to these histone marks and factors. In the

undifferentiated time points, myoblasts and 3hrs, we generally observe three main classes of peaks: 1. A group of GPS2 peaks with proximal MyoD binding, and enrichment for both H3K27ac and H3K4me1, 2. A group of GPS2 peaks that are enriched for H3K4me3 and H3k4me1, and 3. A group of GPS2 peaks primarily marked by H3K4me1 that also contains a subset of peaks that are not in close proximity to enriched signal for any of these marks.

In differentiated myoblasts, at times d2 and d4, we also see three main clusters of peaks. There is a pronounced cluster of GPS2 peaks that show proximal MyoD binding and very strong enrichment for both H3K27ac and H3K4me1. The second cluster generally comprises GPS2 peaks with more diffuse MyoD binding but also enriched for H3K27ac and H3K4me1. The third cluster is mostly marked by H3K4me1 alone but includes a large subset of peaks that bind distally to regions marked by either H3k27ac or H3k4me1.

To better determine potential regulatory targets, we isolated the peaks from these clusters and performed both motif analysis and functional enrichment. We focused on peaks found in myoblasts and d2 of differentiation to combine with RNAseq data we generated at these same time points. The GREAT tool for ChIPseq incorporates distal binding sites and a genomic region-based enrichment test to better annotate cis-regulatory elements. In myoblasts for cluster 1, using the GREAT algorithm, we can see that the peaks are enriched for genes related to RNA processing, ncRNA processing, ribosome biogenesis and tRNA processing (**Figure 20a)**. Motif enrichment reveals no significantly enriched

motifs found in the binding locations (**Figure 20b)**. It is interesting to note that this first cluster contains 263 out of the 521 GPS2 peaks annotated to promoter-TSS regions suggesting that GPS2 is binding to active promoters or in proximity to genes actively transcribed.

The same analyses performed for cluster 2 show a functional enrichment for genes related to focal adhesions, extracellular matrix organization, and regulation of cell shape (**Figure 21a)**. These terms correspond to genes responsible for the patterning of cells into myotubes during the differentiation process. We find a number of enriched motifs corresponding to AP-1 complex binding factors including c-Jun, Fra, and BATF (**Figure 21b)**. We can see that this cluster of peaks also shows p300 binding shifted up- and downstream in a similar pattern as seen for MyoD binding and confirming their status as enhancers.

Interestingly, we also observe a large proportion of peaks that are primarily enriched for H3K4me1 in cluster 3 and we speculate that these peaks may represent a subset of enhancers which are “poised” and become active later in the differentiation time course. Supporting this idea, the genes marked strongly by GPS2, H3K27ac, H3K4me1, and MyoD (cluster 1) at d2 are largely found to be marked by H3K4me1 in myoblasts. Out of the 1226 peaks (K27ac+, K4me1+, MyoD+) at d2, we find 784 are marked by GPS2 peaks enriched near regions of H3k4me1 and 643 that are in close proximity to enriched regions for H3K27ac, H3K4me1, and MyoD in myoblasts.

We performed the same set of analyses for the peaks found at each timepoint. Specifically focusing on day 2, we can see a similar pattern of GPS2 binding in relation to known enhancer marks. We again performed k-means clustering to generate three main groups. We can see GPS2 peaks that show a strong enrichment for H3K27ac and H3K4me1 at their peak center with MyoD binding shifted ~500bp up- and downstream. The second cluster of peaks show strong enrichment for H3K27ac and H3K4me1 at their peak center but display more diffuse MyoD binding. The third cluster largely represents peaks that are primarily marked by H3K4me1.

Functional enrichment for the first cluster of peaks at d2 returns classical muscle differentiation terms such as myofibril assembly, muscle cell development, and actomyosin structure (**Figure 22a)**. We speculate given the strong pattern of MyoD binding and enrichment for enhancer marks that these genes represent a subset of the classical program of muscle genes that become active in the course of differentiation. Motif finding for these peaks demonstrate no significantly enriched motifs at our GPS2 peak locations (**Figure 22b)**.

The second cluster of peaks at d2 is also enriched for certain muscle differentiation terms such as sarcomere organization, and actomyosin structure, but also contains more general terms related to processes such as apoptosis signaling, cell cycle, and regulation of gene expression (**Figure 23a)**. Motif finding on the peaks in this cluster reveal an enrichment for AP-1 complex factors (**Figure 23b).**

RNAseq data demonstrates GPS2 is important for muscle-specific differentiation programs

Our binding data from the ChIPseq experiments suggest that GPS2 binds to a subset of MyoD-bound enhancers throughout differentiation. These peaks were shown to be annotated and functionally enriched for critical muscle development terms and processes such as myofibril assembly, cell cycle and actomyosin organization. To determine which genes are directly and indirectly regulated by GPS2, we performed RNAseq analysis on WT and GPS2-KO C2C12 cells. We chose two time points, myoblasts and 2 days post differentiation, to match to the corresponding time points from our ChIPseq data. We primarily focused on three main sets of differentially expressed analyses: differentially expressed genes at each time point between conditions, genes that changed in a condition-specific manner over time, and genes that naturally become differentially expressed over the course of differentiation.

First, we performed differential expression on our WT samples over time. In agreement with established literature, we observe an upregulation of canonical muscle differentiation terms and a robust downregulation of cell cycle terms in differentiated cells compared with undifferentiated (data not shown). Next, we performed differential expression testing between undifferentiated WT and KO cells to determine what genes are changed before the start of differentiation due

to loss of GPS2. We find 1717 DEGs at a FDR cutoff of < .1 with 984 upregulated and 733 downregulated in our KO C2C12s. Amongst the upregulated genes, we observe enrichment for pathways related to extracellular matrix organization, and PI3K-AKT signaling. The downregulated genes show enrichment for axon guidance and semaphorin-related pathways (**Figure 24a)**. Performing the same analysis between the WT and KO C2C212s at d2 reveals 2635 DEGs. The 1422 upregulated genes at d2 show enrichment for PI3-AKT signaling, adipogenesis, and transcription-related pathways while the 1213 downregulated genes are enriched primarily for cell cycle, and cell cycle regulation terms (**Figure 24b)**.

In addition to examining the transcriptional profile at each time point between WT and KO cells, we also wanted to determine if there were any genes changed in a condition-specific manner over time. Given our knowledge of GPS2 and its role in chromatin remodeling, we were particularly interested in genes that showed differential expression over time between WT and KO. To test this, we performed a LRT test between the full DESeq2 model and a reduced model excluding time as a variable. This analysis resulted in 882 genes identified that display a condition-specific effect over time. Functional pathway enrichment of these genes reveals a strong association with a number of different cell cycle and cell cycle checkpoint terms. In order to visualize the differing expression profiles of these genes over time, we converted the normalized counts for each sample into z-scores, grouped the genes using hierarchical clustering, and plotted them on a heatmap. We can see the profile of the normalized counts for each

differentially expressed gene across time and condition visualized by heatmap (**Figure 25)**. The columns of the heatmap represent every replicate sample taken from WT, KO1, KO2 from myoblasts to day 2. The first eight columns demonstrate how the expression of these genes change over time in the wild-type C2C12s with the first four representing samples from myoblasts and the last four representing samples from d2. This is repeated for each knockout line to demonstrate how the expression of these genes display a condition-specific effect over time in our GPS2-KO lines.

Overlap between ChIPseq and RNAseq identifies enhancers and their regulated genes at each time point

To better determine direct targets of GPS2 regulation, we sought to overlap the peaks we clustered into groups based on their pattern of histone mark enrichment with our differentially expressed genes discovered through RNAseq. For both time points, myoblasts and d2, we used the DEG generated by comparing between WT and KO within that timepoint. For example, the DEG in myoblasts are genes that are differentially expressed between WT and KO C2C12s even before differentiation due to loss of GPS2. Similarly, the DEG list at d2 represents the differences between WT and KO cells both differentiated for two days in parallel.

In myoblasts, we find a total of 1171 peaks that are annotated to genes that are differentially expressed between WT and GPS2-KO C2C12s (**Figure**

26a). We then performed functional enrichment using GREAT on the peaks for each cluster. Cluster 1 returns no significantly enriched terms. Cluster 2 is enriched for terms related to cell motility, cell migration, and cellular movement while cluster 3 displays terms related to muscle fiber development, action potentials, and angiogenesis (**Figure 26b)**.

Repeating the same analysis for the GPS2 peaks found at d2 reveals 1528 peaks annotated to DEG between WT and KO (**Figure 27a)**. Cluster 1 shows no enrichment terms while Cluster 2 is heavily enriched for classical muscle differentiation terms such as actin organization, striated muscle tissue development, and muscle morphogenesis. Cluster 3 displays enrichment of mixed terms relating to neuronal guidance, synapse, and response to norepinephrine/catecholamines (**Figure 27b)**.

We also overlapped the peaks found in myoblasts with genes that were determined to be differentially expressed over time in a condition-specific manner. We postulated that these genes were candidates for loci where remodeling of the chromatin environment in a GPS2-dependent manner was necessary for their proper activation or repression through the course of differentiation. We found that 576 out of the 882 differentially expressed genes across time also display an annotated GPS2 binding event in myoblasts (**Figure 28a)**. We performed functional enrichment for the peak annotations of these overlaps and found that cluster 2 was highly enriched for classical muscle development terms. This cluster of peaks display canonical enhancer signatures

and suggests that the binding of GPS2 in myoblasts to these regions is important for the regulation of genes critical for the core muscle differentiation process over time (**Figure 28b)**. Overlapping peaks found in cluster 3 are enriched for cell migration, cell locomotion and cell movement pathways. This cluster of peaks was speculated to represent poised enhancers and the loss of GPS2 appears to disrupt the proper expression of genes related to early differentiation processes like cell migration and cell movement (**Figure 28c)**.

Loss of GPS2 via siRNA knockdown or CRISPR-Cas9 mediated KO leads to impaired differentiation

Prior work in the laboratory demonstrated that C2C12 differentiation was impaired after siRNA mediated knockdown of GPS2. In order to further validate the importance of GPS2 in muscle differentiation, we developed two pooled CRISPR-Cas9 mediated C2C12 cell lines with genomic deletions for GPS2. We employed two separate strategies, one utilizing two guides targeting both exon 1 and exon 6 (KO1) and another that incorporated a single guide targeting a region in exon 6 of GPS2 (KO2). We confirmed the deletions via deep sequencing of the targeted regions, which confirmed that around 30% of amplicons found in KO1 corresponded to a frameshift mutation resulting in a stop codon in exon 6 and 40% of the remaining amplicons resulting in a mutation that alters a stretch of 4 amino acids at the C-terminal end of GPS2. We speculate that this change results in a conformational change that destabilizes GPS2. For KO1, we observe

around 70% of contigs that correspond to a frameshift mutation resulting in a premature stop codon in exon 1.

In order to investigate to which extent the loss of GPS2 leads to an impairment in muscle differentiation, we conducted a 6-day time course comparing WT and KO lines. Immunofluorescence for GPS2 and Myosin Heavy Chain (MyHC) in differentiated myotubes at 6 days revealed significant impairment in differentiation as seen by the reduced staining for MyHC in both the KO lines. Due to the pooled nature of our cell lines, we expect a certain proportion of cells to remain wild-type due to either successful repair of the genomic editing or the presence of potential synonymous mutations or cases where the editing did not occur in the first place. Interestingly, the myotubes that successfully differentiated appear to be GPS2-positive in both KO cell lines. However, these few GPS2-positive myotubes are surrounded by a large number of undifferentiated cells that lack staining for GPS2. These results clearly demonstrate that loss of GPS2 impairs differentiation and formation of myotubes (**Figures 29, 30)**.

Discussion

In this chapter, we have examined the role of GPS2 in muscle differentiation through ChIPseq and RNAseq with the goal of identifying GPS2-mediated regulation of the chromatin environment and gene expression. We first profiled the genomic occupancy of GPS2 throughout the differentiation time

course. Through simple annotation of our discovered peaks, we observed a high proportion of peaks (> 90% in each timepoint) annotated to distal locations within introns or intergenic regions. This pattern of binding suggested to us that GPS2 may be binding to and influencing enhancer regions, which are known to regulate target gene expression through long range interactions. To directly answer this question, we processed publicly available data for known histone marks and regulatory factors important for muscle-specific enhancers. We generated normalized coverage tracks for each of these marks and visualized their signal centered around our GPS2 peaks via heatmaps. Throughout all the time points, we can see a clear subset of our GPS2 peaks binding to regions strongly enriched for H3K27ac and H3K4me1. This suggests that GPS2 is binding to and mediating the activity of enhancers throughout the differentiation process.

Given our previous knowledge of the role of GPS2 in regulating mitochondrial genes in other models, we wished to investigate whether we see a similar mechanism in C2C12s. We downloaded the MitoCarta3.0 database and overlapped known mitochondrial genes with our annotated list of peaks. Out of the 1140 identified mitochondrial genes, we see 461 and 367 GPS2 peaks in myoblasts and d2 that are annotated to these same genes. Interestingly, only a small fraction of these bound genes appear to be differentially expressed in our GPS2-KO C2C12 lines at these time points. However, it has been observed that during muscle differentiation there is a massive upregulation of mitochondrial encoded protein expression in absence of markedly increased mtDNA copy

number or mitochondrial gene transcription[113]. This suggests that post-transcriptional mechanisms may be more important for the regulation of mitochondrial genes in this setting. Importantly, we have recently demonstrated that GPS2 plays a key role in regulating mitochondrial-associated protein translation through K63 ubiquitination. We speculate that this non-transcriptional mechanism may be the major means of GPS2-mediated mitochondrial gene regulation in this setting.

Focusing on our profiling of GPS2 binding and expression data in myoblasts, we can see three main groups of regulatory regions marked by GPS2 binding in relation to the patterning of known histone marks. The first cluster is marked primarily by H3K27ac, H3k4me1 enriched at the peak center flanked by enrichment for H3k4me3. H3K4me3 is typically seen marking the promoter region and transcription start sites at a genome-wide level. Interestingly, if we subset these peaks to the promoter peaks, we do not see any functional or pathway enrichment for the associated annotated genes suggesting they regulate a diverse set of processes.

The second cluster of GPS2 peaks appear to be classical enhancers marked by H3K27ac, H3K4me1, MyoD, p300 and c-Jun and linked to many genes related to ECM, cell adhesion, and cell shape. At these peak locations, we can clearly see GPS2 enriched for H3K27ac and H3K4me1 but shifted from the center of binding for the factors MyoD, p300 and c-Jun. These peaks appear to represent enhancers that support the regulation of genes important for the core

differentiation process as they are enriched for terms such as myofibril assembly, and actomyosin organization.

The last cluster of peaks appear to contain a large number of poised enhancer sites as evidenced by their enrichment for H3K4me1. Functional enrichment of these peaks reveals no significant pathways or terms. When we subset these peaks by whether their expression is changed in our GPS2-KO C2C12s, we observe enrichment for various pathways related to muscle differentiation including action potentials, cell migration, and muscle development. This suggests that loss of GPS2 at these potential poised enhancers is sufficient to alter their regulation even in the undifferentiated state.

To determine whether these GPS2 peaks could be linked to the regulation of the expression of their annotated genes, we overlapped our ChIPseq data with our RNAseq data. Our first set of analyses were focused on looking at changes in gene expression that occurred between WT and KO cells at a particular time point. For instance, our overlap in myoblasts was looking at genes where loss of GPS2 prior to differentiation already led to changes in gene expression. We can see that peaks in cluster 2 of myoblasts appear to be direct targets of GPS2 and regulate genes responsible for cell motility and migration. Peaks in cluster 3 are also generally enriched for muscle contraction, synapse organization and muscle development. Dysregulation of these classes of genes prior to the induction of differentiation may be the main contributor to our phenotype of impaired myotube formation upon loss of GPS2.

At d2, we find no functional enrichment for overlapping peaks found in our ChIPseq and RNAseq in cluster 1 that are strongly marked by MyoD, K27ac, and K4me1. For cluster 2, which is marked by K27ac, K4me1 and more diffuse MyoD binding, we observe enrichment for core muscle development terms like muscle development and muscle tissue morphogenesis. This suggests that GPS2 is still bound to enhancer regions linked to active muscle gene programs at d2 as these genes are also differentially expressed between WT and GPS2-KO C2C12s at this same time.

Intriguingly, the enrichment of overlapping peaks and genes for cluster 3 at d2 show enrichment for synapse assembly, response to norepinephrine and neuronal projection. As these peaks show enrichment for H3K4me1 with diffuse binding for both MyoD and H3K27ac, these may represent a subset of enhancers that regulate later stage differentiation programs. Mature muscle fibers are highly innervated and norepinephrine is a major signaling molecule controlling force contraction. Fitting with the poised, but not active state of these enhancers, we observe only minimal differentiation and formation of myotubes at d2. We would expect these gene programs to be more fully active in cells differentiated for a longer period of time.

Open questions and Future Directions

Our current studies have focused on examining the function of GPS2 at enhancer locations during muscle differentiation. While we have incorporated a

number of different identifying marks for enhancers such as p300, H3K27ac, H3k4me1, we would like to more precisely characterize these long-range interactions using both eRNA analysis as well as 4C circularized chromosome confirmation capture. Recent literature has begun to elucidate the critical role of eRNA in nearly all aspects of enhancer function including modifying chromatin accessibility, recruitment of transcriptional core complexes, and promoter-enhancer looping. Although we performed RNAseq experiments, we were primarily interested in expression profiles during differentiation and steady-state RNA assays like bulk RNAseq are not optimized for analyzing eRNAs. Future experiments would incorporate GROseq experiments to capture and characterize nascent eRNA transcription. We are also planning to better characterize potential long-range enhancer interactions using 4C. Our current peak annotation strategy relies upon using nearest feature annotation, which lacks the resolution and specificity of chromosome confirmation techniques. 4C experiments would allow for more precise mapping of looped, interacting loci and may help explain the shifted pattern of binding we observe for GPS2 and other muscle regulatory factors. 4C would enable us to determine if our GPS2 peaks are physically interacting on a 3D scale with those loci located in close proximity and marked by MyoD, p300 and c-Jun.

Although we have overlapped our data with a number of known histone marks and performed preliminary ChIP-qPCR experiments, we would like to choose candidate enhancer regions to delete using CRISPR-Cas9 and assay the

effects on target gene expression and the surrounding local chromatin environment. This would serve as additional evidence as to their specificity of gene regulation as well as further our understanding of how enhancers impact the surrounding chromatin environment. A related study of GPS2 and enhancers performed similar deletion experiments and observed alterations in the chromatin environment at multiple GPS2-bound enhancers in close proximity upon loss of one.

Although we have begun to test candidate regions using ChIP-qPCR in our C2C12 GPS2-KO lines. We would also like to expand the range of factors tested to better determine the mechanism by which GPS2 binding is influencing the chromatin environment. We have performed initial ChIP-qPCR experiments using our GPS2-KO C2C12 lines and early results suggest that loss of GPS2 at these locations leads to reduced MyoD binding and reduced H3K27ac. We need to perform further experiments to obtain statistical significance. Given these findings, we would like to also perform ChIP experiments for H3K27me3 and H3K9me3 at these same regions. We have seen previously that GPS2 inhibits the activity of known lysine demethylase KDM4A and these additional ChIPs would begin to answer whether GPS2 is regulating methylation status at these peaks. Although the previous mechanism was largely focused on H3K9me3 at the promoter level, we would like to see if we see similar regulation of methylation by GPS2 at our enhancer regions. Although GPS2 was previously shown to regulate KDM4A, other lysine demethylases are known to be more

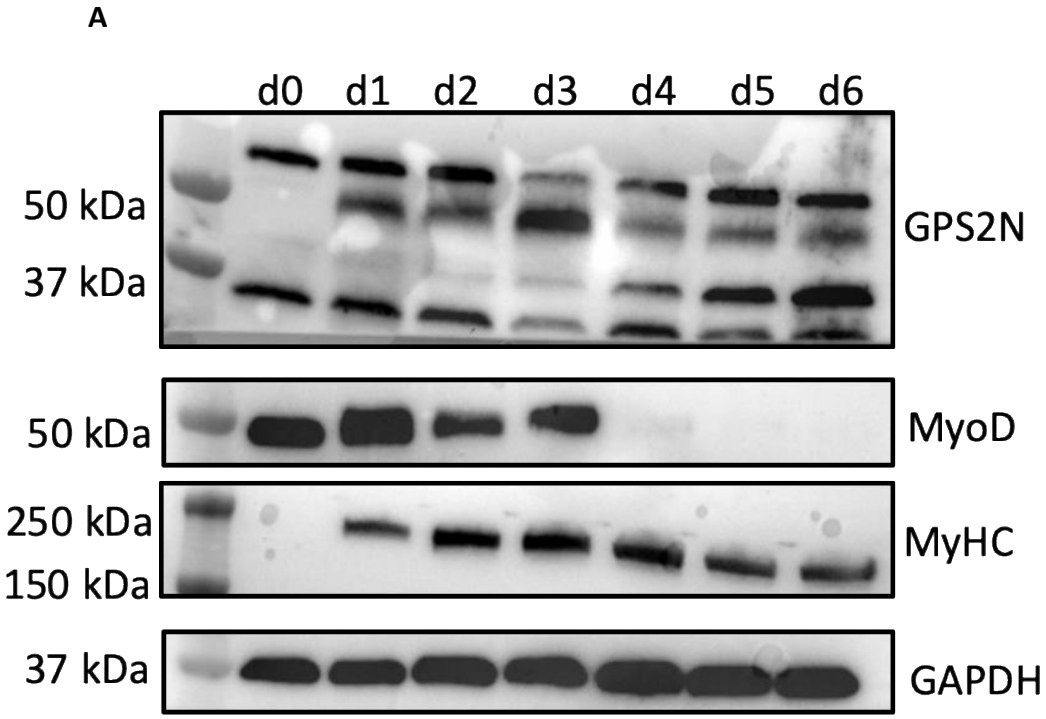
active during muscle differentiation. We will also assay using ChIP-qPCR the binding of the more active lysine demethylases at known GPS2 binding sites and determine if loss of GPS2 affects their recruitment / presence.

One important caveat of our work is that the histone mark, ATACseq, and MyoD binding data are all taken from different published sources. Given the dynamic nature of the chromatin environment and transcription factor binding, it would be more appropriate to have been able to perform these ChIPseq experiments ourselves concurrent with our own GPS2 ChIPseq. However, the C2C12 cell line and the transcriptional cascade occurring differentiation has been well-characterized and shown to be consistent across numerous different studies. As well, we have attempted to include multiple replicates for each factor where possible.

We have so far restricted our studies to the C2C12 model cell line. This was done to leverage the wealth of publicly available genomic data. However, we would like to extend these studies by exploring the effects of GPS2 loss in muscle tissue in an *in vivo* setting. We would generate a GPS2-Acta1 Cre genetic mouse model by crossing our GPS2-Floxed mice with the B6.Cg-Tg(ACTA1-cre)79Jme/J line from Jackson labs. We would be particularly interested in examining skeletal muscle regeneration after injury in WT and GPS2-KO mice. This genetic model would also enable us to dissect differences in the role of GPS2 in different muscle tissues with different fiber types. Given our prior knowledge of the role of GPS2 in mitochondrial function, we would be

interested to look at specific differences in different fiber types, which are known to differ greatly in their metabolism and mitochondrial makeup.

Figures and figure legends



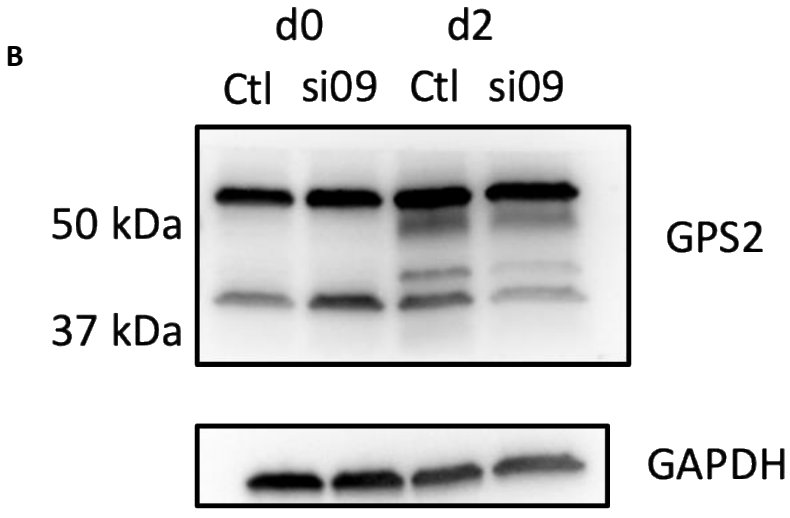


Figure 13. GPS2 is expressed in C2C12s and becomes stabilized or upregulated over the differentiation time course

(a) Western blot of GPS2 protein expression profiled every day from d0 to d6 of differentiation. Characteristic expression of MyoD (early to mid-stage regulator) and MyHC (marker of terminal differentiation) is also observed.

(b) Western blot of GPS2 siRNA treated C2C12s demonstrates specific bands for GPS2

	Myoblast	3hrs	d2	d4
Promoter	521	364	355	315
Intronic	13404	6667	9868	5770
Intergenic	15104	7819	11682	6723
TTS	360	160	259	163
Exon	320	165	203	144
5’	20	12	13	15
3’	260	133	193	112
Non-Coding	82	28	57	29
Total	30,071	15,348	22,630	13,271

Figure 14. A majority of GPS2 peaks are annotated to intronic or intergenic regions throughout all 4 time points

Greater than 90% of annotated peaks are shown to be located in distal regions.

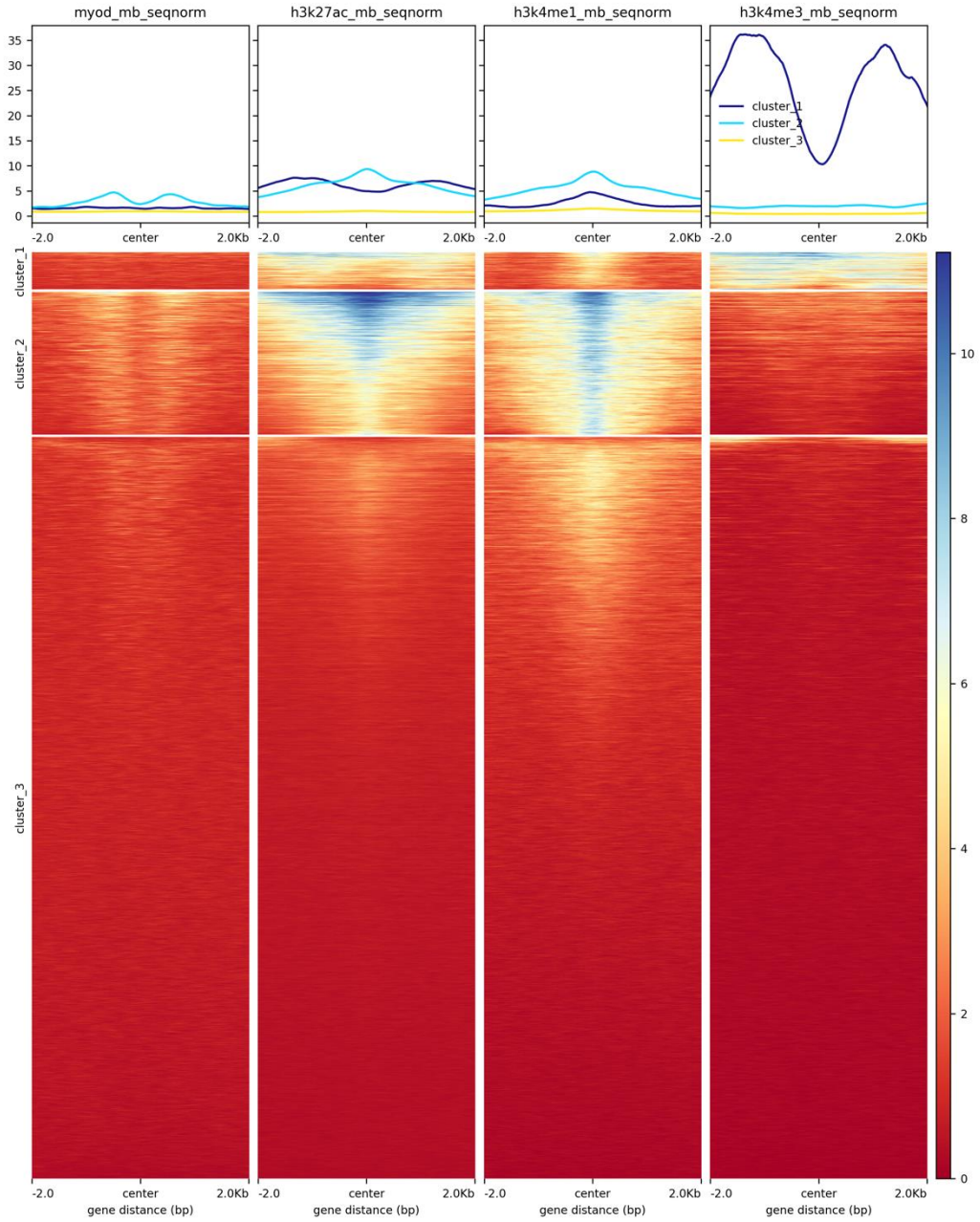


Figure 15. Coverage tracks and signal enrichment for MyoD, H3K27ac, H3K4me1, and H3K4me3 relative to GPS2 peaks found in myoblasts reveals binding in classical enhancer regions

Coverage tracks for histone marks relative to GPS2 peak centers. Heatmap represents normalized signal enrichment after scores were calculated per genomic bin (10bps)

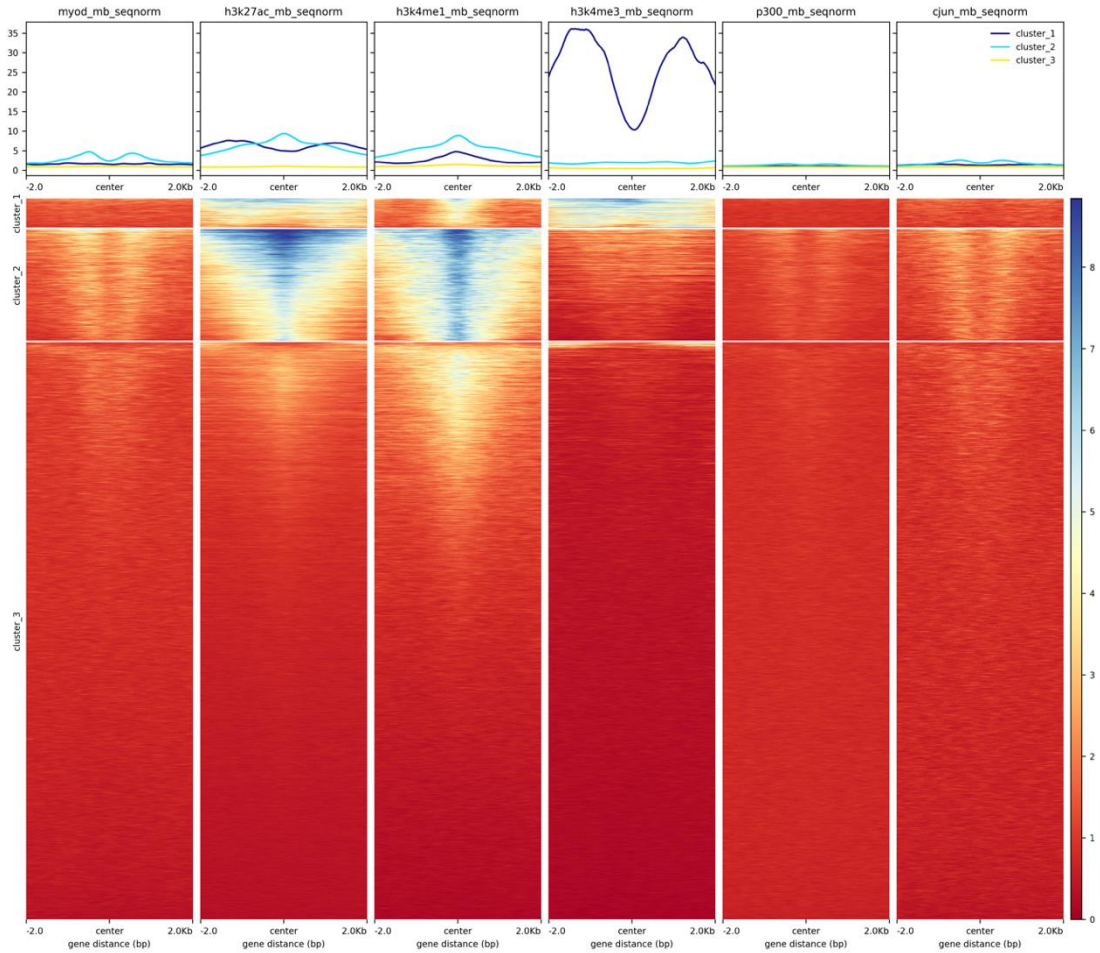


Figure 16. Coverage tracks and signal enrichment for MyoD, H3K27ac, H3K4me1, H3K4me3, p300 and c-Jun relative to GPS2 peaks found in myoblasts reveals binding in classical enhancer regions

Binding pattern of c-Jun and p300 confirm the enhancer status of the second cluster of GPS2 peaks. Both factors display a similar shifted binding pattern from the center of GPS2 peaks and overlap with regions of known MyoD binding.

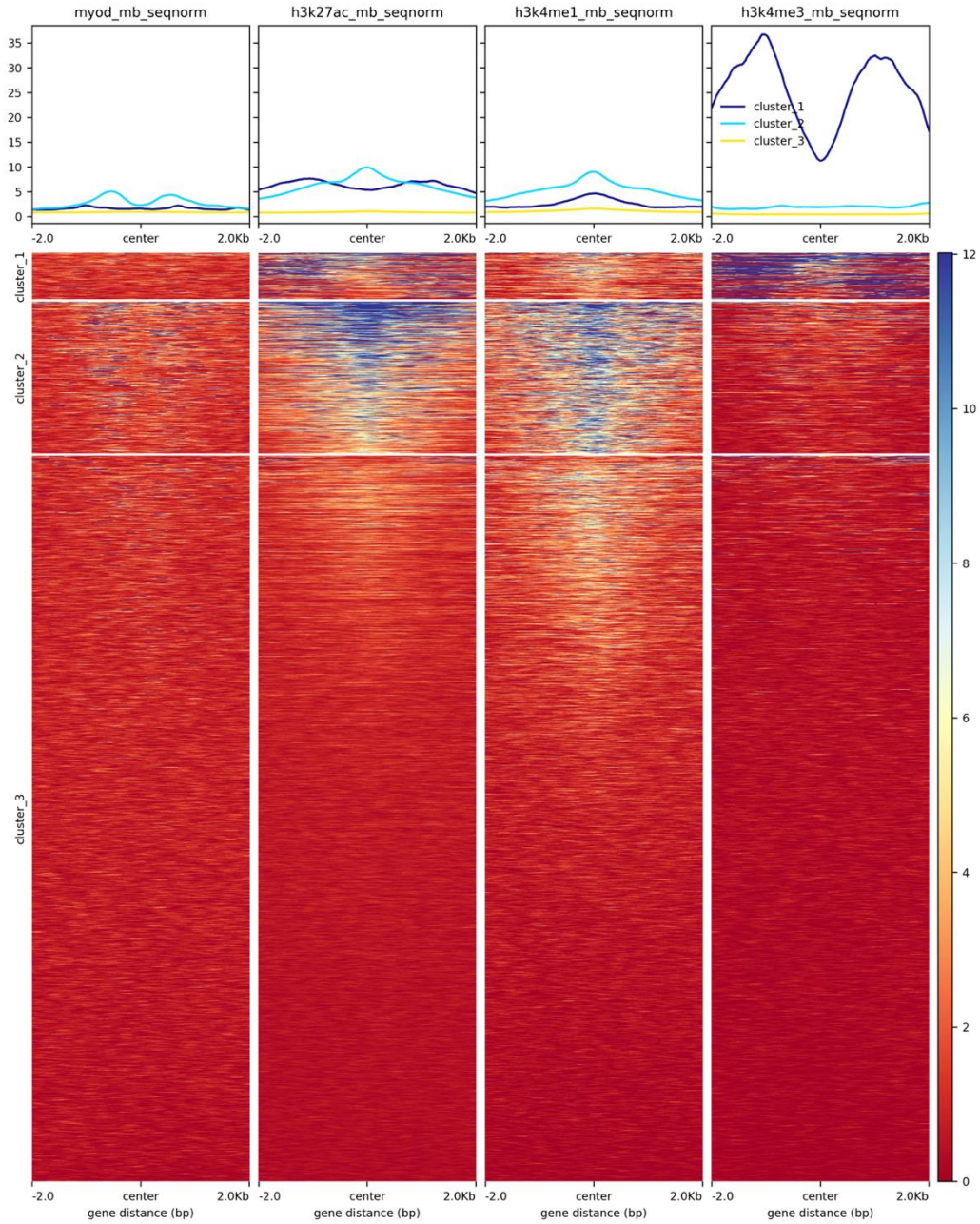


Figure 17. Coverage tracks and signal enrichment for MyoD, H3K27ac, H3K4me1, and H3K4me3 relative to GPS2 peaks found at 3hr post differentiation

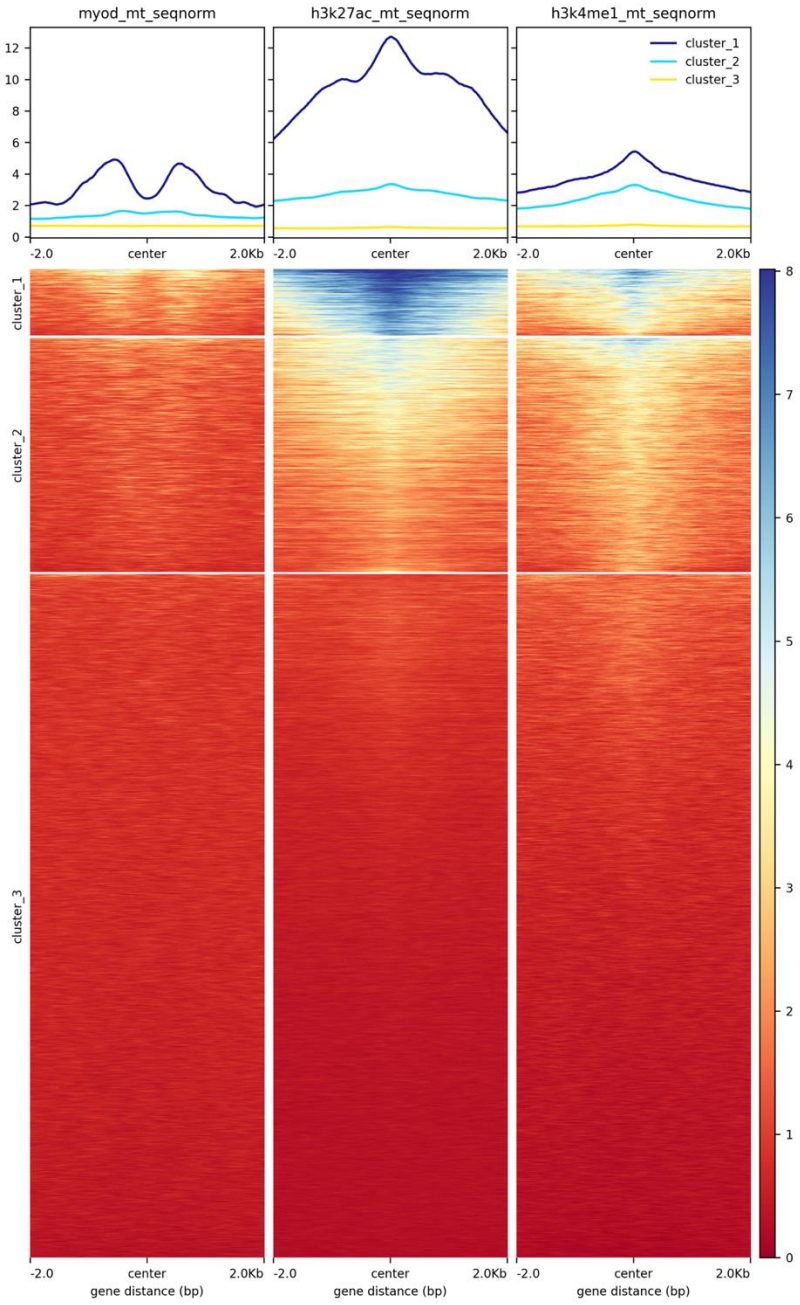


Figure 18. Coverage tracks and signal enrichment for MyoD, H3K27ac, H3K4me1 relative to GPS2 peaks found at 2d post differentiation

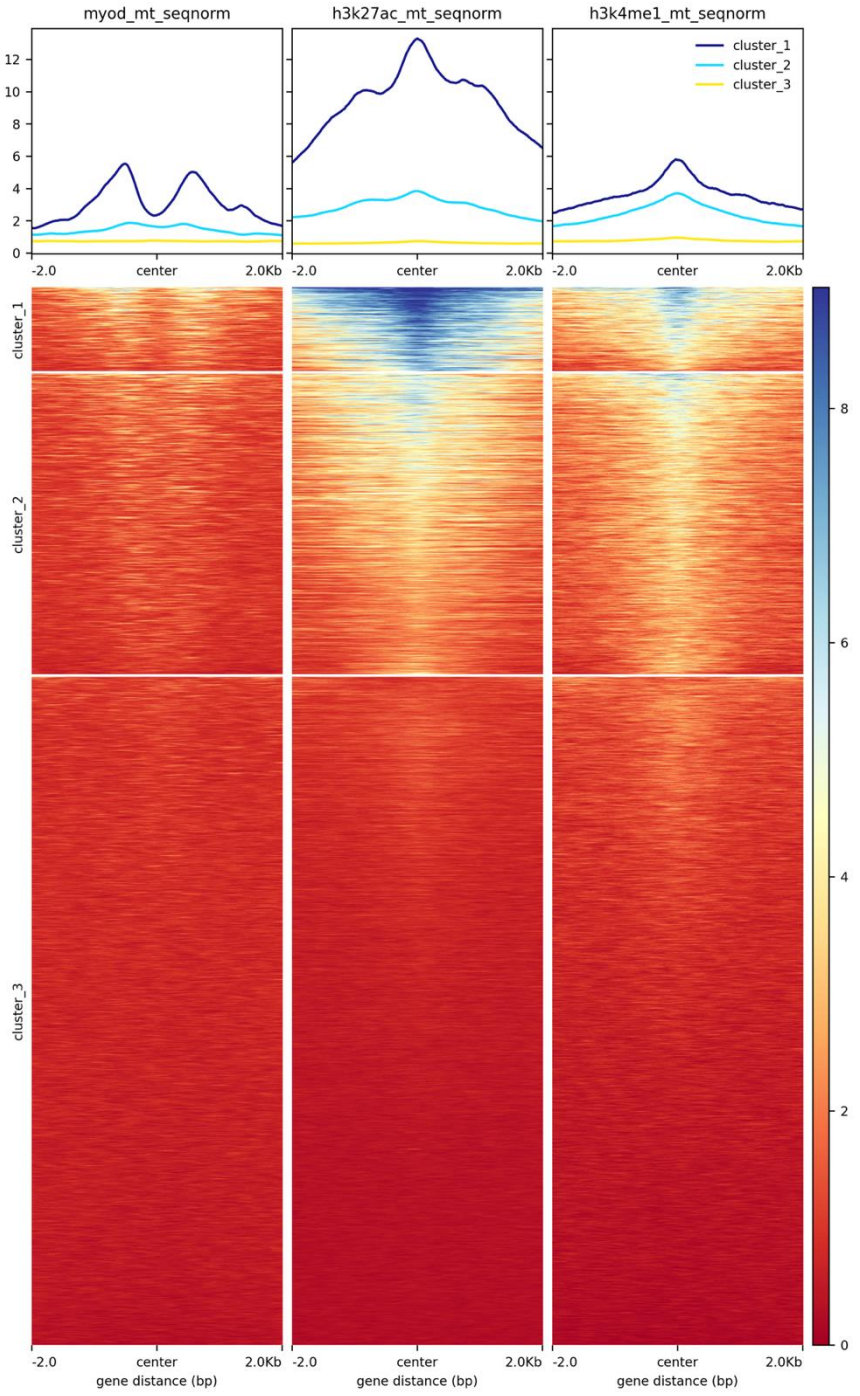
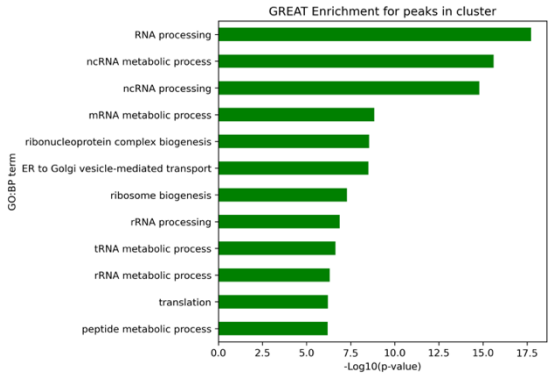
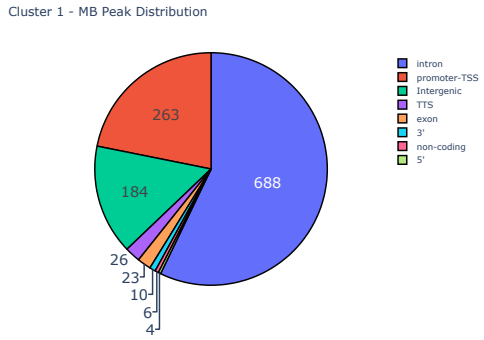


Figure 19. Coverage tracks and signal enrichment for MyoD, H3K27ac, H3K4me1 relative to GPS2 peaks found at 4d post differentiation

A



B



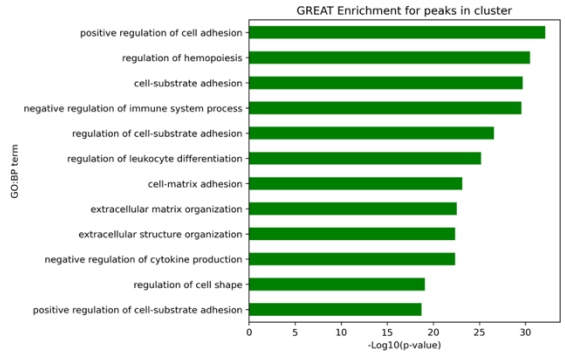
C

Rank	Motif	Name	P-value	log P-pvalue	q-value (Benjamini)	# Target Sequences with Motif	% of Targets Sequences with Motif
1		GFY-Staf(?,Zf)/Promoter /Homer	1e-2	-6.302e+00	0.7585	8.0	0.78%
2		Ronin(THAP)/ES-Thap11-ChIP-Seq(GSE51522)/Homer	1e-2	-5.995e+00	0.7585	6.0	0.58%
3		THRa(NR)/C17.2-THRa-ChIP-Seq(GSE38347)/Homer	1e-2	-5.647e+00	0.7585	41.0	3.98%
4		Brn2(POU,Homeobox)/NPC-Brn2-ChIP-Seq(GSE35496) /Homer	1e-2	-5.552e+00	0.7585	18.0	1.75%

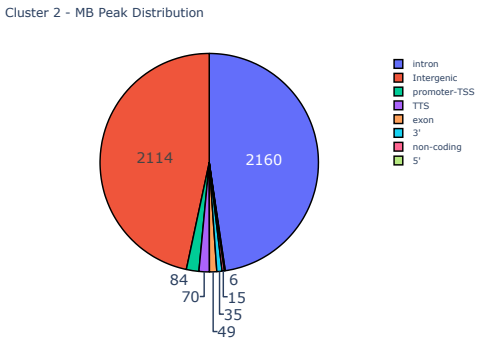
Figure 20. Functional enrichment for peaks found in heatmap cluster 1 in myoblasts reveals enrichment for various RNA processing terms and no significant motif enrichment

(a) GREAT enrichment for ChIPseq peaks found in heatmap cluster 1
(b) Proportion of peaks annotated to various genomic regions
(c) Motif enrichment using HOMER using center of peaks

A



B



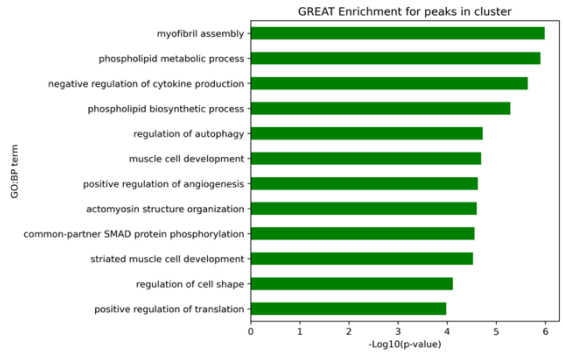
C

Rank	Motif	Name	P-value	log P-pvalue	q-value (Benjamini)	# Target Sequences with Motif	% of Targets Sequences with Motif
1		BATF(bZIP)/Th17-BATF-ChIP-Seq(GSE39756)/Homer	1e-7	-1.836e+01	0.0000	293.0	7.60%
2		AP-1(bZIP)/ThioMac-PU.1-ChIP-Seq(GSE21512)/Homer	1e-7	-1.783e+01	0.0000	330.0	8.56%
3		Fra2(bZIP)/Striatum-Fra2-ChIP-Seq(GSE43429)/Homer	1e-7	-1.767e+01	0.0000	211.0	5.48%
4		Atf3(bZIP)/GBM-ATF3-ChIP-Seq(GSE33912)/Homer	1e-7	-1.761e+01	0.0000	290.0	7.53%
5		RUNX(Runt)/HPC7-Runx1-ChIP-Seq(GSE22178)/Homer	1e-7	-1.643e+01	0.0000	330.0	8.56%
6		JunB(bZIP)/DendriticCells-Junb-ChIP-Seq(GSE36099)/Homer	1e-7	-1.634e+01	0.0000	244.0	6.33%
7		Fra1(bZIP)/BT549-Fra1-ChIP-Seq(GSE46166)/Homer	1e-7	-1.621e+01	0.0000	242.0	6.28%
8		Fosl2(bZIP)/3T3L1-Fosl2-ChIP-Seq(GSE56872)/Homer	1e-5	-1.244e+01	0.0002	144.0	3.74%

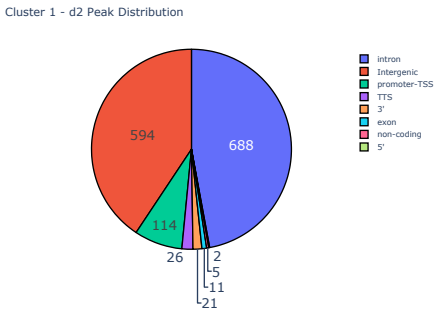
Figure 21. Functional enrichment for peaks found in heatmap cluster 2 in myoblasts reveals enrichment for terms related to ECM, cell shape and cell adhesion and motif finding shows a variety of AP-1 complex motifs enriched at peak centers

(a) GREAT enrichment for ChIPseq peaks found in heatmap cluster 2
(b) Proportion of peaks annotated to various genomic regions
(c) Motif enrichment using HOMER using center of peaks

A



B



C

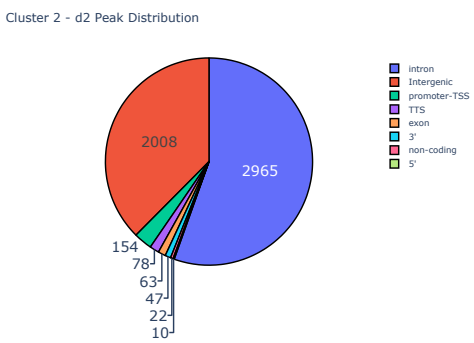
Rank	Motif	Name	P-value	log P-pvalue	q-value (Benjamini)	# Target Sequences with Motif
1		WT1(Zf)/Kidney-WT1-ChIP-Seq(GSE90016)/Homer	1e-2	-6.426e+00	0.6702	68.0
2		PRDM15(Zf)/ESC-Prdm15-ChIP-Seq(GSE73694)/Homer	1e-2	-5.499e+00	0.8467	138.0
3		Foxo1(Forkhead)/RAW-Foxo1-ChIP-Seq(Fan_et_al.)/Homer	1e-2	-5.162e+00	0.8467	264.0
4		GFY(?)/Promoter/Homer	1e-2	-4.985e+00	0.8467	12.0
5		c-Myc(bHLH)/mES-cMyc-ChIP-Seq(GSE11431)/Homer	1e-2	-4.784e+00	0.8467	54.0

Figure 22. Functional enrichment for peaks found in heatmap cluster 1 at d2 reveals enrichment for terms related to myofibril assembly, muscle development, actomyosin organization and motif finding finds no enriched motifs

(a) GREAT enrichment for ChIPseq peaks found in heatmap cluster 2
(b) Proportion of peaks annotated to various genomic regions
(c) Motif enrichment using HOMER using center of peaks

A

B



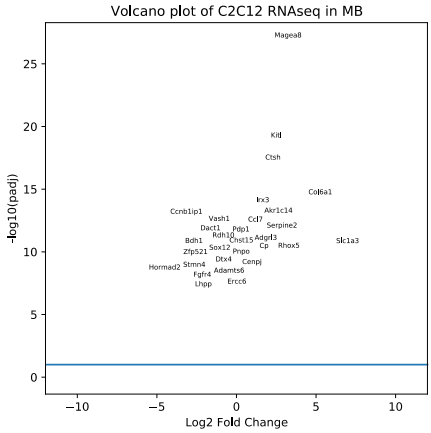
C

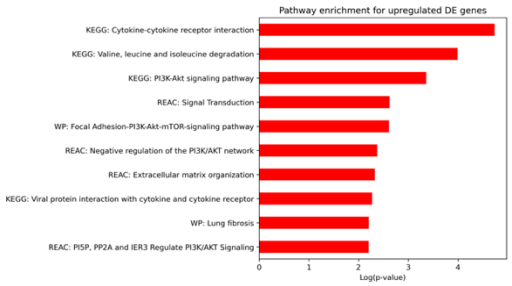
Rank	Motif	Name	P-value	log P-pvalue	q-value (Benjamini)	# Target Sequences with Motif	% of Targets Sequences with Motif
1		Fra1(bZIP)/BT549-Fra1-ChIP-Seq(GSE46166)/Homer	1e-10	-2.339e+01	0.0000	292.0	6.48%
2		BATF(bZIP)/Th17-BATF-ChIP-Seq(GSE39756)/Homer	1e-9	-2.271e+01	0.0000	342.0	7.59%
3		Atf3(bZIP)/GBM-ATF3-ChIP-Seq(GSE33912)/Homer	1e-9	-2.203e+01	0.0000	338.0	7.50%
4		AP-1(bZIP)/ThioMac-PU.1-ChIP-Seq(GSE21512)/Homer	1e-8	-1.993e+01	0.0000	372.0	8.25%
5		Mef2c(MADS)/GM12878-Mef2c-ChIP-Seq(GSE32465)/Homer	1e-7	-1.644e+01	0.0000	261.0	5.79%
6		Mef2b(MADS)/HEK293-Mef2b.V5-ChIP-Seq(GSE67450)/Homer	1e-6	-1.471e+01	0.0000	445.0	9.87%
7		Fosl2(bZIP)/3T3L1-Fosl2-ChIP-Seq(GSE56872)/Homer	1e-6	-1.455e+01	0.0000	159.0	3.53%
8		Fra2(bZIP)/Striatum-Fra2-ChIP-Seq(GSE43429)/Homer	1e-6	-1.435e+01	0.0000	227.0	5.04%

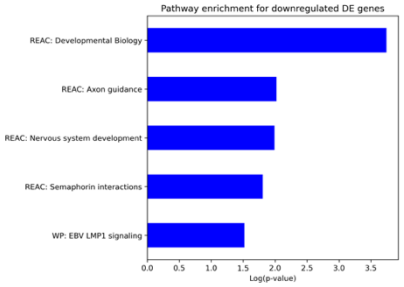
Figure 23. Functional enrichment for peaks found in heatmap cluster 2 at d2 reveals enrichment for terms related to cell cycle, actomyosin structure, sarcomere organization and motif finding shows a variety of AP-1 complex motifs enriched at peak centers

(a) GREAT enrichment for ChIPseq peaks found in heatmap cluster 2
(b) Proportion of peaks annotated to various genomic regions
(c) Motif enrichment at center of GPS2 peaks using HOMER

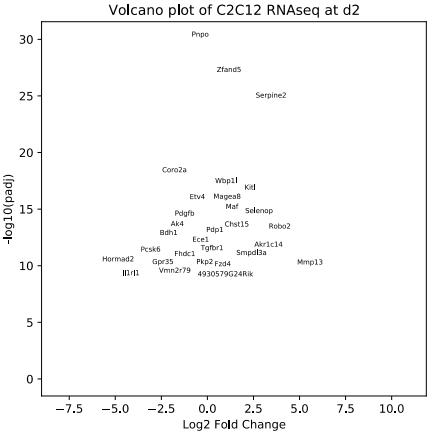
A

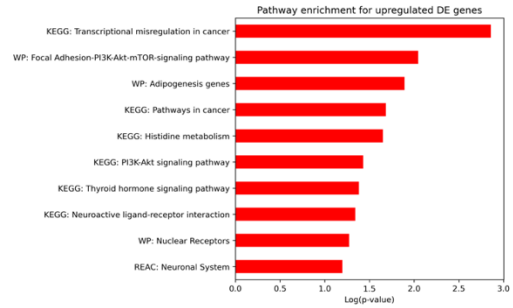

Volcano plot of C2C12 RNAseq in MB
-log10(padj)
Log2 Fold Change
Magea8
Kitl
Ctsh
Col6a1
Ccnb1ip1
Vash1
Ccl7
Akr1c14
Serpine2
Dact1
Pdp1
Rdh10
Chst15
Adgrl3
Bdh1
Sox12
Cp
Rhox5
Slc1a3
Zfp521
Pnpo
Dtx4
Stmn4
Cenpj
Hormad2
Adamts6
Fgfr4
Lhpp
Ercc6


Pathway enrichment for upregulated DE genes
KEGG: Cytokine-cytokine receptor interaction
KEGG: Valine, leucine and isoleucine degradation
KEGG: PI3K-Akt signaling pathway
REAC: Signal Transduction
WP: Focal Adhesion-PI3K-Akt-mTOR-signaling pathway
REAC: Negative regulation of the PI3K/AKT network
REAC: Extracellular matrix organization
KEGG: Viral protein interaction with cytokine and cytokine receptor
WP: Lung fibrosis
REAC: PI5P, PP2A and IER3 Regulate PI3K/AKT Signaling
Log(p-value)


Pathway enrichment for downregulated DE genes
REAC: Developmental Biology
REAC: Axon guidance
REAC: Nervous system development
REAC: Semaphorin interactions
WP: EBV LMP1 signaling
Log(p-value)

B


Volcano plot of C2C12 RNAseq at d2
-log10(padj)
Log2 Fold Change
Pnpo
Zfand5
Serpine2
Coro2a
Wbp1l
Kitl
Etv4
Magea8
Maf
Selenop
Pdgfb
Ak4
Chst15
Robo2
Bdh1
Pdp1
Ece1
Tgfbr1
Akr1c14
Pcsk6
Fhdc1
Smpdl3a
Hormad2
Gpr35
Pkp2
Fzd4
Mmp13
Il1rl1
Vmn2r79
4930579G24Rik


Pathway enrichment for upregulated DE genes
KEGG: Transcriptional misregulation in cancer
WP: Focal Adhesion-PI3K-Akt-mTOR-signaling pathway
WP: Adipogenesis genes
KEGG: Pathways in cancer
KEGG: Histidine metabolism
KEGG: PI3K-Akt signaling pathway
KEGG: Thyroid hormone signaling pathway
KEGG: Neuroactive ligand-receptor interaction
WP: Nuclear Receptors
REAC: Neuronal System
Log(p-value)

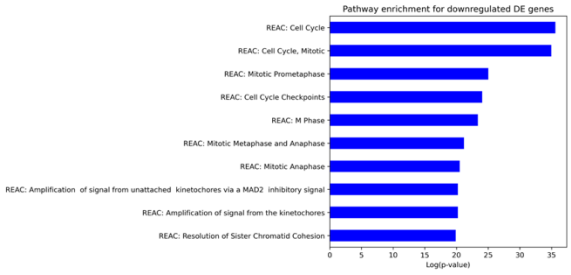

Pathway enrichment for downregulated DE genes
REAC: Cell Cycle
REAC: Cell Cycle, Mitotic
REAC: Mitotic Prometaphase
REAC: Cell Cycle Checkpoints
REAC: M Phase
REAC: Mitotic Metaphase and Anaphase
REAC: Mitotic Anaphase
REAC: Amplification of signal from unattached kinetochores via a MAD2 inhibitory signal
REAC: Amplification of signal from the kinetochores
REAC: Resolution of Sister Chromatid Cohesion
Log(p-value)

Figure 24. Differential expression in myoblasts and d2 between WT and GPS2-KO C2C12s visualized by volcano plot and functional enrichment

(a) DEG analysis in myoblasts results in 1717 significant genes at a FDR < .1. Functional enrichment analysis for DEG separated by directionality of log2FC reveal an enrichment for PI3-Akt signaling and ECM remodeling related terms for the upregulated genes. Downregulated genes demonstrate enrichment for axon guidance, and general development pathways.

(b) DEG analysis at d2 results in 2635 significant genes at a FDR < .1. Functional enrichment analysis for DEG separated by directionality of log2FC reveal an enrichment for PI3-Akt signaling, adipogenesis genes, and transcription related terms for the upregulated genes. Downregulated genes show strong enrichment for cell cycle and cell cycle regulation terms

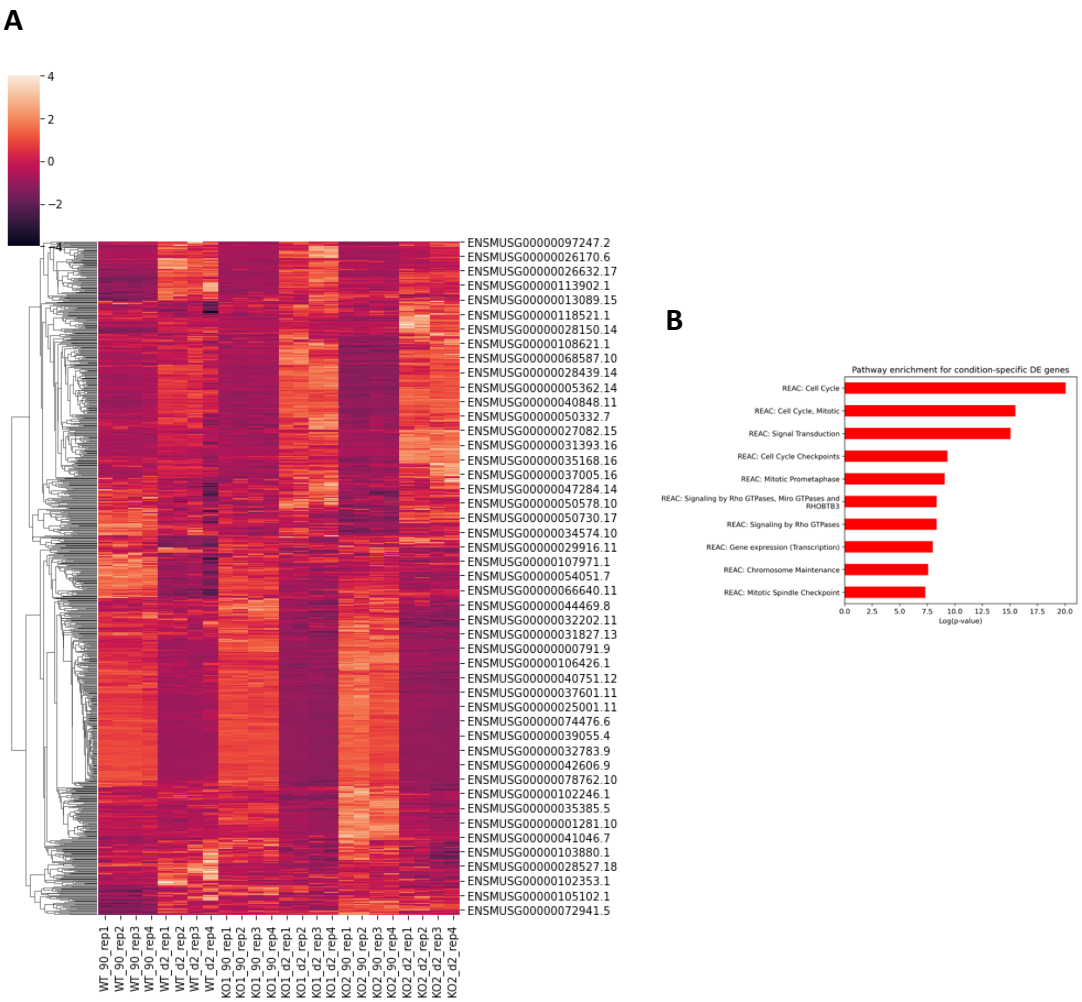
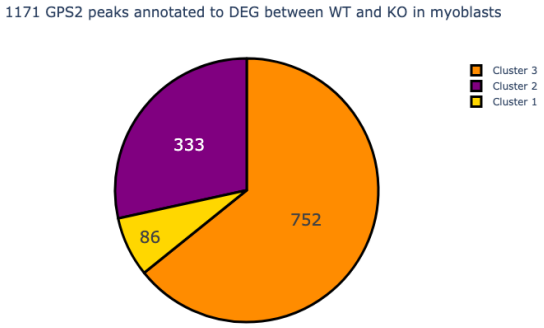


Figure 25. 882 genes are differentially expressed in a condition-specific manner over time. Genes are functionally enriched for cell cycle and cell cycle regulation pathways. Heatmap represents normalized counts transformed to z-scores with rows clustered by hierarchical clustering.

(a) Heatmap of z-scores for normalized counts for all samples across time

(b) Pathway enrichment of genes changed over time

A



B

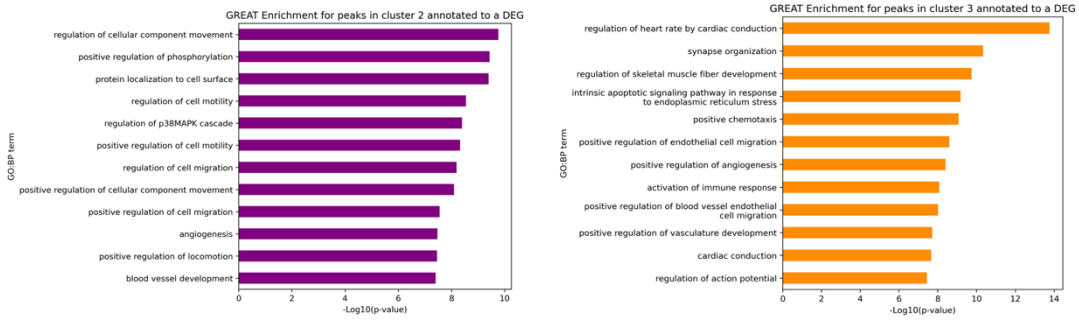
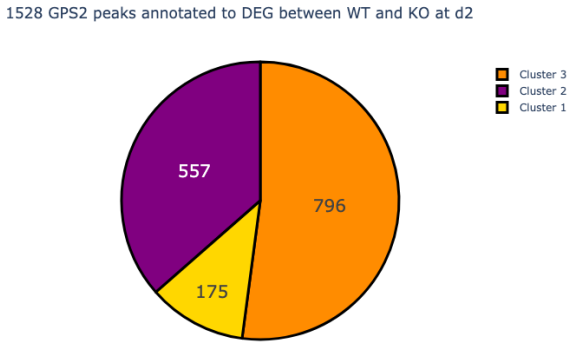


Figure 26. 1171 GPS2 peaks in myoblasts are also annotated to genes that are differentially expressed between WT and GPS2-KO C2C12s in myoblasts. Functional enrichment using GREAT reveals terms related to cell motility, cell shape, and cell migration in cluster 2. Cluster 3 is enriched for terms related to skeletal muscle development, synapse organization and action potential terms.

(a) Proportion of overlapping peaks found in ChIPseq and RNAseq in heatmap clusters

(b) GREAT enrichment for peaks in cluster 2 and 3

A



B

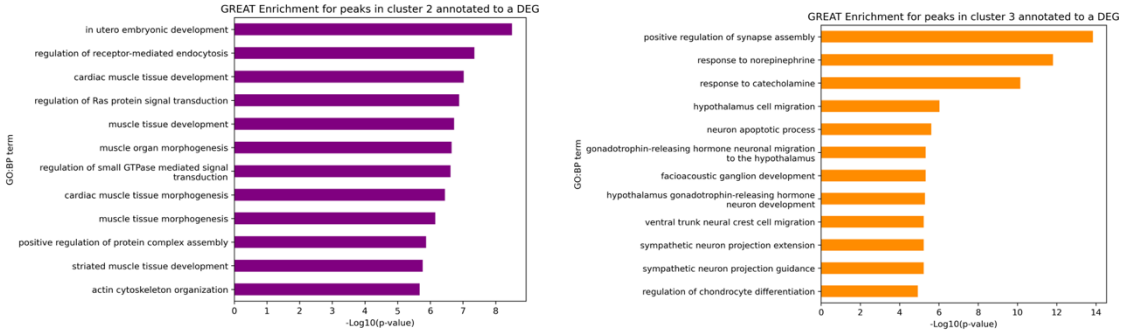


Figure 27. 1528 GPS2 peaks in myoblasts are also annotated to genes that are differentially expressed between WT and GPS2-KO C2C12s at d2. Functional enrichment using GREAT reveals terms related to skeletal muscle development, muscle morphogenesis, actin organization in cluster 2. Cluster 3 is enriched for terms related to synapse assembly, response to norepinephrine, and neuron guidance.

(a) Proportion of overlapping peaks found in ChIPseq and RNAseq in heatmap clusters

(b) GREAT enrichment for peaks in cluster 2 and 3

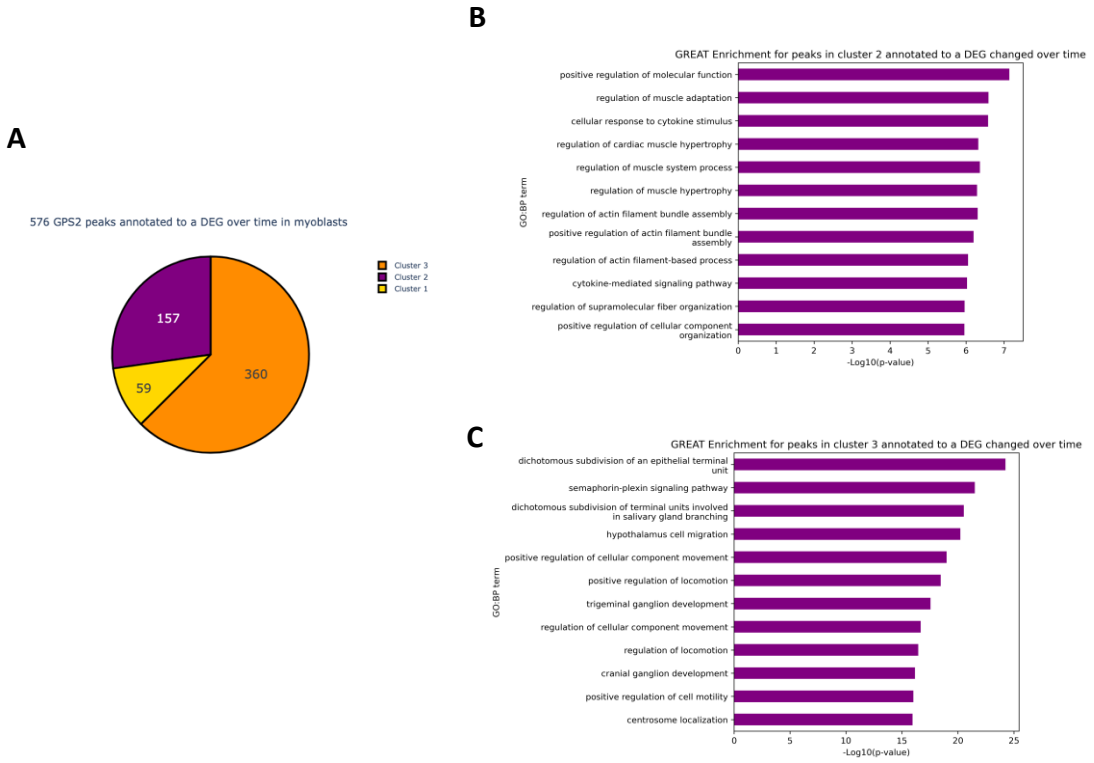


Figure 28. 576 out of 882 differentially expressed genes over time also contain a corresponding annotated GPS2 binding event. Peaks found in cluster 2 displaying known enhancer signatures are enriched for classical muscle differentiation functions. Peaks found in cluster 3 overlapping with DEG are enriched for term related to cellular movement, locomotion and migration.

(a) Proportion of overlapping peaks found in each cluster of peaks from MB with genes differentially expressed over time determine via RNAseq

(b) GREAT enrichment for overlapping peaks in cluster 2 and 3

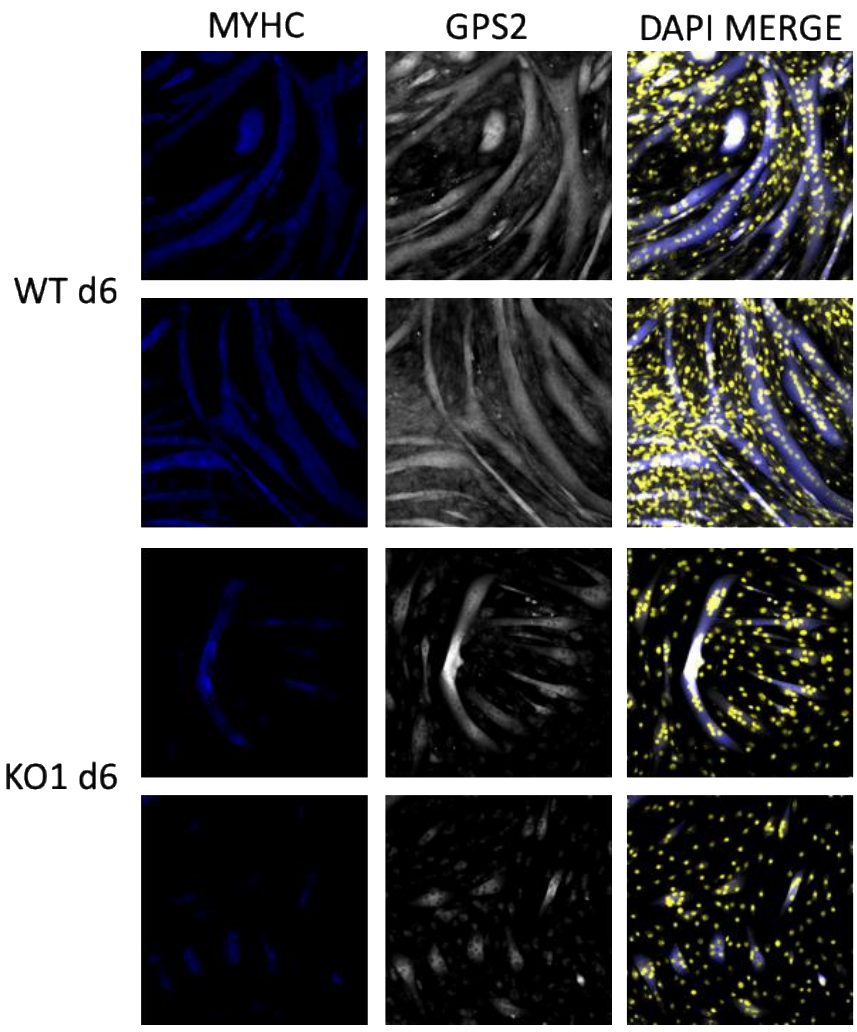


Figure 29. Impairment of differentiation in first C2C12 GPS2-KO cell line demonstrated by lack of staining for MyHC

Two replicates shown WT and KO IF staining for MyHC and GPS2 after 6 days of differentiation. Impairment of differentiation demonstrated by decreased staining for MyHC. Knockout of GPS2 is also confirmed through loss of GPS2 staining.

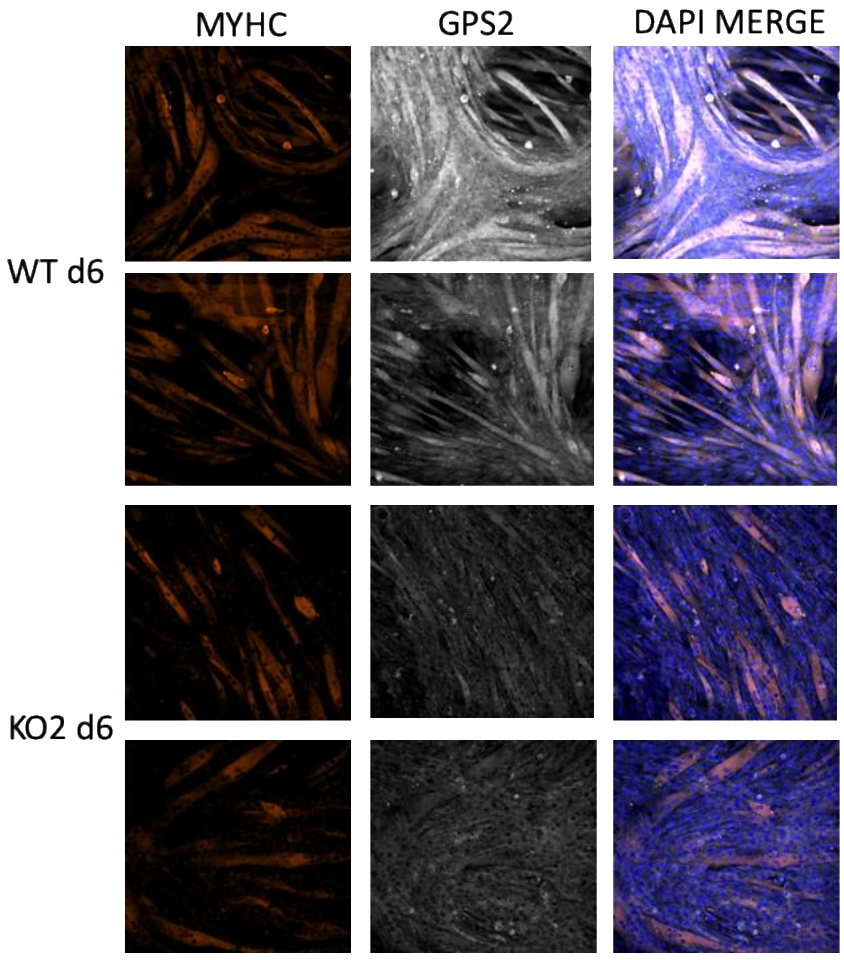


Figure 30. Impairment of differentiation in second C2C12 GPS2-KO cell line demonstrated by lack of staining for MyHC

Two replicates shown WT and KO IF staining for MyHC and GPS2 after 6 days of differentiation. Impairment of differentiation demonstrated by decreased staining for MyHC. Knockout of GPS2 is also confirmed through loss of GPS2 staining.

CHAPTER 4

Final Conclusion and Future Directions

Summary of Key Findings

In summary, there are three main aspects of work that comprise this thesis. The first chapter focuses on elucidating the transcriptional profile of both wild-type and GPS2-AKO adipocytes after being subjected to high fat diet stress. The second chapter continues this work by extending our characterization to include the cellular composition of the underlying SVF of the adipose tissue. The third chapter seeks to understand if a previously demonstrated mechanism for GPS2 regulation of transcription is conserved from adipocytes to a similar metabolic organ, muscle.

In the first chapter, we profile isolated adipocytes from GPS2-AKO and WT mice. In our previous metabolic characterization of our genetic mouse model, we have shown that under conditions of chow diet, GPS2-AKO mice become obese relative to their WT counterparts but display improved insulin resistance and no inflammation[30]. Interestingly, a study by Jones et al. performed a series of experiments profiling the response of wild-type adipocytes subjected to 8 weeks, 20 weeks or 34 weeks of HFD. Interestingly, they observe an upregulation for ECM and ECM-remodeling related genes at each time point. And at 20 weeks of HFD and beyond, they also observe a strong downregulation of mitochondrial gene pathways. After initial characterization of our mouse line, we isolated

adipocytes from our mice under a chow diet and utilized RNAseq to profile their changes in gene expression. Surprisingly, we note the opposite changes in our KO adipocytes as observed generally for wild-type adipocytes subjected to relatively long-term diets. Namely, we see an upregulation of mitochondrial genes specifically related to mitochondrial function and a downregulation of ECM and remodeling terms. However, given the context of our whole-body phenotype, these results are consistent with the fact that the GPS2-AKO mice display no signs of associated metabolic complications or inflammation despite being in a condition of obesity. Our RNAseq studies in the isolated adipocytes also suggests that at a molecular level, there exists a protection from the effects of obesity on the function of the adipocytes themselves. This unique metabolic phenotype under chow is not seen in a parallel GPS2-AKO mouse model and we can only speculate that the differences are due to the differences in the design of the genetic deletion or other experimental factors that will need further study.

However, both GPS2-AKO mouse models show similar phenotypes and responses to high fat diet stress. We can see in 16 week high fat diet studies of both the epididymal and subcutaneous WAT that GPS2-deleted adipocytes display sensitized responses to HFD stress compared to their WT counterparts. In our own studies, we can clearly see pathways associated with adipocyte dysfunction far more enriched in our KO adipocytes and even note potential depot-specific differences. In particular, the epididymal depot displays upregulation of genes related to neutrophils and downregulation of cell migration

pathways. By contrast, the subcutaneous depot displays enrichment for a wide variety of inflammatory pathways amongst the upregulated genes and many different catabolic and anabolic process pathways amongst the downregulated. At this time point, GPS2-KO isolated adipocytes from both depots appear to show heightened responses to HFD stress.

This interesting alteration between metabolic states from chow and HFD in our GPS2-AKO mouse line motivated us to explore earlier time points to better determine when the switch occurs and what underlying changes mediate it. We employed scRNAseq to profile the underlying cell populations in the SVF. Our most important finding was the discovery of a potential increase in a population of APCs in the KO epididymal WAT after a short 5-week exposure to HFD. We also observe several other cell populations that may be changed between our WT and KO mice across both depots including smooth muscle cells, and a population of macrophages robustly expressing collagen and ECM genes. These findings have motivated both a set of flow cytometry assays to physically confirm any proportional changes as well as suggested the existence of an altered secretome present in our GPS2-AKO adipocytes. As GPS2 is only deleted in mature adipocytes in this line, being driven by an adiponectin-Cre, any changes in the underlying SVF must likely be mediated through altered paracrine signaling. We are actively investigating potential changes in the secretome of our GPS2-KO adipocytes in exosomes and conditioned media using mass spectrometry.

In the third chapter, we aimed to determine if GPS2 was remodeling the chromatin environment for MyoD similar to its role as a pioneering factor for PPARγ in adipocytes. We also had previously demonstrated the role of GPS2 as an activator of nuclear encoded mitochondrial genes in another mitochondrial rich tissue and hypothesized it may be acting similarly in muscle development. We performed ChIPseq at four different time points covering undifferentiated myoblasts, myoblasts three hours post differentiation, myoblasts two days post differentiation and myoblasts four days post differentiation. One of the striking results from this experiment was the extraordinarily high proportion of binding annotated to distal regions across all four time points. To determine if this binding was in potential enhancer regions, we generated coverage tracks of major histone marks (H3K27ac, H3K4me1, H3K4me3) and transcription factors (MyoD, p300, c-Jun) and visualized their enrichment relative to the center of our GPS2 peaks. In general, we observe several notable patterns in regards to the binding of these factors. We can see a high proportion of GPS2 peaks in regions of strong signal for H3K27ac and H3K4me1, two commonly used canonical markers of enhancer regions. A subset of these H3K27ac+ and H3K4me1+ peaks also appear to overlap with peaks of MyoD binding, albeit shifted around 500bp up- or downstream of the GPS2 peak center. We observe these same general patterns in both the ChIPseq performed in undifferentiated myoblasts as well as in myoblasts differentiated for two days.

Looking closer at the H3K27ac+, H3K4me1+ and H3K4me3+ peaks in the myoblasts, we see a strong enrichment for various RNA processing pathways and terms (ncRNA processing, tRNA processing, mRNA processing). It has been well-established that muscle tissue is a highly dynamic tissue and many RNA processing functions such as alternative splicing, or alternative polyadenylation are major means which mediate skeletal muscle development and adaptation. In particular, differentiation was shown to be associated with a large number of alternative splicing events and these transitions were shown to be dependent on the myogenic differentiation process. It is interesting to note that we see binding of these GPS2 peaks annotated to these genes prior to the induction of differentiation. We can also see that they appear to be flanked by enriched areas of H3K4me3, which is typically a mark for active or poised promoters, but has also been shown to more generally be associated with regions of active transcription even in non-promoter regions[114]. This suggests a potential role for GPS2 in regulating various RNA processing pathways even prior to the activation of the myogenic program.

We observe a cluster of peaks that show strong enrichment for H3K27ac, H3K4me1, and MyoD binding in myoblasts. Functional enrichment of these peaks reveals pathways related to cell adhesions, cell shape, and ECM organization. Motif enrichment for these peaks show a number of AP-1 complex members including Fra, c-Jun, BATF. We can also see binding of p300 overlapping with MyoD and speculate that this group of peaks represent active

enhancers regulating genes and pathways responsible for the eventual organization and structural alignment of the formed myotubes. Morphologically, one of the earliest observable changes seen in differentiating myoblasts is the alignment and patterning of cells before their eventual fusion. This regulation is also supported by our RNAseq studies, which demonstrated that peaks from this cluster that are annotated to DEG between WT and KO C2C12 myoblasts also share this enrichment for cell shape, cell motility, and cellular movement. This suggests that the binding of GPS2 in these putative enhancer regions are directly regulating genes critical to these pathways.

The last cluster of GPS2 peaks in myoblasts seem to be primarily marked by H3K4me1 or located in distal regions away from any significant signal for H3K27ac or H3K4me1. These peaks may represent GPS2 binding to poised enhancers, which are not yet active or kept inactive in myoblasts. If we overlap the peaks found in this cluster with our RNAseq data again, we observe enrichment for various terms related to action potentials, synapse organization and skeletal muscle development. These peaks may represent a subset of muscle specific enhancers that become activated in the course of differentiation as they represent some of the later stage processes in myogenesis.

We performed the same set of analyses for our peaks found two days post differentiation. The first cluster of peaks are strongly enriched for H3K27ac, H3K4me1 and MyoD binding and display functional enrichment for classical

muscle differentiation terms such as myofibril assembly, actomyosin organization, and striated muscle development.

The second cluster displays the same pattern of marks and shows enrichment for some muscle specific terms like sarcomere organization or actomyosin organization but also general pathway terms relating to regulation of apoptosis and cell cycle. Motif finding again shows enrichment for various members of the AP-1 complex. Overlapping this set of peaks with our RNAseq data at d2, we see that the terms are almost exclusively related to muscle specific pathways such as myofibril assembly and muscle morphogenesis. This suggests that GPS2 is directly regulating the expression of muscle specific genes at this time and also display binding of GPS2 in putative active enhancer regions also marked by MyoD.

The last cluster of peaks at d2 seem to be mostly marked by H3K4me1. Functional enrichment of these peaks that are also annotated to DEG at this time point reveals many pathways related to synapse organization and action potentials. We speculate that this represents another set of poised enhancers that regulate the activity of gene programs that become active in the later to terminal stages of differentiation. Early myogenesis is large characterized by reorganization, alignment and fusion of individual cells into myotubes. Later myogenesis develops many of the muscle-specific machinery that will eventually drive the contractile apparatus and other functions. Consistent with this idea,

synapse organization and action potential pathways are known to be more related to the overall function of mature myotubes/myofibrils.

To expand this work, we would like to perform 4C experiments to characterize the potential 3D interactions between regulatory regions of the genome. These studies would be especially interesting in light of our observation of GPS2 binding shifted from the binding site for known transcription factors MyoD, c-Jun and p300. We are also interested in assaying these genomic regions physically using ChIP-qPCR using our GPS2-KO C2C12 lines to observe how loss of GPS2 is modulating the histone marks seen at these peak locations and/or the recruitment of MyoD and other factors.

MATERIALS AND METHODS

mRNAseq Analysis – Isolated Adipocytes

mRNAseq libraries were subjected to basic quality control and read trimming using the FASTQC and Trimmomatic packages. Surviving reads were analyzed using Salmon to obtain mRNA abundance estimates using the Gencode M10 gene annotation. Abundance estimates were concatenated into a single matrix and filtered by removing genes with a row mean of < 10. Normalization and differential expression of the abundance matrix was performed using DESeq2 adjusting for litter status. Differentially expressed genes were considered significant with a FDR <0.1.

Overrepresentation Analysis – Isolated Adipocytes

Overrepresentation analysis was performed using either the EnrichR webtool or the gProfiler API. Terms were considered significant with FDR <.1 or gSCS < .1.

Gene Set Enrichment Analysis

Gene set enrichment analysis was performed using the fgsea R package in Bioconductor against the MSigDB C2 Canonical Pathway database. Gene lists were first converted to their human orthologues using BiomaRt. Normalized Enrichment Scores were calculated against each gene list sorted by descending Log2FC, and gene sets were considered significant with a FDR <.25.

Isolation of Stromal Vascular Fraction (SVF) From Mouse Adipose Tissue

Upon completion of a 5 week 60% high-fat diet (HFD), C57BL6/J WT and GPS2-AKO mice were euthanized according to IACUC standards, and adipose tissue depots excised. The tissue was then minced in cold PBS in a tissue culture dish, and filtered through a 250 micron mesh. Following the filtration, the minced tissue was resuspended with 25 mL of collagenase digestion buffer. For subcutaneous fat, the buffer is 25 mL DMEM media (with 4.5 g/L glucose & L-glutamine, without sodium pyruvate) along with 0.05 g collagenase Type II and 0.5 g BSA, and for epididymal adipose tissue the formulation is the same except 0.025 g collagenase type II is utilized. The tissue is then placed in a shaker for 1 hour at 37 degrees, moving at 120 RPM. Following digestion, the milieu is transferred to a new tube, once again filtering through 250 micron mesh. 5ml of FBS was added to neutralize digestion, and the lipid-laden adipocytes float to the top of the suspension. This layer of cells is removed, and the remaining volume is spun down at 500 xg for 10 minutes. The pellet that is then generated is the stromal vascular fraction (SVF).

Single Cell RNAseq

FASTQ files were generated by demultiplexing the Illumina base call files (BCL) via the Cellranger mkfastq command. Individual FASTQ files are then aligned to the MM10 reference transcriptome provided by 10x Genomics through the Cellranger count command. Samples are then aggregated together via the

Cellranger Aggr command to generate a single matrix for downstream analysis using a combination of Scanpy and Seurat.

Cell quality control was performed by plotting distributions for counts per barcode, number of genes per barcode, and percent of reads aligned to mitochondrial genes. These distributions were used to filter out potential outliers by visual thresholding. Gene counts were then normalized to 10,000 reads per cell using a counts per million normalization method. The data was then logarithmized using the log1p scanpy transformation function. Highly variable genes were selected by their mean to variance ratios using the Scanpy highly variable genes function with default parameters. The effects of total counts per cell and percent mitochondrial genes expressed were then regressed out of the data and each gene was scaled to unit variance.

Neighborhood graphs using the PCA representation of the matrix were generated using the scanpy Neighbors function, and embedded in two dimensions using the UMAP projection implemented in Scanpy. Clustering of the neighborhood graph was performed via the Leiden Graph Clustering Method. Marker genes for cluster identification were determined by a Wilcoxon rank sum test that compared the expression of genes in one cluster vs all of the other clusters combined.

scRNAseq Differential Expression

Differential expression within clusters between WT and KO cells was performed using MAST with default parameters.

Western Blot

Cultured cells were pelleted via centrifugation at 1000rpm for 5 minutes at 4C. Whole cell extracts were prepared by lysis in IPH buffer (50mM Tris HCL pH 8.0, 250mM NaCl, 5mM EDTA, 0.5%NP40) supplemented with 1x Protease Inhibitors (Roche). Cells were resuspended in IPH buffer for 20 minutes on ice and spun at max speed for 10 minutes at 4C. The resultant supernatant was collected and its protein concentration determined via the Bradford assay using a standard curve generated through BSA standards (Bio-Rad). Consistent amounts were loaded for each experiment and extracts boiled with NuPAGE LDS Laemmli buffer and 100mM DTT for 10 minutes prior to loading on 4-15% Mini-PROTEAN TGX gels (Bio-Rad). Gels were electrophoresed in SDS-PAGE Running Buffer (25mM Tris, 192mM glycine, .1% SDS) (Boston Bioproducts) at 180V for 40 minutes. Proteins were transferred onto .45um low fluorescence PVDF membranes (Bio-Rad) using the TransBlot Turbo System (Bio-Rad). Membranes were blocked in 5% non-fat milk in PBST (137 mM NaCl,5.4 mM KCl, 16.2mM Na2HPO4, 2.9 mM KH2PO4, 0.1% Tween-20) for one hour at room temperature with agitation before incubation overnight with primary antibody at 4C with agitation. Membranes were then washed three times for ten minutes in PBST and

incubated with appropriate secondary antibodies for one hour at room temperature with agitation, followed by three washes for ten minutes in PBST. Blots were visualized using the ChemiDoc MP imaging system and Image Lab Software (Bio-Rad). Protein molecular weights were determined via comparison to the Precision Plus Standard Protein Ladder (Bio-Rad).

ChIPSeq

ChIPseq samples were generated via the ChIP-qPCR protocol described in this methods section. Library preparation was performed at the Tufts University Core Facility for Genomics using the Illumina TruSeq Library Preparation kit. Libraries were sequenced using a Illumina HiSeq2500 high output single-End 50 flowcell. ChIPseq libraries were subjected to basic quality control using FastQC and trimmomatic. Surviving reads were aligned to the appropriate reference genome using BowTie2. SAM files were sorted, indexed and converted to BAM files before peak calling through either MACS2 or Homer using default parameters. Peak files were filtered against known blacklist regions using BedTools and any peaks intersecting these regions were removed from the final analysis. Peak files were annotated to genomic regions using the annotatePeaks.pl function in Homer. For visualization purposes, BigWig files were generated directly from BAM files using DeepTools and normalized to 1x Reads Per Genomic Coverage. BigWig files were uploaded to the CyVerse web hosting service and visualized on the UCSC genome browser.

C2C12 RNAseq

RNAseq samples were analyzed for basic quality metrics via Bioanalyzer. Library preparation was performed at the BU School of Medicine Microarray and Sequencing Core using the Kapa RNA HyperPrep with RiboErase kit. Libraries were sequenced using the Illumina NextSeq 2000. mRNA libraries were analyzed and subjected to basic quality control using FastQC and Trimmomatic. Trimmed reads were then aligned to the mouse Gencode M23 genome and counted using Verse. Counts were normalized using DESeq2 and differential expression was performed using several contrasts. We determined DEGs over time for both our WT samples and KO samples. We also performed DE testing between WT and KO at both time points. Finally, we tested using a reduced model excluding time vs. the full model to determine condition-specific DEG over time.

ChIP-qPCR

Cells were crosslinked in the presence of 1% formaldehyde for 10 minutes at room temperature. Cells were pelleted in PBS with a protease inhibitor cocktail before being frozen at -80C. Frozen pellets were resuspended in hypotonic lysis buffer () for 10 minutes on ice before centrifugation at max for 10 minutes at 4C. The supernatant was discarded and the pellet resuspended in SDS lysis buffer () for 10 minutes on ice. Samples were sonicated in a Diagenode … for 2 cycles of 10 minutes at (settings). Sonicated samples were cleared of debris by

centrifugation at 13000rpm for 10 minutes at 4C. The supernatant was diluted 10x in ChIP dilution buffer and incubated with antibodies at 4C with rotation overnight. 50ul of protein 4B sepharose was added to each sample and rotated for 1hr at 4C. Sepharose beads were pelleted by centriguation for 5 minutes at 1000rpm at 4C. Beads were washed with the following buffers in sequential order: Low Salt Buffer (), High Salt Buffer (), LiCl Buffer, and TE Buffer. DNA-protein complexes were eluted with EB for 30 minutes with shaking at room temperature. Samples were de-crosslinked at 65C for 6 hours. After crosslinking, a 1:1 volume of Phenol:Chloroform:Isoamyl alcohol was added to each sample before vortexing for 10 seconds. 3ul of glycogen () was added to each sample and incubated for 1 minute. 30ul of 3M Sodium acetate was added to each sample and incubated for 1 minute. 750ul of 100% cold ethanol was added to each sample, and samples were stored at -80C for 2 hours. DNA was pelleted by centriguation at max speed for 20 minutes at 4C. DNA pellets were then washed with 1ml of 70% cold ethanol, and centrifuged at max speed for 5 minutes at 4C. DNA pellets were air-dryed before resuspension in nuclease free water.

RT-qPCR

Total RNA was harvested from cultured cells using the Qiagen RNeasy Mini Kit according to manufacturer's instructions. RNA concentration was quantified using a NanoDrop 2000 Spectrophotometer (ThermoFisher). First strand cDNA synthesis was performed with the iScript cDNA synthesis kit) using 1000 ng of

total RNA and following the manufacturer's protocol. RT-qPCR reactions were performed using Fast Sybr Master Mix (ABI) on a ViiA7 Real-Time PCR system. Relative mRNA expression was determined after normalization to the housekeeping gene cyclophilin A.

Cell lines

C2C12 cells were obtained from the ATCC. Cells were cultured in DMEM (10-017-CV) and 10% Fetal Bovine Serum (). For maintenance, cells were never allowed to reach greater than 70% confluence and only utilized for experiments within the first 15 passages. Cells were induced to differentiate by changing the growth media to 2% Horse Serum, DMEM (10-017-CV). Differentiation media was changed daily throughout all time courses.

CRISPR KO Line Generation

We utilized the pSpCas9(BB)-2A-Puro (PX459) V2.0 plasmid available on AddGene to generate our KO lines. Separately, we designed sgRNAs using the online software ChopChop to efficiently generate deletion mutants in the GPS2 genomic locus. We took the designed guides and synthesized matching oligos with the appropriate BbsI overhangs and 5' phosphorylation modifications. After digestion of 1ug of px459 plasmid with BbsI, we gel purified the digested plasmid using the QIAquick Gel Extraction Kit and eluted the vector in nuclease free. We annealed our phosphorylated guides in a thermocycler using the following

parameters: 37C for 30mins, 95C for 5min ramping down to 25C at 5C/min. Digested and purified vector was ligated with annealed guides and ligation mixture was transformed into DH5a competent E. coli. Positive colonies were selected and placed in liquid culture for outgrowth. Cultures were sequenced for proper insertion of the guides before preparation in a Midi Prep Kit and a final round of sequencing to verify proper cloning. A low passage number of C2C12 cells were grown to 50% confluence before transfection with sequenced vector. 2ug of vector was transfected using the JetPrime transfection reagent. Cells were kept in transfection media for 24hrs before changing to growth media for another 24hrs. Cells were then switched to selection media (growth media + 2ug/ml puromycin) and regularly changed for a week after complete death of WT cells in selection media. Knockout lines were kept as pooled lines and subjected to verification of genomic editing and efficient GPS2 knockout via deep sequencing, and western blot. KO cell lines were maintained in selection media and similar to the WT line used for a maximum of 15 passages.

Immunofluorescence

Cells were grown on Collagen 1 coated coverslips (Corning® BioCoat™ Collagen I) and differentiation was induced upon reaching >90% cell confluence. At measured timepoints, cells were fixed in 4% paraformaldehyde for 15 minutes at room temperature, washed 3 times for 5 minutes in PBS, and permeabilized in 0.3% Triton X-100 in PBS for 5 minutes at room temperature. Coverslips were

blocked in 0.5% BSA, Normal Donkey Serum (1:50) in PBS for 30 minutes at room temperature before incubation with primary antibodies GPS2 (custom) and MHC (RD MAB4470) at dilutions of 1:300 and 1:500 in blocking solution, respectively. Coverslips were then washed 3 times for 5 minutes in a 1:5 dilution of blocking solution (Working solution) before incubation with secondary antibodies, FITC () and RD () for 1 hour at room temperature. Coverslips were washed 3 times for 5 minutes in working solution followed by a single wash for 5 minutes in PBS, and a final wash for 5 minutes in ddH20. Coverslips were mounted onto slides via ProLong Gold Antifade Mountant with DAPI (P10144).

REFERENCES

1. Suter, D. M. Transcription Factors and DNA Play Hide and Seek. *Trends in Cell Biology* vol. 30 491–500 (2020).
2. Lambert, S. A. *et al.* The Human Transcription Factors. *Cell* **172**, 650–665 (2018).
3. Rosenfeld, M. G., Lunyak, V. V. & Glass, C. K. Sensors and signals: A coactivator/corepressor/epigenetic code for integrating signal-dependent programs of transcriptional response. *Genes and Development* vol. 20 1405–1428 (2006).
4. Horwitz, K. B. *et al. Nuclear Receptor Coactivators and Corepressors*. https://academic.oup.com/mend/article/10/10/1167/2713277 (1996).
5. Dai, Z., Ramesh, V. & Locasale, J. W. The evolving metabolic landscape of chromatin biology and epigenetics. *Nature Reviews Genetics* **21**, 737–753 (2020).
6. Hyun, K., Jeon, J., Park, K. & Kim, J. Writing, erasing and reading histone lysine methylations. *Experimental and Molecular Medicine* vol. 49 324 (2017).
7. Grewal, S. I. S. & Moazed, D. Heterochromatin and Epigenetic Control of Gene Expression. *Science* **301**, 798–802 (2003).
8. Saksouk, N., Simboeck, E. & Déjardin, J. Constitutive heterochromatin formation and transcription in mammals. *Epigenetics & Chromatin 2015 8:1* **8**, 1–17 (2015).

9. Sterner, D. E. & Berger, S. L. Acetylation of Histones and Transcription-Related Factors. *Microbiology and Molecular Biology Reviews* **64**, 435–459 (2000).

10. Soshnev, A. A., Josefowicz, S. Z. & David Allis, C. Molecular Cell Review Greater Than the Sum of Parts: Complexity of the Dynamic Epigenome. (2016) doi:10.1016/j.molcel.2016.05.004.

11. Lee, K. K. & Workman, J. L. Histone acetyltransferase complexes: one size doesn't fit all. *Nature Reviews Molecular Cell Biology 2007 8:4* **8**, 284–295 (2007).

12. Husmann, D. & Gozani, O. Histone lysine methyltransferases in biology and disease. *Nature Structural & Molecular Biology 2019 26:10* **26**, 880–889 (2019).

13. Gallinari, P., Marco, S. Di, Jones, P., Pallaoro, M. & Steinkühler, C. HDACs, histone deacetylation and gene transcription: from molecular biology to cancer therapeutics. *Cell Research 2007 17:3* **17**, 195–211 (2007).

14. Dimitrova, E., Turberfield, A. H. & Klose, R. J. Histone demethylases in chromatin biology and beyond. *EMBO reports* **16**, 1620–1639 (2015).

15. Yap, K. L. & Zhou, M.-M. Keeping It in the Family: Diverse Histone Recognition by Conserved Structural Folds. *Critical Reviews in Biochemistry and Molecular biology* **45**, 488 (2010).

16. Zhu, J. & Emerson, S. G. Hematopoietic cytokines, transcription factors

and lineage commitment. *Oncogene* **21**, 3295–3313 (2002).

17. Farmer, S. R. Transcriptional control of adipocyte formation. *Cell Metabolism* **4**, 263–273 (2006).

18. Sartorelli, V. & Caretti, G. *Mechanisms underlying the transcriptional regulation of skeletal myogenesis*.

19. Dilworth, F. J., Seaver, K. J., Fishburn, A. L., Htet, S. L. & Tapscott, S. J. In vitro transcription system delineates the distinct roles of the coactivators pCAF and p300 during MyoD/E47-dependent transactivation. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 11593–11598 (2004).

20. Mal, A., Sturniolo, M., Schiltz, R. L., Ghosh, M. K. & Harter, M. L. A role for histone deacetylase HDAC1 in modulating the transcriptional activity of MyoD: inhibition of the myogenic program. *The EMBO Journal* **20**, 1739 (2001).

21. Chan, S. S.-K. & Kyba, M. What is a Master Regulator? *Journal of Stem Cell Research & Therapy* **3**, (2013).

22. Spain, B. H. *et al.* Two human cDNAs, including a homolog of Arabidopsis FUS6 (COP11), suppress G-protein- and mitogen-activated protein kinase-mediated signal transduction in yeast and mammalian cells. *Molecular and Cellular Biology* **16**, 6698–6706 (1996).

23. Jin, D. Y. *et al.* A human suppressor of c-Jun N-terminal kinase 1 activation by tumor necrosis factor α. *Journal of Biological Chemistry* **272**, 25816–

25823 (1997).

24. Zhang, J., Kalkum, M., Chait, B. T. & Roeder, R. G. The N-CoR-HDAC3 nuclear receptor corepressor complex inhibits the JNK pathway through the integral subunit GPS2. *Molecular Cell* **9**, 611–623 (2002).

25. Jakobsson, T. *et al.* GPS2 Is Required for Cholesterol Efflux by Triggering Histone Demethylation, LXR Recruitment, and Coregulator Assembly at the ABCG1 Locus. *Molecular Cell* **34**, 510–518 (2009).

26. Peng, Y.-C., Breiding, D. E., Sverdrup, F., Richard, J. & Androphy, E. J. AMF-1/Gps2 Binds p300 and Enhances Its Interaction with Papillomavirus E2 Proteins. *Journal of Virology* **74**, 5872–5879 (2000).

27. Toubal, A. *et al.* SMRT-GPS2 corepressor pathway dysregulation coincides with obesity-linked adipocyte inflammation. *The Journal of Clinical Investigation* **123**, 362 (2013).

28. Fan, R. *et al.* Loss of the co-repressor GPS2 sensitizes macrophage activation upon metabolic stress induced by obesity and type 2 diabetes. *Nature Medicine 2016 22:7* **22**, 780–791 (2016).

29. K, D. *et al.* GPS2 Deficiency Triggers Maladaptive White Adipose Tissue Expansion in Obesity via HIF1A Activation. *Cell Reports* **24**, 2957-2971.e6 (2018).

30. Cederquist, C. T. *et al.* Systemic insulin sensitivity is regulated by GPS2 inhibition of AKT ubiquitination and activation in adipose tissue. *Molecular Metabolism* **6**, 125–137 (2017).

31. Komander, D. & Rape, M. The ubiquitin code. *Annual Review of Biochemistry* **81**, 203–229 (2012).

32. Yau, R. & Rape, M. The increasing complexity of the ubiquitin code. *Nature Cell Biology* **18**, 579–586 (2016).

33. Grice, G. L. & Nathan, J. A. The recognition of ubiquitinated proteins by the proteasome. *Cellular and Molecular Life Sciences* **73**, 3497 (2016).

34. Swatek, K. N. & Komander, D. Ubiquitin modifications. *Cell Research* vol. 26 399–422 (2016).

35. Luo, L. & Liu, M. Adipose tissue in control of metabolism. *Journal of Endocrinology* vol. 231 R77–R99 (2016).

36. Lee, M. J., Wu, Y. & Fried, S. K. Adipose tissue heterogeneity: Implication of depot differences in adipose tissue for obesity complications. *Molecular Aspects of Medicine* vol. 34 1–11 (2013).

37. Bjørndal, B., Burri, L., Staalesen, V., Skorve, J. & Berge, R. K. Different adipose depots: Their role in the development of metabolic syndrome and mitochondrial response to hypolipidemic agents. *Journal of Obesity* vol. 2011 (2011).

38. Cypess, A. M. *et al.* Identification and Importance of Brown Adipose Tissue in Adult Humans. *New England Journal of Medicine* **360**, 1509–1517 (2009).

39. Lee, P., Greenfield, J. R., Ho, K. K. Y. & Fulham, M. J. A critical appraisal of the prevalence and metabolic significance of brown adipose tissue in

adult humans. *American Journal of Physiology - Endocrinology and Metabolism* **299**, (2010).

40. Kajimura, S., Spiegelman, B. M. & Seale, P. Brown and beige fat: Physiological roles beyond heat generation. *Cell Metabolism* vol. 22 546–559 (2015).
41. Ailhaud, G. Adipose tissue as a secretory organ: from adipogenesis to the metabolic syndrome. *Comptes Rendus - Biologies* vol. 329 570–577 (2006).
42. Kershaw, E. E. & Flier, J. S. Adipose Tissue as an Endocrine Organ. *The Journal of Clinical Endocrinology & Metabolism* **89**, 2548–2556 (2004).
43. Margetic, S., Gazzola, C., Pegg, G. & Hill, R. Leptin: a review of its peripheral actions and interactions. *International Journal of Obesity 2002 26:11* **26**, 1407–1433 (2002).
44. Ouchi, N., Parker, J. L., Lugus, J. J. & Walsh, K. Adipokines in inflammation and metabolic disease. *Nature Reviews Immunology 2011 11:2* **11**, 85–97 (2011).
45. Lefterova, M. I. *et al.* PPARγ and C/EBP factors orchestrate adipocyte biology via adjacent binding on a genome-wide scale. *Genes and Development* **22**, 2941–2952 (2008).
46. Darlington, G. J., Ross, S. E. & MacDougald, O. A. The Role of C/EBP Genes in Adipocyte Differentiation *. *Journal of Biological Chemistry* **273**, 30057–30060 (1998).

47. Lefterova, M. I., Haakonsson, A. K., Lazar, M. A. & Mandrup, S. PPARg and the global map of adipogenesis and beyond. *Trends in Endocrinology & Metabolism* **25**, 293–302 (2014).

48. Cardamone, M. D. *et al.* GPS2/KDM4A pioneering activity regulates promoter-specific recruitment of PPARγ. *Cell Reports* **8**, 163–76 (2014).

49. Cardamone, M. D. *et al.* GPS2 regulates mitochondria biogenesis via mitochondrial retrograde signaling and chromatin remodeling of nuclear-encoded mitochondrial genes. (2017) doi:10.1101/162297.

50. A, A. K. *et al.* Comparative Secretome Analyses of Primary Murine White and Brown Adipocytes Reveal Novel Adipokines. *Molecular & Cellular Proteomics : MCP* **17**, 2358–2370 (2018).

51. Jones, J. E. C. *et al.* The Adipocyte Acquires a Fibroblast-Like Transcriptional Signature in Response to a High Fat Diet. *Scientific Reports 2020 10:1* **10**, 1–15 (2020).

52. AD, B. *et al.* Subcutaneous adipose tissue accumulation protects systemic glucose tolerance and muscle metabolism. *Adipocyte* **7**, 261–272 (2018).

53. Zuriaga, M. A., Fuster, J. J., Gokce, N. & Walsh, K. Humans and Mice Display Opposing Patterns of “Browning” Gene Expression in Visceral and Subcutaneous White Adipose Tissue Depots. *Frontiers in Cardiovascular Medicine* **4**, 1 (2017).

54. Sun, K., Kusminski, C. M. & Scherer, P. E. Adipose tissue remodeling and obesity. *The Journal of Clinical Investigation* **121**, 2094 (2011).

55. Heijden, R. A. van der *et al.* High-fat diet induced obesity primes inflammation in adipose tissue prior to liver in C57BL/6j mice. *Aging (Albany NY)* **7**, 256 (2015).

56. T, S., M, T. & Y, O. Adipose tissue inflammation and ectopic lipid accumulation. *Endocrine journal* **59**, 849–857 (2012).

57. Mirza, M. S. Obesity, Visceral Fat, and NAFLD: Querying the Role of Adipokines in the Progression of Nonalcoholic Fatty Liver Disease. *ISRN Gastroenterology* **2011**, 1–11 (2011).

58. Bays, H., Mandarino, L. & DeFronzo, R. A. Role of the Adipocyte, Free Fatty Acids, and Ectopic Fat in Pathogenesis of Type 2 Diabetes Mellitus: Peroxisomal Proliferator-Activated Receptor Agonists Provide a Rational Therapeutic Approach. *Journal of Clinical Endocrinology and Metabolism* vol. 89 463–478 (2004).

59. Lim, S. & Meigs, J. B. Links between ectopic fat and vascular disease in humans. *Arteriosclerosis, Thrombosis, and Vascular Biology* **34**, 1820–1826 (2014).

60. Bourin, P. *et al.* Stromal cells from the adipose tissue-derived stromal vascular fraction and culture expanded adipose tissue-derived stromal/stem cells: A joint statement of the International Federation for Adipose Therapeutics and Science (IFATS) and the International Society for Cellular Therapy (ISCT). *Cytotherapy* **15**, 641–648 (2013).

61. Pincu, Y. *et al.* Diet-induced obesity regulates adipose-resident stromal cell

quantity and extracellular matrix gene expression. *Stem Cell Research* **17**, 181–190 (2016).

62. Planat-Benard, V. *et al.* Plasticity of Human Adipose Lineage Cells Toward Endothelial Cells: Physiological and Therapeutic Perspectives. *Circulation* **109**, 656–663 (2004).

63. Hausman, G. J. & Dodson, M. V. Stromal Vascular Cells and Adipogenesis: Cells within Adipose Depots Regulate Adipogenesis. *Journal of Genomics* **1**, 56–66 (2013).

64. Lee, B. C. & Lee, J. Cellular and molecular players in adipose tissue inflammation in the development of obesity-induced insulin resistance. *Biochimica et Biophysica Acta - Molecular Basis of Disease* vol. 1842 446–462 (2014).

65. Burl, R. B. *et al.* Deconstructing Adipogenesis Induced by β3-Adrenergic Receptor Activation with Single-Cell Expression Profiling. *Cell Metabolism* **28**, 300-309.e4 (2018).

66. PC, S. *et al.* A stromal cell population that inhibits adipogenesis in mammalian fat depots. *Nature* **559**, 103–108 (2018).

67. C, H. *et al.* Identification of functionally distinct fibro-inflammatory and adipogenic stromal subpopulations in visceral adipose tissue of adult mice. *eLife* **7**, (2018).

68. Jaitin, D. A. *et al.* Lipid-associated macrophages control metabolic homeostasis in a Trem2-dependent manner. *Cell* **178**, 686 (2019).

69. P, R. *et al.* Single cell analysis reveals immune cell-adipocyte crosstalk regulating the transcription of thermogenic adipocytes. *eLife* **8**, (2019).

70. Deutsch, A., Feng, D., Pessin, J. E. & Shinoda, K. The Impact of Single-Cell Genomics on Adipose Tissue Research. *International Journal of Molecular Sciences* **21**, 1–13 (2020).

71. Sun, K., Tordjman, J., Clément, K. & Scherer, P. E. Fibrosis and Adipose Tissue Dysfunction. *Cell metabolism* **18**, 470 (2013).

72. Suganami, T., Nishida, J. & Ogawa, Y. A Paracrine Loop Between Adipocytes and Macrophages Aggravates Inflammatory Changes. *Arteriosclerosis, Thrombosis, and Vascular Biology* **25**, 2062–2068 (2005).

73. Itoh, M., Suganami, T., Hachiya, R. & Ogawa, Y. Adipose Tissue Remodeling as Homeostatic Inflammation. *International Journal of Inflammation* **2011**, 1–8 (2011).

74. Muir, L. A. *et al.* Adipose tissue fibrosis, hypertrophy, and hyperplasia: correlations with diabetes in human obesity. *Obesity (Silver Spring, Md.)* **24**, 597 (2016).

75. Zheng, G. X. Y. *et al.* Massively parallel digital transcriptional profiling of single cells. *Nature Communications 2017 8:1* **8**, 1–12 (2017).

76. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY : large-scale single-cell gene expression data analysis. *Genome Biology 2018 19:1* **19**, 1–5 (2018).

77. Traag, V., Waltman, L. & van Eck, N. J. From Louvain to Leiden: guaranteeing well-connected communities. *Scientific Reports* **9**, (2018).

78. Gupta, O. T. & Gupta, R. K. Visceral Adipose Tissue Mesothelial Cells: Living on the Edge or Just Taking Up Space? *Trends in Endocrinology & Metabolism* **26**, 515–523 (2015).

79. Rana, B. M. J. *et al.* A stromal cell niche sustains ILC2-mediated type-2 conditioning in adipose tissue. *The Journal of Experimental Medicine* **216**, 1999 (2019).

80. Merrick, D. *et al.* Identification of a mesenchymal progenitor cell hierarchy in adipose tissue. *Science (New York, N.Y.)* **364**, 2501 (2019).

81. Pham, T. D. *et al.* High-fat diet induces systemic B-cell repertoire changes associated with insulin resistance. *Mucosal Immunology 2017 10:6* **10**, 1468–1479 (2017).

82. Büttner, M., Ostner, J., Müller, C., Theis, F. & Schubert, B. scCODA: A Bayesian model for compositional single-cell data analysis. *bioRxiv* 2020.12.14.422688 (2020) doi:10.1101/2020.12.14.422688.

83. Finak, G. *et al.* MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biology 2015 16:1* **16**, 1–13 (2015).

84. Wolbert, J. *et al.* Redefining the heterogeneity of peripheral nerve cells in health and autoimmunity. *Proceedings of the National Academy of Sciences* **117**, 9466–9476 (2020).

85. Sárvári, A. K. *et al.* Plasticity of Epididymal Adipose Tissue in Response to Diet-Induced Obesity at Single-Nucleus Resolution. *Cell Metabolism* **33**,

437-453.e5 (2021).

86. Cho, D. S., Lee, B. & Doles, J. D. Refining the adipose progenitor cell landscape in healthy and obese visceral adipose tissue using single-cell gene expression profiling. *Life Science Alliance* **2**, (2019).

87. Jeffery, E. *et al.* The Adipose Tissue Microenvironment Regulates Depot-Specific Adipogenesis in Obesity. *Cell Metabolism* **24**, 142–150 (2016).

88. LF, L. *et al.* Adipose tissue macrophages impair preadipocyte differentiation in humans. *PloS one* **12**, (2017).

89. Luecken, M. D. & Theis, F. J. Current best practices in single-cell RNA-seq analysis: a tutorial. *Molecular Systems Biology* **15**, e8746 (2019).

90. Wang, Q. A. & Scherer, P. E. The AdipoChaser mouse: A model tracking adipogenesis in vivo. *Adipocyte* **3**, 146 (2014).

91. Drareni, K. *et al.* Adipocyte Reprogramming by the Transcriptional Coregulator GPS2 Impacts Beta Cell Insulin Secretion. *Cell Reports* **32**, 108141 (2020).

92. Mukund, K. & Subramaniam, S. Skeletal muscle: A review of molecular structure and function, in health and disease. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* vol. 12 (2020).

93. Gan, Z., Fu, T., Kelly, D. P. & Vega, R. B. Skeletal muscle mitochondrial remodeling in exercise and diseases. *Cell Research* vol. 28 969–980 (2018).

94. Davis, R. L., Weintraub, H. & Lassar, A. B. Expression of a single

transfected cDNA converts fibroblasts to myoblasts. *Cell* **51**, 987–1000 (1987).

95. Tapscott, S. J. *et al.* MyoD1: A nuclear phosphoprotein requiring a Myc homology region to convert fibroblasts to myoblasts. *Science* **242**, 405–411 (1988).

96. Hernández-Hernández, J. M., García-González, E. G., Brun, C. E. & Rudnicki, M. A. The myogenic regulatory factors, determinants of muscle development, cell identity and regeneration. *Seminars in Cell and Developmental Biology* vol. 72 10–18 (2017).

97. Relaix, F. & Zammit, P. S. Satellite cells are essential for skeletal muscle regeneration: The cell on the edge returns centre stage. *Development (Cambridge)* vol. 139 2845–2856 (2012).

98. Asfour, H. A., Allouh, M. Z. & Said, R. S. Myogenic regulatory factors: The orchestrators of myogenesis after 30 years of discovery. *Experimental Biology and Medicine* **243**, 118–128 (2018).

99. Kim, J.-W. *et al.* Tip60 regulates myoblast differentiation by enhancing the transcriptional activity of MyoD via their physical interactions. *The FEBS Journal* **278**, 4394–4404 (2011).

100. Gerber, A. N., Klesert, T. R., Bergstrom, D. A. & Tapscott, S. J. Two domains of MyoD mediate transcriptional activation of genes in repressive chromatin: A mechanism for lineage determination in myogenesis. *Genes and Development* **11**, 436–450 (1997).

101. Forcales, S. V. *et al.* Signal-dependent incorporation of MyoD-BAF60c into Brg1-based SWI/SNF chromatin-remodelling complex. *EMBO Journal* **31**, 301–316 (2012).

102. Yaffe, D. & Saxel, O. Serial passaging and differentiation of myogenic cells isolated from dystrophic mouse muscle. *Nature* **270**, 725–7.

103. Burattini, S. *et al.* C2C12 murine myoblasts as a model of skeletal muscle development: morpho-functional characterization. *European Journal of Histochemistry* **48**, 223–234 (2009).

104. Dedieu, S., Mazères, G., Cottin, P. & Brustis, J. J. Involvement of myogenic regulator factors during fusion in the cell line C2C12. *International Journal of Developmental Biology* **46**, 235–241 (2002).

105. Rajan, S. *et al.* Analysis of early C2C12 myogenesis identifies stably and differentially expressed transcriptional regulators whose knock-down inhibits myoblast differentiation. *Physiological Genomics* **44**, 183–197 (2012).

106. Blum, R. Activation of Muscle Enhancers by MyoD and epigenetic modifiers. *Journal of Cellular Biochemistry* **115**, 1855 (2014).

107. Wardle, F. C. Master control: transcriptional regulation of mammalian Myod. *Journal of Muscle Research and Cell Motility* **40**, 211 (2019).

108. Cabane, C., Englaro, W., Yeow, K., Ragno, M. & Dérijard, B. Regulation of C2C12 myogenic terminal differentiation by MKK3/p38α pathway. *https://doi.org/10.1152/ajpcell.00078.2002* **284**, 658–666 (2003).

109. S, D. *et al.* The Histone Variant MacroH2A1.2 Is Necessary for the Activation of Muscle Enhancers and Recruitment of the Transcription Factor Pbx1. *Cell Reports* **14**, 1156–1168 (2016).

110. K, M. *et al.* eRNAs promote transcription by establishing chromatin accessibility at defined genomic loci. *Molecular Cell* **51**, 606–617 (2013).

111. Blum, R., Vethantham, V., Bowman, C., Rudnicki, M. & Dynlacht, B. D. Genome-wide identification of enhancers in skeletal muscle: the role of MyoD1. *Genes & Development* **26**, 2763–79 (2012).

112. Khilji, S., Hamed, M., Chen, J. & Li, Q. Loci-specific histone acetylation profiles associated with transcriptional coactivator p300 during early myoblast differentiation. *Epigenetics* **13**, 642 (2018).

113. Zhang, X. *et al.* MicroRNA Directly Enhances Mitochondrial Translation during Muscle Differentiation. (2014) doi:10.1016/j.cell.2014.05.047.

114. Cruz, C. *et al.* Tri-methylation of histone h3 lysine 4 facilitates gene expression in ageing cells. *eLife* **7**, (2018).

CURRICULUM VITAE