Deciphering the Function of Obesity-associated Regulatory Elements within *FTO*

Samantha Laber

DPhil Candidate

Department of Physiology, Anatomy and Genetics
Wadham College, University of Oxford

Genetics of Type 2 Diabetes,
MRC Mammalian Genetics Unit, Harwell

MRC | Harwell

DEPARTMENT OF PHYSIOLOGY, ANATOMY AND GENETICS
Medical Sciences Division

Wadham College

Deciphering the Function of Obesity-associated Regulatory Elements within *FTO*

Samantha Laber

Department of Physiology, Anatomy and Genetics

Wadham College, University of Oxford

Genetics of Type 2 Diabetes, Mammalian Genetics Unit, MRC Harwell Institute

A thesis presented for the degree of Doctor of Philosophy, Trinity Term, 2017

Abstract

Genome-wide association studies have repeatedly shown that the strongest association with obesity arises from variants in the first intron of *FTO*. The intronic *FTO* variant rs1421085 is within an adipocyte-specific enhancer and that risk allele carriers have increased *IRX3* and *IRX5* expression in early adipogenesis (Claussnitzer *et al.*, 2015). Additionally, the same human risk variant was linked to decreased *AKTIP, RPGRIP1L* and *FTO* expression in iPSC-derived neurons (Stratigopoulos *et al.*, 2016). These data point towards several likely causal transcripts and tissues at the *FTO* locus and essentially, several likely mechanisms. Importantly, whether any of the high-risk variants at the *FTO* locus has any effect on the organismal level has not been addressed so far. The aim of my DPhil project was to use novel gene manipulation strategies *in vivo* to mechanistically dissect the *Fto* regulatory circuitry in mouse to pinpoint causal transcripts their effector tissues and to unravel their physiological role in body weight regulation . Using publicly available as well as my own genomic data (ATAC-seq) revealed that the intronic *FTO* regulatory element in human adipocytes is conserved in mouse pre-adipocytes. Manipulation of the corresponding motif in mouse (by deleting 82 nucleotides at the mouse orthologous region around rs1421085) resulted in depot- and sex-specific alteration of target genes *Irx3* and *Irx5* in pre-adipocytes. In addition to recapitulating many of the human findings in mouse, my results further unravelled a new level of regulatory complexity at the *FTO/Fto* locus. When these mutant mice were put on a high fat diet, I found a reduction on overall fat-mass that could be linked to altered mRNA levels of *Irx3* and *Irx5* in pre-adipocytes. Using a number of genetic techniques, I further showed that *Irx3* regulates several processes during adipocyte

development, amongst which is modulation of mitochondrial function. In summary, my findings provide new insight into how variants in *FTO* intron 1 affect adipocyte development and more specifically how *IRX3* affects early adipocyte differentiation.

Word count: approx. 45,000

Acknowledgements

Foremost, I want to thank my supervisor Roger Cox for his support, for helping me develop my ideas and the many opportunities I was given over the past four years. I would like to extend my thanks to my second supervisor James Cantley for his mentorship and guidance throughout my DPhil.

For their many discussions, for their time, for their advice and support, I am deeply thankful to Melina Claussnitzer and Cecilia Lindgren, who became more than mentors to me and who have shaped the way I approach scientific questions.

I also want to express my gratitude to Liz Bentley, for her valuable input and help on the project and beyond, and for facilitating my work throughout my time at MRC Harwell.

I can never thank my mentors enough, which have guided and supported me throughout my studies. Thanks to Mary Cotter and Henriette Kirchner.

Big thanks to the Cox team members, past and present, for their company and help including Chris Church, Ying Bai, Shahana Sengupta, Lee Moir and Alison Hough. Thanks also to the great staff in the animal facility, including Sara Wells, Lucie Vizor, Tamzin Osborne, Gemma Frake, and also Jeremy Sanderson and Mark Gardiner for their training. Thanks to Abraham Acevedo and his team for a great first lab rotation. I would also like to thank Juan Fernandez, Iain Williamson, Tom Nicol, Joffrey Mianne, Li Wang, Nick Sinnott-Armstrong and Adam Buckle for the many valuable discussions and help.

Thanks to Debbie Ashworth and Helen Ward for all their administrative help.

I am also grateful to the Medical Research Council for funding my DPhil.

My special gratitude goes to Thomas Agnew, who has been there for me throughout my DPhil and inspired me in many ways.

Finally, this thesis would have not been possible without the love and support of my parents Tanja and Rudi and my siblings Desiree and Michael, to whom I dedicate this thesis.

Danksagung und Widmung

Ein besonderer Dank gilt meinen Eltern Tanja und Rudi und meinen Geschwistern Desiree und Michael, die mich auf meinem Weg durch das Studium begleitet haben. Ich widme diese Doktorarbeit meinen Eltern und Geschwistern.

Publications arising from DPhil

Laber S, Cox RD (2017). Mouse models of human GWAS hits for obesity and diabetes in the post genomic era: time for re-evaluation. *Front Endocrinol.* 8:11.

Laber S, Cox RD (2015). Commentary: FTO obesity variant circuitry and adipocyte browning in humans. *Front Genet.* 6:318.

Merkestein M*, **Laber S***, McMurray F, Andrew D, Sachse G, Sanderson J, Li M, Usher S, Sellayah D, Ashcroft FM, Cox RD (2015). FTO influences adipogenesis by regulating mitotic clonal expansion. *Nat Commun.* 6:6792. **denotes shared, equal contribution*

Conference abstracts

Laber S, Agnew T, Cox RD. IRX3 regulates Adipocyte Browning via Mitochondrial Gene Clusters (2016). DOI: 10.1530/endoabs.44.OC5.6

Laber S, Merkestein M, Andrew D, Li M, McMurray F, Usher S, Sellayah D, Ashcroft FM, Cox RD. A novel role for FTO in adipogenesis. Obesity (2014) Proc Physiol Soc 32, PC012

Table of Contents

Table of Figures

Table of Supplementary Figures

Table of Tables

Table of Supplementary Tables

Abbreviations

18S	18S Ribosomal RNA
3C	Chromosome conformation capture
3D	3 dimensional
4C	Circularized chromosome conformation capture
ADCY5	Adenylate cyclase type 5
Adipoq	Adiponectin
Adrb3	Adrenergic receptor, beta 3
adWAT	Abdominal dermal white adipose tissue
AgRP	Agouti-related protein
Aktip	AKTIP AKT interacting protein
ARAP1	ArfGAP With RhoGAP Domain, Ankyrin Repeat And PH Domain 1
ATAC	Assay for transposase-accessible chromatin
B6J	C57BL/6J mouse strain
B6N	C57BL/6NJ mouse strain
BAT	Brown adipose tissue
BCA	Bicinchoninic acid
BMI	Body mass index
bp	Basepairs
BSA	Bovine serum albumin
C3H	C3H/HeH mouse strain
CAMK1D	Calcium/calmodulin-dependent protein kinase type 1D
Canx	Calnexin
CaptureC	3C technique that is both high resolution and high throughput
CDC123	Cell division cycle protein 123 homolog
cDNA	Complementary DNA
Cebpa	CCAAT/Enhancer Binding Protein Alpha
ChIA-PET	Chromatin Interaction Analysis by Paired-End Tag Sequencing
ChIP	Chromatin Immunoprecipitation
CI	Complex I, NADH-dehydrogenase
Cidea	Cell Death-Inducing DFFA-Like Effector A
CII	Complex II, succinate dehydrogenase
CIII	Complex III, ubiquinone-cytochrome c oxidoreductase

CIV	Complex IV, Cytochrome oxidase
CLAMS	Comprehensive Lab Animal Monitoring System
CO_2	Carbon Dioxide 2
Cox7a1	Cytochrome C Oxidase Subunit 7A1
Cox8b	Cytochrome C Oxidase Subunit 8B
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CRMs	*cis*-regulatory modules
Ct	Cycle threshold
CTCF	CCCTC-binding factor
Cux1	Cut-like homeobox 1
CV	Complex V, ATP synthase
DEL82	Mouse mutant harbouring a 82-nucleotide deletion in Fto intron 1
Dexa	Dexamethasone
DGKB	Diacylglycerol kinase beta
DHS	DNase I hypersensitive site
Dio2	Deiodinase, Iodothyronine, Type II
DMEM	Dulbecco's Modified Eagle's Medium
DNA	Deoxyribonucleic acid
DPBS	Dulbecco’s phosphate-buffered saline
ECAR	Extracellular acidification rate
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
EE	Energy expenditure
EF	Ejection fraction
Elovl3	Elongation of very long chain fatty acids protein 3
ENCODE	The Encyclopedia of DNA Elements
ENU	N-ethyl-N-nitrosourea
eQTL	Expression quantitative trait loci
ESCs	Embryonic stem cells
Essra	Estrogen-related receptor alpha
ETC	Electron transport chain
EtOH	Ethanol
Fabp4	Fatty-acid binding protein 4
FAIRE	Formaldehyde-Assisted Isolation of Regulatory Elements
Fasn	Fatty acid synthase
FBS	Foetal bovine serum

FCCP	Carbonyl cyanide-4-(trifluoromethoxy)phenylhydrazone
FFA	Free fatty acids
FISH	Fluorescent in situ hybridisation
Fto	Fat mass and obesity-associated gene
FTO-DEL10	Mouse mutant harbouring a 10-nucleotide deletion in Fto intron 1
FTO-RS1421085	Mouse mutant harbouring a point-mutation in Fto intron 1
Gapdh	Glyceraldehyde-3-Phosphate Dehydrogenase
gDNA	Genomic DNA
GWAS	Genome-wide association Study
gWAT	Gonadal white adipose tissue
H&E	Hematoxylin and eosin
H_2O	Dihydrogen monoxide
hASC	Human adipocyte-derived stem cells
HCNEs	Highly conserved non-coding elements
HDL	High-density lipoprotein
HDR	Homology Directed Repair
HEPES	Zwitterionic biological buffer
HFD	High-fat diet
HNF4A	Hepatocyte nuclear factor 4-alpha
HRP	Horseradish peroxidase
IBMX	3-Isobutyl-1-Methylxathine
Ins2	Insulin-2 precursor
IP	Immunoprecipitation
IPGTT	Intraperitoneal Glucose Tolerance Test
iPSC	Induced pluripotent stem cell
Irx3	Iroquois Homeobox 3
Irx5	Iroquois Homeobox 5
Irx6	Iroquois Homeobox 6
iWAT	Inguinal white adipose tissue
JAZF1	Juxtaposed with another zinc finger protein 1
Kb	Kilobase
KCNK16	Potassium two pore domain channel subfamily K member 16
KCNK17	Potassium two pore domain channel subfamily K member 17
KCNQ1	Potassium voltage-gated channel subfamily Q member 1
kg	Kilogram
KI	Knock-In

KLF14	Krüppel-like factor 14
KO	Knock-out
LD	Linkage disequilibrium
LDL	Low-density lipoprotein
LFD	Low-fat diet
M	Molarity
Mb	Megabase
m^6A	N6-methyadenosine
MEFs	Mouse embryonic fibroblasts
MRI	Magnetic resonance imaging
mRNA	Messenger RNA
mtDNA	Mitochondrial genome
MTNR1B	Melatonin receptor 1B
MW	Molecular weight
NaCl	Sodium Chloride
nDNA	Nuclear genome
Ndufb3	NADH:ubiquinone oxidoreductase subunit B3
Ndufb7	NADH:ubiquinone oxidoreductase subunit B7
Ndufb8	NADH:ubiquinone oxidoreductase subunit B8
Ndufs1	NADH:ubiquinone oxidoreductase core subunit S1
Ndufv1	NADH:ubiquinone oxidoreductase core subunit V1
Npy	Neuropeptide Y
O_2	Oxygen
OCR	Oxygen consumption rate
OE	Overexpression
OXPHOS	Oxidative phosphorylation
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
Pdgfra	Platelet derived growth factor receptor alpha
Pgc1a (Ppargc1a)	Pparg coactivator 1 alpha
pH	Potential of hydrogen
Plin1	Perilipin 1
PMCA	Phylogenetic Module Complexity Analysis
Pomc	pro-opiomelanocortin
Ppara	Peroxisome Proliferator Activator Receptor Alpha
Pparg	Peroxisome Proliferator Activator Receptor Gamma

Prdm16	PR Domain Zinc Finger Protein 16
Prrx1	Paired related homeobox 1
PS	Penicillin Streptomycin
pWAT	Perirenal white adipose tissue
QC	Quality control
qRT-PCR	Quantitative real-time PCR
RER	Respiratory exchange ratio
RNA	Ribonucleic acid
Rpgrip1l	Retinitis pigmentosa GTPase regulator-interacting protein-1 like
RPKM	Reads Per Kilobase per Million mapped reads
SD	Standard deviation
sdWAT	Subcutaneous dermal white adipose tissue
SEM	Standard error mean
sgRNA	Single guide RNA
Sim1	Single-minded family bHLH transcription factor 1
SNP	Single nucleotide polymorphism
SNV	Single nucleotide variants
ssODN	single-stranded oligodeoxynucleotides
STARD10	START domain-containing protein 10
SVF	Supra-vascular fraction
SWAT	Subcutaneous white adipose tissue
T2D	Type 2 Diabetes
TAD	Topologically associated domain
TBST	Tris-buffered saline, 0.1% Tween 20
Tbx15	T-box transcription factor 15
TF	Transcription factor
Tfam	Transcription Factor A, mitochondrial
TFBS	Transcription factor binding site
THADA	Thyroid adenoma-associated protein
TNFα	Tumor necrosis factor alpha
Ubc	Polyubiquitin-C
Ucp1	Uncoupling protein 1
VCO_2	CO_2 production
VO_2	Oxygen consumption
VWAT	Visceral white adipose tissue
Wars2	Tryptophanyl tRNA Synthesase 2, mitochondrial

WAT	White adipose tissue
WGS	Whole genome sequencing
WHR	Waist-to-hip ratio
WHRadjBMI	Waist-to-hip ratio adjusted for body mass index
WT	Wildtype
ZMIZ1	Zinc finger MIZ domain-containing protein 1

1 Introduction

1.1 The genetic underpinnings of obesity

The prevalence of obesity has reached pandemic proportions worldwide. More than 25% of the current world's population, more than 1.9 billion individuals, can be classified as either overweight or obese, defined by a body mass index (BMI) ≥ 25 kg/m^2 or BMI ≥ 30 kg/m^2, respectively (Afshin *et al.*, 2017). Although obesity has been closely tracked in the developed nations, almost two-thirds of the world's obese population reside in the developing world and particularly urban environments (Ng *et al.*, 2014; Prentice, 2006), putting a tremendous health burden on society worldwide.

Obesity is linked to increased morbidity and mortality in the human population (Heitmann & Lissner, 2011) and has systemic health implications by affecting many organs in the body (Fig. 1.1). For example, obese individuals are at higher risk of developing type 2 diabetes (T2D), neurodegenerative diseases such as dementia and Alzheimer's disease, hypertension, stroke and cardiovascular disease, and several types of cancer; many of which have inflammatory components (Kanneganti & Dixit, 2012; Ng *et al.*, 2014; Prentice, 2006). Furthermore, obesity has been implicated in fertility and pregnancy complications including gestational diabetes (Nicholas *et al.*, 2013; Oben *et al*., 2010). Maternal (as well as paternal) obesity can have long-term consequences on metabolic health in the offspring (Ferrara, 2007; Kamimae-Lanning *et al.*, 2015; Lecoutre *et al.*, 2016; Penfold & Ozanne, 2015).

Pathogenesis of obesity and T2D is complex and not yet fully understood and is influenced by behavioural, environmental and genetic factors. Globally, environmental factors are certainly driving the increased incidence of T2D and obesity. Within a population, however, where there is a great individual variance in body weight (BW) despite similar environmental factors, heritability of obesity susceptibility is estimated to be between 40% and 70% (O'Rahilly & Farooqi, 2000). It is well established that obesity and body fat distribution have

a strong genetic component. One of the experimental approaches to identify genetic variants associated with certain disease traits is genome-wide association studies (GWAS).

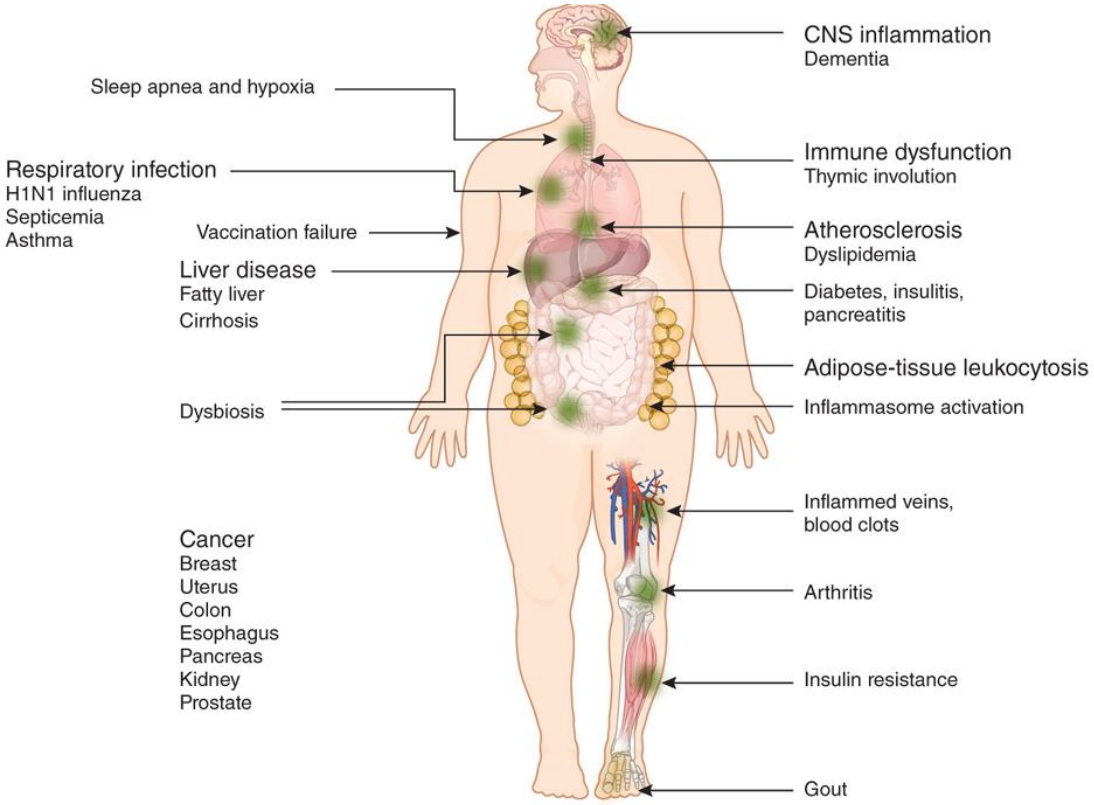


Fig. 1.1 | Obesity-associated comorbidities.
Obese individuals are at higher risk of a number of diseases including Type 2 Diabetes, hypertension, stroke, cardiovascular disease, and several types of cancer (Figure from Kanneganti & Dixit, 2012).

1.1.1 GWAS and obesity

In the past 10 years, GWAS have improved our knowledge of the polygenic nature of complex metabolic diseases including obesity. GWAS have identified hundreds of loci and thousands of single nucleotide polymorphisms (SNPs) associated with T2D and obesity traits, such as BMI and waist-to-hip-ratio (WHR) in the human population (DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium *et al.*, 2014; Locke *et al.*, 2015; Morris *et al.*, 2012; Shungin *et al.*, 2015). Although these different measures interrogate largely the same trait and are concerned with the same problem (i.e. obesity), progressively larger GWAS in obesity-related traits have shown that each trait measured possesses a distinct genetic signature, indicating that different phenotypic measures detect different underlying biology of this highly complex phenotype. Most strikingly perhaps, are the insights gained from the two most recent GWAS in BMI and WHRadjBMI (Locke *et al.*, 2015; Shungin *et al.*, 2015) that were performed on a largely overlapping cohort of >330000 individuals collected by the GIANT consortium. Firstly, there was very little overlap of loci that were significant for both traits. Secondly, of the 49 WHRadjBMI-associated loci, the majority were sexually dimorphic, while out of the 97 BMI-associated loci, only 3 were specific to sex. Using a tool for predicted gene function, called Data-driven Expression Prioritized Integration for Complex Traits (DEPICT), that integrates RNA-seq and GWAS summary statistics (Pers *et al.*, 2015) further revealed that the underlying pathways of the various loci were different for the traits. This analysis revealed a role for the central nervous system and mesenchymal stem cells in the biology of BMI, and a role for mesenchymal stem cells and adipocytes for body fat distribution (Locke *et al.*, 2015; Shungin *et al.*, 2015). Consistent with this, GWAS for volume of visceral white adipose tissue (VWAT) and subcutaneous white adipose tissue (SWAT) (measured using imaging biomarkers) found that adipose tissue development and differentiation are key pathways affected in fat distribution phenotypes (Chu *et al.*, 2017). This suggests that while common overall obesity (as measured via BMI) might be driven by altered feeding behaviour originating from

neuronal signals (consistent with monogenic forms of obesity such as SIM1 (Ramachandrappa *et al.,* 2013)), both BMI and body fat distribution are governed by biological mechanisms originating in adipose tissue and there must be adipose-intrinsic molecular processes that regulate obesogenic WAT expansion.

The first obesity locus identified through GWAS was published in 2007, when SNPs within intron 1 of the fat mass and obesity associated (*FTO*) gene were found to be strongly associated with BMI (Frayling *et al.*, 2007; Scuteri *et al.*, 2007). Originally found to associate with T2D, when corrected for BMI, this association disappeared, suggesting that the effect of *FTO* on T2D risk was due to increased body weight. It was found that those homozygous for the *FTO* risk allele (A-allele variant rs9939609) were approximately 3kg heavier than individuals with the protective allele, with approximately 44% of the population harbouring the risk allele. Since then, GWAS have repeatedly confirmed that the strongest genetic association with risk to polygenic obesity arises from SNPs in intronic regions within the *FTO* locus, independent of sex, ethnicity or age (Frayling *et al.*, 2007; Locke *et al.*, 2015; Scuteri *et al.*, 2007; Yang *et al.*, 2012). In addition to being strongly associated with increased BMI, risk variants in *FTO* are associated with a number of adipose tissue related traits, including increased risk for metabolic syndrome (Locke *et al.*, 2015), adiposity (Locke *et al.*, 2015), WHR (weak but significant) (Heid *et al.*, 2010) and increased SWAT mass (most significant predictor) (Fox *et al.*, 2012).

Over 90% of disease-associations identified through GWAS map to the non-coding genome and do not directly alter protein coding sequences (Zhang *et al.*, 2014). Instead, GWAS loci are enriched in regulatory elements as predicted by histone modifications and DNase accessible chromatin (ENCODE Project Consortium, 2012), suggesting that risk variants may disrupt *cis*-regulatory networks and more broadly speaking, that common obesity risk is mediated via gene modulation or fine-tuning. Regulatory elements are highly cell type- and tissue-specific and therefore functional approaches require the access of relevant human samples. Furthermore, non-coding GWAS signals characteristically harbor dozens of

significant SNPs in linkage disequilibrium (LD), adding further interpretational challenges. Not surprisingly, until today GWAS variants have very rarely been resolved into causal variants or the mechanism by which they affect disease susceptibility. This included the 89 BMI-associated variants in *FTO* intron 1 and 2. Due to the complexity that accompanies most GWAS signals as outlined above, the causal variants in *FTO* and their mechanistic underpinning have only recently started to be unravelled (Claussnitzer *et al.*, 2015; Smemo *et al.*, 2014; Stratigopoulos *et al.*, 2016). Important questions to answer include the identification of causal variants, their regulatory circuitries, the relevant tissues, and the important developmental stages when these variants exert their effect. To explain the mechanisms of disease-associations arising from GWAS, it is therefore essential to understand the chromatin biology of a gene regulatory network together with the relevant tissue-specific knowledge of cell development and function. Establishing the relevant context is crucial in identifying how a variant disrupts or modulates cellular and organismal biology.

1.2 Adipose tissue development and expansion in obesity

While GWAS have revealed a large number of associations in obesity and related traits, we are just at the beginning of unravelling their mechanistic underpinnings. A body of evidence suggests that genetic risk for obesity and associates traits, particularly for body fat distribution, is heavily driven by mesenchymal lineage cells and that adipose tissue is a key effector of disease risk (Dahlman *et al.*, 2016; Shungin *et al.*, 2015; Voight *et al.*, 2010). Almost half of the 96 candidate genes in loci associated with fat distribution have been linked to abdominal SWAT parameters (in female subjects), including fat cell size and number (morphology) and/or adipose tissue function (lipolysis or insulin-stimulated lipogenesis) (Dahlman *et al.*, 2016). To understand the genetic contribution to disease risk, we need to understand the processes that govern WAT development and expansion *in vivo*, including the origin of different WAT depots, their contribution to disease risk, and the way they respond to obesogenic stimuli. On top of that, the sexual dimorphism unravelled for body fat distribution, but not overall obesity (as measured via BMI), remains to be explained.

In humans and mice, WAT is mainly found in two distinct depot types: subcutaneous and visceral (intra-abdominal) adipose tissue. SWAT can further be divided into deep and superficial SWAT, which relates to its anatomical location in the body. Furthermore, excess circulating fat can deposit ectopically, with the major sites of surplus fat deposits being the liver (Bille *et al*., 2012) and skeletal muscle (Fonvig *et al*., 2012). Whether muscle cells take up lipids or whether there is a distinct set of adipocyte-like progenitors within muscle that can develop into intra-muscular fat cells is unknown *(*Arrighi *et al.*, 2015; Uezumi *et al.*, 2014; Wang *et al*., 2017).

1.2.1 White adipose tissue distribution

Excess energy is stored as triglycerides within WAT, leading to the expansion of the tissue (Spiegelman & Flier, 2001). Although overall obesity is associated with cardiometabolic disease risk, it is believed that the distribution of excess fat into SWAT and VWAT depots contributes differentially to this risk. Excess VWAT has been linked to increased metabolic disease risk across ethnicities (Coutinho *et al.*, 2011; Emerging Risk Factors Collaboration *et al.*, 2011; Lee & Lee, 2014; Nazare *et al.*, 2012; Phillips & Prins, 2008; Pischon *et al.*, 2008) whereas elevated SWAT mass has no effect or potentially protective properties on cardiometabolic health (Misra *et al.*, 1997; Snijder *et al.*, 2003; Tran *et al.*, 2008). Specifically, increased SWAT is associated with improved blood pressure, plasma lipid profiles, insulin sensitivity, and atherosclerosis (Appleton *et al.*, 2013; Heitmann & Lissner, 2011; Manolopoulos *et al.*, 2010; Snijder *et al.*, 2003). Together these reports suggest that the mechanisms that govern body fat distribution and depot-specific expansion of WAT have direct implications on cardiometabolic health.

Intrinsic differences in VWAT and SWAT have traditionally been explained partly by their anatomical location in the body and the resulting implications this has on venous drainage of the tissue (Wajchenberg, Giannella-Neto, da Silva, & Santos, 2002). However, we now know that VWAT and SWAT display an array of differences including gene expression of important

development and patterning genes such as KLFs (Cantile *et al.*, 2003; Cohen *et al.*, 2014; Gesta *et al.*, 2006; Grove *et al.*, 2010; Karastergiou *et al.*, 2013; Yamamoto *et al.*, 2010), developmental lineage (Berry *et al.*, 2014; Chau *et al.*, 2014; Gesta*et al*., 2007; Krueger *et al.*, 2014; Palmer & Clegg, 2015; Sanchez-Gurmaches & Guertin, 2014), metabolic features (Karastergiou *et al.*, 2013; Tran *et al.*, 2008), levels of adipokine secretion, lipolysis rate, mitochondrial content (Siersbaek *et al.*, 2012) and tendency to develop inflammation (Wajchenberg *et al.*, 2002). Furthermore, there are differential fat distribution patterns between females and males. Males tend to accumulate fat in visceral depots, whereas pre-menopausal women preferentially accumulate SWAT (Palmer & Clegg, 2015). In humans, this differential fat distribution pattern has been described as “apple-shaped” and “pear-shaped” obesity, for males and females respectively (Gesta *et al.*, 2007). In post-menopausal women, WAT distribution resembles the male visceral obesity pattern, suggesting that sex hormones drive the differential WAT expansion (Palmer & Clegg, 2015). Nonetheless, there is currently little known about the cellular and molecular mechanisms mediating depot- and sex-specific WAT growth and/or development.

1.2.2 Differences in WAT plasticity across depots in obesity

There are two distinct mechanisms by which WAT can expand *in vivo*, which are increasing size of existing adipocytes via increased lipid filling (hypertrophy) or increasing cell number by formation of new adipocytes (hyperplasia). Storing excess energy in existing adipocytes will lead to enlargement of cells and can trigger inflammatory responses. Obese individuals with T2D typically have larger adipocytes in both SWAT and VWAT (Fang *et al.,* 2015), whereas distributing fat into numerous small cells has in general been linked to decreased systemic disease risk. Enlargement of adipocytes, but not overall adipose tissue mass, is associated with increased risk of developing T2D and is further linked to glucose intolerance and hyperinsulinaemia (Weyer *et al.*, 2000). Mature adipocytes are post-mitotic and therefore “healthy WAT expansion” in form of adipocyte hyperplasia depends on recruitment

and differentiation of adipocyte precursors that are located around the vasculature in WAT (Berry *et al.*, 2014).

The turn-over of adipose tissue is estimated to be 10% per year, based on carbon-14 “birth-dating” experiments (Spalding *et al.*, 2008). Briefly, Spalding *et al.* (2008) established the dynamics of adipocyte turnover in adults by analysing carbon-14 that had integrated into genomic DNA during nuclear bomb tests during the cold war. These analyses revealed that approximately double the amount of adipocytes are turning over annually in obese compared to lean individuals, which suggests that obese subjects have twice as many adipocytes (Spalding *et al.*, 2008). Following substantial weight loss in obese individuals, adipocyte number remains higher, implicating that obesity-induced hyperplasia in adult life has a lasting effect on adipose tissue homeostasis. The development of obesity and expansion of WAT is further accompanied by increased angiogenesis, most likely to maintain oxygen supply to the increased demand of the growing tissue. It has been reported that this response is triggered by macrophages that are recruited following local hypoxia, and that hypoxia is the major microenvironmental trigger of blood vessel growth in adipose tissue (Michailidou *et al.*, 2015). Mice that lack proinflammatory cytokine TNFα in adipose tissue present with reduced angiogenesis and are resistant to adipose tissue expansion following a HFD (Wernstedt-Asterholm *et al*., 2014), suggesting that inflammation and angiogenesis play a critical role in obesogenic WAT expansion. Intriguingly, in addition to adipogenesis, angiogenesis is one of the pathways enriched in GWAS loci identified for body fat distribution (Shungin *et al.*, 2015).

Male VWAT has been reported to expand mainly through hyperplastic growth in humans and mice (Arner *et al.*, 2010; Jeffery *et al.*, 2015; Wang *et al.*, 2013), whereas male SWAT seems to be resistant to hyperplasia following dietary challenges such as a high-fat-diet (HFD) (Jeffery *et al.*, 2015; Wang *et al.*, 2013). To study the development and growth of different WAT depots *in vivo*, Wang *et al.* (2013) created the AdipoChaser mouse model. Briefly, the AdipoChaser mouse expresses lacZ selectively in adipocytes upon induction of

doxycycline in the food. This is achieved by generating mice that harbour three transgenes: i) transgene for the tet-on transcription factor rtTA under the control of the adiponectin promoter ii) transgene for doxycyclin-reponsive TRE-Cre transgene, and iii) the transgenic cassette Rosa26-loxP-stop-loxP-lacZ. This triple transgenic mouse allows for an inducible, permanent labelling of mature adipocytes. Using this system, Wang *et al.* (2013) found that the main contributor to SWAT and VWAT expansion in the first month of a HFD is hypertrophy. A prolonged exposure to the obesogenic environment resulted in *de novo* adipogenesis specifically in VWAT, but not in SWAT (Wang *et al.*, 2013). This is consistent with a recent report by Jeffery *et al.* (2016), where 8-week long exposure to a HFD lead to hypertrophy in both SWAT and VWAT, but only VWAT underwent additional hyperplastic growth (Jeffery *et al.*, 2015). Interestingly, in young animals a HFD leads to SWAT and VWAT hyperplasia. In contrast, older mice lose the ability to expand SWAT via hyperplasia, suggesting that with age, SWAT progenitor plasticity declines (Kim *et al.*, 2014).

The mechanisms governing depot-specific pre-adipocyte recruitment and differentiation *in vivo* are currently unknown. A widely observed phenomenon, however, is that the adipogenic capacity *in vitro* varies drastically between adipocyte progenitors originating from different depots. Pre-adipocytes isolated from subcutaneous depots have a higher differentiation capacity compared to pre-adipocytes from visceral depots. The observation that progenitors from VWAT are more resistant to adipogenic stimuli could explain why VWAT is resistant to hyperplasia and instead increases preferentially through hypertrophic growth; as a result of the limited capacity of VWAT-derived adipocyte progenitor recruitment and differentiation (Pellegrinelli *et al.,* 2016). In SWAT, on the other hand, higher adipocyte precursor number and/or higher differentiation capacity could explain the hyperplastic growth of this tissue. The abundance of pre-adipocytes in SWAT has indeed been proposed to be a determinant of SWAT expandability in humans (Tchoukalova *et al.*, 2010) and the capacity for SWAT to undergo *de novo* adipogenesis in adult life seems to protect from metabolic disease risk. However, how an individual's genetic make-up contributes to the capability to expand WAT

and to respond to environmental triggers in the development of obesity is largely unknown. As discussed above, GWAS support the notion that body fat distribution has a strong genetic underpinning. For example, SNPs in T-box 15 (TBX15) are associated with WHRadjBMI (Shungin *et al.*, 2015) and TBX15 is a developmental transcription factor (TF) that is expressed more in SWAT than VWAT and is a regulator of adipocyte differentiation and function (Gesta *et al.*, 2011; Yamamoto *et al.*, 2010).

1.2.3 Differences in WAT expansion between sexes

As briefly mentioned above, obesogenic WAT expansion exhibits sexual dimorphism, and this has direct implications in associated comorbidities and health risks. A recent study by Jeffery *et al.* (2016) has shown that a HFD challenge in mice leads to sex-specific patterns of adipogenesis and that this response contributes to the differential fat distribution observed between females and males (Fig. 1.2). 8 week-long HFD feeding triggered a hyperplastic response in male VWAT, but not SWAT, whereas in female mice both VWAT and SWAT showed increased adipogenesis following a HFD (Jeffery *et al.*, 2016). Jeffery *et al.* (2016) then found that the lack of hyperplastic growth in male SWAT was due to the lack of oestrogen, as adipocyte progenitor transplant studies demonstrated that the fat depot niche mediated the tissues' adipogenic response to obesogenic stimuli. Rather than adipocyte-intrinsic mechanisms, the depot-specific differences in HFD-induced hyperplasia were due to the adipocyte microenvironment in which the progenitors resided, suggesting that pre-adipocytes are plastic and can modify their behaviour depending on cues in their direct surroundings (Jeffery *et al.*, 2016). Human data supports this notion, as it was shown that there is no sex-specific difference in adipogenesis of pre-adipocytes originating from SWAT, however, substantially more adipocyte progenitors are believed to reside in female SWAT than male (Tchoukalova *et al.*, 2010). In conclusion, sex-specific pre-adipocyte activation and eventually WAT expandability and distribution are mediated via the sex-specific niche the tissue resides in.

1.2.4 Developmental origin of WAT

In humans, adipose tissue develops during the second trimester of pregnancy (Pellegrinelli *et al.*, 2016). Mature adipocytes derive from adipocyte precursors that reside at the stromal vasculature within WAT, which themselves originate from mesenchymal stem cells. The mechanisms orchestrating pre-adipocyte differentiation are well understood, however, not much is known about the embryonic origin of adipocytes. Lineage-tracing studies to date suggest that adipocytes originate from the mesoderm and the neural crest in the early developing embryo (Billon *et al.*, 2007).

The AdipoChaser mouse model, allowed for the first time to track adipocytes through development and to visualize the temporal pattern of WAT formation in subcutaneous and visceral depots (Wang et al., 2013). Wang *et al.* (2013) revealed that VWAT develops postnatally and adipocytes in this depot keep forming until sexual maturation, whereas adipocytes that make up SWAT differentiate during E14-E18 of embryogenesis. Interestingly, no additional adipocytes were formed outside this window. There was no difference in adipose development between sexes. This study very elegantly uncovers the discrete developmental timeframes of SWAT and VWAT formation *in vivo*.

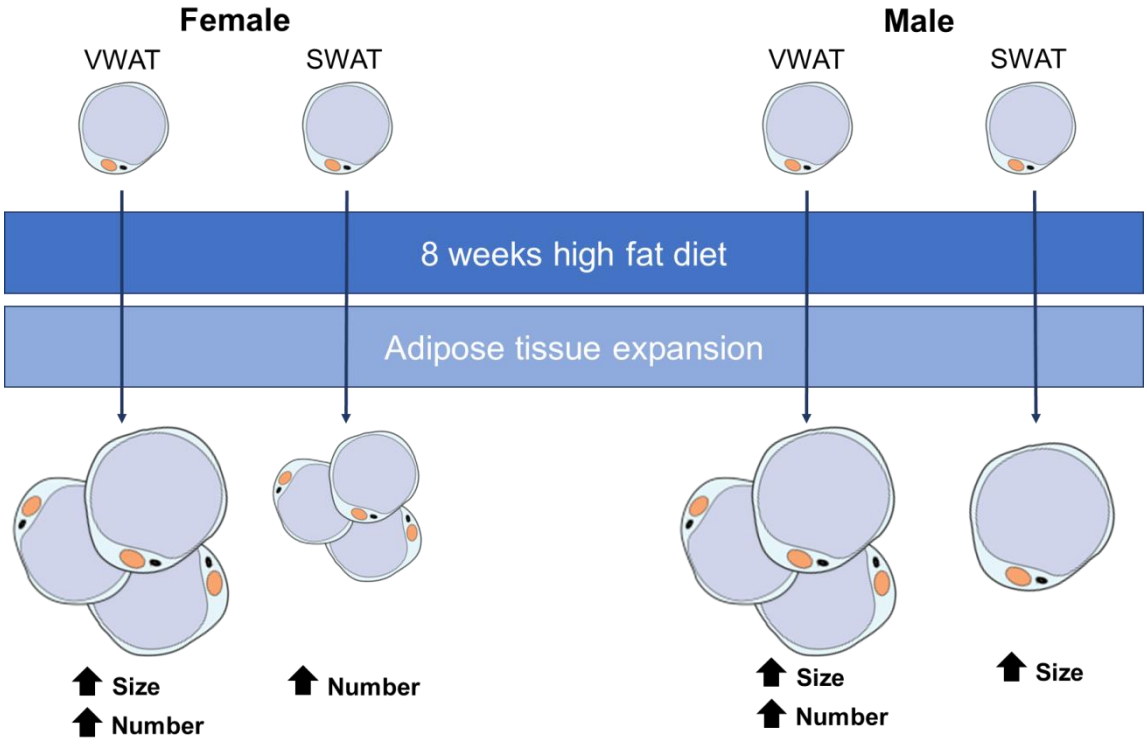


Fig. 1.2 | The sex- and depot-specific nature of HFD-induced obesity in mice.
Studies by Jeffery et al. (2016) and Wang et al. (2013) revealed a unique pattern of HFD-induced adipose tissue expansion in different depots and between the sexes. Female mice on a HFD for 8 weeks showed adipose tissue hyperplasia in VWAT and SWAT, but only VWAT also underwent hypertrophic growth. In male mice, both VWAT and SWAT displayed increased hyperplasia, however, only VWAT underwent additionally hyperplastic growth.

WAT browning

Brown adipose tissue (BAT) is a functionally and anatomically distinct class of adipose tissue and is characterised by small, multi-locular adipocytes with high amounts of mitochondria and the capacity to produce heat (Harms & Seale, 2013). White adipose tissue is plastic and can be triggered to become more brown-like by acquiring some of the properties of brown fat, a process termed “browning” or “beiging” of WAT. There is currently evidence for both, *trans*-differentiation of mature white adipocytes into brown-like ones or recruitment and differentiation of inducible brown pre-adipocytes residing within WAT depots (reviewed in Berry *et al.*, 2014). Adipocyte browning and accordingly elevated thermogenic capacity within WAT is triggered by sympathetic activation or cold-exposure. Although SWAT is the predominant depot that undergoes browning, there is also evidence for browning to occur in VWAT (Kim *et al.*, 2016). Both stimuli (adrenergic and cold) trigger a cascade of gene expression changes in adipocytes that lead to increased mitochondrial uncoupling, which is largely dependent on increased UCP1 protein expression. Since adipocyte browning is linked to reduced metabolic disease risk and obesity, it has been considered as being of therapeutic potential (Giordano *et al.*, 2016).

1.3 *FTO* locus regulatory circuitry

Due to the complexity of GWAS signals, progress has been slow in explaining the mechanistic underpinnings of disease-associations. The challenges are to identify causal variants, effector transcripts and their spatial, temporal, environmental, and epigenetic context. Once a variant is found to have a certain function, the question of causality and whether the mechanism discovered accounts in total for the disease association remains to be answered. Progress in identifying causal mechanism for the *FTO* association was certainly hampered by the complexity of the signal and our limited understanding of how the non-coding genome contributes to disease risk. There is currently a lack of comprehensive maps connecting distal elements that harbor risk SNPs with their target transcripts in relevant tissues and developmental time-points. Also, efforts in extensive fine-mapping of

risk associations are necessary to narrow down disease associations to their likely causal variants that can then be experimentally and functionally dissected (Liu *et al.*, 2014). As a result, many studies were performed with the assumption that the gene closest to a GWAS signal accounts for the disease association, including for BMI-associated variants in *FTO* (Church *et al.*, 2009, 2010, Gulati *et al.*, 2013, 2014; Karra *et al.*, 2013; McMurray *et al.*, 2013; McTaggart *et al.*, 2011; Merkestein *et al.*, 2015; Zhang *et al.*, 2015; Zhao *et al.*, 2014). Historically, candidate target genes in close proximity to a GWA signal were selected for *in vitro* and *in vivo* manipulations and functional characterization was undertaken on germline and/or tissue-specific knock-out or knock-in models (Yazdi, Clee, & Meyre, 2015). Although this approach has led to important discoveries in individual gene function, it has potentially led to many causal genes being overlooked, such as in the case of the *FTO* locus (Claussnitzer *et al.*, 2015; Smemo *et al.*, 2014). Advancements in methodology, mainly chromatin biology and genome editing, together with informatics integration have recently yielded new insight into the relationship between non-coding *FTO* risk-variants, target gene expression and disease phenotype.

The first report suggesting that SNPs in *FTO* are within a region regulating *IRX3* was published by Ragvin *et al.* (2010). They found that the 47 kb linkage disequilibrium (LD) block that harbors the obesity-associated variants rs8050136, rs9939609, rs1421085, and rs17817449 within *FTO* contains several highly conserved non-coding elements (HCNEs). Ragvin *et al.* (2010) used enhancer reporter assays in zebrafish to test expression pattern of the most deeply conserved HCNEs in the *FTO* LD block. They found that both elements drove expression patterns that are consistent with that of *IRX3*, a gene that is over 500 kb distal to *FTO*. One of the tested elements drove reporter expression in the kidneys; the second element drove GFP expression in the notochord, both of which are domains of IRX3 gene expression in development. Ragvin *et al.* (2010) conclude that HCNEs within the *FTO* LD block are most likely regulating *IRX3*. However, none of the two HCNEs tested harbored any disease-associated variants. Interestingly, when Ragvin *et al.* (2010) repeated their

experiment using elements that contained risk variants rs1421085 and rs9939609, they observed that those elements drove reporter expression in the pancreas. In conclusion, although Ragvin *et al.* (2010) did not explicitly identify any functional risk SNPs or directly link a specific variant to an effector transcript, their findings provide evidence that the LD block of 47 kb within *FTO* contains evolutionary conserved regulatory elements that likely act on the distal gene *IRX3*.

To directly study the DNA interactions formed within the *FTO* locus and to decipher its *cis*-regulatory circuitry (including *RPGRIP1L, FTO* and *IRX3*), Smemo *et al.* (2014) used an array of chromosome conformation capture (3C) techniques in different cell types from different species. Performing 4C-seq in whole mouse embryos (E9.5) and in adult (8 weeks) mouse brains, they found that the promoter of *Irx3* forms broad-spread long-range interactions spanning almost 2 megabases and including strong interactions with BMI-associated regions in *FTO*. However, the *FTO* promoter was mainly interacting with regions close to its promoter and in mouse embryos, this included *FTO* intronic regions. The interactions between *IRX3* and *FTO* obesity-associated regions were further confirmed using publicly available (ENCODE) and generated chromatin interaction data sets in adult mouse brains (3C), MCF-7 cells (ChIA-PET), human fibroblasts (Hi-C) and zebrafish embryos (4C-seq). Taken together, these results clearly point towards DNA looping between the *IRX3* promoter and obesity-associated *FTO* intronic regions in human, mouse, and zebrafish genomes. Screening through chromatin annotation maps revealed that *FTO* risk variants are within regulatory elements, making it interesting to hypothesize that allelic variants possibly disrupt transcription factor binding sites (TFBSs) and enhancer activity, which could result in altered expression of their target gene(s). eQTL mapping in human cerebellum and hypothalamus then confirmed a weak, but statistically significant association between 11 BMI-associated SNPs and *IRX3*, but not *FTO*, expression. In summary, Smemo *et al.* (2014) provide evidence for noncoding risk variants in *FTO* to drive brain-specific modulation of *IRX3*. Unfortunately, *IRX5* (or any additional gene within topologically associated domain

(TAD) boundaries) was outside their analysis window of genomic interactions and therefore escaped from their study design.

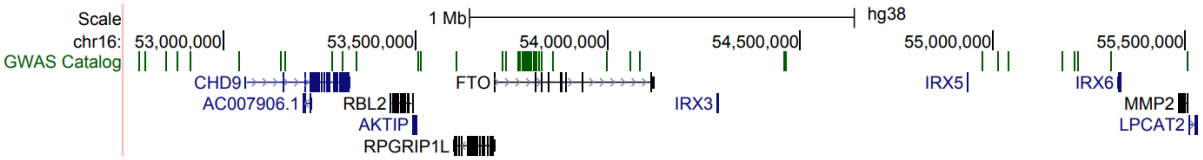


Fig. 1.3 | *FTO* Locus.
Screenshot from the UCSC Genome Browser (Kent *et al.*, 2002).

It became increasingly clear that profiling genomic interactions and integrating bioinformatics information in GWAS dissection is key to unlocking regulatory complexity. To predict *cis*-regulatory variants that have high probability to affect gene expression, Claussnitzer *et al.* (2014) developed a method that uses the power of integrating cross-species analysis of TFBS patterns with *cis*-regulatory modules (CRMs). Their discovery tool, called Phylogenetic Module Complexity Analysis (PMCA), can be used to systematically identify *cis*-regulatory variants within GWAS risk loci that overlap conserved TFBSs. Variants within these conserved modules are likely to disrupt or alter the affinity of TFs to their respective sites which could explain underlying disease association. Of the 89 *FTO*-intronic SNPs in high LD, the variant with highest regulatory probability score was rs1421085. To experimentally validate whether this variant has an effect on *cis*-regulatory activity, Claussnitzer *et al.* (2015) performed luciferase assays of the risk and non-risk allele in relevant cell types (INS-1 pancreatic β-cells, differentiated 3T3-L1 adipocytes, C2C12 myocytes, and Huh7 cells) and revealed that this variant alters allele-specific enhancer activity only in adipocytes, once more stressing the regulatory dependency of tissue and cell type (Claussnitzer *et al.*, 2014).

A comprehensive study examining a GWAS signal concerns the systematic dissection of the *FTO* regulatory circuitry in human adipose tissue (Claussnitzer *et al.*, 2015). Using chromatin annotation maps across 127 cell types (Roadmap Epigenomics Consortium *et al.*, 2015), Claussnitzer *et al.* (2015) showed that the BMI-associated region within *FTO* intron 1 contains a potent super-enhancer which is stretched in mesenchymal stem cells and adipocytes. Hi-C data in myofibroblasts suggested that the *FTO* LD block is within a TAD that harbours 8 genes, indicating that – given the relevant epigenomic context - *FTO* intronic elements could be in genomic contact with any of those. In other words, regulatory alterations within this TAD could in theory affect gene expression of any gene within its boundaries. Using primary human cells derived from risk- and non-risk carriers, Claussnitzer *et al.* (2015) revealed that the *FTO* obesity risk genotype activates both *IRX3* and *IRX5* expression in adipocyte precursors, and that this activation is lost in mature tissue, positing

the developmental aspect of this genetic risk association. Genome editing of the identified causal variant rs1421085 from risk towards non-risk achieved the reversal of gene expression changes, which suggested that this variant is responsible for the effect on *IRX3* and *IRX5* expression.

Notably, the distances between the adipocyte super-enhancer containing the GWAS signal and the *IRX3* and *IRX5* gene promoters are 516 kb and 1164 kb, respectively. *FTO* and *RPGRIP1L* promoters are closer to the BMI-associated variants, however, neither of those genes were altered in adipocytes of risk allele carriers. Global analysis of The Genotype-Tissue Expression (GTEx) project data (Ardlie *et al.*, 2015) suggests that for all GWAS traits investigated, the nearest gene is rarely implicated as the target (Barbara Engelhardt, personal communication).

Following on from identifying a causal variant, the target genes and the relevant cell type, Claussnitzer *et al.* (2015) went on to pinpoint an upstream regulator of rs1421085 and the downstream cellular and organismal consequences of alteration of this variant as discussed in the next section (Fig. 1.4). Their study showcases the use of genomic information in a functional setting and how it can guide variant, tissue and target prioritization and dissection.

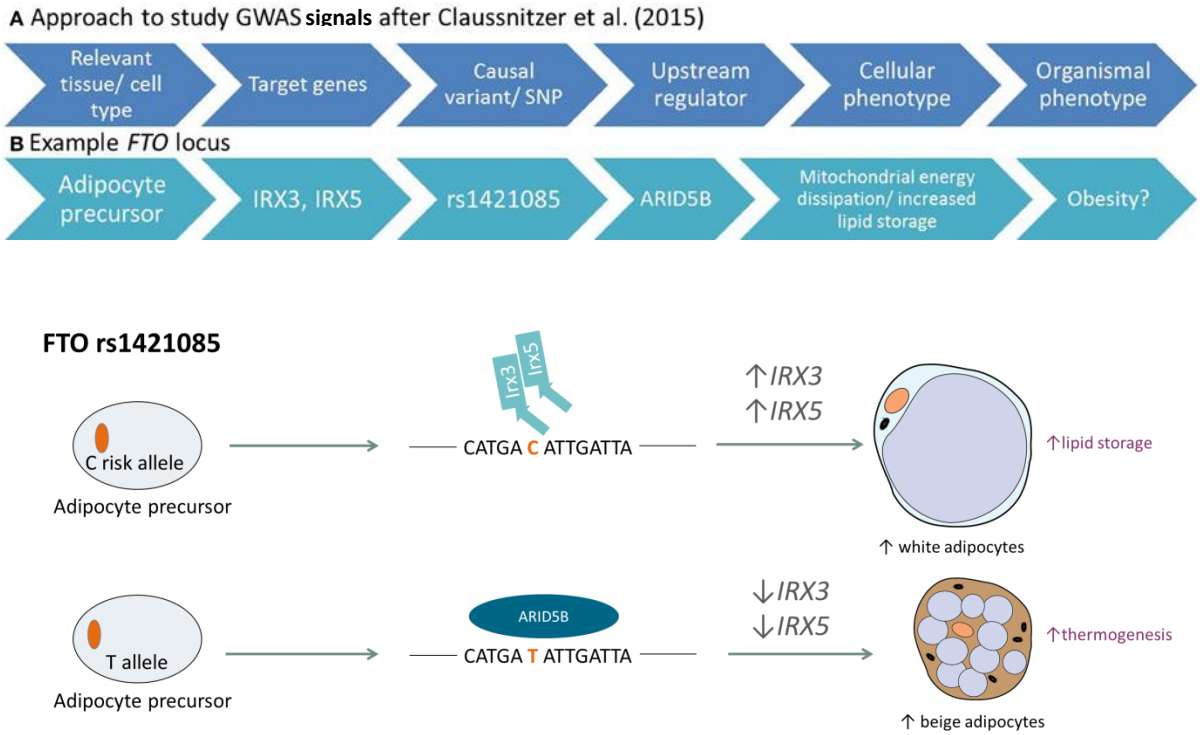


Fig. 1.4 | Risk variant rs1421085 and adipocyte browning.
Showcased with their work on the *FTO* locus, Claussnitzer *et al.* (2015) present a new strategy to decipher the function of non-coding disease-associated variants identified by GWAS. They present a model that deciphered causal variant, target genes, target cell type, upstream regulator and the cellular and organismal phenotypic consequences. Claussnitzer *et al.* established that the 10 kb enhancer element contains variant rs1421085 in a conserved motif for ARID5B. Carriers of *FTO* risk variant rs1421085 have a motif disruption caused by this variant that leads to decreased binding of ARID5B to this element, resulting in increased expression of *IRX3* and *IRX5* specifically in pre-adipocytes. Elevated *IRX3* and *IRX5* expression in risk C allele carriers then shifts the development of adipocytes toward the "white program," characterized by increased lipid storage in the mature fat tissue. In contrast, the protective T allele causes pre-adipocyte to develop into brown-like adipocytes, apparent via increased mitochondrial function and thermogenic capacity.

Finally, rs1421085/rs8050136 have recently been proposed to selectively alter *AKTIP*, *RPGRIP1L* and *FTO* expression in human iPSC-derived neurons (Stratigopoulos *et al.*, 2016). Stratigopoulos *et al.* (2016) cultured primary fibroblasts from individuals harboring heterozygous or homozygous alleles for the risk or protective alleles at rs1421085/rs8050136 (both variants are in perfect LD in these individuals), turned these into iPCS and subsequently into neuronal cells. Per risk allele, *AKTIP, RPGRIP1L* and *FTO* mRNA expression was reduced by approx. 15%, 25%, and 30%, respectively, suggesting a dose-dependent effect of the risk alleles on target gene expression. They propose that the rs8050136 is within a DNA-binding motif for the cut-like homeobox 1 (CUX1) and while the protective allele is preferentially occupied by CUX1 isoform P110, and acts as a transcriptional activator of *RPGRIP1L* and *FTO*, the risk allele is bound by CUX1 isoform P200, which acts as a transcriptional repressor for *RPGRIP1L* and *FTO in vitro* (Stratigopoulos *et al.*, 2014, 2016). Additionally, the C risk allele at rs1421085 was proposed to have a higher affinity for P200, whereas the protective T allele preferentially binds P110 (Stratigopoulos *et al.*, 2016). Stratigopoulos *et al.* (2016) propose a model in which the *FTO* obesity-risk alleles rs8050136 and rs1421085 alter CUX1 binding affinity that results in downregulation of hypothalamic *RPGRIP1L/FTO*, reduced leptin responsiveness and ultimately increased appetite and food consumption.

From the 8 genes within interaction domains at the *FTO* locus, there has certainly been some degree of evidence for the majority of these to be implicated in disease risk. Gene expression analysis suggested that *FTO/FTM* are altered in fibroblasts of rs8050136 risk allele carriers (Stratigopoulos *et al.*, 2008), rs8050136 is linked to altered *RBL2* expression in lymphocytes (Jowett *et al.*, 2010), and rs9939609 has been reported to alter *FTO* expression in blood and fibroblasts (Berulava & Horsthemke, 2010). Table 1.1 summarizes the available publications addressing the question of function of BMI-associated non-coding regions in *FTO* intron 1. The relationship between the SNPs in table 1.1 (rs1421085, rs8050136, rs9939609, rs9930506) is illustrated in Fig.1.5.

Table 1.1 | Current literature concerning the function of BMI-associated regions in *FTO* intron 1.

Variant(s)/ Region	Implicated target tissue(s) (organism)	Implicated target gene(s)	Evidence	Reference
rs8050136	Fibroblasts (human)	*FTO/FTM*	Silencing of predicted regulator CUTL1	(Stratigopoulos *et al.*, 2008)
Region containing rs1421085 and rs9939609	Pancreas (zebrafish)	*IRX3*	Reporter assays	(Ragvin *et al.*, 2010)
rs8050136	Lymphocytes (human)	*RBL2*	eQTL	(Jowett *et al.*, 2010)
rs9939609	Blood (human)	*FTO*	Ratio of allelic transcript levels	(Berulava & Horsthemke, 2010)
rs8050136	Hypothalamic neurons (mouse), N2A neuroblastoma (human)	*FTO, RPGRIP1L*	Inferred from EMSAs and reporter assays	(Stratigopoulos *et al.*, 2011)
rs9930506	Cerebellum, hypothalamus (human, mouse)	*IRX3*	4C, eQTL	(Smemo *et al.*, 2014)
Regions containing rs1421085	3T3-L1 adipocytes (mouse)	-	PMCA[1], Reporter assays	(Claussnitzer *et al.*, 2014)
rs1421085	Primary adipocytes (human)	*IRX3, IRX5*	Epigenomic annotations, Gene expression and genome editing of SNP	(Claussnitzer *et al.*, 2015)
rs1421085, rs8050136	iPSC-derived neurons (human)	*FTO, RPGRIP1L, AKTIP*	Gene expression	(Stratigopoulos *et al.*, 2016)

[1]Phylogenetic Module Complexity Analysis

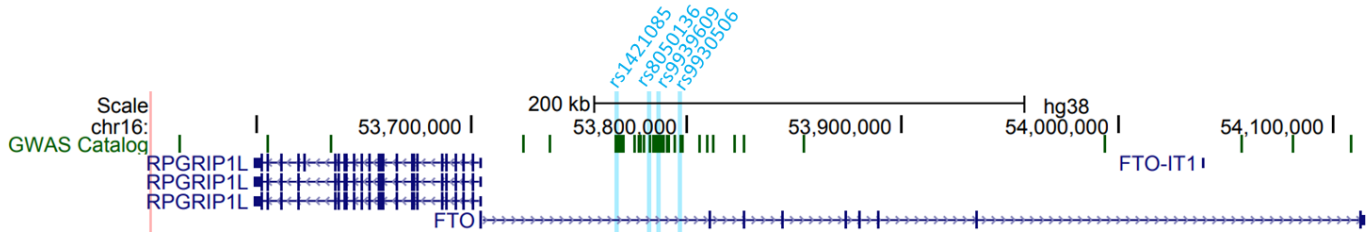


Fig.1.5 I Relationship between rs1421085, rs8050136, rs9939609 and rs9930506.
Screenshot from the UCSC Genome Browser (Kent *et al.*, 2002). Blue vertical lines represent location of SNPs mentioned in Fig.1.1.

In summary, multiple lines of evidence suggest that several genes in the *FTO-IRX* interval fall within a single regulatory block and that *IRX3* expression relies on an extensive *cis*-regulatory circuitry that extends into *FTO* (Ragvin *et al.*, 2010; Smemo *et al.*, 2014). This evidence further includes the high density of deeply conserved non-coding elements (Ragvin *et al.*, 2010), as well as patterns of CCCTC-binding factor (CTCF) binding and chromatin interactions (Dixon *et al.*, 2012) at the *FTO-IRX* interval. Data from several studies suggest long-range interaction relationships between *FTO* risk alleles and the megabase-distant *IRX3* and *IRX5* transcripts (Claussnitzer *et al.*, 2015; Ragvin *et al.*, 2010; Smemo *et al.*, 2014). Additional evidence suggests a role for *FTO* and *RPGRIP1L* in the effect of *FTO* risk alleles (Berulava & Horsthemke, 2010; Stratigopoulos *et al.*, 2016). Adipose tissue and brain have both been implicated in being the affected organs, pointing towards a multi-organ and multi-variant involvement in explaining disease risk at the *FTO* locus.

1.4 *FTO* function

FTO is located on mouse chromosome 8 and human chromosome 16, spans around 418 kb and is encoded by nine exons. It is ubiquitously expressed with highest expression in the hypothalamus, a region crucial for control of food intake and energy homeostasis (Gerken *et al.*, 2007, Fredriksson *et al.*, 2008). *FTO* belongs to the AlkB family of non-heme Fe (II)/dioxygenases and although the function of FTO remains to be fully understood, the crystal structure of FTO revealed that it is a demethylase for single stranded DNA or RNA (Gerken *et al.*, 2007, Jia *et al.*, 2008) and binds 3-meT (Han *et al.*, 2010). *In vivo*, however, FTO has only been linked to demethylation of N6-methyladenosine (m^6A) in RNA (Jia *et al.*, 2011).

Since the discovery of *FTO*, animal studies in mice and rats have established that manipulations of *Fto* result in body weight changes that are inconsistent in direction (Church *et al.*, 2010; Stratigopoulos *et al.*, 2008, 2011; Tung *et al.*, 2010; Wang *et al.*, 2011). Furthermore, no reliable effects on body weight are detected in humans heterozygous for *FTO* null alleles (Meyre *et al.*, 2009). Our group has previously shown that mice overexpressing *Fto* by harbouring 2 additional alleles of *Fto* in the Rosa26 locus, termed FTO-4, have increased body weight compared to WT mice on a standard chow diet (Church *et al.*, 2010), which is due to increased fat mass as a result of increased food intake. These differences become more obvious when mice are on a HFD (45% Kcal fat). Global FTO-KO mice have an 'obesity resistant' phenotype (Fischer *et al.*, 2009; McMurray *et al.*, 2013) with a decrease in lean and fat mass, increased energy expenditure as well as high embryonic lethality and post-natal growth retardation. In contrast, adult-onset FTO-KO mice are rescued from growth retardation and present with increased fat mass and gonadal WAT (gWAT) weight (due to adipocyte hypertrophy) and a decrease in food intake (McMurray *et al.*, 2013). FTO-KO in the adult mediobasal hypothalamus by stereotactical injection of adeno-associated viral vectors encoding *Cre* recombinase could not fully recapitulate the effect of adult-onset loss of *Fto.* McMurray *et al.* (2013) observed a small decrease in food

intake and weight gain without having an effect on energy expenditure or body composition in hypothalamic FTO-KO. These findings suggest that although hypothalamic *Fto* impacts feeding behaviour, the effect of loss of *Fto* on body composition is due to actions elsewhere and possibly includes several tissues. It is interesting to speculate that the effect of *Fto* on controlling fat mass and food intake is mediated independently. Interestingly, there are reports indicating that global *Fto* hypomorphism affects body composition in the same direction as *Fto* overexpression (Stratigopoulos *et al.*, 2016). Stratigopoulos *et al.* (2016) measured body composition of 14-weeks old mice heterozygous for *Fto* exon 3 deletion and found that loss of one *Fto* copy results in approximately 16% heavier mice compared to WT animals on a chow diet. Remarkably, heterozygous deletion for *Fto* caused 71% more fat mass than WT mice. Together, these results suggest that although germline loss of both *Fto* copies results in severe loss of lean and fat mass in mice and humans (Boissel *et al.*, 2009; Daoud *et al.*, 2016; Gao *et al.*, 2010), *Fto* hypomorphism increases adiposity in mice (Stratigopoulos *et al.*, 2016).

Several tissue-specific mouse models of FTO-4 or FTO-KO have since been generated in our group (for adipose (Adipoq-Cre), pre-adipocytes and mesenchymal stem cells (Prrx1-Cre, Pdgfra-Cre) and brain (Sim1-Cre)) attempting to elucidate the functional role of *Fto* in relevant tissues without fully replicating the global *Fto* mouse models (unpublished).

Functionally, *FTO* has been linked to both adipose and brain. Fasting has been shown to downregulate *Fto* expression in the arcuate hypothalami of WT and leptin-deficient mice and administration of leptin in turn upregulated *Fto* in these animals, indicating that *Fto* may regulate appetite via the canonical leptin signalling pathway (Stratigopoulos *et al.*, 2014). Based on findings that MEFs derived from FTO-KO embryos showed decreased leucyl-tRNA synthetase (LRS) levels compared to WT MEFs, Gulati *et al.* (2013) postulated that FTO acts a cellular amino acid sensor (Gulati *et al.*, 2013).

Findings by us and others revealed that FTO-dependent demethylation activity regulates adipogenesis as decreasing levels of *Fto* in 3T3-L1 pre-adipocytes (Zhao *et al.*, 2014) or MEFs (Merkestein *et al.*, 2015) inhibited their potential to differentiate, whereas increased *Fto* expression enhanced this process (Merkestein *et al.*, 2015). Additional reports show that *Fto* deficiency results in upregulation of UCP-1 expression and mitochondrial uncoupling in adipocytes (Tews *et al.*, 2013). These findings have now been replicated and numerous studies report an effect of *Fto* on adipocyte development, thermogenic capacity and function in a number of different model organisms (Ben-Haim *et al.*, 2015; Chen *et al.*, 2016, 2017; Merkestein *et al.*, 2015; Ronkainen *et al.*, 2015, 2016; Tews *et al.*, 2013; Tews, Fischer-Posovszky & Wabitsch, 2011; Wang *et al.*, 2015; Zhang *et al.*, 2015; Zhao *et al.*, 2014). It is still unclear, however, whether these observations have any direct physiological implications in human obesity and whether they account for the BMI-association at *FTO*.

1.5 *RPGRIP1L* function

Retinitis pigmentosa GTPase regulator-interacting protein-1 like (*RPGRIP1L*) is located <100 bp upstream of *FTO* and transcribed in opposite direction than *FTO*. *RPGRIP1L* encodes a protein that localises at the transition zone of primary cilia (Liu *et al.*, 2011; Williams *et al.*, 2011) and ciliopathies such as Bardet-Biedl and Alström syndromes are associated with obesity in humans (Baker & Beales, 2009).

Mice homozygous for the null allele of *Rpgrip1l* are embryonically lethal whereas mice hypomorphic for *Rpgrip1l* are hyperphagic and fatter than their controls (Stratigopoulos *et al.*, 2014). The authors attributed the effect of increased food intake in heterozygous RPGRIP1L-KO mice to the finding that these mice had reduced number of ciliated cells in the hypothalamus and a diminished response to leptin in arcuate neurons after intraperitoneal administration (Stratigopoulos *et al.*, 2014).

Based on their finding that the BMI-risk alleles at rs8050136 and rs1421085 are associated with decreased expression of both *Rpgrip1l* and *Fto* in iPSC-derived neurons, Stratipogoulos

et al. (2016) crossed RPGRIP1L-KO and FTO-KO lines to examine the effects on weight gain and fat mass in a *Fto/Rpgrip1l* compound hypomorphism (Stratigopoulos *et al.*, 2016).

At 14-weeks of age, male compound heterozygous RPGRIP1L/FTO-KO mice fed standard chow had 130% increased total fat mass compared to WT mice, which was higher than for hypomorphic *Rpgrip1l* and *Fto* mice alone, which is evidence for *Rpgrip1l* and *Fto* to have additive effects on adiposity. To examine leptin sensitivity in these compound heterozygous animals, Startigopoulos *et al.* (2016) treated mice with intraperitoneal leptin at 5 weeks of age, before any body mass or composition phenotype was observed. Indeed, food intake was reduced in compound heterozygous RPGRIP1L/FTO-KO mice and this effect was solely due to the loss of one *Rpgrip1l* copy, as heterozygous FTO-KO animals did not present with altered feeding behaviour. Hyperphagia in compound heterozygous RPGRIP1L/FTO-KO was accompanied by increased expression of neuropeptide Y (*Npy*) and decreased expression of pro-opiomelanocortin (*Pomc*) in the hypothalamus, which was in accordance with diminished leptin sensitivity in these mice.

Finally, brain-specific (Nestin-Cre) adult-onset (tamoxifen injections at 12-weeks of age) Rpgrip1l hypomorphism resulted in 70% decreased *Rpgrip1l* expression in nestin-positive hypothalamic neurons compared to control animals and at 21-weeks of age, mutant animals had a 10% increase in body weight and a 60% increase in fat massdue to a 20% increase in food intake (measured as 24-hour cumulative) (Stratigopoulos *et al.*, 2016). Therefore, decreased levels of *Rpgrip1l* in the adult CNS are sufficient in upregulating food intake in mice. In summary, these data point towards distinct roles for *Rpgrip1l* and *Fto* in energy homeostasis; *Rpgrip1l* mediates food intake through the leptin pathway, whereas *Fto* controls appetite and energy expenditure in a leptin-independent fashion.

1.6 *IRX3* and *IRX5* function

IRX3 and *IRX5* are transcription factors that belong to the family of 6 Iroquois homeobox genes, all of which were first identified in 1997 to be involved in embryo development (Bosse

et al., 1997). Like all Iroquois homeobox family members, *IRX3* and *IRX5* were found to be expressed in highly specific patterns (Christoffels *et al.*, 2000; Houweling *et al.*, 2001). *IRX3* and *IRX5* have since been found to be involved in neurogenesis, where they are expressed in distinct spatio-temporal patterns (Cohen *et al.*, 2000). Further, *IRX3* is differentially up-regulated in female gonads during sex determination (Jorgensen & Gao, 2005) and pattern of *IRX3* directs nephron segment identity (Reggiani *et al.*, 2007) as well as heart morphogenesis and postnatal physiology (Gaborit *et al.*, 2012). Although *IRX3* and *IRX5* are found to be co-regulated (Houweling *et al.*, 2001), the function of these two genes has been described as being agonistic as well as antagonistic in cardiac development and function (Gaborit *et al.*, 2012). Cumulative, data suggest that *IRX3* and *IRX5* play a role in early embryonic organ development for a number of tissues derived from ectoderm, mesoderm and endoderm such as eye, limb, brain, neural tube, and branchial arches (Christoffels *et al.*, 2000; Houweling *et al.*, 2001). However, the role of *IRX3* and *IRX5* in the development of obesity or in metabolism has not been addressed yet and there is currently only limited literature describing roles for *IRX3* and *IRX5* in obesity. The next section is a summary of the available reports that link *IRX3* and *IRX5* to obesity.

To discover new genes involved in obesity-related adipose tissue function, Dankel *et al.* (2010) set out to profile global gene expression changes in abdominal subcutaneous adipose tissue obtained from individuals before and one year after bariatric surgery (Dankel *et al.*, 2010). In their study, Dankel *et al.* (2010) found that *IRX3* and *IRX5* expression are significantly lower in SWAT of obese individuals. One year after bariatric surgery and profound fat loss, the gene expression levels of *IRX3* and *IRX5* were restored to that of healthy controls. The fold-change post/pre operation was 2.1-fold and 1.5-fold for *IRX3* and *IRX5*, respectively. This suggests that *IRX3* and *IRX5* are transcription factors in the regulation of adipose tissue function (Dankel *et al.*, 2010).

Following on from their discovery that SNPs within *FTO* are functionally connected with regulation of *IRX3* expression, Smemo *et al.* (2014) metabolically characterised a mouse

homozygous for the global *Irx3* null allele (IRX3-KO). IRX3-KO animals present a striking metabolic phenotype independent of diet or gender, including decreased overall body size, glucose intolerance and insulin resistance, decreased overall body weight that is due to decreased fat mass to BW ratio and resistance to HFD-induced weight and fat mass gain (Smemo *et al.*, 2014). IRX3-KO mice have smaller adipose tissue depots, reduced adipocyte size and exhibit browning of WAT (Smemo *et al.*, 2014). Consistent with higher gene expression of browning markers (*Ucp1, Ppargc1a, Prdm16*, and *Cidea*) in gonadal WAT (gWAT), deficiency in *Irx3* lead to increased energy expenditure and elevated adipose levels of *Adrb3*, indicating higher responsiveness to sympathetic activation. Smemo *et al.* (2014) suggested that the body weight phenotype, including increased thermogenic capacity in WAT, was driven by hypothalamic neurons. In line with this idea, *Irx3* was found to be expressed in the arcuate nucleus and median eminence of the hypothalamus, two regions in the brain that are known to be critical for regulation of food intake and energy homeostasis (Kong *et al.*, 2012; Shen *et al.*, 2014). To test their hypothesis, they next set out to cross an Ins2-Cre allele with a dominant negative form of *Irx3* with the aim to generate a conditional, hypothalamus-specific KO allele. This model recapitulated every aspect of the global *Irx3* null, including increased energy expenditure and browning of WAT. Their data suggests that hypothalamic expression of *Irx3* regulates energy homeostasis and body composition (Smemo *et al.*, 2014).

As discussed previously, using an array of bioinformatics and genetics approaches, Claussnitzer *et al.* (2015) identified that the *FTO* risk allele disrupts TF binding of a transcriptional repressor (ARID5B) that leads to the activation of a potent enhancer in pre-adipocytes and consequently over-expression of *IRX3* and *IRX5*. Higher levels if *IRX3* and *IRX5* expression in pre-adipocytes of risk allele carriers then lead to decreased mitochondrial energy production and increased lipid accumulation in mature cells. This cellular phenotype was reversed following genome editing of rs1421085 from risk to non-risk in human pre-adipocytes. Decreasing the levels of *IRX3* and *IRX5* in human pre-adipocytes

caused a cell-autonomous shift towards increased mitochondrial energy dissipation and mitochondrial uncoupling (Claussnitzer *et al.*, 2015). This effect was maintained in mature adipocytes, where risk allele carriers showed decreased mitochondrial copy number, oxygen consumption rate and increased adipocyte size. At the organismal level, they show that aP2-Cre–driven knockdown of *Irx3* in mice leads to 57% reduced fat mass ratio and resistance to HFD-induced obesity. In comparison, the conditional Ins2-IRX3-KO generated by Smemo *et al.* (2014) lead to 19% reduced fat mass ratio, having a significantly smaller effect. In summary, the levels if *IRX3* and *IRX5* in pre-adipocytes were shown to regulate basic anabolic function of adipocytes and to determine the relative extent of fuel dissipation through increased mitochondrial uncoupling. Importantly, modulation of mitochondrial activity of human WAT is a cellular phenotype consistent with obesity and a process that has previously been suggested to have therapeutic potential (Harms & Seale, 2013).

Together, these results by Claussnitzer *et al.* (2015) suggest that a single-nucleotide change from T-to-C at rs1421085 can alter cellular metabolism in a direction that could account for the association between the first intron of *FTO* and obesity in humans. Additionally, this study proposes adipocyte precursors as one of the primary effectors of *FTO* locus activity.

1.7 Epigenomic conservation between human and mouse

As discussed previously, the vast majority of trait-associated variants identified by genome-wide efforts map to the non-coding genome and are enriched for regulatory annotations such as histone modifications and DNase accessibility (ENCODE Project Consortium, 2012). Characteristically, GWAS loci contain dozens of statistically significant variants in LD and are within TADs that harbor many potential effector genes (Dixon *et al.*, 2012; Gómez-Marín *et al.*, 2015). In the case of *FTO*, there is currently evidence for a number of variants and causal genes that might affect several tissues. In theory, mouse models that could assist in identifying functional variants, their target genes and relevant cell types and tissues and that capture their relevant spatiotemporal effects would be valuable in dissecting the mechanistic

underpinnings of human disease-associated sequence variants. However, whether modeling regulatory variation in mouse has potential will depend on the conservation of functional elements between human and mouse.

Large-scale genome-wide comparative analysis between mouse and human epigenomes have highlighted that *cis*-regulatory elements are (in contrast to *trans*-regulatory elements) the major site of species divergence (Vierstra *et al.*, 2014). At the same time, studies comparing the regulatory landscape between human and mouse have yielded new insight into pathways involved in diabetes and obesity and cross-species conservation has successfully been exploited and used as criteria in deciphering functional elements that are associated with metabolic diseases (Claussnitzer *et al.*, 2014; Dickel *et al.*, 2016; Multhaup *et al.*, 2015).

1.7.1 Sequence vs functional conservation

Differences in gene expression between and within a species are often mediated via changes in TFBSs which undergo rapid evolutionary turnover (Ward & Kellis, 2012). Gene regulatory regions in eukaryotes are believed to consist of complex pattern of co-occurring TFBSs, termed *cis*-regulatory modules (CRMs), which allows for TFs to bind in different combinations (Arnone & Davidson, 1997; Pennacchio *et al.*, 2006; Visel *et al.*, 2013). It has been shown that disease-associated variants within *cis*-regulatory regions can be identified by falling into conserved modules of co-occurring TFBSs that allow for combinatorial TF binding (Claussnitzer *et al.*, 2014). Furthermore, phylogenetic conservation of CRMs revealed that homeobox TFBSs are a *cis*-regulatory feature of T2D risk loci (Claussnitzer *et al.*, 2014).

1.7.2 Insights from the Mouse ENCODE Project

Comparing gene expression data from human and mouse revealed that there are sets of genes that cluster more by species and others more by tissue (Yue *et al.*, 2014). However, a later report that re-analyzed these data sets suggests that the level of gene expression is

dominated by tissue, not by species (Gilad & Mizrahi-Man, 2015), pointing towards conserved biological mechanisms of gene regulation within tissues.

Single cell transcriptomic analysis of human and mouse pancreatic alpha and beta cells revealed that overall, correlation of transcriptomes was high between species, while still highlighting the presence of cross-species differences (Xin *et al.*, 2016). This is consistent with a report by Breschi *et al.* (2016) that describes how transcriptomes demonstrate a variety of species-dominated and organ-dominated clustering. Notably, GWAS signals tend to be within organ-dominated clusters (Breschi *et al.*, 2016), an important observation when considering to model GWA signal in the mouse.

Some of the key findings that arose from the Mouse ENCODE project were that the *trans*-regulatory landscape, meaning TF-mediated gene regulation, is substantially more conserved between human and mouse than *cis*-regulatory networks (Stergachis *et al.*, 2014; Yue *et al.*, 2014). The majority of regulatory plasticity between human and mouse can be accounted for by divergence of the *cis*-regulome (Stergachis *et al.*, 2014; Yue *et al.*, 2014). However, the degree of *cis*-regulatory divergence differs between different types of elements and tissue contexts (Stergachis *et al.*, 2014; Vierstra *et al.*, 2014). Yue *et al.* (2014) revealed that there are sequence orthologous regions in the human genome for 66.7% of TFBSs and 79.3% of predicted mouse enhancers (based on regulatory annotations and histone modifications). Using reporter assays, Yue *et al.* (2014) were able to show that 61.5% of tested candidate mouse-specific enhancers also demonstrate enhancer activity in human embryonic stem cells (ESCs), suggesting a degree of functional conservation between human and mouse gene regulation. This degree of conservation allows for the speculation that mouse chromatin annotation maps could partially guide the identification of potential sites for functional characterization in mouse for human GWAS hits.

Yue *et al.* (2014) mapped 4265 human disease-associated SNPs onto the mouse genome and overlapped those regions with chromatin states of 15 mouse samples of different

tissues. They found that specific histone modifications were enriched for disease traits in relevant mouse tissues, or in other words, there was a trait-dependent link between human variants and specific histone modifications in relevant mouse tissues (Yue *et al.*, 2014). H3K4me1 was enriched for variants associated with urate levels and metabolites in mouse kidney; H3K36me3 was enriched for high-density lipoprotein (HDL) cholesterol and triglyceride levels in mouse liver. Cumulatively, 55% of variants overlapped with at least one histone modification mark in mouse (Yue *et al.*, 2014), suggesting that comparing human-mouse chromatin annotations could potentially be informative for human risk variant prioritization and/or functional dissection. Furthermore, disease-associated variants are found more frequently in TFBSs (Boyle *et al.*, 2012). Consistent with this, GWAS variants fall more often into conserved regulatory sequences in the genome that are occupied by orthologous TFs than expected by chance (Cheng *et al.*, 2014). Interestingly, this observation was dependent on the trait investigated. For Type 1 Diabetes, 13 out of 20 variants were in conserved TFBSs, however, none of the variants associated with pulmonary function were within conserved regulatory elements (Cheng *et al.*, 2014). Whether obesity-associated SNPs are enriched in conserved regulatory elements is yet to be determined.

Systematic comparisons between human and mouse regulome and transcriptome in adipose tissue (at different stages of development and in both sexes) are currently lacking. With the continuous increase in available global mouse data sets and the ongoing efforts of consortia such as ENCODE (ENCODE Project Consortium, 2012) and the Epigenomic Roadmap (Roadmap Epigenomics Consortium *et al.*, 2015), it will become possible in the future to draw conclusions about the degree of human-mouse conservation of TFBSs and enrichment of GWAS SNPs in adipose tissue. Mikkelsen *et al.* (2010) generated comparative histone modification maps from human adipose tissue and mouse 3T3-L1 at four time-points during differentiation and revealed that 15-30% of regulatory regions were shared between species, pointing towards largely species-specific gene regulation in adipose development. Importantly, this study compared the epigenome of a mouse cell line and primary human

tissue-derived adipocytes. The associated ontogenetic differences between the two could potentially affect full interpretation.

In summary, although the *cis*-regulatory landscape is the major site of cross-species divergence on a global level, human disease-associated variants fall more frequently into sites that are conserved between human and mouse for the majority of GWAS traits examined.

1.8 FTO-IRX locus – one of the most conserved regions in the genome

Non-coding regions in the *FTO-IRX* locus show a striking degree of evolutionary conservation that places them amongst the top 2% of comparably sized genomic regions (Smemo *et al.*, 2014). *In vivo* enhancer expression assays, together with CTCF and chromatin interaction data, support the idea of an ancient, extensive *cis*-regulatory circuitry at the *FTO* locus as well as the presence of deep, evolutionary conserved, non-coding elements that regulate *IRX3* (Bosse *et al.*, 1997; Claussnitzer *et al.*, 2014; Ragvin *et al.*, 2010; Smemo *et al.*, 2014; Visel *et al.*, 2007).

1.9 Approaches for establishing the function of regulatory variants

Non-coding variants can modulate *cis*-regulatory machineries through a number of mechanisms, such as DNA methylation, disrupting or creating TFBSs, altering DNA-DNA interactions and chromatin looping, or miRNA recruitment (reviewed in Zhang *et al.*, 2014). Continuous advancement in large-scale and high-throughput sequencing together with the development of new methodologies is advancing our ability to study and understand chromatin biology and genome function. Fig. 1.6 summarizes currently used experimental approaches that allow the dissection of the regulatory landscape. For example, TFBSs and chromatin states (histone modifications) can be assessed using ChIP-seq, chromatin interactions can be identified using chromosome conformation capture techniques (e.g. 3C, 4C, 5C, CaptureC, Hi-C, ChIA-PET, etc.), and open chromatin, indicative of active regulatory regions, can be studied using DHS-seq, FAIRE-seq or ATAC-seq (Buenrostro *et al.*, 2015;

Dekker, Marti-Renom, & Mirny, 2013; Hughes *et al.*, 2014; Knight, 2014) (Fig. 1.7). Additionally, computational tools are emerging that exploit the power of integrating genomic information to assess and predict the impact of disease-associated variants.

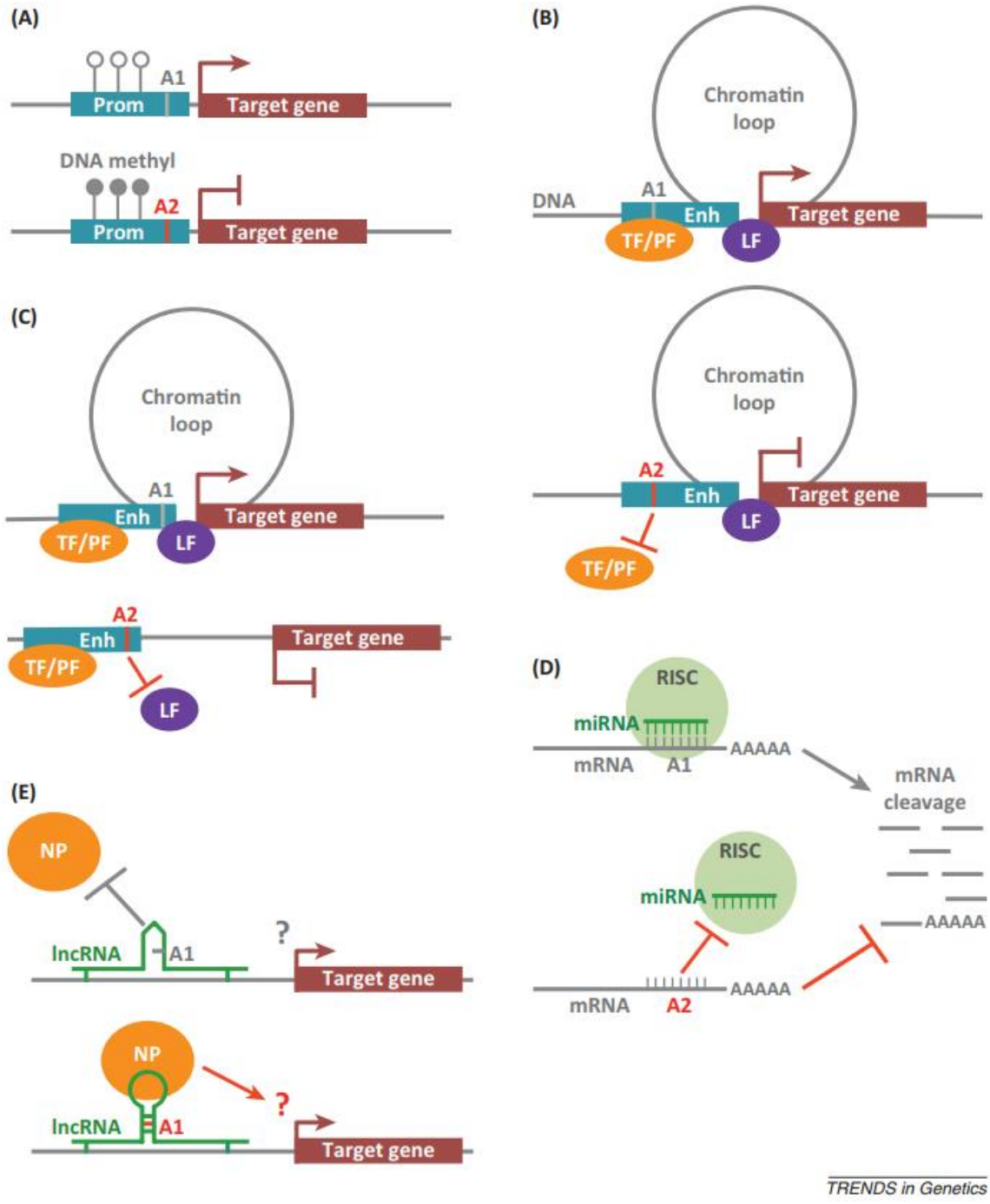


Fig. 1.6 | Potential mechanisms for non-coding disease-associated variants.
Depending on their location, SNPs could (A) interrupt TFBSs in promoters, or alter promoter methylation; disrupt binding sites in regulatory elements for (B) TFs or pioneering factors or (C) looping factors with the result of altered target gene expression; (D) alter miRNA recruitment; or (E) modulate RNA structure. Figure from Zhang *et al.* (2014).

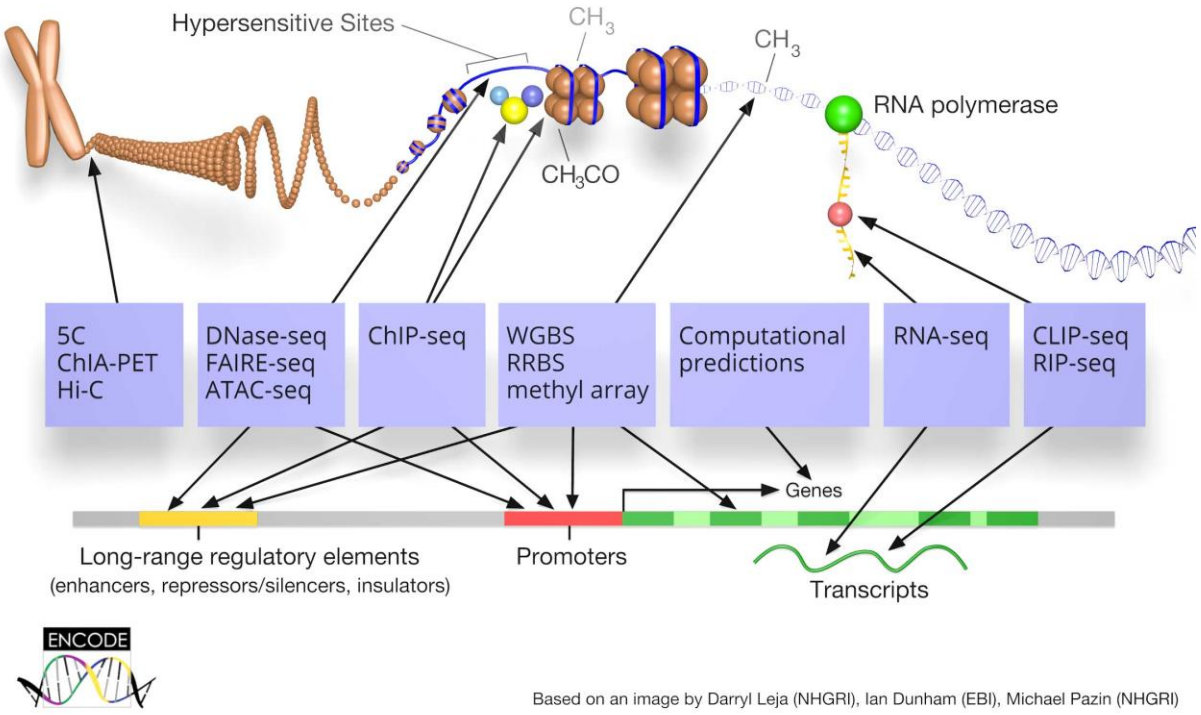


Fig. 1.7 | Methods to study the epigenome.
Figure from the ENCODE project website (Yue *et al.*, 2014).

For example, HaploReg (Ward & Kellis, 2012) and RegulomeDB (Boyle *et al.*, 2012) allow to predict changes in TF affinity caused by genetic risk variants, the GTEx portal (https://gtexportal.org/home/) (Ardlie *et al.*, 2015; Carithers *et al.*, 2015) can be used to predict target genes for genetic risk variants (eQTLs), and The Islet Regulome Browser (Pasquali, *et al.*, 2014) and Enlight (Guo, Conti, & Wang, 2015) can be used to visualize GWAS signals embedded into their epigenomic and cell type-specific context. These computational and genomics tools for the post-GWAS era are summarized in Flannick & Florez (2016) and Zhang *et al.* (2014). The list of data repositories and analysis platforms for GWAS hits is growing rapidly and by allowing the integration of GWAS results with genetic and epigenetic annotations, they are valuable instruments in the dissection of the gene regulatory networks that underpin genomic association signals.

Comparative analysis of such regulatory information in tissues relevant for T2D, namely adipose, muscle, liver and pancreas, revealed that risk variants for T2D are enriched in enhancer elements in pancreatic islets (Fuchsberger *et al.*, 2016; Pasquali *et al.*, 2014). Identification of variants that overlap epigenomic annotations in pancreatic islets together with fine-mapping could in turn reveal the most likely causal variant at a given locus. For T2D, the list of examples of loci for which variants are known to affect regulatory activity in the relevant tissue (pancreatic islet) is increasing and includes *MTNR1B, JAZF1, CDC123/CAMK1D, ZFAND3, WFS1, KCNK16/17* and *KCNQ1* (Gaulton, 2017). Additionally, *cis*-expression mapping has clearly identified biological candidates for a number of T2D GWAS signals, including *DGKB, ADCY5, WARS* and *ZMIZ1* in pancreatic islets, where the major impact of these SNPs is to reduce insulin secretion (van de Bunt *et al.*, 2015); and for *KLF14,* an eQTL provided evidence for an effect of risk alleles on *KLF14* expression in adipose tissue – and fittingly KLF14 is associated with insulin resistance and hyperlipidemia (Small *et al.*, 2011). According to a recent review that marked the 10th anniversary of GWAS discoveries (Visscher *et al.*, 2017), it is possible to assign a convincing effector transcript to

around one-third of the 100 T2D risk loci. These target genes can be taken forward into mechanistic dissection and translation.

Demonstrated by their work on the *PPARG* (Claussnitzer *et al.*, 2014) and *FTO* (Claussnitzer *et al.*, 2015) loci, Claussnitzer *et al.* (2014, 2015) present a model for translating information arising from GWAS that integrates such genomic data with experimental discovery. Applying PMCA to the T2D-associated *PPARG* locus for variant prioritization followed by a series of genetic approaches revealed that the rs4684847 risk variant interrupts a TFBS for PRRX1 and consequently affects *PPARG2* gene expression (5-fold reduction) and insulin sensitivity (Claussnitzer *et al.*, 2014). They show that rs4684847 (which is in perfect LD with Pro12Ala), over-compensates for the paradoxically increased activity effect of Pro12Ala on *PPARG2*, suggesting that at this locus, there are multiple variants at play per haplotype. Claussnitzer *et al.* (2015) next applied and extended their approach at the *FTO* locus, where they identified and validated the cellular function of the identified variant rs1421085. Their approach exploits the toolbox of methods that can be used to functionally dissect genetic risk associations and adds to the methods used previously, such as gene knockdown or overexpression, mouse models of candidate genes, gene expression or biochemistry approaches; methods that have been used to explain risk associations for *FTO, ATM, CDKAL1, SLC2A2* and *GCKR* (Beer *et al.*, 2009Church *et al.*, 2009, 2010; Flannick & Florez, 2016; McCulloch *et al.*, 2011; McMurray *et al.*, 2013; Wei *et al.*, 2011; Zhou *et al.*, 2011). Another example of a comprehensive study that elucidates the mechanistic underpinnings of a diabetes risk locus has recently been published for the *ARAP1* (*CENTD2*) and *STARD10* locus (Carrat *et al.*, 2017). This study very elegantly combines genetic fine-mapping, epigenomic annotation data, eQTL mapping, and animal models to identity the causal variant, the effector gene as well as the downstream mechanism in the relevant context. Using 3C, Carrat *et al.* (2017) identified a 5kb region that is within an islet-specific stretch-enhancer and further went on to show that the rs140130268 risk allele selectively reduces *STRAD10* mRNA expression in β-cells which consequently

results in impaired insulin secretion. In contrast, a previous study has shown that that the T2D-risk and proinsulin-decreasing allele (C) for rs11603334 is associated with increased *ARAP1* mRNA levels (Kulzer *et al.*, 2014), and *ARAP1* in turn, has been linked to defects in insulin processing (Dimas *et al.*, 2014), providing evidence for two effector transcripts with relevant functional effects on glycaemic traits at this locus.

With the integration of new technologies in chromatin biology and genome editing, the complexity appears to slowly disentangle into clear pathways and mechanism for BMI-associated variants in *FTO*. However, examples of unravelling non-coding disease associations like *PPARG*, *FTO* and *STARD10* are currently sparse and there is certainly a need for new systematic approaches that have the potential to unravel disease associations of complex traits.

One example of a novel approach towards context-embedded systematic dissection of GWAS candidate genes has recently been published for T2D using the human β-cell line EndoC-βH1 (Thomsen *et al.*, 2016). Risk for developing T2D is largely mediated via the pancreatic islet, which is supported by physiological (e.g. candidate effector transcripts affect glycemic traits) (Dimas *et al.*, 2014) as well as epigenetic (e.g. disease-associated variants fall within islet-specific regulatory elements) evidence (Gaulton *et al.*, 2015; Pasquali *et al.*, 2014). Thomsen *et al.* (2016) developed a gene silencing screen that tested for four β-cell-relevant phenotypes (insulin secretion by several stimulants) and proliferation, across 300 positional candidates (genes located within 1 Mb of association signal) selected from 75 T2D loci. For more than half of these loci, they could identify likely target genes based on the finding that silencing of those genes affected at least one of the β-cell phenotypes examined (Thomsen *et al.*, 2016). Functional follow-up studies for *ARL15, ZMIZ1, THADA,* and *HNF4A* further confirmed a function of these genes in core islet biology. Although this approach does not directly link non-coding variants to target genes, but rather infers them based on function, it provides a powerful strategy for target prioritization. Expanding such approaches by for example including all possible target genes within TAD boundaries and integrating

genomic annotations would further strengthen the power of this approach. Undoubtedly, such systematic studies provide insight into disease pathology and will certainly guide future mechanistic studies of identified genes.

Interestingly, in the case of the *FTO* locus, there is currently evidence for four out if the potentially eight possible effector transcripts to be involved in adipocyte and/or brain function. Alterations in *FTO, IRX3* and *IRX5* have all been liked to altered adipocyte development and function and *RPGRIP1L, FTO* and *IRX3* have been functionally linked to actions in the hypothalamus. However, whether all of these are causal for the BMI-association is highly questionable. Indeed, there have been reports suggesting that functionally related genes (defined as genes that are involved in the same pathways) tend to cluster (Hurst *et al*., 2004; Lee & Sonnhammer, 2003), suggesting that it is not unlikely that there are “neighborhoods of genes” that for example affect adipogenesis or insulin secretion in the relevant cell type. To avoid miss-identification of effector transcripts of disease risk, assigning candidacy on the basis of known or discovered biology cannot undoubtedly establish causality of a given gene for a given association.

The discoveries made in GWAS follow-up studies to date clearly unraveled the function of individual target genes and variants. However, they rarely exclude the effect of other SNPs and whether a number of variants act together and in what tissues they are implicated in often remains a mystery. The continuous growth of tissue- and cell type-specific genomic data sets (including ATAC, ChIP, CRISPR/Cas9 libraries, 3C-based techniques) in both sexes together with further fine-mapping of loci will ease functional investigation of disease-associated non-coding variants in the future.

1.9.1 Towards new avenues in mouse model engineering

With our current knowledge of the regulatory-dependency of context (cell type, developmental stage, etc.) and the many layers of complexity of most GWAS signals, it becomes increasingly clear that deciphering the underpinnings of an association signal in

the relevant human cell type is critical. On an organismal level, the function of a candidate gene has traditionally been studied by generating global knockout and overexpression models (Yazdi *et al.*, 2015). To draw conclusions on the tissue-specific nature of genetic variation, CRE-drivers have been used to achieve tissue-specific target gene manipulation. Although a powerful tool in principle, there is currently a lack of reliable pre-adipocyte-specific CRE lines that could be used to functionally dissect identified target genes in cases of pre-adipocyte-specific signals. New methods are required that overcome some of those challenges. Recent progress in genome engineering approaches utilizing CRISPR/Cas9 (Aida *et al.*, 2014) makes it possible to conveniently edit any regulatory sequence of interest. In theory, altering TFBSs and regulatory elements in the mouse has the potential to create *in vivo* models of human risk associations that i) are cell type-specific, ii) alter all target genes; iii) alter target genes at the relevant level and direction; and iv) alter target genes at the relevant time of development. In many ways, however, we are currently just at the beginning of understanding how such models could be exploited, what the translational utility will be and what information would be gained. The success of such an approach will depend on the conservation of the regulatory circuitry of a given GWAS signal. From the current literature, I surmise that conservation is less on the nucleotide level, but on clusters of regulatory modules. With the work presented in this thesis, I set out to generate mouse models that harbor mutations in *Fto* intronic regions - orthologous to human rs1421085 and surrounding sequences - with the aim of using these models to further unravel the mechanism of this human obesity-associated variant. The present thesis will focus predominantly on one of those mouse lines named DEL82, which harbors a 82-nucleotide deletion around orthologous region rs1421085.

1.10 Thesis Aims and Objectives

The aim of this thesis is to decipher the *FTO* regulatory circuity in adipocytes using novel mouse models that harbour mutations at the mouse orthologous region around human rs1421085 and to phenotypically characterise the *in vivo* consequences of altered regulatory activity. Further, this thesis aims to unravel the function of genes within the *FTO* locus, namely *IRX3* and *IRX5*, in adipocyte development and function. The objective of this thesis is to answer the following questions:

1. Is the *FTO* regulatory circuitry functionally conserved in mouse? **Chapter 3**
2. What are the *in vivo* phenotypic consequences of DEL82 (mice that harbour an 82 nucleotide deletion in a regulatory element around orthologous region rs1421085, located in *Fto* intron 1) on the organismal level? **Chapter 4**
3. What are the target genes and tissues of the *Fto* mouse regulatory element around orthologous region rs1421085 and what is the spatio-temporal expression pattern of identified effector transcripts? **Chapter 5**
4. What are the cellular and molecular mechanisms of altered regulatory activity in DEL82? **Chapter 5**
5. What is the function of *Irx3* and *Irx5* in pre-adipocytes? **Chapter 6**
6. What are the consequences of T-to-C editing at mouse orthologous at rs1421085 (FTO-RS1421085) on body weight, fat mass and candidate gene expression in pre-adipocytes? **Chapter 7**
7. What is the spatio-temporal expression pattern of identified effector transcripts in FTO-RS1421085? **Chapter 7**
8. Are there previously unknown sex- and depot-specific effects of *Fto* intronic regulatory elements? **Chapter 5 + 7**

2 Materials and Methods

2.1 Cell Biology Techniques

2.1.1 Adipocyte Precursor Isolation, Culture and Differentiation

Mouse primary adipocytes were isolated essentially as previously described (C. D. Church, Berry, & Rodeheffer, 2014). Briefly, mouse WAT depots were excised from 6-10 weeks old C57BL/6NJ (B6N) mice, individual adipose tissues were excised from 6-10 weeks old B6N mice and placed in PBS. Tissues were minced and digested in 5ml per depot digestion buffer (sterile Hank's Balanced Salt Solution (HBSS; H8264), 0.8 mg/ml collagenase type 2 (Worthington Biochemical Corporation, NJ, USA; LS004174), 3% bovine serum albumin (BSA) (with fatty acids) and incubated in a 37°C water bath for 60-75min with shaking every 10min by hand. When digestion was complete, tubes were centrifuged for 3min at 300xg to separate floating mature adipocytes from the supra-vascular fraction (SVF), containing adipocyte precursors. The supernatant containing the floating adipocyte fraction was removed and the cell pellet re-suspended in pre-warmed growth media consistent of DMEM GlutaMax (# 10569010 DMEM, high glucose, GlutaMAX™ Supplement, pyruvate) supplemented with 10% Fetal Bovine Serum (Gibco 10082-147) and 1% Penicillin-Streptomycin (5,000 U/mL) (Gibco, 15070063) and grown at 37°C and 5% CO_2. The cell solution was subsequently filtered through a 40µm nylon mesh and plated on a 10cm dish. Media was replaced the next day.

For differentiation experiments, attached pre-adipocytes were trypsin-treated using TrypLE™ Express Enzyme (1X) (Gibco 12605-010), inactivated with growth media and counted using either thr Scepter™ 2.0 Handheld Automated Cell Counter (Millipore) with 60µm tips (Millipore) or the Countess Automated Cell Counter (ThermoFisher). 100K cells/ml were plated in 12-well or 6-well dishes and grown to confluence. 2 days post-confluence, pre-adipocytes were induced to differentiate using induction medium (DIFF1) containing growth

medium supplemented with 0.5mM 3-Isobutyl-1-Methylxathine (IBMX) (Sigma, I5879), 1µM Dexamethasone (Sigma, D2915), and 5µg/ml human insulin (Sigma, I9278). A summary table outlining the ingredients of the differentiation cocktails can be found below. At day 4 of differentiation, media was changed to maintenance media (DIFF2) containing growth medium supplemented with insulin only. Media was changed every 2 days and differentiation complete at day 7-9 post adipogenic induction. Where stated, β-adrenergic stimulation was performed by overnight treatment of cells with 1µmol/l isoproterenol (Sigma, I6504).

Reagent	**Stock solution**	**Stock added to 50ml growth media***		**Final concentration**	
		DIFF1	**DIFF2**	**DIFF1**	**DIFF2**
Insulin (Sigma, I9278)	10mg/ml stock solution stored at 4°C, stable for 3-4 months	25µl	25µl	5 µg/ml	5 µg/ml
IBMX (3-Isobutyl-1-Methylxathine) (Sigma, I5879)	0.5M IBMX stock solution was prepared by dissolving 100mg of IBMX powder in 900µl DMSO. Stock solution was aliquoted in 100-300µl and stored at -20°C. aliquots are stable for up to 4 months	50µl	-	0.5mM	-
Dexamethasone (Sigma, D2915)	2mM stock solution was prepared by dissolving 25mg of dexamethasone powder in 31.8ml 100% EtOH. Solution was stored at -20°C	25µl	-	1µM	-
β-adrenergic stimulation					
(−)-Isoproterenol hydrochloride (Sigma I6504)	1mM stock was prepared by adding 0.0123g isoproterenol powder (247.72g/mol) to 50ml sterile H_2O	50µl	50µl	1µM	1µM

* Growth media consistent of DMEM, supplemented with 10% FBS and 1% PS antibiotics

A schematic overview of a typical primary pre-adipocyte isolation and differentiation protocol is shown below.

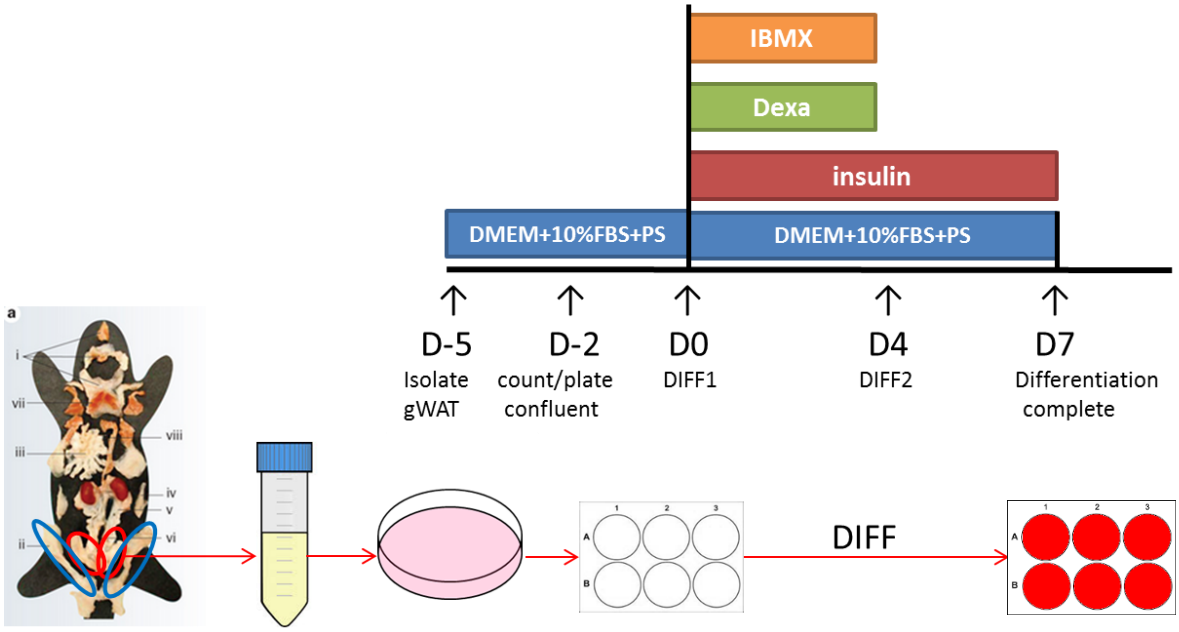


Fig. 2.1 | Schematic overview of a typical primary pre-adipocyte isolation and differentiation protocol.
Image of mouse adipose tissues on the left was taken from (Tran & Kahn, 2010) with the following legend:

Subcutaneous white adipose tissue
i) Anterior white adipose tissue
ii) Inguinal white adipose tissue (iWAT)
Visceral white adipose tissue
iii) Mesenteric adipose tissue
iv) Retroperitoneal white adipose tissue
v) Perirenal adipose tissue
vi) Gonadal white adipose tissue (gWAT)
Brown adipose tissue
vii) Interscapular brown adipose tissue
viii) Mediastinal adipose tissue

2.1.2 MEFs Isolation and Differentiation

To obtain MEFs, embryos were harvested at E12.5−E14.5 and placed in ice-cold PBS. Individual embryos were removed from the placenta and membranous sac. Following removal of the head, limbs, liver, heart and inner organs, individual embryos were dissociated in 0.25% trypsin (Gibco, with 25200056 or without 25050030 EDTA) using a sterile scalpel. Once minced, embryos were incubated in a 37°C water bath for approximately 15min and subsequently cultured in growth media consistent of DMEM

GlutaMax (# 10569010 DMEM, high glucose, GlutaMAX™ Supplement, pyruvate) supplemented with 10% Fetal Bovine Serum (Gibco, 10082-147) and 1% Penicillin-Streptomycin (5,000 U/mL) (Gibco, 15070063) at 37°C and 5% CO_2. Differentiation was performed as described in section 'Primary pre-adipocyte differentiation'.

2.1.3 3T3-L1 Culture and Differentiation

3T3-L1 pre-adipocytes were maintained in DMEM GlutaMax (# 10569010 DMEM, high glucose, GlutaMAX™ Supplement, pyruvate) supplemented with 10% newborn calf serum (Gibco 26010-074) and 1% Penicillin-Streptomycin (5,000 U/mL) (Gibco, 15070063) at 37°C and 5% CO_2. For differentiation experiments, attached pre-adipocytes were trypsin-treated using TrypLE™ Express Enzyme (1X) (Gibco 12605-010), inactivated with growth media and counted using either thr Scepter™ 2.0 Handheld Automated Cell Counter (Millipore) with 60µm tips (Millipore) or the Countess Automated Cell Counter (ThermoFisher) with . 60-80K cells/ml were plated in 12-well or 6-well dishes and grown to confluence. Two days post confluence, differentiation was induced as described in section 'Primary pre-adipocyte differentiation'.

2.1.4 hWAT Culture and Differentiation

Primary SVF cells were received from Prof Yu-Hua Tseng, Ph.D. at Harvard medical School, Joslin Diabetes Center, One Joslin Place, Boston, MA 02215. The cells were previously isolated and immortalized from human brown fat (hBAT) and white fat (hWAT) of a female subject, aged 56 with a BMI = 30.8. Culture and differentiation was performed following the protocol from the originating lab as described in Xue *et al.*, 2015. Briefly, pre-adipocytes are cultured in DMEM GlutaMax (# 10569010 DMEM, high glucose, GlutaMAX™ Supplement, pyruvate) supplemented with 10% Fetal Bovine Serum (Gibco 10082-147) and 1% Penicillin-Streptomycin (5,000 U/mL) (Gibco, 15070063) at 37°C and 5% CO_2. For differentiation, cells were treated with 0.25% trypsin (Gibco, with 25200056 or without 25050030 EDTA), counted as previously mentioned and 50K cells per well were seeded in a 24-well plate. Cells were grown until they are fully confluent. Once cells reached confluency, differentiation was

induced by adding freshly prepared induction medium to cells. The ingredients of the differentiation cocktail can be found below. Induction media was replaced with fresh induction media every three days for 21-24 days, until fully differentiated. The reagent stocks for differentiation medium were prepared as follows:

- Biotin: 0.08g in 10ml 0.1N NaOH (30mM)
- Insulin: 10 mg/ml in 0.01M HCl
- Pantothenate: 8.5mM in H2O (500x)
- Dexamethasone: 0.002g in 1ml EtOH (2mg/ml; 5mM)
- Isobutyl methylxanthine (IBMX): 50 mM stock in 0.1M KOH
- 3,3′,5-Triiodo-L-thyronine (T_3): add 1.0ml of 1.0N NaOH to 1mg, gently swirl to dissolve the powder and add 152.6ml of sterile medium (10μM)
- Indomethacin: Dissolve 447.25mg Indomethacin in 10ml EtOH (0.125M). Heat to 75°C.
- Aliquot stock solutions and store them at -20°C. Avoid multiple freeze/thaw cycles.

Human pre-adipocyte induction medium was prepared as described below.

Ingredients	Molarity	Volume
DMDM/H		48.5ml
FBS	2%	1ml
Pen/Strep	0.1mM, 0.06mM	500μl
Biotin	33μM	50μl
Human Insulin	0.5μM	15μl
Pantothenate	17μM	100μ
Dexamethasone	0.1μM	1μl
T3	2nM	10μl
IBMX	500μM	500μl
Indomethacin	30μM	12.5μl

2.1.5 RNA Silencing and Plasmid Transfection

Pre-adipocytes were plated in 12-well or 6-well plates at a cell density of 80-100K or 160-200K cells per well, respectively, to reach 60-70% confluence the next day. If silencing was performed in other vessel sizes, cells were seeded to be 60-70% confluent at time of transfection. Silencing was performed using Lipofectamine® RNAiMAX Transfection Reagent (ThermoFisher #13778075) and following the manufacturer's protocol. Briefly, Lipofectamine® RNAiMAX Reagent was diluted in Opti-MEM medium. At the same time, siRNA was diluted in Opti-MEM medium. Then, diluted siRNA was added to the diluted Lipofectamine® RNAiMAX reagent at a ratio 1:1 and incubated for 5min.

All silencing experiments were performed on 3 technical triplicates. Per 3 wells of a 6-well plate, 11.25µl of silencing oligo or negative control duplex (see table below for details) were diluted in 0.3ml of Opti-MEM medium (Invitrogen). At the same time, 45µl of lipofectamin RNAimax reagent (master mix for 6 wells) was diluted in 0.6ml of Opti-MEM medium as shown in the table below. The diluted lipofectamin RNAimax mix was mixed by vortex and then added to the tubes containing oligo mix and the complexes incubated at room temperature for 5min. During the incubation, cells were washed twice with PBS and 1.8ml DMEM GlutaMax containing 10% serum only (no antibiotics were added to medium) was added to each well. After 5min incubation, 200µl siRNA/Lipofectamine complex was gently added to each well. The plate was gently swirled and placed in a 37°C incubator at 5% CO_2 for 48 hours. Cells were then induced to differentiate following the standard differentiation cocktail or harvested for gene expression analysis and to assess knockdown efficiency. Silencing efficiency was compared between experiments to analyze data collected from repeated silencing experiments.

6 well plate	volume/ well	Mix and incubate at RT	Volume of complex	Growth medium (DMEM+10% NBCS)	Final volume (each well)	Final concentration of siRNA
siRNA oligo (20µM)	3.75µl	5min	200µl	1800µl	2ml	40nM
Optimem Medium	100µl					
Lipofectamin RNAimax	7.5µl					
Optimem Medium	100µl					

Silencing oligo	**Species**	**Cat #**	**Sequence (5'-3')**	**GC %**
siCon medGC siCon highGC	Mouse	12935-300 12935-400	for 45-55% GC for 55-70% GC	48 68
siIrx3 oligo 1	Mouse	10620318 Irx3MSS275126 10620319 Irx3MSS275126	CGCUUGGACCAACAGGCCUUUCCCA UGGGAAAGGCCUGUUGGUCCAAGCG	60
siIrx3 oligo 2	Mouse	10620318 Irx3MSS275128 10620319 Irx3MSS275128	GAUACCAGUACAUCCGCCCUCUCUA UAGAGAGGGCGGAUGUACUGGUAUC	52
siIrx3 oligo 3 (detects TV2 only)	Mouse	10620319 Irx3MSS275127 10620318 Irx3MSS275127	UGCCGAGCCGGAGAGUGGAACAGAU AUCUGUUCCACUCUCCGGCUCGGCA	64
siIrx5	Mouse	10620318 Irx5MSS225572 10620319 Irx5MSS225572	ACUAUGGCUCCUUCGGACAUCUUCA UGAAGAUGUCCGAAGGAGCCAUAGU	50

siIRX5	Human	10620318 HSS115689 10620319 HSS115689	CCAGUCUCAGCUAGACCUGUGCAAA UUUGCACAGGUCUAGCUGAGACUGG	52
siIRX3	Human	10620318 HSS128350 10620319 HSS128350	ACUGACGAGGAGGGAAACGCUUAUG CAUAAGCGUUUCCCUCCUCGUCAGU	52

2.1.6 Seahorse

Oxygen Consumption and Bioenergetics Profile was measured using the XF24 extracellular flux analyzer from Seahorse Bioscience. The protocol used in this assay was adapted from Gesta *et al.*, 2011. For this assay, pre-adipocytes were counted using the Countess (ThermoFisher) and 20K cells per well were seeded on to seahorse 24 well plates in 100µl of growth media and left to adhere overnight. The next morning, 100µl of growth media was added to each well. Two or three days later (depending on the depot of origin/ cell type), cells ere induced to differentiate within the seahorse plate following the adipogenic differentiation protocol as described previously. Each cell type was run in 10 replicates and four wells evenly distributed within the plate were left without cells to use for correction of temperature variation. When the cells were terminally differentiated at day 7-9 post adipogenic induction, the assay was performed. For isoproterenol stimulation experiment, cells were treated with 1µmol/l isoproterenol (Sigma, I6504) overnight (approx. 16-20h depending on time of assay) before the assay was performed. The evening before the assay, the seahorse XF-24 instrument cartridge was loaded with seahorse calibrant and placed in a CO_2-free incubator at 37°C overnight.

On the day of the assay, cells were washed in XF Assay Media, L-glutamine 2mM, sodium pyruvate 2mM, and glucose 10mM (pH was measured and adjusted to pH7.4 at 37°C). The seahorse plate containing the differentiated adipocytes was then incubated for at least 1

hour at 37°C in a CO_2-free incubator to allow CO_2 to diffuse out of solution. According to the manufactures protocol, the ports of the seahorse XF-24 analyser cartridge were then loaded with the following compounds:

- Port A: Oligomycin (complex 1 inhibitor)
- Port B: FCCP (carbonyl cyanide-p-trifluoromethoxyphenylhydrazone; mitochondrial uncoupler)
- Port C: Rotenone and Antimycin (inhibitors of electron transfer)

Before running the assay, the XF-24 instrument cartridge was calibrated.

For total oxygen consumption rate (OCR) measurements, the minimum OCR reading after Rotenone / Antimycin A treatment was subtracted from the initial untreated level, following the manufacturer's protocol. To directly measure mitochondrial thermogenesis, uncoupled respiration (proton leak) was measured by subtracting the minimum OCR level after Rotenone / Antimycin from the minimum level after oligomycin treatment. Oxygen concentrations were measured over time periods of 4 min with 2 min waiting and 2 min mixing.

The protocol for a standard bioenergetics profile is composed of basal mitochondrial respiration, ATP turnover, proton leak and mitochondrial respiratory capacity. First, OCR in basal conditions was determined and used to calculate the basal mitochondrial respiration. After this, 2 µM oligomycin was injected from the first port to inhibit ATP synthase, resulting in an accumulation of protons in the mitochondrial intermembrane space and a reduced activity of the electron transport chain (ETC). The resulting decrease in OCR reveals the respiration driving ATP synthesis in the cells, indicating ATP turnover. Residual oxygen consumption capacity can be attributed to the proton leak maintaining a minimal ETC and non-mitochondrial respiration. Next, 2 µM of the mitochondrial uncoupler FCCP was injected which results in an increase in OCR as the proton gradient across the inner mitochondrial membrane is dissipated and ETC resumed. This measurement reflects the maximal

mitochondrial respiratory capacity. Finally, 2 µM Rotenone / Antimycin A are injected to completely stop ETC activity and the OCR reading at this phase reflects non-mitochondrial respiration.

For 3T3-L1 siIrx3 experiments, data was normalized to the relative number of live cells in each well of the 24 well seahorse plate directly after the seahorse assay. Media was removed and each well was washed with DPBS twice. Pre-adipocytes were then simultaneously treated with 2 µM Calcein AM (live cells) and 4 µM EthD-1 (dead cells) (Invitrogen) followed by 15 min incubation at room temperature, protected from light. A Optima Plate reader (BMG labtech, Offenburg, Germany) was used to measure absorbance with filters for live cells (Calcein AM) excitation 485-12nm and emission 520-10nm and dead cells (EthD-1) excitation 520-10nm and emission 650-10nm.

Tom Agnew and Tom Nicol helped with the Seahorse experiments.

2.1.7 Oil-Red-O and DAPI Staining

Cells were washed with PBS and fixed in 4% PFA for 1h at room temperature. Cells were washed with ddH2O and the treated with 60% isopropanol for 5min. After additional washes with ddH2O, cells were stained with a filtered Oil Red-O (Sigma) solution (0.5% Oil Red-O in isopropyl alcohol) for 30min at room temperature (covered with aluminum foil under gentle rotation). The Oil Red O stain was removed with several washes of PBS until no dye came off the cells anymore and then cells were either stained for DAPI or NucBlue stain or directly visualized under an inverted microscope.

2.2 Gene Expression Analysis

2.2.1 RNA Isolation

Total RNA from all cells and tissues (adipose tissue and brain) was extracted using Trizol reagent (ThermoFisher, #15596026). For cells grown on a monolayer in a 12-well and 6-well plate, between 350ul – 1ml of Trizol depending on vessel size was added directly to the cells. Using a P1000, cells were scraped and washed of the wells and placed into a clean

Eppendorf tube. Homogenates were then incubated for 5min at room temperature. For tissues, 1ml of Trizol Reagent was added to 50-100 mg of tissue and homogenized using ceramic beads (Precellys) in a Precellys-24 automated homogenizer (Precellys). The following two kits and protocols were used for RNA extractions:

RNeasy Mini Spin Column Kit

200µl chloroform per 1ml Trizol was added and tubes vigorously shaken for 15 secs, followed by 3min incubation at room temperature. Tubes were then centrifuged at 13,000rpm for 15min at 4°C. Following phase separation, the colorless upper phase containing RNA was transferred to a clean tube and an equal volume of 70% EtOH added. After thorough mixing (brief vortex) to precipitate the RNA, the sample was transferred into a RNAeasy spin column and consecutive steps were following the manufactures protocol (RNeasy Mini Kit, Qiagen #74106). For DNase digestion, each sample was treated with use 10µl DNaseI stock plus 70µl Buffer RDD (1500 Kunitz units) for 15min at room temperature (RNase-Free DNase Set, Qiagen #79254). Contaminant DNA removal was performed between Buffer RW1 washes (350µl before and 350µl added just after genomic DNA digestion). RNA was eluted in 30µl of RNase-free water. Extracted RNA was quantified described in nucleic acid quantification.

Direct-zol™ RNA MiniPrep Plus Kit

100% EtOH was directly added to the samples in Trizol following a Direct-zol™ RNA MiniPrep Plus kit protocol (Zymo research #R2072). RNA was eluted in 30µl of RNase-free water. Isolated RNA was quantified described in nucleic acid quantification.

2.2.2 Nucleic Acid Quantification

Purified DNA and RNA were quantified using an Epoch Microplate Spectrophotometer (BioTek). 2µl of DNA or RNA solution was placed on the Take3 Micro Volume Plate (BioTek) and the absorbance of UV-light at 260nm and 280nm or at 260nm, 280nm and 230nm for

DNA and RNA, respectively, was measured. The nucleic acid concentrations were calculated from the measured absorbance at 260nm using the Beer Lambert Law. DNA or RNA purity was assessed using the 260nm/280nm ratio; a 260nm/280nm ratio of ~1.8 or ~2.0 was accepted as 'pure' DNA or RNA, respectively. The purity of RNA was additionally evaluated using the 260nm/230nm ratio and a 260nm/230nm ratio of 2.0 – 2.2 was accepted as 'pure'.

2.2.3 cDNA Synthesis

cDNA synthesis was performed using a High-Capacity cDNA Reverse Transcription Kit (ThermoFisher #4368814) according to the manufacturer's protocol. Briefly, 1 µg of RNA was reverse-transcribed by adding 10µl of 10ng/ul RNA to a master mix containing all reagents for the reverse transcription reaction listed in the following table:

Component	**Volume**
10X RT Buffer	2.0 µl
25X dNTP Mix (100 mM)	0.8 µl
10X RT Random Primers	2.0 µl
MultiScribe™ Reverse Transcriptase	1.0 µl
Nuclease-free H_2O	4.2 µl
Total per reaction	10.0 µl

A thermal cycler was programmed as follows:

Settings	**Temp.**	**Time**
Step 1	25°C	10 minutes
Step 2	37°C	120 minutes
Step 3	85°C	5 minutes
Step 4	4°C	∞

2.2.4 qRT-PCR

Gene expression was measured using quantitative real-time PCR real time fluorescence detection. TaqMan Gene Expression Assay reagents and TaqMan FAM dye-labelled probes (ThermoFisher, table below) were used to set up appropriate reactions according to the manufacturer's protocol. Briefly, per well, a 20µl reaction was prepared consistent of 10µl PCR Master Mix (2x) (ThermoScientific K0171), 1µl probe, 4µl nuclease-free H2O and 5µl cDNA (4ng/µl) to make a total of 20ng DNA/reaction. RT-PCR reactions were then analysed using an ABIPRISM 7500 Fast Real-Time PCR System (Applied Biosystems) using the manufacturer's settings (40 cycles). Data was normalized to house-keeping genes specific to the tissue or cell type. To determine the most suitable house-keeping gene to use, GeNORM analysis (kit from PrimerDesign) of 12 house-keeping genes was perform in advance. If not stated otherwise, *Canx* was determined as most-stable gene between samples and conditions and used as housekeeping gene.

Table 2.1 – Table of taqman gene expression assays.

Gene name	Taqman assay ID
Canx	Mm00500330_m1
Ubc	Mm01201237_m1
Rpgrip1l	Mm00452421_m1
Fto	Mm00488755_m1
Irx3	Mm00500463_m1
Irx5	Mm00502107_m1
Irx6	Mm01253620_m1
Pparg	Mm00440940_m1
Cebpa	Mm00514283_s1
Fabp4	Mm00445878_m1
Plin1	Mm00558672_m1
Fasn	Mm00662319_m1
Wars2	Mm00840490_m1

Tbx15	Mm00447443_m1
Pgc1a (Ppargc1a)	Mm01208835_m1
Tfam	Mm00447485_m1
Elovl3	Mm00468164_m1
Prdm16	Mm00712556_m1
Adrb3	Mm02601819_g1
Ucp1	Mm01244861_m1
Dio2	Mm00515664_m1
Cidea	Mm00432554_m1
Ppara	Mm00440939_m1
Cox7a1	Mm00438297_g1
Cox8b	Mm00432648_m1
Ndufb3	Mm00835179_g1
Ndufb7	Mm00788005_s1
Ndufv1	Mm00504941_m1
Ndufs1	Mm00523640_m1
Gapdh	Mm99999915_g1
18S	Primer design probe (taqman style): HK-DD-mo-300
Essra	Mm00433143_m1

2.2.5 Statistical Analysis

All values are expressed as Mean ± SD or SEM. First, the average of the Ct value of technical qPCR replicates was taken. The gene of interest was normalized to the house keeping gene or the geometric mean of two housekeeping genes if stated. Data were analysed using the comparative $2^{-\Delta\Delta Ct}$ method derived by Applied Biosystems to determine the difference in sample groups relative to control samples. Relative gene expression changes and SD/SEM were then calculated as follows:

$\Delta Ct = Ct_{target} - Ct_{HKG}$

$\Delta\Delta Ct = \Delta Ct_{(target\ treated)} - \Delta Ct_{(HKG\ treated)}$

Relative expression level = $2^{-\Delta\Delta Ct}$

$\%KD = 100 \times (1 - 2^{-\Delta\Delta Ct})$

Data was only included if the house-keeping gene was comparable (all data points within Ct=1) between individual data points within the same data set. The appropriate statistical test for evaluating significance was chosen based on the comparisons made and measurements taken and is indicated below the figures in the text. All differences were analyzed by ANOVA, or Student's t-test if data was normally distributed. A Wilcoxin t-test was used for data that were not normally distributed. If necessary, a Grubbs' test was performed to identify outliers, which were removed if significance was $P<0.05$. Statistical significance is indicated by using: *$P<0.05$, **$P<0.01$, ***$P<0.001$.

2.2.6 Irx3 Transcript Analysis

For Irx3 transcript analysis in mouse, primers detecting either transcript were designed and synthesized by PrimerDesign. Primers arrived optimized for amplification specificity and efficiency. The primer pairs used to assess Irx3 transcripts are the following:

Primer	Sequence	Product length
Irx3_TV1 sense	AGCACCTGCTGGGACTTCC	115
Irx3_TV1 antisense	ACTACAGCGATATTTTTCTTGTCACC	
Irx3_TV2 sense	GCCGGAGAGTGGAACAGAT	108
Irx3_TV2 antisense	GCAGCGTCCAGATGGTTC	

A qPCR was run using the following protocol:

Component	Volume (µl)
PrimerDesign Master Mix (2X)	10
Resuspended primer mix	1
cDNA (4ng/µl)	5
RNAse/DNAse free H_2O	5
Total	20

qRT-PCR was performed using the Fast SYBR Green system from Applied Biosystems on a ABIPRISM 7500 Fast Real-Time PCR Machine (Applied Biosystems) using the standard program for SYBR Green assays.

2.3 DNA

2.3.1 Mitochondrial DNA Copy Number Assay

The ratio of mitochondrial DNA (mtDNA) to genomic DNA (gDNA) was assessed using qRT-PCR to quantitatively assess the ratio of genes on the mitochondrial and nuclear genome. To amplify the genomic gene GAPDH and the mitochondrial gene mtND1, primers for both genes (see table below for details) were designed using the Primer3Plus online software (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi). Total DNA, including both gDNA and mtDNA, was extracted from cells and tissues using either the DNeasy Blood and Tissue Kit (Qiagen) or Trizol following the manufacturer's protocol. qRT-PCR was performed using the Fast SYBR Green system from Applied Biosystems on a ABIPRISM 7500 Fast Real-Time PCR Machine (Applied Biosystems) using the standard program for SYBR Green assays. Before the assays were run, optimal primer concentrations were determined as the lowest primer concentrations that achieve the lowest Ct value whilst minimizing non-specific amplification. To optimize the primers, a PCR was run with a matrix of primer concentrations ranging from 50–900nM for each primer pair. The optimal primer concentrations can be found below. Following from the optimization PCR, a melting curve was performed to ensure the presence of a single amplicon per primer pair. To ensure that the size of the amplicon

was as predicted, the PCR products were run on a 2% agarose gel. All assays were run in technical triplicates.

Primer	**Sequence**
mtND1 Forward Primer	5’-CCCATTCGCGTTATTCTT-3’
mtND1 Reverse Primer	5’-AAGTTGATCGTAACGGAAGC-3’
GAPDH Forward Primer	5’-CAAGGAGTAAGAAACCCTGGACC-3’
GAPDH Reverse Primer	5’-CGAGTTGGGATAGGGCCTCT-3’

Component	**Volume (µl)**
Fast SYBR Green Master Mix (2X)	10
mtND1 Forward Primer (5µM)	1.2
mtND1 Reverse Primer (5µM)	1.2
DNA (5ng/µl)	2
RNAse/DNAse free H2O	5.6
Total	20

Component	**Volume (µl)**
Fast SYBR Green Master Mix (2X)	10
GAPDH Forward Primer (5µM)	0.2
GAPDH Reverse Primer (5µM)	0.2
DNA (5ng/µl)	2
RNAse/DNAse free H2O	7.6
Total	20

2.4 Protein techniques

2.4.1 Protein Extraction

Tissues were previously snap-frozen and stored at -80°C. To homogenize tissues, a small piece of tissue was placed in 2ml Precelly tubes containing ceramic beads (Bertin Technologies) in pre-prepared lysis buffer CelLytic MT Mammalian Tissue Extraction Reagent (Thermo Fisher Scientific); 10ml of lysis buffer were freshly supplemented with 1 tablet of 1 X complete protease inhibitor (Roche) and 1 tablet 1 X PhosSTOP phosphatase inhibitor (Roche). Tissues were then homogenized using a Precelly-24 automated homogenizer (Bertin Technologies) in a mode of 1:5000-20X20-015, or 2:5000-1X20-005. For soft tissues (e.g. brain, liver, kidney and adipose tissues) we used Precelly CK14 tubes and for fibrous tissue (e.g. muscle and heart) were used Precelly CK28 tubes. Subsequent to homogenization, lysates were centrifuged at 13,000 rpm for 15min at 4°C to pellet cell debris and non-homogenized material. The supernatant containing the protein lysate was placed in 1.5ml Eppendorf tube on ice, making sure that minimal lipid was carried over.

Adherent cells in culture (grown in a monolayer) were either pelleted or lysed directly in the well on ice using M-PER Mammalian Protein Extraction Reagent (Thermo Fisher Scientific), freshly supplemented with 1 tablet of 1 X complete protease inhibitor (Roche) and 1 tablet 1 X PhosSTOP phosphatase inhibitor (Roche). Approximately 0.5 ml of buffer was used per $5x10^6$ cells. Cells were then scraped off the dish using a plastic cell scraper, then gently transferred into a pre-cooled 1.5ml Eppendorf tube. Cells were left on ice for 30min to lyse, with a brief vortex every 10min. Tubes were then centrifuged at 13,000 rpm for 15min at 4°C and the supernatant containing the protein lysate was placed in 1.5ml Eppendorf tube on ice, making sure that minimal lipid was carried over.

2.4.2 Protein Concentration Assay

A Bradford Assay was performed to determine protein concentrations using the BCA (bicinchoninic acid) Protein Assay Reagent (BioRad) according to the manufacturer's

instructions. To achieve an absorbance that was within the linear range of the standard curve, protein samples were diluted appropriately (typically 1:3 – 1:10). A protein standard was prepared using from a pre-prepared solution of 10µg/µl Bovine Serum Albumin (BioLabs #B9001S) and a serial dilution of BSA in ddH2O was used to generate a standard curve from 0.16 – 5µg/µl. 5µl of diluted sample or protein standard were pipetted into a 96 well flat-bottomed optical plate in duplicate for the samples and triplicate for the standard. 5µl of lysis buffer was added to a separate well as control. 20µl Reagent S (BioRAD #500-0115) was added per 1ml reagent A (BioRAD #500-0113) and vortexed. 25µl of Reagent A+S was added to each sample/standard followed by 200µl of Reagent B (BioRAD #500-0114). The plate was then covered in aluminum foil and placed on a plate shaker at 300rpm for 15–30 min. Absorbance was measured at 750nm on an Epoch Microplate Spectrophotometer (BioTek).

2.4.3 Immunoblotting

Protein electrophoresis and immunoblotting was performed using the NuPAGE Electrophoresis System (Invitrogen). Protein samples were diluted to 4µg/µl in lysis buffer and prepared as follows:

Component	**Volume (µl)**
NuPage LDS Sample Buffer (4X)	2.5
NuPage Reducing Agent (10X)	1
ddH_2O	1.5
Protein Sample (4µg/µl)	5
Total	10

The samples were heated at 70°C for 10min to denature and the placed on ice. Gel electrophoresis for protein separation was performed using 4-12% linear gradient Bis-Tris ready polyacrylamide gels with 1 X MOPS electrophoresis running buffer (Invitrogen) and set up in a XCell Surelock Mini Cell tanks (Invitrogen). The inner chamber of tank was filled with 200ml running buffer supplemented with 500µl of NuPage Antioxidant (Invitrogen) to prevent

oxidation of samples during electrophoresis. 10µl of protein sample or ladder were loaded per lane and the gel run at 200V for 50min. Electrontransfer of proteins from the gel onto a PVDF membrane (Hybond – P, GE Healthcare Amersham) was done using an XCell II Blot Module (Invitrogen) in NuPage Transfer Buffer (Invitrogen) containing 10% methanol, run for 1 hour at 30V. PVDF membranes were charged by soaking them in 100% methanol for 1min before assembly. Membranes were blocked in either 5% skimmed milk TBST (non-phosphor antibodies) or 5% Bovine Serum Albumin TBST (phosphor-antibodies) at room temperature for an hour or at 4°C overnight. Membranes were incubated in primary antibodies diluted to optimized concentrations in the TBST-milk overnight at 4°C. After primary antibody incubation, membranes were washed in TBST 3 times for 10min at room temperature followed by incubation of species-specific secondary horseradish peroxidase (HRP) conjugated antibodies in 5% skimmed milk TBST for 1-2 hours at room-temperature. Membranes were washed 5 times for 10min in TBST at room temperature.

To expose membranes, membranes were placed on cling film with the protein facing upwards and 1ml of ECL2 (Amersham, GE Healthcare) was added onto the membrane surface and incubated for 5 mins in the dark according to the manufacturer's instructions. Excess ECL was removed from the membrane and membranes were wrapped in cling film before imaging. To probe the membrane with an appropriate house-keeping antibody, the blots were washed in TBST to remove residual ECL before the process was repeated.

2.5 Chromatin Biology Techniques

2.5.1 ATAC-seq

Mouse primary pre-adipocytes and hWAT during adipogenesis

ATAC-seq was performed by adapting the protocol from Buenrostro *et al.*, 2015. Differentiating cells were lysed directly in cell culture plate. Lysis buffer was added directly onto cells grown in a 12-well plate. Plates were incubated on ice for 10min until cells were permeabilised and nuclei released. Lysis buffer was gently pipetted up and down to wash

nuclei off the well and transferred into a chilled 1.5ml tube to create crude nuclei. Nuclei were spun down at 600 x g for 10 min at 4℃, nuclei pellets were then re-suspended in 40ul Tagmentation DNA (TD) Buffer, nuclei counted using trypan blue (as described in Krishnaswami *et al.*, 2016) and volume of 50.000 nuclei was determined. Transposition reaction was performed following a previously published protocol (Buenrostro *et al.*, 2015).

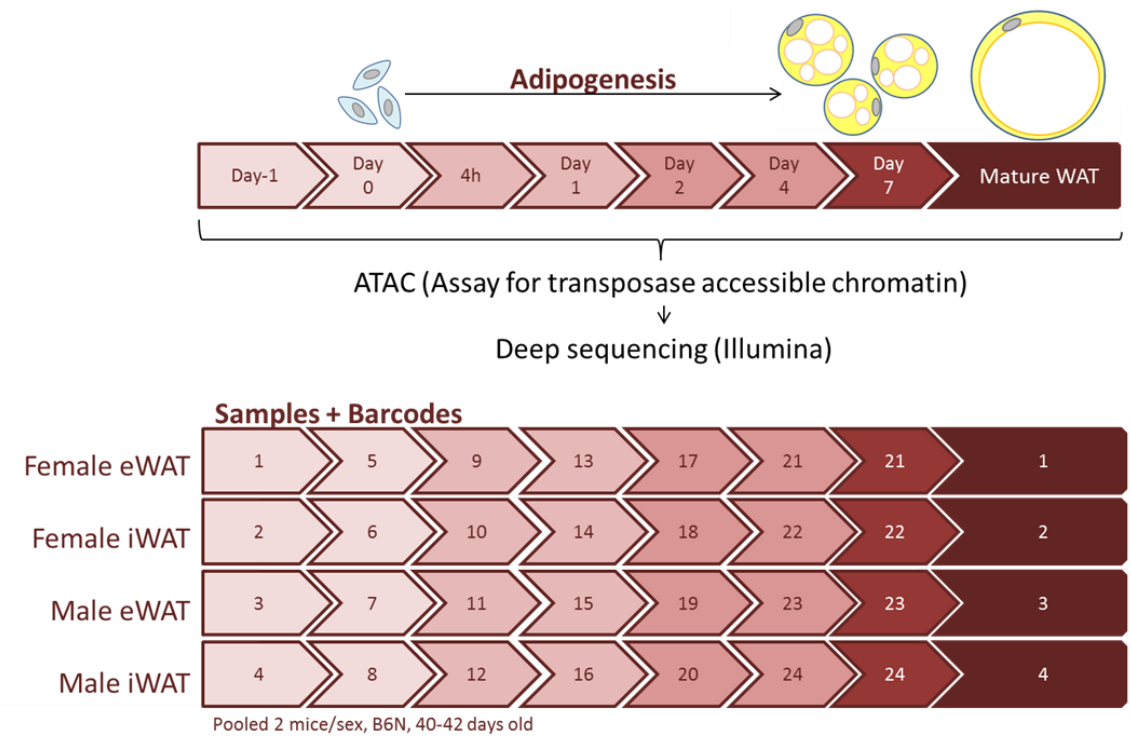
Mature mouse white adipose tissue (WAT) depots

Mature white adipose tissues (WAT) from inguinal (iWAT) and gonadal (gWAT) depots were digested and mature adipocyte nuclei isolated following a published protocol (Church *et al.*, 2014). Once cells were dissociated, ATAC lysis buffer was added to the mature adipocytes. See counting protocol from above. Tagmentation was performed following a previously published protocol (Buenrostro *et al.*, 2015).

All tagmented DNA was PCR amplified for 11-13 cycles for mouse libraries, and 8 cycles for hWAT libraries. Quality was assessed using a DNA1000 Chip (Applied Biosystems) and run on a Bioanalyzer (Applied Biosystems). The profiles showed that all libraries had a mean fragment size of ~190bp and characteristic nucleosome patterning, indicating good quality of the libraries. Libraries were sequenced at the Wellcome Trust Centre for Human Genetics in Oxford on a HiSeq4000 Illumina generating 50mio reads/sample, 75bp paired end. To reduce bias due to PCR amplification of libraries, duplicate reads were removed. Sequencing reads were aligned to mm10 and hs37d5, for mouse and human, respectively. BWA-MEM was used for mapping.

Analysis was performed by Juan Fernandez from the Wellcome Trust Centre for Human Genetics in Oxford, using their pipelines.

Schematic overview of ATAC-seq during adipogenesis in mouse primary pre-adipocytes

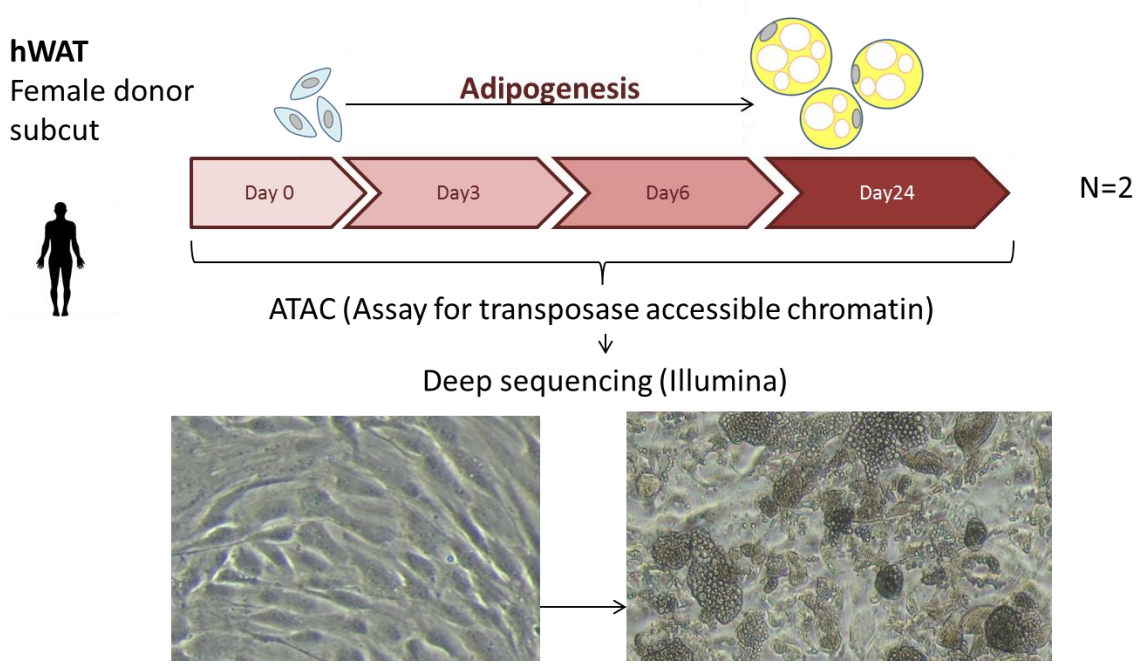


Barcode information

sample			barcode		
day	sex	depot	#	full sequence	pool
D-1	female	gWAT	1	CAAGCAGAAGACGGCATACGAGATTCGCCTTAGTCTCGTGGGCTCGGAGATGT	1
	female	iWAT	2	CAAGCAGAAGACGGCATACGAGATCTAGTACGGTCTCGTGGGCTCGGAGATGT	1
	male	gWAT	3	CAAGCAGAAGACGGCATACGAGATTTCTGCCTGTCTCGTGGGCTCGGAGATGT	1
	male	iWAT	4	CAAGCAGAAGACGGCATACGAGATGCTCAGGAGTCTCGTGGGCTCGGAGATGT	1
D0	female	gWAT	5	CAAGCAGAAGACGGCATACGAGATAGGAGTCCGTCTCGTGGGCTCGGAGATGT	2
	female	iWAT	6	CAAGCAGAAGACGGCATACGAGATCATGCCTAGTCTCGTGGGCTCGGAGATGT	2
	male	gWAT	7	CAAGCAGAAGACGGCATACGAGATGTAGAGAGGTCTCGTGGGCTCGGAGATGT	2
	male	iWAT	8	CAAGCAGAAGACGGCATACGAGATCCTCTCTGGTCTCGTGGGCTCGGAGATGT	2
T4	female	gWAT	9	CAAGCAGAAGACGGCATACGAGATAGCGTAGCGTCTCGTGGGCTCGGAGATGT	2
	female	iWAT	10	CAAGCAGAAGACGGCATACGAGATCAGCCTCGGTCTCGTGGGCTCGGAGATGT	2
	male	gWAT	11	CAAGCAGAAGACGGCATACGAGATTGCCTCTTGTCTCGTGGGCTCGGAGATGT	2
	male	iWAT	12	CAAGCAGAAGACGGCATACGAGATTCCTCTACGTCTCGTGGGCTCGGAGATGT	2
D1	female	gWAT	13	CAAGCAGAAGACGGCATACGAGATATCACGACGTCTCGTGGGCTCGGAGATGT	2
	female	iWAT	14	CAAGCAGAAGACGGCATACGAGATACAGTGGTGTCTCGTGGGCTCGGAGATGT	2
	male	gWAT	15	CAAGCAGAAGACGGCATACGAGATCAGATCCAGTCTCGTGGGCTCGGAGATGT	2
	male	iWAT	16	CAAGCAGAAGACGGCATACGAGATACAAACGGGTCTCGTGGGCTCGGAGATGT	2
D2	female	gWAT	17	CAAGCAGAAGACGGCATACGAGATACCCAGCAGTCTCGTGGGCTCGGAGATGT	2
	female	iWAT	18	CAAGCAGAAGACGGCATACGAGATAACCCCTCGTCTCGTGGGCTCGGAGATGT	2

	male	gWAT	19	CAAGCAGAAGACGGCATACGAGATCCCAACCTGTCTCGTGGGCTCGGAGATGT	2
	male	iWAT	20	CAAGCAGAAGACGGCATACGAGATCACCACACGTCTCGTGGGCTCGGAGATGT	2
D4	female	gWAT	21	CAAGCAGAAGACGGCATACGAGATGAAACCCAGTCTCGTGGGCTCGGAGATGT	3
	female	iWAT	22	CAAGCAGAAGACGGCATACGAGATTGTGACCAGTCTCGTGGGCTCGGAGATGT	3
	male	gWAT	23	CAAGCAGAAGACGGCATACGAGATAGGGTCAAGTCTCGTGGGCTCGGAGATGT	3
	male	iWAT	24	CAAGCAGAAGACGGCATACGAGATAGGAGTGGGTCTCGTGGGCTCGGAGATGT	3
D7	female	gWAT	21	CAAGCAGAAGACGGCATACGAGATGAAACCCAGTCTCGTGGGCTCGGAGATGT	2
	female	iWAT	22	CAAGCAGAAGACGGCATACGAGATTGTGACCAGTCTCGTGGGCTCGGAGATGT	2
	male	gWAT	23	CAAGCAGAAGACGGCATACGAGATAGGGTCAAGTCTCGTGGGCTCGGAGATGT	2
	male	iWAT	24	CAAGCAGAAGACGGCATACGAGATAGGAGTGGGTCTCGTGGGCTCGGAGATGT	2
mature	female	gWAT	1	CAAGCAGAAGACGGCATACGAGATTCGCCTTAGTCTCGTGGGCTCGGAGATGT	2
	female	iWAT	2	CAAGCAGAAGACGGCATACGAGATCTAGTACGGTCTCGTGGGCTCGGAGATGT	2
	male	gWAT	3	CAAGCAGAAGACGGCATACGAGATTTCTGCCTGTCTCGTGGGCTCGGAGATGT	2
	male	iWAT	4	CAAGCAGAAGACGGCATACGAGATGCTCAGGAGTCTCGTGGGCTCGGAGATGT	2

Schematic overview of ATAC-seq during adipogenesis in hWAT



Barcode information

sample		barcode	
day	**rep**	**#**	**full sequence**
D0	1	5	CAAGCAGAAGACGGCATACGAGATAGGAGTCCGTCTCGTGGGCTCGGAGATGT
	2	6	CAAGCAGAAGACGGCATACGAGATCATGCCTAGTCTCGTGGGCTCGGAGATGT
D3	1	7	CAAGCAGAAGACGGCATACGAGATGTAGAGAGGTCTCGTGGGCTCGGAGATGT
	2	8	CAAGCAGAAGACGGCATACGAGATCCTCTCTGGTCTCGTGGGCTCGGAGATGT
D6	1	9	CAAGCAGAAGACGGCATACGAGATAGCGTAGCGTCTCGTGGGCTCGGAGATGT
	2	10	CAAGCAGAAGACGGCATACGAGATCAGCCTCGGTCTCGTGGGCTCGGAGATGT
D24	1	13	CAAGCAGAAGACGGCATACGAGATATCACGACGTCTCGTGGGCTCGGAGATGT
	2	14	CAAGCAGAAGACGGCATACGAGATACAGTGGTGTCTCGTGGGCTCGGAGATGT

2.5.2 ChIP-seq

Primary pre-adipocytes were isolated and differentiated as described in section 'Adipocyte Precursor Isolation, Culture and Differentiation'. Adipose depots from 6-12 animals, aged 41 days, were pooled and digested. Cells were grown in 10cm dishes. Crosslinking was performed by adding 1ml of freshly-prepared formaldehyde solution to 10ml of the existing media in each dish of cells. The dishes were agitated for exactly 15min at room temperature. The fixation was stopped (quenching) by adding 0.55ml of 2.5M glycine solution to the existing media and the dishes left at room temperature for 5 minutes. After the glycine incubation the cells were scraped of the culture surface and washed by transferring the contents of each dish to a 50ml falcon tube. Samples were kept on ice for the remainder of the procedure. The tubed were centrifuged at 800 x g at 4°C for 10min to pellet the cells. The supernatant was removed and the cell pellet re-suspend in 10ml chilled PBS-Igepal solution by pipetting up and down. At this step, cells grown on multiple 10cm dishes were combined. Cells were pelleted by centrifugation as before. After the supernatant was discarded, 10 ml chilled PBS-Igepal was added to each tube and 100µl PMSF (100mM in ethanol; final concentration will be 1 mM) was added to each tube the pellet re-suspended by pipetting up and down. Tubes were centrifuged again at 800 x g at 4°C for 10min and the supernatant removed completely. Cell pellets were then snap-frozen on dry ice and stored at -80°C before shipment to an US-based company called Active Motif that performed the IP and subsequent sequencing and data quality control (QC).

For the histone modification analysis, published ChIP-seq data from Mikkelsen *et al.* (2010) was downloaded from its online repository. Binarised 200bp density maps from the bed files of the 6 ChIP-seq marks were created using a Poisson distribution implemented in the BinaryBed function of the ChromHMM software as described in (Ernst *et al.*, 2011; Ernst & Kellis, 2012). From these epigenomic density maps, 11 ChIP-only chromatin states were derived using a multivariate Hidden Markov Model implemented in the Learnmodel function (standard settings, hg19 genome) of the software ChromHMM (Ernst and Kellis, 2012).

Chromatin-state annotation maps and peaks were visualized on the igv and the UCSC genome browsers (Kent *et al.*, 2002).

Solutions

Reagents	Final concentration	
1. Formaldehyde Solution (to be prepared fresh before use)		**Per 20 ml**
37% Formaldehyde (e.g. Sigma #F-8775)	11%	6 ml
5 M NaCl	0.1 M	0.4 ml
0.5 M EDTA, pH 8.0	1 mM	40 µl
1 M HEPES, pH 7.9	50 mM	1 ml
H2O (Note: NaCl, EDTA, and HEPES should be molecular biology grade.)		to 20 ml
2. Glycine Solution		**Per 20 ml**
Glycine, MW 75 (e.g. Sigma #G-7403)	2.5 M	3.75 g
H2O		to 20 ml
3. PBS-Igepal		**Per 100 ml**
PBS, pH 7.4 (e.g. ThermoFisher #10010023)	~1X	100 ml
100% Igepal CA-630 (e.g. Sigma #I-8896)	0.5%	0.5 ml
4. PMSF (e.g. Sigma #P-7626)		
Prepare at 100 mM in ethanol and store at -20°C. (Note: PMSF Phenylmethanesulfonyl fluoride.)		

2.5.3 3D FISH

Pre-adipocytes were isolated as described in section 'Adipocyte Precursor Isolation, Culture and Differentiation'. Adipocytes were plated on Superfrost Plus slides (Thermo Scientific REF J1800AMNZ) placed in quadriperm dishes. Cells were either fixed 1 day after they attached (D0) or induced to differentiate for 2 days (D2). For fixation, cells were wash 3

times in PBS and fixed in 4% PFA for 10mins. After 3 washes in PBS, cells were permeabilised in 0.5% Triton X for 10min, followed by 3 washes in PBS. Slides were then air dried for approximately 30min and frozen at -80°C. Slides were then shipped to Iain Williamson (Wendy Bickmore group) at The MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, where the in situ hybridization was performed with the following fosmids:

(mm10 coordinates)

chr8	91783407	91817571	Irx3
chr8	92341437	92380953	Irx5
chr8	91371910	91392325	Fto-en

DEL82 location:

chr8	91374370	91374451	DEL82

(mm9 coordinates)

chr8	94307306	94341470	Irx3
chr8	94865336	94904852	Irx5
chr8	93895809	93916224	Fto-en

DEL82 location:

chr8	93898269	93898350	DEL82

Briefly, 80-100ng of fosmid, 100ng plasmid, 200ng paint, and 100-200ng BAC were dispensed per slide. Cot I DNA DNA was added (~3-4ug per fosmid, ~7ug per paint arm, ~5-10ug per BAC) and 5ug sonicated salmon sperm DNA per slide. EtOH was added and the probe air dried. Hybridisation mix containing deionised formamide, 20 x SSC, 50% dextran sulphate and Tween 20 was added to the slides for ~1h. Slides were treated with RNAse for 1hour at 37°C in 2XSSC (100ug/ml). Denaturing was performed by heating the samples at 70°C for 5min and then for 15mins at 80°C. Slides were blocked in blocking buffer (4 x SSC, 5% Marvel) at 37°C for ~5 min. The following antibody dilutions were made:

Avidin FITC (Vector) 1:500

Avidin Texas Red (Vector) 1:500

Biotinylated anti-avidin (Vector) 1:100

FITC anti-sheep (Vector) 1:100

Texas Red anti-sheep 1:100

Rhodamine anti-dig FAB fragments (Roche) 1:20

Fluorescein anti-dig FAB fragments (Roche) 1:20

Streptavidin Cy5 1:20

The slides were incubated with antibody in a moistened chamber at 37°C for 30-60min. Slide was washed and consecutive antibody incubation performed. For nuclear staining, slides were treated with 1:1000dilution of DAPI (stock 50ug/ml) for 5min before mounting in Vectashield under a coverslip.

2.6 Mouse Methods

2.6.1 Animal Husbandry

Animal studies were conducted in accordance with the UK Animals (Scientific Procedures) Act (1986). Mouse studies were performed under the Home Office License 30/2642. Mice were housed according to UK Home Office welfare guidelines in a 12hr light / dark cycle at a temperature of 21± 2°C and humidity of 55 ± 10%. Mice were fed ad libitum and had free access to water (25 ppm chlorine). The research diets used in this study indicated in the figures. Mice that were used for primary pre-adipocyte isolation were fed SDS maintenance chow (RM3, 3.6 kcal/g). The high-fat and low-fat research diets used for the phenotyping cohorts were the following matched diets from ResearchDiets.Inc:

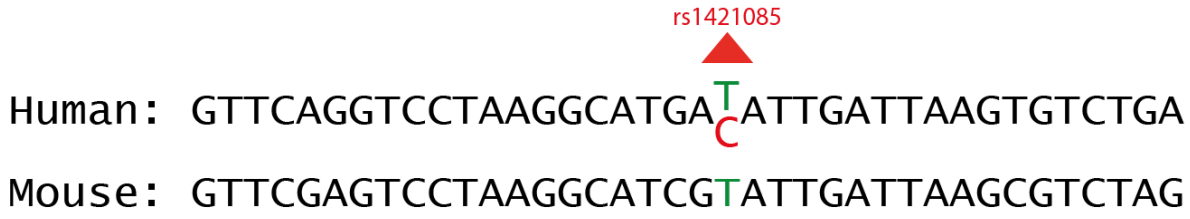
Product #	**HFD** High fat (60%) low sucrose (7%)		**LFD** low fat (10%) low sucrose (7%) control	
	D12492		D12450J	
Protein	gm%	kcal%	gm%	kcal%
Carbohydrate	26	20	19.2	20
Fat	26	20	67.3	70
Total	35	60	4.3	10
kcal/gm	5.24	100	3.85	100

High fat diets are usually not grain-based such as chow. Matched, Purified Ingredient Diets decrease the possibility that diet-induced differences are due to differences in fibre type and amount, source of carbohydrate, and variation in plant chemicals such as phytoestrogens between the diets.

For the phenotyping cohorts, mice were weaned and randomized onto their respective diets at 3 weeks of age. All phenotyping was performed blind, following the ARRIVE guidelines for animal research.

2.6.2 Genetic Mouse Model Generation

Using CRISPR/Cas9 technology, the mouse models described in this thesis were generated by the CRIPSR/Cas9 core facility at MRC Harwell essentially as published in Mianné *et al.*, (2017). The aim was to create a point mutation at the mouse orthologue of human rs1421085, namely to alter the T to a C, mimicking the human BMI-risk allele:

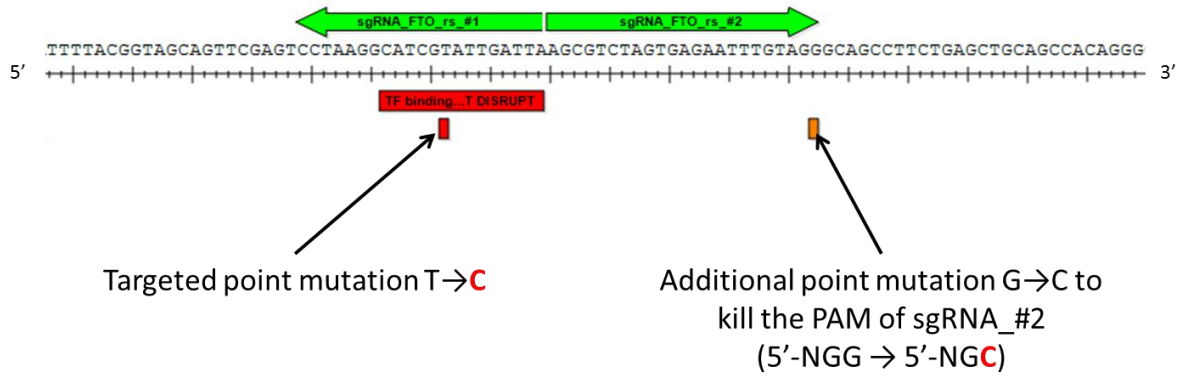


2.6.2.1 Design tool and choice of guide

Single guide RNA (sgRNA) sequences were designed using the following online tools:

- http://tefor.net/crispor/crispor.cgi
- http://crispr.mit.edu/
- http://www.sanger.ac.uk/htgt/wge/

The main criteria during sgRNA sequence selection were to minimise predicted off-target events, particularly on the same chromosome. Due to several rounds of crosses and the selection for a given mutation, the chances of off-target events were minimised, however, cannot be ruled out with certainty. Two sgRNA sequences were selected for cloning:



Additional silent point mutation were created within the PAM of the guide used to prevent the system to “re-process/re-cut” the allele repair through Homology Directed Repair (HDR) with the donor oligo (= expected allele).

2.6.2.2 Donor oligonucleotides

Donor sequence templates were designed with homology arms of 80nt in size flanking the intended point mutation. An additional silent point mutation was introduced to the ssODN sequence to disrupt the PAM sequence and prevent re-cutting of the modified allele by Cas9. The donor sequences were ordered as Ultramer™ DNA oligonucleotides from IDT. The sequences of oligonucleotides, protospacers and donor DNAs used within the examples presented in this study are shown in the following table:

Oligo name	5'-3' sequence [PAM]
sgRNA_#1 sequence	TAATCAATACGATGCCTT **AGG**
sgRNA_#2 sequence	AGCGTCTAGTGAGAATTTGT **AGG**
WT sequence	GATAATCTCATTGTTCCTCCTGCTACTTAAAATAAAGGTAATATTGATTTTACGGT AGCAGTTCGAGTCCTAAGGCATCGTATTGATTAAGCGTCTAGTGAGAATTTGTA GGGCAGCCTTCTGAGCTGCAGCCACAGGGCATCTCCCCACTGGGCCAGGCCT CTGTGCTG
Mutated sequence design 1 (if use of sgRNA_1)	GATAATCTCATTGTTCCTCCTGCTACTTAAAATAAAGGTAATATTGATTTTACGGT AGCAGTTCGAGTCCTAAGGCATCG**C**ATTGATTAAGCGTCTAGTGAGAATTTGTA GGGCAGCCTTCTGAGCTGCAGCCACAGGGCATCTCCCCACTGGGCCAGGCCT CTGTGCTG
Mutated sequence design 2 (if use of sgRNA_2)	GATAATCTCATTGTTCCTCCTGCTACTTAAAATAAAGGTAATATTGATTTTACGGT AGCAGTTCGAGTCCTAAGGCATCG**C**ATTGATTAAGCGTCTAGTGAGAATTTGTA G**C**GCAGCCTTCTGAGCTGCAGCCACAGGGCATCTCCCCACTGGGCCAGGCCT CTGTGCTG

2.6.2.3 Pronuclear microinjections of zygotes

Pronuclear microinjection into C57BL/6NTac embryos was performed as described in (Gardiner & Teboul, 2009). Cas9 mRNA, sgRNAs and ssODN were injected using a FemtoJet system (Eppendorf) at concentrations of 200 or 100 ng/µl, 100 or 50 ng/µl each and 50 or 20 ng/µl, respectively. Injected embryos were re-implanted in CD1 pseudo-pregnant female hosts which were allowed to litter and rear F_0 progeny.

2.6.2.4 Resulting alleles

CRISPR/Cas9-mediated genome editing resulted in the following lines:

Allele	Sequence
WT	GTTCGAGTCCTAAGGCATCGTATTGATTAAGCGTCTA
FTO-DEL82-EM1-B6N	GTTCGAGTCCTAAGGCAT*[82ntdel]*CTCTGTGCTGA
FTO-DEL10-EM7-B6N	GTTCGAGTCCTA*[10ntdel]*TTGATTAAGCGTCTA
FTO-RS1421085-EM5-B6N	GTTCGAGTCCTAAGGCATCGCATTGATTAAGCGTCTA

2.6.2.5 Genotyping strategy

All genotyping was performed by the GEMS core facility at MRC Harwell.

i) FTO-DEL82-EM1-B6N Genotyping Strategy

A PCR-based assay was designed to detect the deleted sequence. DNA for genotyping was derived from mouse ear clips.

FTO-DEL82-LOA-WT1 assay (FAM labelled probe)

Fragment Sequence (Sequence between<> represents the 82nt deletion in Fto gene)

GCAGACTTGTAAGCAACAAGATAATCTCATTGTTCCTCCTGCTACTTAAAATAAAGGTAA

TATTGATTTTACGGTAGCAGTTCGAGTCCTAAGGCAT<CGTATTGATTAAGCGTCTAGTG

AGAATTTGTAGGGCAGCCTTCTGAGCTGCAGCCACAGGGCATCTCCCCACTGGGCCAG

GC>CTCTGTGCTGACCCTACTATTAG

FTO-DEL82-LOA-WT1 primers and probes

Primer 1 = AGCGTCTAGTGAGAATTTGTAGGGCAG

Primer 2 = CTGGCCCAGTGGGGAGATG

Probe = CCTTCTGAGCTGCAGCCACAGG

Dot1l internal control (VIC labelled)

TCATAGGGTGACTGGCCAACCCAGGGAAGCCGGAGTGCTGCGTCTTCTGTTTCCTTGT

TCTTTTCCCCTCTAGTCGTTTTCTGTTAGTAGTTGGCATCCTTATGCTTCATCTTACAGTC

GACTTGAGAGCTGGCCCTGAATGGTCGTGCTGGGGCAAGGCTTTATTTCAGGCGTAGC

ACACATGGTGGCCAATGGGACTCTGTAGGATCTGCCCACACCCATCAG

Dot1l primers and probes

Primer 1 = GCCCCAGCACGACCATT

Primer 2 = TAGTTGGCATCCTTATGCTTCATC

Probe = CCAGCTCTCAAGTCG

qPCR master mix

- ABI GTX Taqman master mix 5µl
- Primers Dot1L_2F (20µM) 0.225µl

- Primers Dot1L_R (20μM) 0.225μl
- Probe DotL_2M (5μM) 0.2μl
- FAM Assay (probe 5μM & primers 15μM each) 0.3μl
- ddH20 1.55μl
- DNA (1/10 dilution of ABI Sample-to-SNP prep) 2.5μl

ii)FTO-RS1421085-EM5-B6N Genotyping Strategy

An Allelic Discrimination assay was performed to detect the two possible variants of this SNP. This assay used two taqman probes, one detecting the WT and the other the variant sequence. DNA was derived from mouse ear clips.

WT allele: T

TTCCTGAGCTAGTGTGTGTACCCTCAACTCTCCCTGTCTGTGGCCCTGCTGCCAGGCT
GCCTCCGAGCCCAGCAAACTCATTCCTCCCAGGCAAAAGCAAAAGGTGACATACACCA
GGAGCCAGATTGTCCACAGCAGACTTGTAAGCAACAAGATAATCTCATTGTTCCTCCTG
CTACTTAAAATAAAGGTAATATTGATTTTACGGTAGCAGTTCGAGTCCTAAGGCATCGTA
TTGATTAAGCGTCTAGTGAGAATTTGTAGGGCAGCCTTCTGAGCTGCAGCCACAGGGC
ATCTCCCCACTGGGCCAGGCCTCTGTGCTGACCCTACTATTAGAAGTGTTTAGGCTCTG
GTCAGAACATTTATTTTACCTGATGGATGGCTTAATCCAGCCCGTCACCTTAATCTGACC
TTAAAGATGGCAGTCCAACTCTTTTGTTCTGAGCGAATGTGGAGACTGATTTTGGCAAT
GTTGAGAGCTGTGTTATCAGCA

FTO-RS1421085 allele: C

TTCCTGAGCTAGTGTGTGTACCCTCAACTCTCCCTGTCTGTGGCCCTGCTGCCAGGCT
GCCTCCGAGCCCAGCAAACTCATTCCTCCCAGGCAAAAGCAAAAGGTGACATACACCA
GGAGCCAGATTGTCCACAGCAGACTTGTAAGCAACAAGATAATCTCATTGTTCCTCCTG
CTACTTAAAATAAAGGTAATATTGATTTTACGGTAGCAGTTCGAGTCCTAAGGCATCGCA
TTGATTAAGCGTCTAGTGAGAATTTGTAGGGCAGCCTTCTGAGCTGCAGCCACAGGGC

ATCTCCCCACTGGGCCAGGCCTCTGTGCTGACCCTACTATTAGAAGTGTTTAGGCTCTG
GTCAGAACATTTATTTTACCTGATGGATGGCTTAATCCAGCCCGTCACCTTAATCTGACC
TTAAAGATGGCAGTCCAACTCTTTTGTTCTGAGCGAATGTGGAGACTGATTTTGGCAAT
GTTGAGAGCTGTGTTATCAGCA

Primers and Probes

Primer 1 ACGGTAGCAGTTCGAGTCCT

Primer 2 AGAAGGCTGCCCTACAAATTCTCA

Allele 1 (WT) probe (FAM-Labelled) CGCTTAATCAATACGATGC

Allele 2 (Mut) probe (TET-Labelled) CTTAATCAATGCGATGCC

qPCR master mix

- ABI GTX Taqman master mix 5µl
- Assay (Probes 5µM each & Primers 15µM each) 20uM 2µl
- ddH2O 0.5µl
- DNA (1/10 dilution of ABI Sample-to-SNP prep) 2.5µl
- No need to run the samples in duplicates.
- Allele 1 = WT on 7500 FAM-labelled. Allele 2 = MUT on 7500 TET-labelled.

2.6.3 Mouse Phenotyping Techniques

All *in vivo* phenotyping (including data presented in this thesis for Echo-MRI, IPGTT, bleeds) was performed by Liz Bentley. Liz Bentley further assisted in tissue collection.

2.6.3.1 Body composition

Total body weight was measured every two weeks on a scale calibrated to 0.01 g. Body composition measurements were performed every two weeks using an Echo-MRI (Echo-MRI-100, Echo-MRI, Texas, U.S.A.). The read-outs were total fat mass (g), lean mass (g) and free water in live, non-anaesthetized mice.

2.6.3.2 Intra-Peritoneal Glucose Tolerance Test

Intra-Peritoneal Glucose Tolerance Test (IPGTT) was performed in the morning after an overnight (up to 16h) fast. Body weight of the animals was measured and a local anaesthetic was administered to the mouse tail (EMLA cream, Eutectic mixture of Local Anaesthetics Lidocaine / Prilocaine, AstraZeneca, UK). To establish the baseline glucose level at time point zero, a blood sample was collected from the tail in a Lithium-Heparin microvette tubes (CB30, Sarstedt, Numbrecht, Germany). Subsequently, 2g glucose / kg body weight (20% glucose in 0.9% NaCl) was injected intra-peritoneally and blood samples taken 30, 60 and 120min post-injection. At each time point, blood glucose levels were measured immediately using the handheld Alphatrak (Abbott) glucose monitor with Alphatrak strips (Abbott). A fresh strip was used for each reading.

2.6.3.3 Comprehensive Lab Animal Monitoring System

Mice were placed in individual caged into a Comprehensive Lab Animal Monitoring System (CLAMS) for a total period of 72h. Body weight measurement and Echo-MRI was performed immediately before and after animals being placed in the CLAMS system. This system allowed the simultaneously measurement of energy expenditure, activity, feeding and drinking at the home cage temperature (22°C) in individual mice according to standard protocols. Briefly, oxygen (O_2) and carbon-dioxide (CO_2) in-flow and out-flow concentrations were automatically monitored and recorded along with food consumption and water intake throughout a 72 hour period. To allow mice to acclimatize to the new environment and the stress of single housing, recordings from the first 24 hours were discarded and only data collected in the subsequent 24 hour period were used for analysis. Oxygen consumption (VO_2), CO_2 production (VCO_2), respiratory exchange ratio (RER) and heat production were calculated using the O_2 and CO_2 concentration measurements in the following calculation:

$VO_2 = ViO_2i - VoO_2o$

$VCO_2 = ViCO_2i - VoCO_2o$

$RER = VCO_2 / VO_2$

Heat = CV x VO_2 (where CV = Calorific value = 3.815 + 1.232 x RER)

o = outflow, i = inflow

The readings for VO_2, VCO_2 and EE were normalized by adjusting for differences in individual lean mass using multiple linear regression analysis (ANCOVA) as previously described (McMurray *et al.*, 2013).

2.6.3.4 Retro-orbital bleed

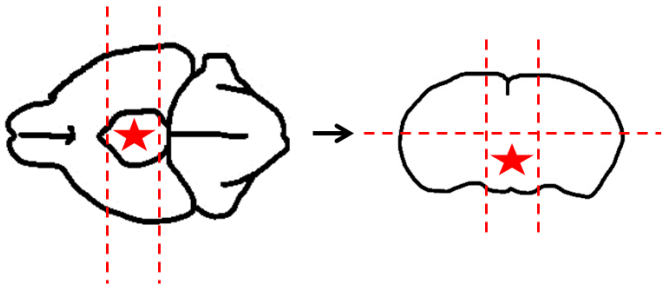
Animals were euthanized by intra-peritoneal injection of an over-dose of anaesthetic (0.2 mL of pentobarbitone) in accordance with home office procedures. Once fully anaesthetized, a glass capillary was inserted into the anterior corner of the mouse eye to perforate the membrane of the retro-orbital sinus. Blood was collected from the capillary in Lithium-Heparin microvette tubes (CB30, Sarstedt, Numbrecht, Germany). Once enough blood was collected, animals were dissected and tissues were taken for weight, histology and gene expression analysis.

2.6.3.5 Plasma Biochemistry

The blood samples collected via retro-orbital bleeding were kept on ice before centrifuged for 10min at 8000xg at 8°C. The supernatant blood plasma was removed and immediately analyzed on a board Beckman Coulter AU680 clinical chemistry analyzer using reagents and settings recommended by the manufacturer. The blood measurements taken were Total Cholesterol, HDL, LDL, Glucose, Triglycerides, Glycerol and Free Fatty Acids.

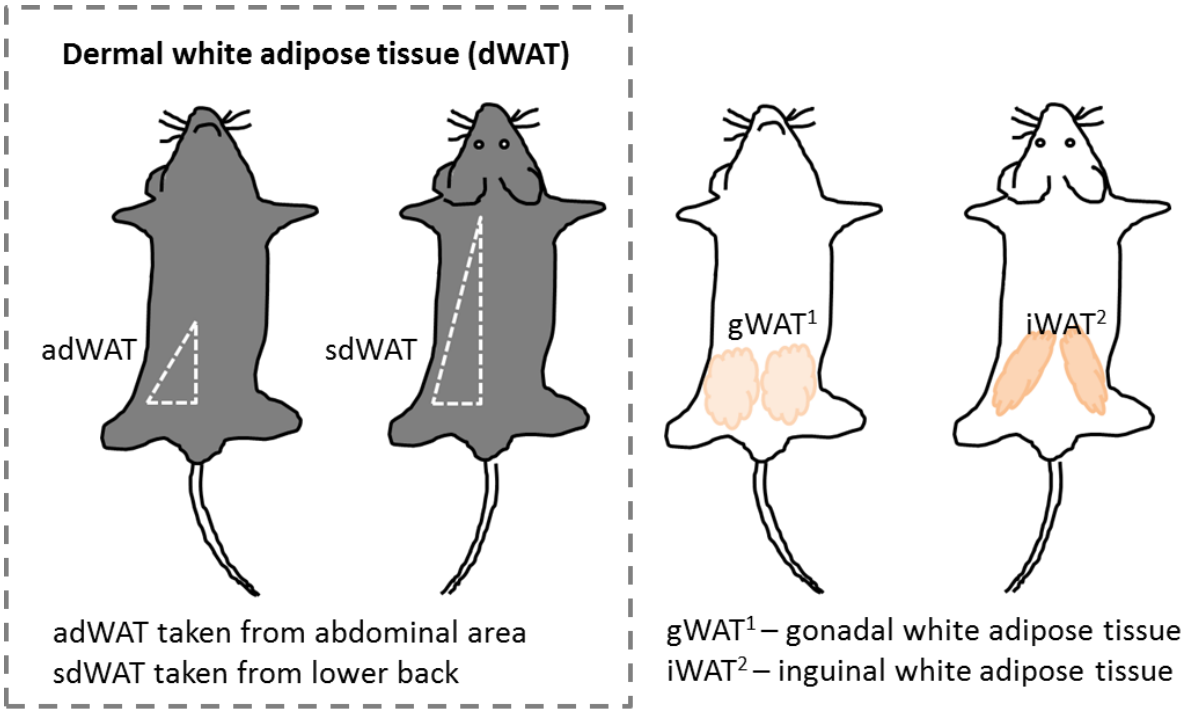
2.6.4 Tissue Collection and Organ Weights

Mice were culled after the phenotyping pipeline was completed and following confirmation of death, several tissues were dissected such as heart, liver, kidney, inguinal white adipose tissue, gonadal white adipose tissue, brown adipose tissue, skin from the back, skin from the abdomen, skeletal muscle and brain. After the brain was dissected out, it was placed into a matrix and hypothalamus (★) was excised using two razor blades as indicated in the following schematic:



Hindbrain, forebrain, cerebellum and the rest were separated and stored at -80°C for later use.

Adipose tissues were placed on a scale and measurements taken. One WAT depot per mouse was placed into a 10% formalin pot for Hematoxylin and Eosin (H&E) staining and one depot was snap-frozen in cryotubes (Nunc, Thermo Fisher Scientific) for subsequent protein, RNA and DNA analysis. Tissue samples were stored long-term at -70 °C.



2.6.5 Histology

After adipose tissues and skin were fixed in neutral buffered 4% formaldehyde for at least 24 hours at room temperature, tissues were paraffin embedded and cut into 8µm sections. Serial sections were then stained with H&E according to standard protocols (perform by the Histology Department, MRC Harwell). Stained sections were imaged using the Hamamatsu NaoZoomer slide scanner and NDP 2 serve software was used to export the images at appropriate magnifications. For analysis of dWAT thickness, 3 measurements per sample (blinded) were taken randomly on the slide and average calculated.

2.6.6 Statistical Analysis

All phenotyping data is presented as mean ± standard deviation (SD) or standard error mean (SEM). Statistical analysis was performed using GraphPad Prism software. For data with more than two variables, such as body weight measurements over time, statistical analysis was performed using repeated measures two-way ANOVA and when applicable, Bonferroni's post hoc test was used to correct for multiple comparisons. In cases of incomplete data sets, instead of repeated measures, the data was analyzed using ANOVA followed by Bonferroni's post hoc test. For data comparing two groups, a Student's t-test was performed, e.g. when comparing blood glucose levels between WT and DEL82. Significance was assigned to data that had a probability of >5% ($P<0.05$). Statistically significance results are indicated with asterisks: *$P<0.05$, **$P<0.01$, ***$P<0.001$.

3 *FTO/Fto* Epigenomic Comparison

3.1 Introduction

The majority of human genetic variants associated with common metabolic disease traits are located within distal regulatory elements, including BMI-associated variants in *FTO*. Non-coding variants can modify *cis*-regulatory pathways through several mechanisms, including the disruption or creation of TFBSs (Kasowski *et al.*, 2010), the modulation of TF affinity to a TFBS (Stratigopoulos *et al.*, 2016), chromatin openness (Degner *et al.*, 2012), and/or by altering chromatin interactions (Li *et al.*, 2012; Pomerantz *et al.*, 2009; Visser *et al.*, 2012). In the case of *FTO*, several risk variants alter TFBSs. For example, the risk variant for rs1421084 interrupts a TFBS for ARID5B that leads to increased expression of *IRX3* and *IRX5* in pre-adipocytes, presumably via increased *IRX3* and *IRX5* promoter binding to the *FTO* intronic enhancer element around rs1421085 (Claussnitzer *et al.*, 2015) and risk variants rs1421085 and rs8050136 are within DNA-binding motifs for CUX1; depending on allele, one of the two isoforms of CUX1 is preferentially bound which leads to altered *RPGRIP1L* and *FTO* expression *(Stratigopoulos et al.,* 2016).

There are currently a number of chromatin biology techniques that can be utilized to assess the regulatory landscape in specific cell types of interest and that can be used to infer the possible function of disease-associated variants. For example, TFBSs and chromatin states (based on histone modifications) can be assessed using ChIP-seq, the spatial organization of DNA can be studied using chromosome conformation capture techniques (e.g. 3C, 4C, 5C, CaptureC, Hi-C, ChIA-PET, etc.), and open chromatin, indicative of regulatory regions, can be studied using DHS-seq, FAIRE-seq or ATAC-seq (Buenrostro *et al.*, 2015; Dekker *et al.*, 2013; Hughes *et al.*, 2014; Knight, 2014). Such data sets provide valuable resources for generating hypotheses about possible function of non-coding variants and connections between distal regulatory elements and their target genes, albeit at low resolution of several hundred base pairs.

The organization of the mammalian genome in the nucleus in three-dimensional (3D) space is crucial in gene regulation. For a distal regulatory element to alter gene expression, it needs to be in physical contact with the target gene promoters. On a larger scale, DNA is organized in TADs, which have been proposed to be the elementary unit of mammalian genome organization (Dixon *et al.*, 2012). During adipocyte development, the genome organization changes and just following induction of adipocyte differentiation, the chromatin remodels to give rise to super-enhancers that are active in early adipocytes and vital in orchestrating adipogenesis (Mikkelsen *et al.*, 2010; Siersbaek *et al.*, 2012; R. Siersbæk & Mandrup, 2011).

A comparative analysis of genome-wide chromatin state maps as well as PPARG (a master regulator of adipogenesis) and CTCF (an important factor in marking chromatin state boundaries) localization maps in human and mouse revealed that the specific locations for the majority of putative pre-adipocyte- and adipocyte-specific *cis*-regulatory elements differed between the species (Mikkelsen *et al.*, 2010). These differences were attributed, in part, to the evolutionary turnover of TF motifs (Mikkelsen *et al.*, 2010). Based on these findings, modelling distal regulatory elements in the mouse to explain human regulatory networks is not an obvious choice. On the other hand, it has also been shown that conservation of co-occurring TFBS can guide identification and functional translation of genetic association signals (Claussnitzer *et al.*, 2014). We reasoned that comparative analysis of human and mouse genome function in the relevant cell type might help to evaluate the feasibility of such an approach at the *FTO* locus. Therefore, I compared data generated by multiple experimental tools that are used to study genome function, including histone modifications, CTCF binding sites, accessible chromatin regions and 3D genome organization.

Categorization of regulatory elements as either species-specific or shared could provide insight into the relative importance of such elements and also help to evaluate the potential of modelling *FTO* risk alleles in mouse. To assess human-to-mouse conservation at the *FTO*

regulatory circuitry I used publicly available as well as my own generated genomic data sets in adipocytes and their progenitors. This chapter aimed to generate a comparative epigenomic profile of the regulatory landscape around *FTO*, with particular focus on rs1421085.

3.2 Results

3.2.1 Histone Modifications

Histone modifications are known to mark regulatory elements, such as enhancers and repressors, in the genome. To compare the chromatin state landscape around *FTO*, I visualized publicly available ChIP-seq data sets from adipocytes for six histone modifications (H3K4me3/me2/me1, H3K27me3/ac and H3K36me3) and CTCF at the *FTO* locus (Fig. 4.1).

Different histone modifications are believed to mark different regulatory functions. For example, H3K4me3 is associated with transcriptional initiation and is found near promoters; H3K4me2/me1 and H3K27ac are associated with accessible chromatin and *cis*-regulatory activity, and are usually found spread out across promoters, intronic and intergenic regions; H3K36me3 is associated with transcriptional elongation and is found across active gene bodies; and H3K27me3 is associated with Polycomb-mediated repression, mainly distributed across inactive regions (Mikkelsen *et al.*, 2010). Combinatorial analysis of these marks can be used to generate chromatin state annotations using ChromHMM (Ernst & Kellis, 2012), which allows the characterization of the regulatory landscape in a single, color-coded track and define regions as e.g. promoters (red) and enhancers (yellow). A colour code for all automated chromatin states can be found in Appendix Fig. S.1. Additionally, the position of CTCF is linked to chromatin states, as CTCF is one of the factors that plays a role in maintaining the boundary between distinct chromatin 'types' (Bannister & Kouzarides, 2011; Dixon *et al.*, 2012; Gómez-Marín *et al.*, 2015).

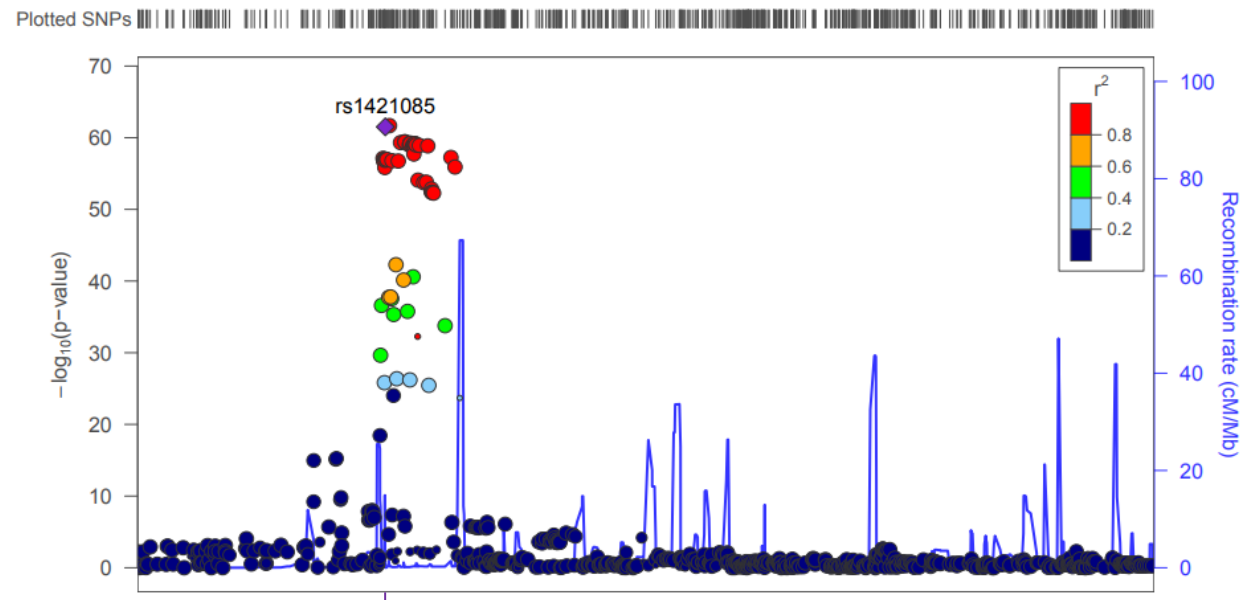
Mikkelsen *et al.* (2010) generated ChIP-seq data at four time points during adipogenesis for human primary pre-adipocytes (hASC) and mouse 3T3-L1 adipocytes. The time-points in

human and mouse were matched to similar stages of adipogenesis based on morphology and lipid droplet accumulation. I downloaded aligned ChIP-seq data files from the respective repository. A combinatorial analysis of all six histone marks was achieved using ChromHMM algorithm (Ernst & Kellis, 2012), resulting annotations were visualized using the UCSC genome Browser and tracks were overlapped with SNPs from the GWAS catalogue. Raw ChIP-seq peaks were further visualized using the igv browser for species-comparisons.

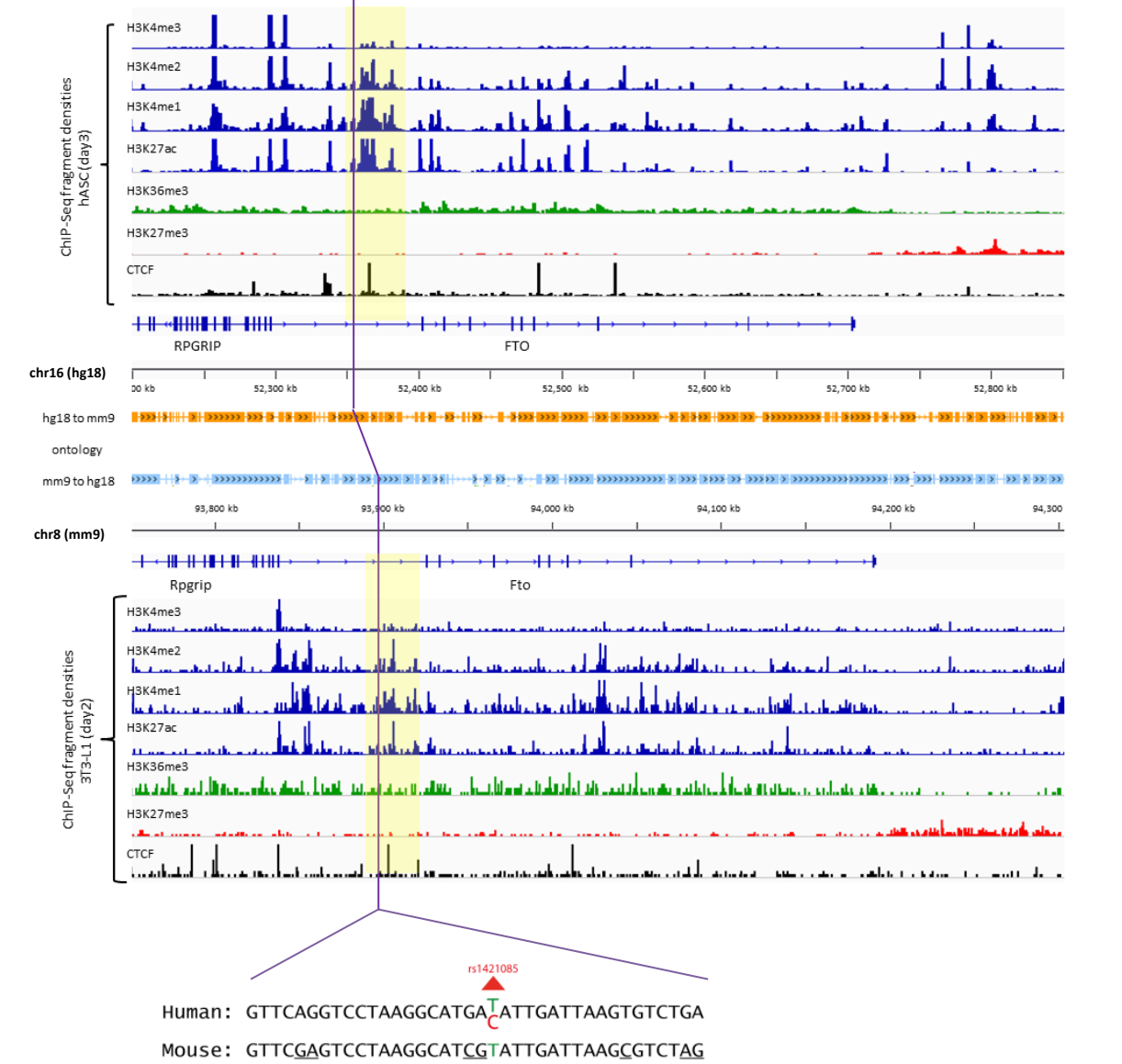
When comparing the chromatin state annotations at all time-points it became obvious that enhancer marks in *FTO* intron 1 are present throughout adipogenesis in human (Fig. 3.2) and mouse (Fig. S.2). However, early differentiating adipocytes present with the strongest enhancer (yellow) signal in BMI-associated regions in both models (Fig. 3.2; Appendix Fig. S.2). Therefore, I chose to compare chromatin maps generated in an early adipocyte differentiation time-point in both models (day 3 for hASCs, day 2 for 3T3-L1s).

The BMI-association signal in *FTO* overlaps marks for H3K4me2/me1 and H3K27ac, which are associated with *cis*-regulatory activity (Fig. 3.1). The orthologous region in mouse shows densities for the same histone marks, with the closest matching peak patterns for H3K4me1. Elongation marks across *FTO* are also present in both species. Repressive marks across regions downstream of *FTO* also appear to be shared between human and mouse. Although there seems to be a conserved CTCF binding site within the disease-associated region in *FTO* intron 1 that is shared between human and mouse, overall, there seems to be little overlap between CTCF sites in human and mouse adipocytes, suggesting divergence of CTCF binding sites between the two species at the wider locus. Alignments of the sequences underlying rs1421085 in human and the orthologous site in mouse shows that the underlying motif is largely conserved at the sequence level (Fig. 3.1). In summary, the BMI-associated signal within *FTO* is within an adipocyte enhancer that is conserved in mouse, both in sequence and in function - based on histone modifications.

A


Plotted SNPs
rs1421085
r²
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)

B


ChIP-Seq fragment densities hASC (day3)
H3K4me3
H3K4me2
H3K4me1
H3K27ac
H3K36me3
H3K27me3
CTCF
RPGRIP
FTO
chr16 (hg18)
hg18 to mm9
ontology
mm9 to hg18
chr8 (mm9)
Rpgrip
Fto
ChIP-Seq fragment densities 3T3-L1 (day2)

rs1421085
Human: GTTCAGGTCCTAAGGCATGATTATTGATTAAGTGTCTGA
Mouse: GTTCGAGTCCTAAGGCATCGTATTGATTAAGCGTCTAG

Fig. 3.1 | Comparison of *FTO*/*Fto* in hASC and 3T3-L1 adipocytes during early differentiation. (A) GWAS variants for BMI (Locke *et al.*, 2015) are visualized using LocusZoom (Pruim *et al.*, 2010) with rs1421085 as the tag SNP. (B) All chromatin state histograms are shown on the same scale and high values were truncated as necessary. Genomic and chromatin state maps from hASC (top) and 3T3-L1 (bottom) adipocytes. Orthology tracks show regions mapped from the mouse to the human genome (blue) and *vice versa* (orange). Purple vertical line highlights position of rs1421085 and orthologous site in mouse. Tracks in black show CTCF ChIP-seq signals. Alignment of the sequences underlying rs1421085 (red) in human and the orthologous site in mouse shows that the underlying motif is largely conserved. Different histone modifications mark different regulatory functions. For example, H3K4me3 marks transcriptional initiation and is found near promoters; H3K4me2/me1 and H3K27ac mark accessible chromatin and *cis*-regulatory activity (usually found across promoters, intronic and intergenic regions); H3K36me3 marks transcriptional elongation and is found across active gene bodies; and H3K27me3 marks Polycomb-mediated repression, mainly distributed across inactive regions (Mikkelsen *et al.*, 2010). Additionally, the position of CTCF is linked to chromatin states, as CTCF is one of the factors that plays a role in maintaining the boundary between distinct chromatin 'types' (Bannister & Kouzarides, 2011; Dixon *et al.*, 2012; Gómez-Marín *et al.*, 2015)

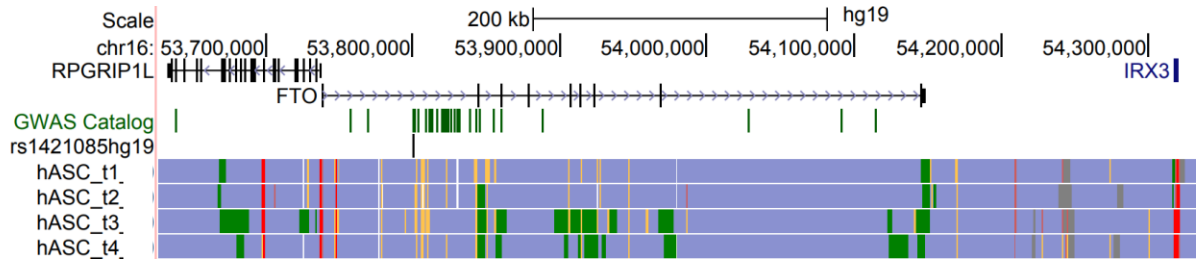


Fig. 3.2 | Chromatin annotation maps during hASC adipogenic differentiation.
ChromHMM (Ernst & Kellis, 2012) was used to annotate adipogenesis in hASCs (Mikkelsen *et al.*, 2010) at 4 time-points during differentiation corresponding to proliferating cells 2 days before adipogenic induction (t1), in confluent cells at the induction of differentiation (t2), in early differentiating cells 3 days post induction (t3) and at terminal differentiation at day 9 (t4). Data was visualized using the UCSC Genome Browser (Kent *et al.*, 2002). Combinatorial analysis of histone marks can be used to generate chromatin state annotations using ChromHMM (Ernst & Kellis, 2012), which allows the characterization of the regulatory landscape in a single, color-coded track and define regions as e.g. promoters (red) and enhancers (yellow). A colour code for all automated chromatin states can be found in Appendix Fig. S.1.

3.2.3 ATAC-seq

To assess the spatial-temporal nature of the regulatory element within *FTO*/*Fto* intron 1, I performed an assay for transposase accessible chromatin followed by deep sequencing (ATAC-seq), a technique recently developed by Buenrostro *et al.* (2013). ATAC-seq enabled me to map open chromatin regions that are accessible for TF binding in both human and mouse. To do so, I adopted the original protocol (Buenrostro *et al.*, 2015) for pre-adipocytes differentiated *in vitro* and mature adipose tissue (see Materials and Methods for details) and added a nuclei preparation step that improved quality of generated data. To enable comprehensive epigenomic profiling of adipocytes undergoing differentiation, I expanded human pre-adipocytes and mouse primary pre-adipocytes and induced differentiation using an adipogenic cocktail as described in Materials and Methods.

Immortalized human pre-adipocytes (hWAT) were obtained from Xue *et al.* (2015) and differentiated following their protocol. Differentiation of these cells was completed at day 24 of differentiation. At four time-points during differentiation, nuclei were isolated and ATAC libraries prepared. These time-points corresponded to cells before adipogenesis was induced (D0), during early (D3) and advanced differentiation (D6), as well as at terminal differentiation (D24) (Fig. 3.3).

For mouse, I isolated primary pre-adipocytes from 6 week old B6N mice by treating mature fat pads with collagenase followed by culturing the SVF containing adipocyte precursors. ATAC libraries were generated for 7 time-points during adipocyte differentiation and for one mature tissue sample originating from iWAT and gWAT from female and male mice (Fig. 3.3).

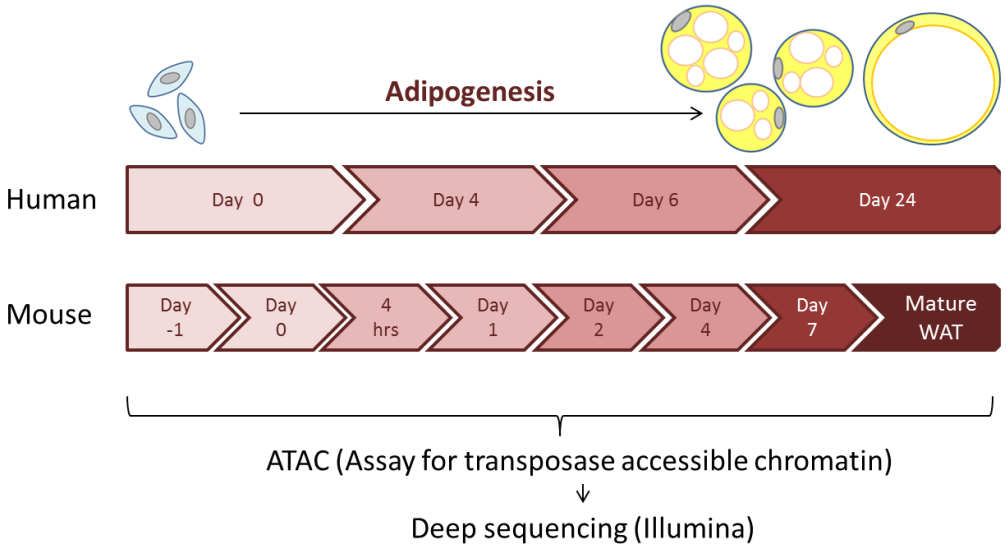


Fig. 3.1 | Experimental design.
ATAC libraries were prepared in human and mouse pre-adipocytes at 4 (human) and 8 (mouse) different developmental stages throughout differentiation, followed by deep sequencing. Terminal differentiation was achieved at day 24 of differentiation in hWAT and day 7 in mouse cells.

QC steps

Libraries were assessed for quality using multiple QC steps including the evaluation of differentiation, nuclei quality and yield as well as the tagmentation efficiency using the Bioanalyzer for DNA size profiling (Fig. 3.4A,B). During the initial method adaptation in mouse primary pre-adipocytes, ATAC libraries were further assessed for tagmentation efficiency using gel electrophoresis and qRT-PCR for two known closed and open regions (Fig. 3.4C,D,E). Following sequencing, data quality was further assessed by determining the complexity (how many reads are duplicated), how many reads map to the mitochondrial genome, how many reads map uniquely to the genome, what the insert size distribution looks like and whether sequenced libraries resemble an expected pattern. All libraries passed QC standards set by the WTCHG (Oxford), where the sequencing was performed. Internal assessment further validated a high quality of generated data sets, based on low duplication rate (average 17%) and an acceptable amount of reads mapping to the mitochondrial genome (between 15-30%). Additionally, the insert size distribution showed the expected nucleosome patterning that is characteristic for ATAC-seq libraries and further showed the optimal size distribution of DNA fragment, with high read counts around 200bp fragments (Fig. 3.5).

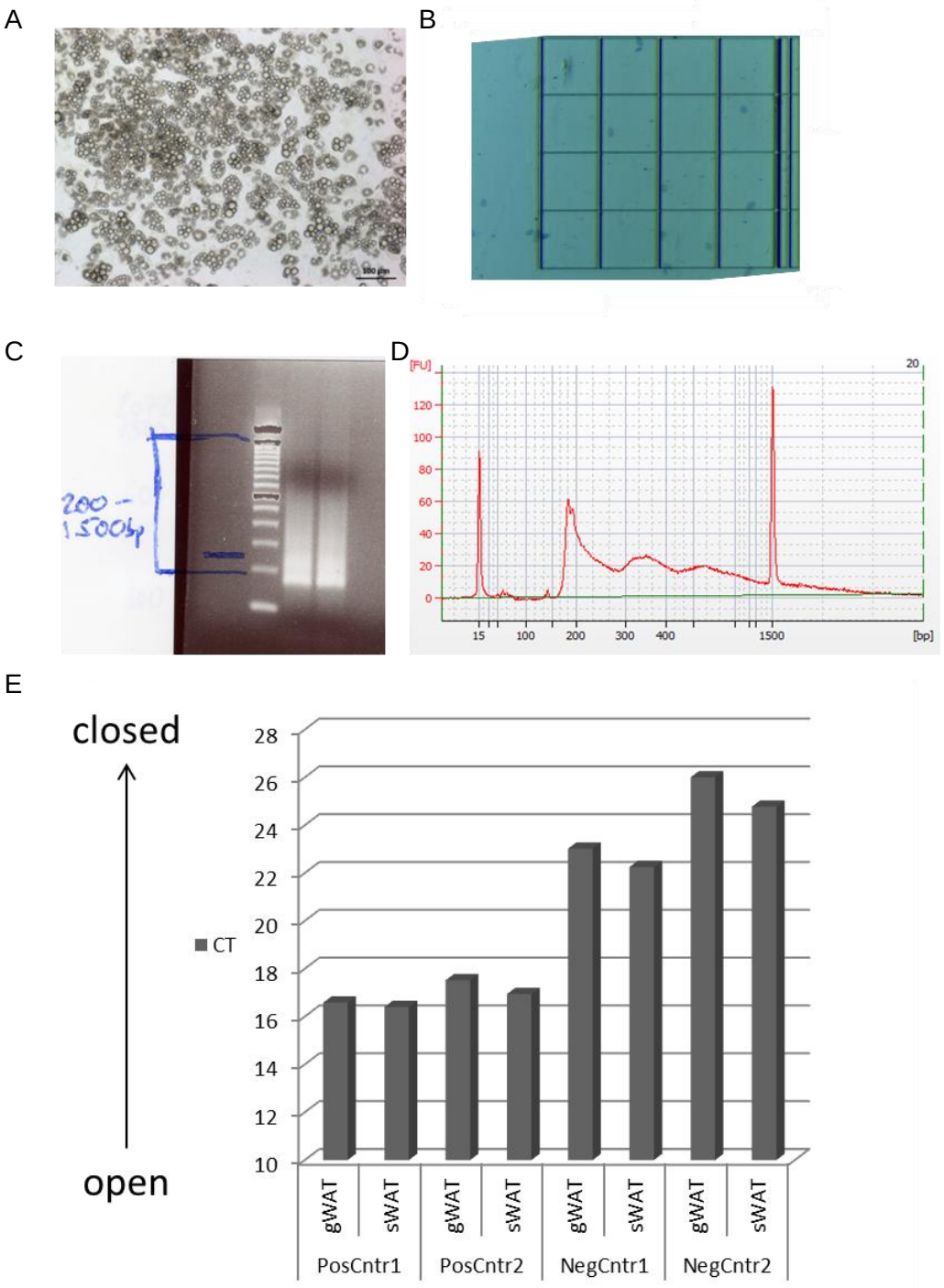


Fig. 3.2 | QC steps during ATAC protocol adaptation for *in vitro* differentiated and mature adipocytes.
(A) Differentiation was successful (B) Nuclei are intact following lysis (nuclei are stained with trypan blue and visualised under a light microscope) (C, D) Tagmentation was successful as DNA fragment size is in expected range as evaluated using gel electrophoresis (C) and Bioanalyzer DNA1000 chip (D) and qPCR confirmed that there are more fragments in the sample pool for two known open regions in the genome compared to two known closed ones.

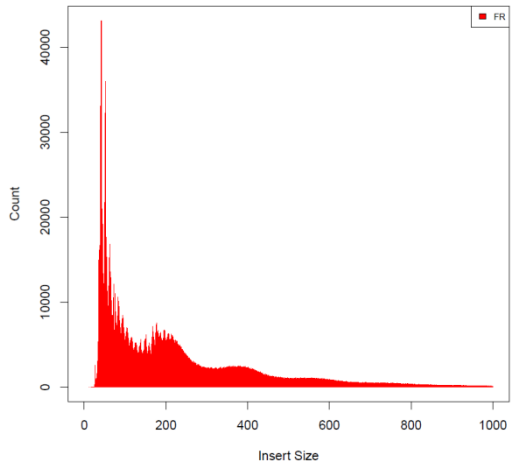
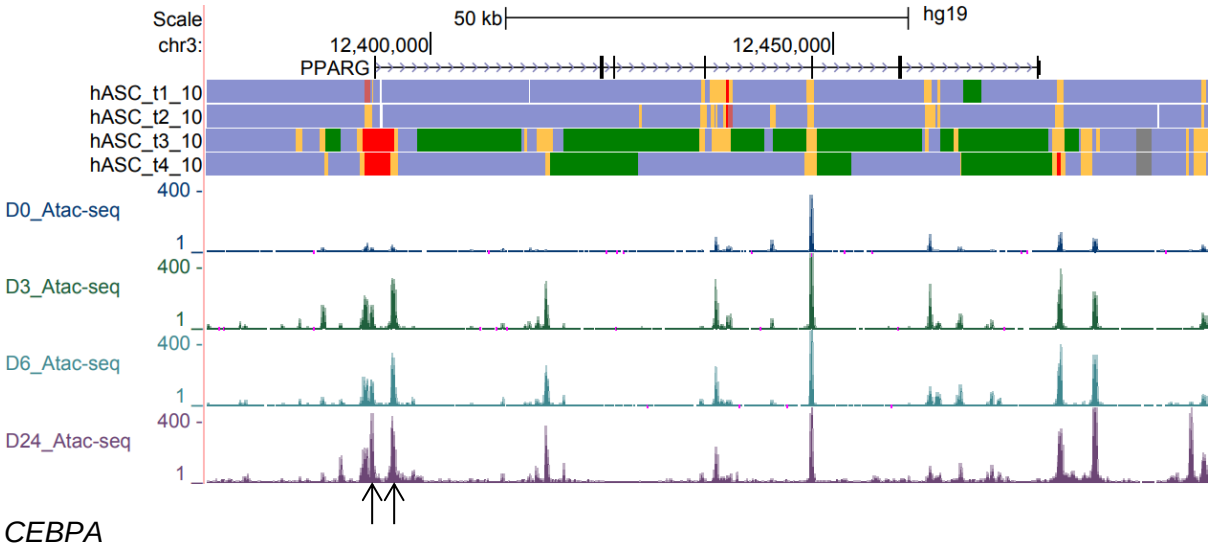


Fig. 3.3 | Representative image of ATAC-seq insert sizes determined by high-throughput sequencing.
Graph showing the fragment size distribution of a representative ATAC-seq library, with the majority of peaks being around 200 base pairs in size. ATAC-seq libraries showed characteristic nucleosome patterning and expected fragment size distribution. More details about the generation of libraries can be found in Materials and Methods.

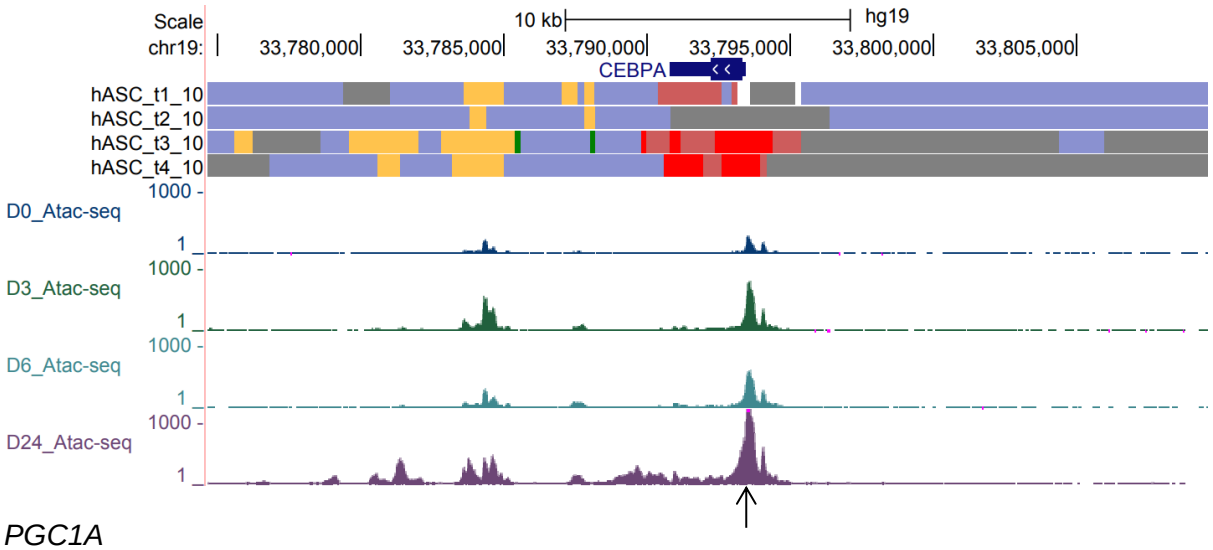
The raw sequencing data was then put through an analysis pipeline as described in Materials and Methods. Once analysed, histograms were generated for known adipogenesis- and adipocyte-related genes (positive controls) and two house-keeping genes (negative controls) to further assess whether biologically meaningful data sets were produced. In this thesis, I am showing the representative results for the hWAT ATAC-seq libraries. ATAC-seq data together with chromatin state annotations using ChromHMM (Ernst & Kellis, 2012) from hASC at 4 time-points during differentiation (Mikkelsen *et al.*, 2010) were visualized using the UCSC Genome Browser. ATAC-seq peaks overlap histone modifications for promoters (red) and regulatory elements such as enhancers (yellow), suggesting that the sequenced libraries indeed are associated with regulatory genomic features (Fig. 3.6). Furthermore, as expected, promoter regions dynamically become accessibly for genes that are upregulated during differentiation, as shown by the increasing promoter peak strengths for *PPARG, CEBPA* and *PGC1A* with the progress of differentiation in hWAT (Fig. 3.6). Furthermore, gene promoters associated with genes important in lipid-filled mature adipocytes (*FASN, PLIN1, ADIPOQ, FABP4*) were only open in the later stages of adipogenesis and closed in pre-adipocytes. As anticipated, the two house-keeping genes (*CANX, EIF3F*) showed open regions around gene promoters in pre-adipocytes that remained open throughout differentiation (Fig. 3.6).

A Genes known to be upregulated during early differentiation

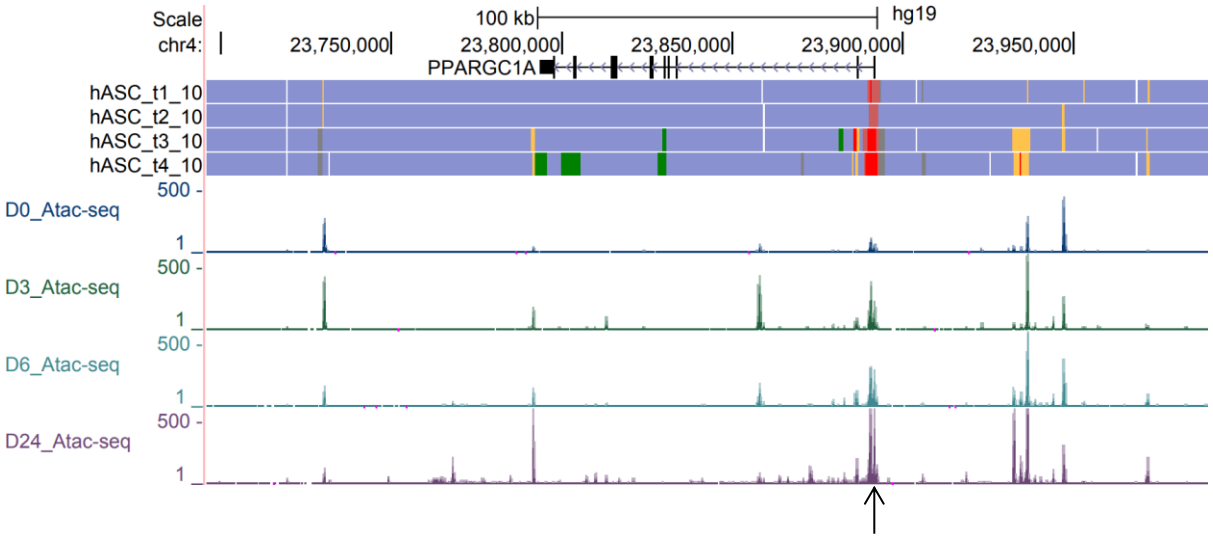
PPARG



CEBPA

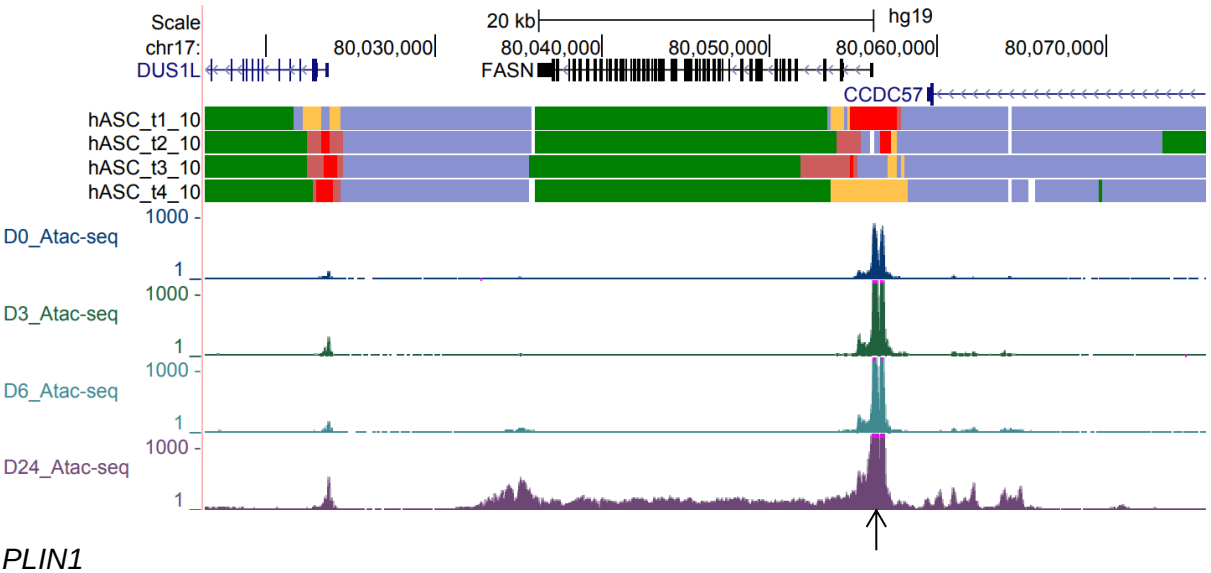


PGC1A

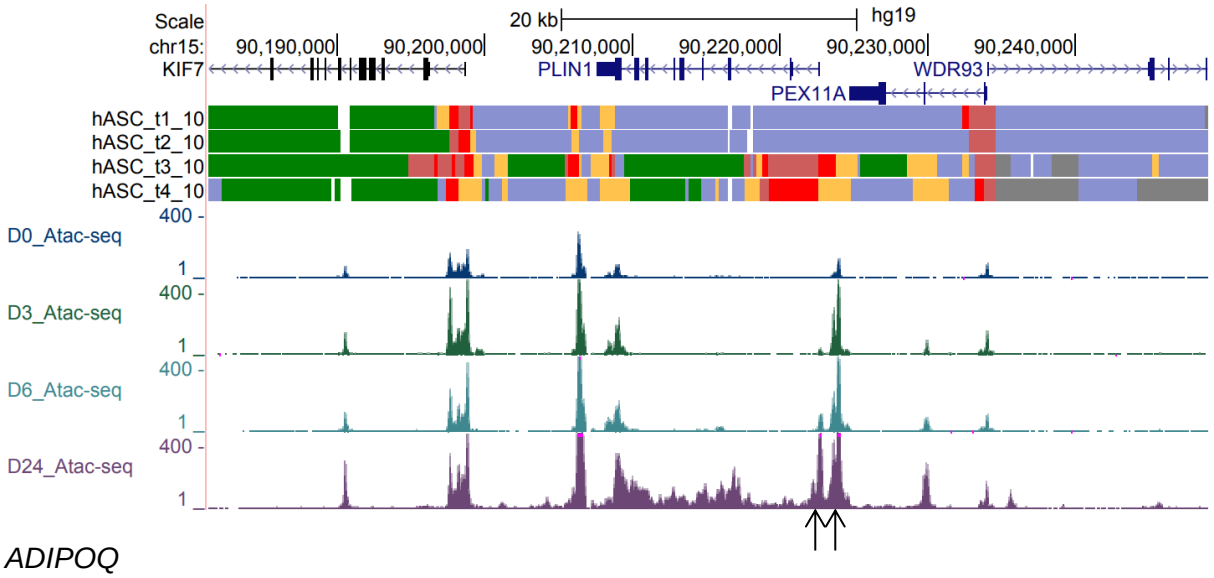


B Genes important in lipid-filled mature adipocytes

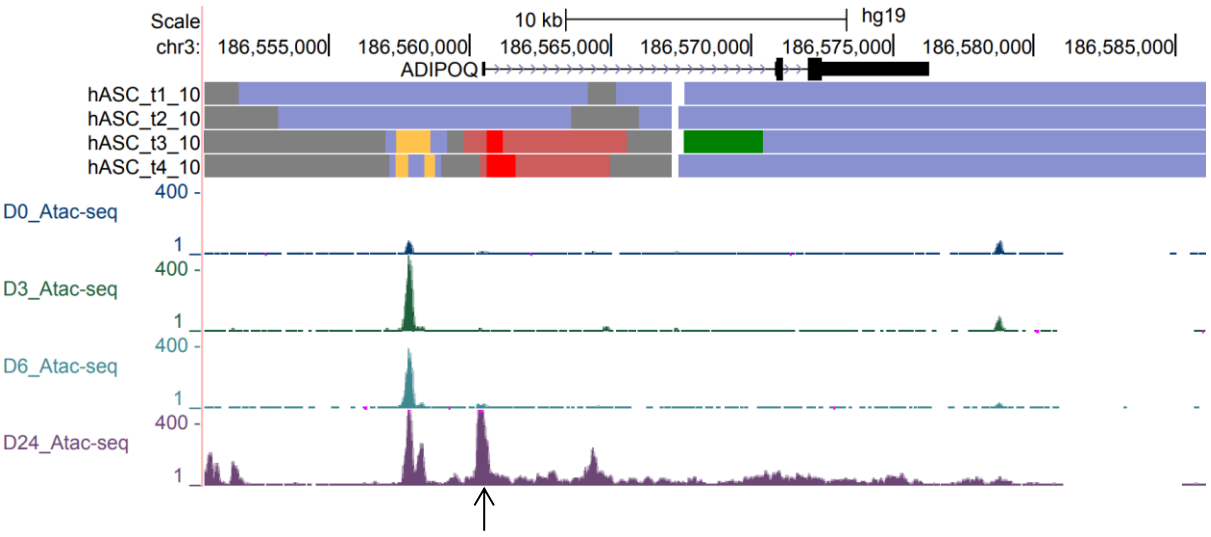
FASN



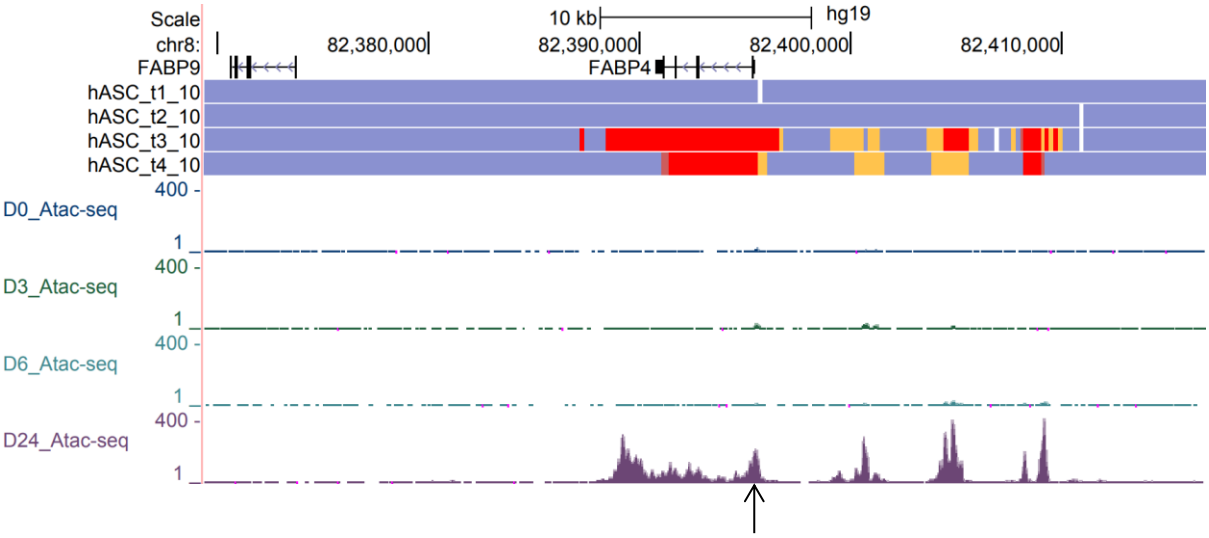
PLIN1



ADIPOQ

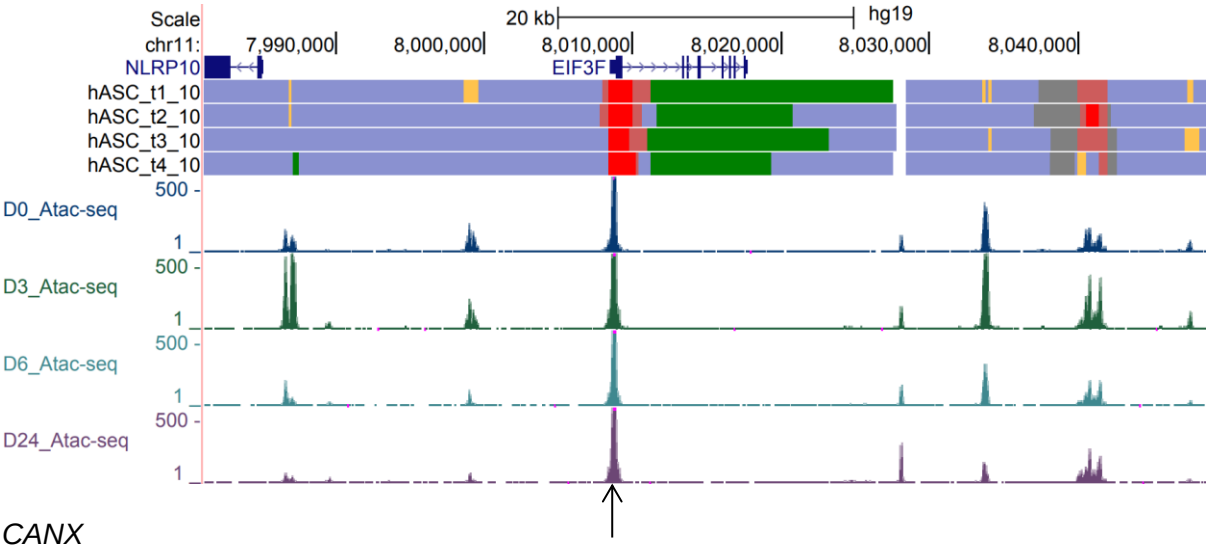


FABP4



C House-keeping genes, expressed throughout differentiation

EIF3F



CANX

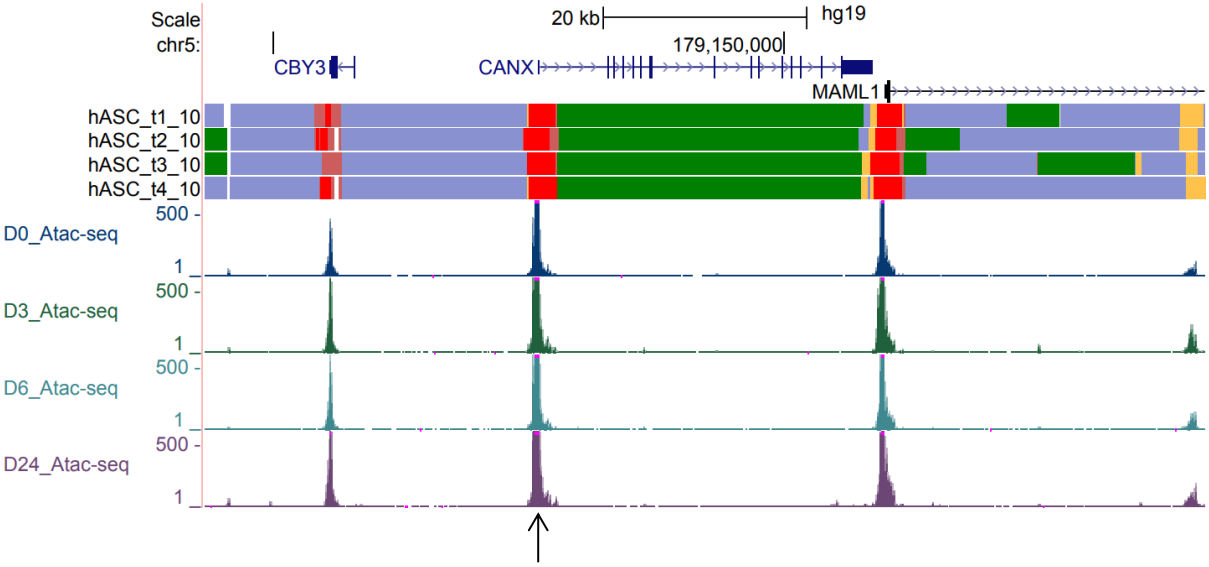


Fig. 3.4 | Regulatory landscape during human adipogenesis.
To assess the quality of the generated ATAC-seq data in hWAT using a biological context, ATAC-seq profiles were visualized over gene promoters during differentiation. ATAC-seq data was visualized using the UCSC Genome Browser (Kent *et al.*, 2002). Chromatin state annotations using ChromHMM from human adipocytes (hASC) during 4 time-points during differentiation (top to bottom: day-2, day0, day3, day9) (Mikkelsen *et al.*, 2010) are shown in the first 4 tracks in each image. The bottom 4 tracks show ATAC-seq histograms from hWAT adipocytes (Xue *et al.*, 2015) during 4 time-points during differentiation (top to bottom: day0, day3, day6, day24). All histograms within each browser screenshot are shown on the same scale and high values were truncated as necessary. Black vertical arrows below each graph point to gene promoter of interest. ATAC-seq peaks overlap histone modifications for promoters (red) and regulatory elements such as enhancers (yellow). (A) shows genes that are known to open following adipogenic induction and that are important regulators throughout differentiation (*PPARG, CEPBA, PGC1A*). (B) shows 4 genes that are known to be expressed in mature adipocytes and that regulate lipid metabolism and adipocyte function (*FASN, PLIN1, ADIPOQ, FABP4*). (C) shows two known house-keeping genes (*EIF3F, CANX*) that are expressed throughout differentiation.

3.2.3.1 Comparison of human and mouse ATAC-seq data

I then set out to investigate the regulatory landscape around rs1421085 at the *FTO* locus. My results show that in hWAT, there is an accessible region within *FTO* intron 1 that is in close proximity to rs1421085 (Fig. 3.7). This regulatory element opens shortly after adipogenesis is induced and becomes inaccessible in fully differentiated adipocytes, suggesting that regulatory activity of this element is developmentally restricted to early differentiation of adipocytes. Additionally, the open chromatin regions around rs1421085 overlap with enhancer marks in adipose as shown using chromatin annotation maps from the Epigenomic Roadmap (Fig. 3.7 top track in human), consistent with the data presented previously in this chapter (3.2).

Interestingly, mouse adipocytes show a similar temporal pattern to human adipocytes. As in human cells, there is a transiently open region in *Fto* intron 1 in mouse adipocytes at the orthologous region. Once differentiation is induced, the corresponding regulatory element close to rs1421085 in mouse opens and then becomes inaccessible as differentiation progresses, being closed in fully differentiated cells at day 7 of differentiation and in mature tissue (Fig. 3.8). Although present in both iWAT- and gWAT-derived pre-adipocytes, the enhancer peaks are stronger in gWAT-derived adipocytes. Data presented in Fig. 3.7 is matched for sex and originates from females for both, human and mouse. These ATAC-seq profiles clearly revealed the presence of a regulatory region around human rs1421085 and mouse orthologous region that overlaps with enhancer marks and that opens during early adipogenesis. This was true for female and male samples in the two major adipose depots examined (Fig. 3.7; Fig. 3.8), although the temporal pattern of this enhancer element was less striking in male adipocytes. Together my results suggest that regulation at this *FTO* regulatory region is independent of species, depot (based on the available mouse ATAC data from gWAT and iWAT) or sex. Together these ATAC-seq results show that there is a conserved regulatory element around rs1421085 and that the histogram profiles a very similar in human and mouse.

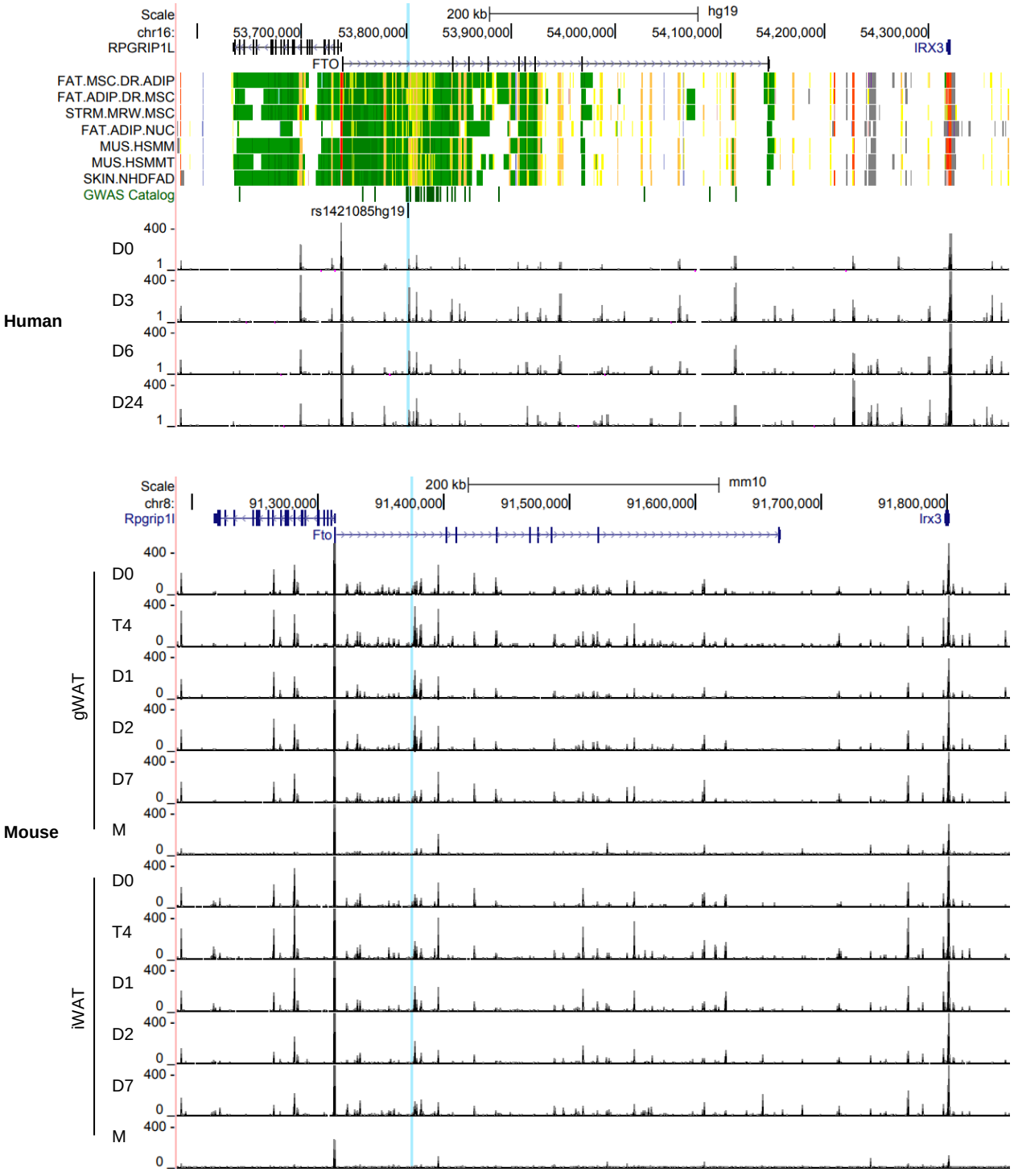


Fig. 3.5 | Accessible chromatin profiles in human and mouse adipocytes at the *FTO* locus.
The top track shows human data (hg19 - chr16:53580916-54376648) and the bottom track shows mouse data (mm10 - chr8:91188110-91850320). Orthologous chromosome positions were determined using the LiftOver tool from UCSC genome browser (Kent *et al.*, 2002). Human ATAC-seq was performed at 4 time-points (day 0, day 3, day 6, day 24) during adipocyte differentiation in hWAT of a female donor (Xue *et al.*, 2015). Mouse ATAC-seq libraries were generated in primary pre-adipocytes originating from female subcutaneous (iWAT) and visceral (gWAT) fat depots. Data presented here is from 5 time-points during differentiation (day 0, 4 hours, day 1, day 2, day 7) and one mature adipose tissue sample (M) for each depot. Blue line represents the position of rs1421085 and the mouse orthologous region. ATAC-seq data shown in this figure is matched for sex, as all ATAC data was derived from females. All histograms are shown on the same scale and high values were truncated as necessary. Additional information can be found in Materials and Methods section.

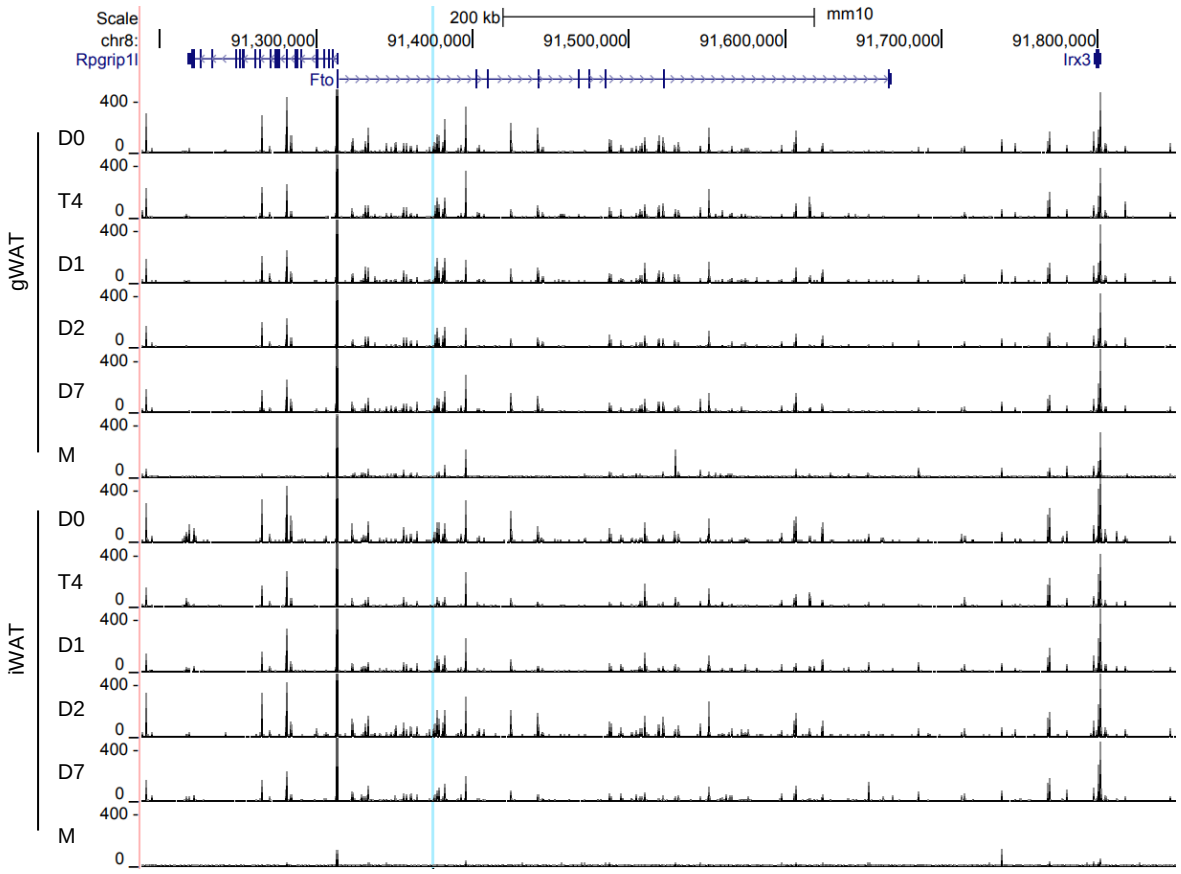


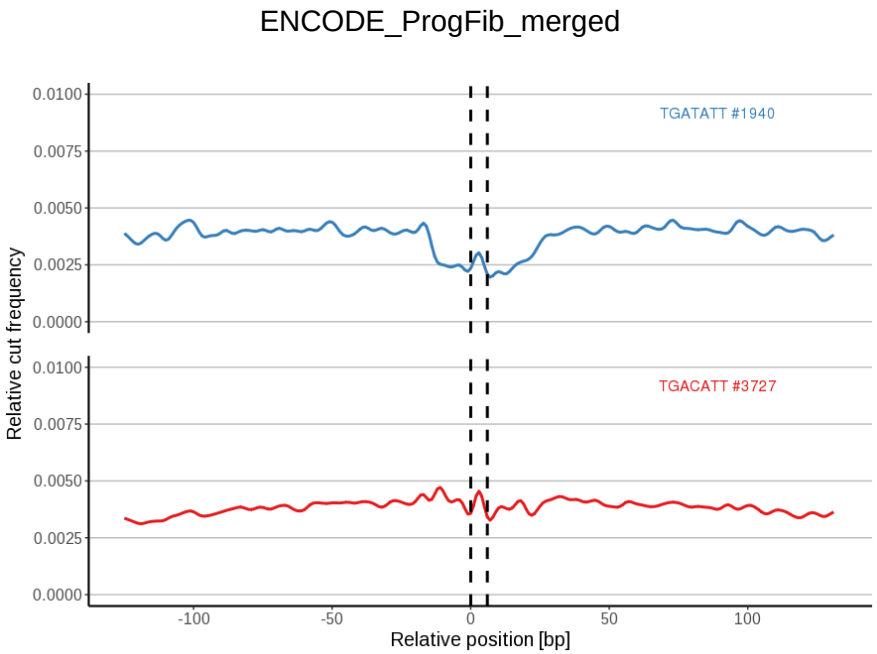
Fig. 3.6 | Accessible chromatin profiles in male mouse adipocytes at the *Fto* locus.
ATAC-seq during time-course differentiation in male adipocytes from subcutaneous and visceral depots at the *Fto* locus (mm10 - chr8:91188110-91850320). Mouse ATAC-seq libraries were generated in primary pre-adipocytes originating from male subcutaneous (iWAT) and visceral (gWAT) fat depots. Data presented here is from 5 time-points during differentiation (day 0, 4 hours, day 1, day 2, day 7 of adipogenic stimulation) and one mature adipose tissue sample (M) for each depot. Dotted line represents the position of the mouse orthologous region around rs1421085. All histograms are shown on the same scale and high values were truncated as necessary. ATAC-seq data was visualized using the UCSC Genome Browser (Kent *et al.*, 2002). Additional information can be found in Materials and Methods section.

3.2.4 DNase1 Footprinting

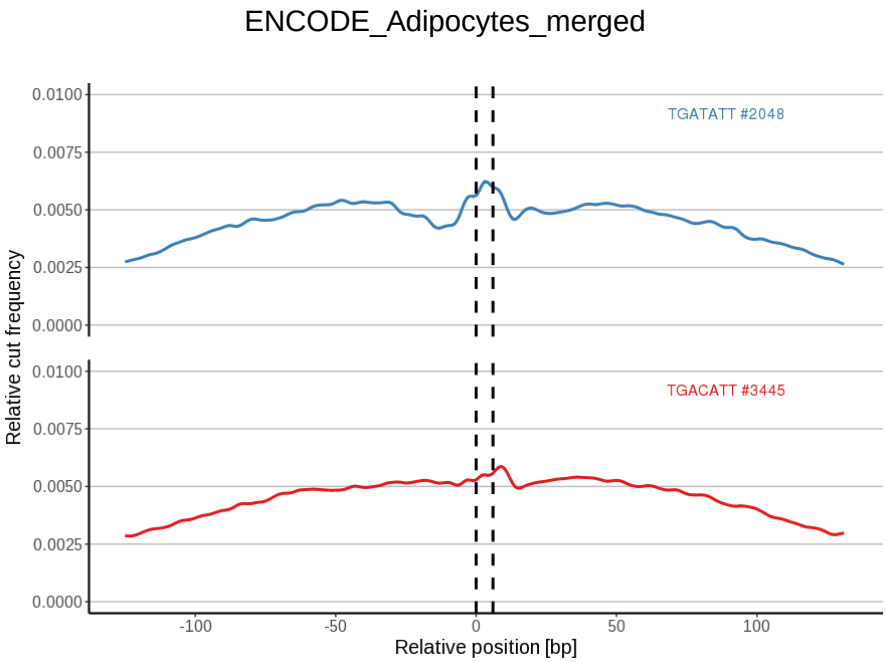
Enzymes that are used to assay accessible chromatin, such as DNase1, cut genomic DNA, but are obstructed by proteins (e.g. nucleosomes or TFs) bound to the DNA. Therefore, DNase-seq can be used to map open regions that are associated with regulatory activity. By increasing the resolution, data generated using DNase1 can be used to map the cut sites at base pair resolution, creating protein footprints. Jim Hughes and his team recently developed a webtool called Sasquatch which is a novel computational approach that uses DNase1 footprint data to estimate and visualize the effects of non-coding variants on TF binding (http://apps.molbiol.ox.ac.uk/sasquatch/cgi-bin/foot.cgi). The webtool incorporates publicly available DNase1-seq data, including data from ENCODE (Yue *et al.*, 2014), making it possible to prioritize SNPs in their tissue-specific context. Due to the current lack of DNase1 data in pre-adipocytes, I focused my analysis on fibroblasts (as undifferentiated pre-adipocytes are fibroblast-like) and mature adipocytes. ATAC-seq derived data sets are not suitable for footprint analysis, as the enzyme used in this assay removes bound TFs from DNA (Schwessinger *et al.*, 2017).

Using this tool, I tested the effect of risk variant rs1421085 within a 7-mer sequence with the SNP in the centre (Fig. 3.9C). In fibroblasts, the T allele shows a footprint, suggesting TF occupancy at this sequence. However, no such pattern is observed in the C allele, suggesting that the risk allele interrupts TF binding (Fig. 3.9). In adipocytes, there was no footprint in either allele (Fig. 3.9), indicating that none of the proteins present in mature adipocytes are binding to the motif around rs1421085. This provides further evidence for a tissue-specific nature and developmental time-point dependency for the effect of rs1421085. This also points towards a function of rs1421085 in which the C risk allele interrupts a TFBS in fibroblasts.

A



B



C

	kmer (ref)	kmer (var)	sfr (ref)	sfr (var)	damage	change (%)
ProgFib	TGA**T**ATT	TGA**C**ATT	1.558	1.127	1.56	77.3
Adipocytes	TGA**T**ATT	TGA**C**ATT	1.055	1.053	0.002	3

Fig. 3.7 | Transcription factor footprints at rs1421085.
TF footprints were visualized using the Sasquatch tool (http://apps.molbiol.ox.ac.uk/sasquatch/cgi-bin/foot.cgi), which uses DNase1 footprint data to estimate the effects of non-coding variants on TF binding. The webtool incorporates publicly available DNase1-seq data, including data from ENCODE (Yue *et al.*, 2014). The effect of risk variant rs1421085 on TF binding was tested within a 7-mer sequence with the SNP in the centre using DNase1-seq data from fibroblasts (A) and adipocytes (B). The risk C allele for rs1421085 interrupts a TFBS in fibroblasts (A,C), but not mature adipocytes (B,C).

3.2.5 Hi-C

Mammalian genomes are organized in 3D space and for distal regulatory elements (e.g. enhancers, repressors, etc) to affect target gene expression, the DNA of regulatory elements needs to be in physical contact with target gene promoters. *Cis*-regulatory networks tend to function within the boundaries of TADs. To compare the spatial organization between human and mouse, I used the 3D Genome Browser (http://3dgenome.org) (Wang *et al.*, 2017) to visualize Hi-C data generated by Dixon *et al.* (2012) in embryonic stem cells from human (H1-ESC) and mouse (mESC) (Fig. 3.10). Hi-C chromatin interactions are shown as heatmaps and the values on a colour scale correspond to the number of times that reads in two 40kb bins were sequenced together (red – stronger interaction, white – little or no interaction). TADs form red triangles.

Hi-C chromatin interaction data shows that in human embryonic stem cells, regions in FTO intron 1 are in physical contact with a number of gene promoters, including *RPGRIP1L, FTO, IRX3, IRX5*, and *IRX6*. In mouse ESCs, there is a comparably sized TAD that also encompasses *Rpgrip1l, Fto, Irx3, Irx5* and potentially *Irx6*. The TAD boundaries in human and mouse appear to mirror-image one another around the greater *FTO* locus.

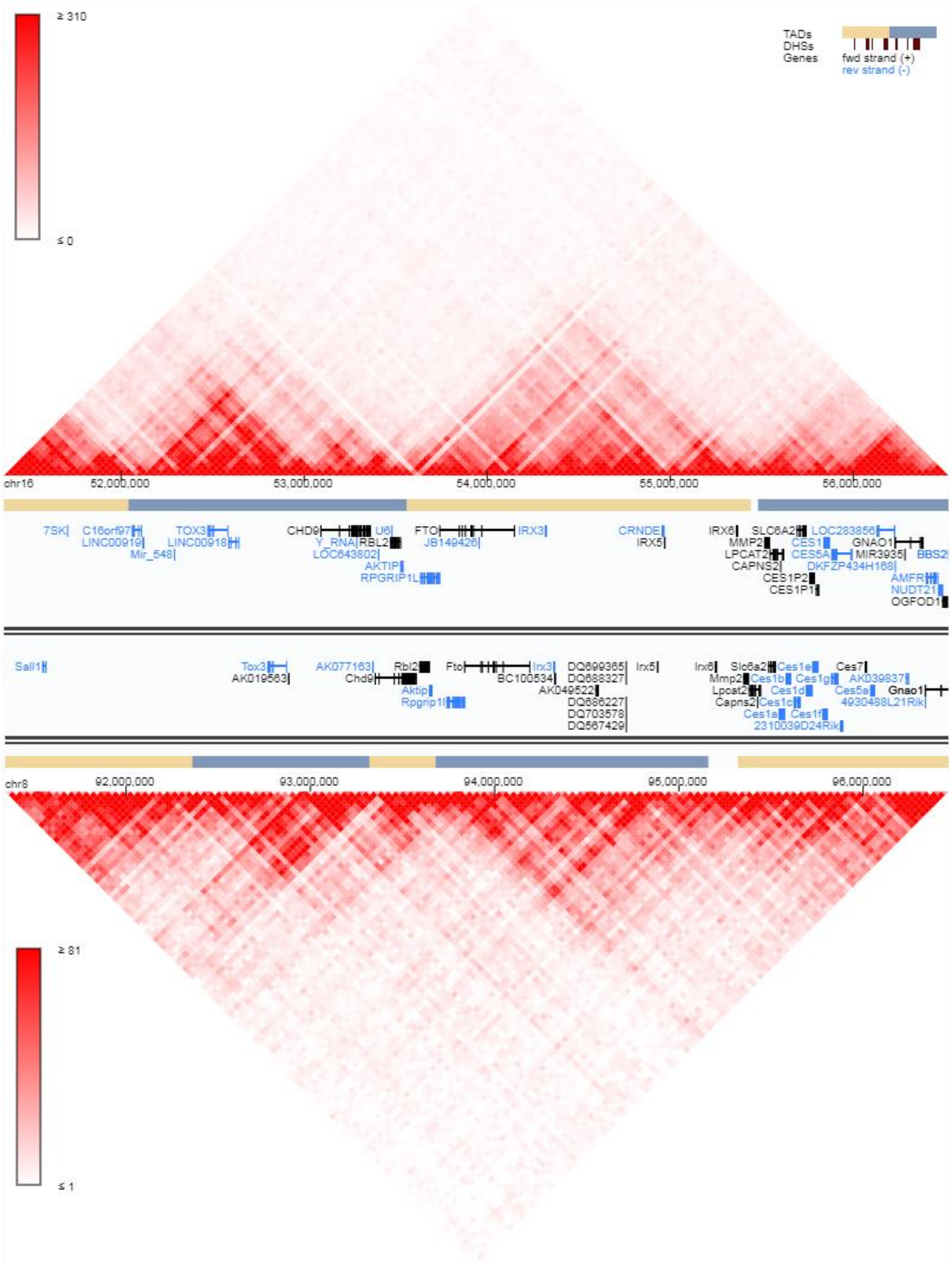


Fig. 3.8 | Comparing 3D chromatin interaction Hi-C data in human and mouse embryonic stem cells (ESC).
Hi-C data from human H1-ESC (top track) and mouse mESC (bottom track) are visualized in the 3D Genome Browser (http://3dgenome.org) (Y. Wang *et al.*, 2017). Hi-C data for H1-ESC and mESC is from (Dixon *et al.*, 2012). The heatmap values on a colour scale correspond to the number of times that reads in two 40kb bins were sequenced together (red – stronger interaction, white – little or no interaction). TADs form red triangles.

3.3 Discussion

Comparative epigenomic profiling of *FTO/Fto* suggests a deep functional conservation at this locus. Firstly, visualizing and comparing chromatin interactions in embryonic stem cells from human and mouse showed striking similarity of TADs across the *FTO* locus. Hypothetically, regulatory alterations in FTO intron 1 could affect the same target genes in human and mouse. Histone modification maps further point towards a similar pattern of the regulatory landscape at the *FTO/Fto* locus. Using publicly available chromatin maps, I found that the disease-associated regions within *FTO* intron 1 are not only highly conserved on the sequence level, but also in regards to histone modifications. This is in line with findings from the Mouse ENCODE Consortium (Yue *et al.*, 2014), which has shown that human GWAS SNPs with high regulatory potential are enriched in conserved regulatory binding sites and supposed that mouse histone modifications can inform about human risk variants.

FTO intronic variants are enriched for enhancer marks, suggesting that they interrupt regulatory activity. For a variant to alter regulatory activity and to interrupt a TFBS, it is believed that the chromatin has to be accessible. However, histone modifications are independent of accessible chromatin and are not restricted to 'open' chromatin. Performing ATAC-seq during adipogenesis revealed that not only do human and mouse share their regulatory regions in *FTO/Fto* intron 1, there is also the same temporal pattern of regulation where the regulatory element around rs1421085 opens after adipogenic induction and closes in fully matured cells.

In summary, comparative epigenomic profiling of *FTO/Fto* revealed the presence of a shared regulatory landscape across the relevant disease-associated intronic region. I hypothesize that manipulation of *FTO/Fto* intronic regulatory elements would result in similar downstream perturbations. These findings open opportunities for novel gene manipulation strategies *in vivo* to mechanistically dissect the *Fto* regulatory circuitry and to study the potential sex- and depot-specific nature of the obesity-signal in mouse.

4 Phenotypic consequences of DEL82

4.1 Introduction

Variants in *FTO* intron 1 are strongly associated with BMI in the human population (Frayling *et al.*, 2007; Locke *et al.*, 2015; Scuteri *et al.*, 2007). These variants are in a non-coding region that is enriched for gene regulatory marks and that forms functional connections to distal gene promoters (Claussnitzer *et al.*, 2015; Ragvin *et al.*, 2010; Smemo *et al.*, 2014), suggesting that SNPs exert their effect through gene regulatory mechanisms. However, out of the 89 SNPs in high LD, the causative variant(s) and their target tissues are challenging to identify. Human studies recently identified rs1421085 as a functional variant in adipocytes (Claussnitzer *et al.*, 2014, 2015), with high phylogenetic module conservation. To assess the regulatory conservation and further dissect the mechanism of this *Fto* intronic variant, we exploited CRISPR/Cas9 genome editing (Ran *et al.*, 2013) to create a mouse model that harbours a point mutation in the mouse orthologous region that corresponds to location of rs1421085. As a welcomed side product, genome editing in the first instance resulted in a mouse mutant that harboured a 82 nucleotide deletion that disrupted the motif around the target nucleotide. Specifically, this mutant harboured a mutation that deletes two nucleotides upstream and 79 nucleotides downstream of the desired SNP location (Fig. 4.1). Given that this additional allele targets the regulatory element of interest, it was taken forward for generation of a cohort. Homozygous animals harbouring this deletion were named DEL82. To assess the phenotypic consequences of alterations of this regulatory module, DEL82 mice underwent a metabolic phenotyping pipeline (Fig. 4.2), followed by terminal tissue and blood collection for further functional analysis. Mice were either fed a HFD or a matched LFD from weaning at 3 weeks of age. Further details about mouse husbandry, randomization and blinding of experiments can be found in the Materials and Methods. Common variants implicated in common disease by GWAS have characteristically a modest to low effect size and effects are often dependent on environmental stimuli and challenges. To calculate the number of mice needed to detect a subtle body weight phenotype, power calculations were

performed based on a B6N body weight data set generated in the group, using the online tool http://hedwig.mgh.harvard.edu/sample_size/js/js_parallel_quant.html. I initially calculated that with a significance level of 0.05 (two-sided) and a standard deviation of 2.5 of the outcome variable (based on previous data) and a power of 0.9 and a minimal detectable difference of 2.3g, I would require 26 animals per genotype to detect statistical significance. Despite this aim, due to breeding delays, the data presented in this thesis is collected from 11-17 animals per genotype. *In vivo* phenotyping in this Chapter (Echo-MRI, IPGTT and retro-orbital bleeds) was performed by Liz Bentley.

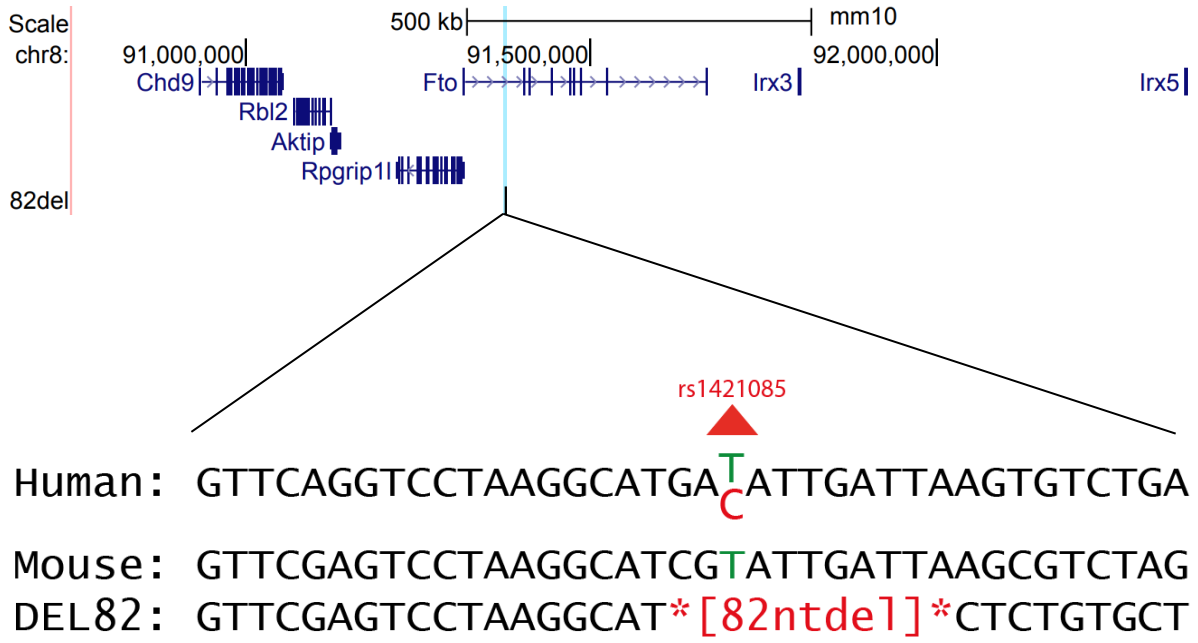


Fig. 4.1 | Schematic of DEL82 mutation in mouse.
Using CRISPR/Cas9 genome editing, a mouse model was created that harbours a deletion of 82 nucleotides overlapping the mouse orthologous region around human obesity-associated variant rs1421085 in intron one of *FTO*. Mice are on a B6N background. Further details about design of guide RNAs and generation of mouse mutant can be found in the Materials and Methods.

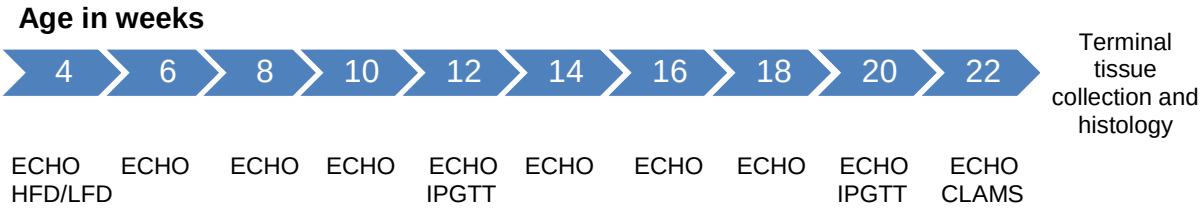


Fig. 4.2 | Metabolic phenotyping pipeline.
Male and female DEL82 WT, HET and HOM mice were put on a high-fat-diet (HFD) or low-fat-diet (LFD) at weaning and subsequently underwent a metabolic phenotyping pipeline. Every two weeks, mice were weighed and body composition measured using Echo-MRI (ECHO). For all animals, glucose tolerance was measured by means of an intra-peritoneal glucose tolerance test (IPGTT) at 12 weeks of age. For males, an additional IPGTT was performed at 20 weeks of age. Whole body energy expenditure was assessed using the CLAMS system (see Materials and Methods for details) for 20-22-week old male DEL82 WT and HOM mice on HFD feeding, as only males on HFD presented with a body weight and fat mass phenotype at this point. Terminal tissue and blood collection for downstream analysis was performed once the phenotyping battery was completed.

4.3 Results

4.3.1 Body Composition

To assess body composition, Echo-MRI was performed on DEL82 WT and HOM animals every two weeks from weaning. By removing 82 nucleotides of a putative enhancer element for *Irx3* and *Irx5*, I expected less promoter binding, reduced *Irx3* and *Irx5* expression in pre-adipocytes, and consequently reduced body weight and fat mass in DEL82 HOM mice. As predicted, male B6N mice harbouring a deletion around human rs1421085 had a reduced response to diet-induced obesity compared to WT, as apparent via decreased body weight gain due to reduced fat mass in males homozygous for DEL82 on a HFD (Fig. 4.3A,C). This was statistically significant between 14-18 weeks of age ($P<0.05$) and 14-16 weeks of age ($P<0.05$) for BW and fat mass, respectively (Fig. 4.3A,C). Lean mass was unaltered between genotypes (Fig. 4.3E). DEL82 did not have an effect on body weight in males that were on a LFD control diet (Fig. 4.3A,C,E), nor did this mutation have a detectable effect on body weight or fat mass in females on either diet (Fig. 4.3B,D,F).

Consistent with a decrease in overall body weight and fat mass, male DEL82 homozygotes had statistically significant decreased body weight gain ($P=0.039$) and fat mass gain ($P=0.019$) on a HFD between 4-18 weeks of age (Fig. 4.4), whereas male DEL82 HOM mice on LFD control diet did not present with a difference in body-weight or fat mass gain compared to WT (Fig. 4.4). Heterozygous DEL82 had an intermediate effect on body weight and fat mass gain on a HFD compared to homozygous alleles (Fig. 4.5), suggesting a dose-dependent effect of the DEL82 allele on body weight regulation and obesity development.

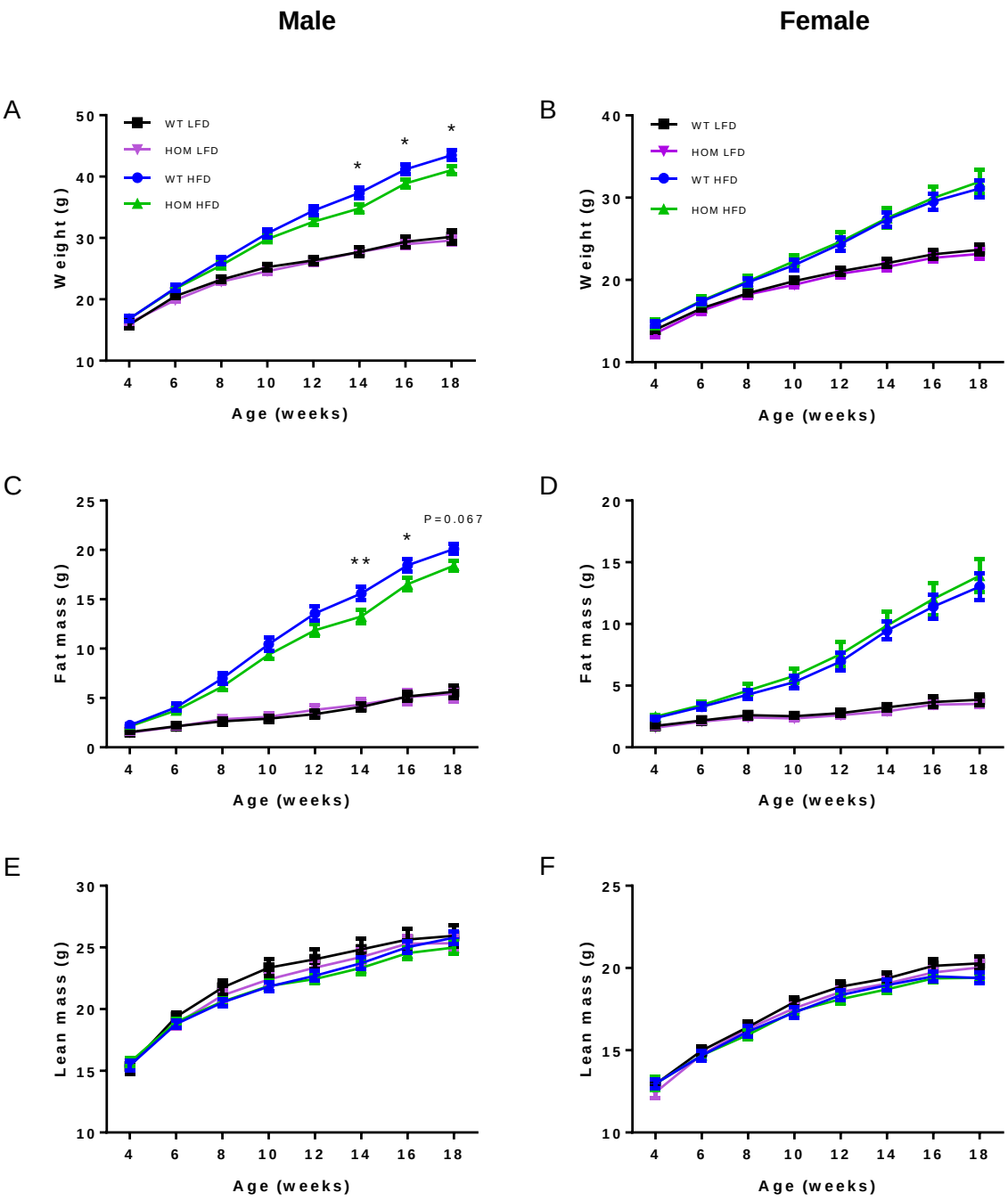


Fig. 4.3 | DEL82 causes reduced body weight and fat mass in males on a HFD.
(A + B) Body weight in male (A) and female (B) WT and DEL82 mice fed a LFD or HFD. (C + D) Fat mass in male (C) and female (D) WT and DEL82 mice on HFD and LFD. (E + F) Lean mass in male (E) and female (F) WT and DEL82 mice on HFD and LFD. N numbers for each group are: male HFD WT (n=17), HOM (n=17); male LFD WT (n=13), HOM (n=15); female HFD WT (n=17), HOM (n=14); LFD WT (n=11), HOM (n=12). Statistical significance was determined using repeated measures two-way ANOVA, Bonferroni's multiple comparisons test, data are expressed as mean ± SEM *P<0.05, **P<0.01.

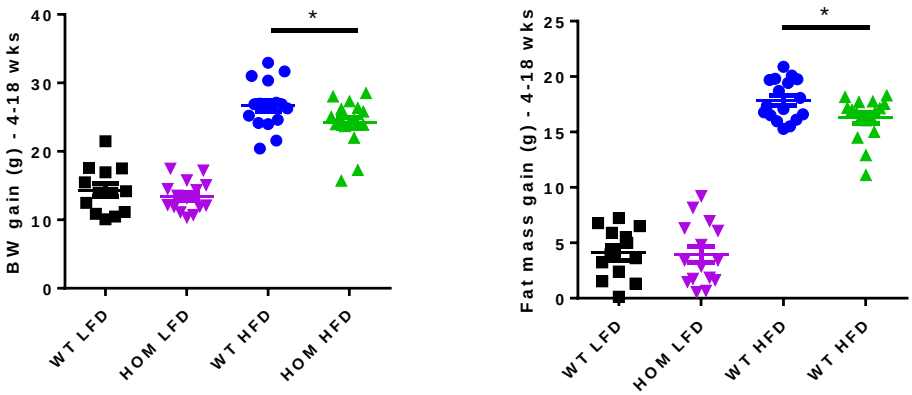


Fig. 4.4 | Decreased body weight and fat mass gain in male DEL82 on a HFD.
Body weight (BW) gain (left) and fat mass gain (right) were calculated between 4-18 weeks of age in male mice that received either a HFD or LFD. Statistical significance was determined using multiple comparisons one-way ANOVA, data are expressed as individual data points and mean ± SEM *P<0.05.

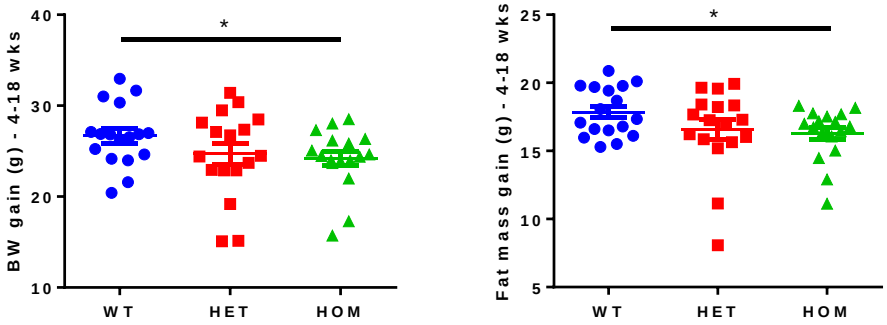


Fig. 4.5 | Heterozygous DEL82 has an intermediate phenotype on body weight and fat mass gain.
Data presented is generated using male mice on a HFD that are homozygous for the WT (blue), HET (red) or HOM (green) allele for the DEL82 mutation. Body weight (BW) gain (left) and fat mass gain (right) were calculated between 4-18 weeks of age. Statistical significance was determined using multiple comparisons one-way ANOVA, data are expressed as individual data points and mean ± SEM *P<0.05.

4.3.3 Glucose tolerance

To assess the effect of DEL82 on glucose tolerance, IPGTTs were performed in male and female mice on both HFD and LFD. In the human population, protective alleles within FTO have a protective effect on T2D risk, due to them having an effect on decreasing BMI. Based on decreased fat mass and body weight gain in male DEL82 HOMs on a HFD, one could expect these mutants to be mildly protected from developing T2D. However, given the subtle body weight difference between mutants and WT animals, I did not expect a significant effect of DEL82 on glucose tolerance.

At all time-points measured, DEL82 had no effect on glucose tolerance (Fig. 4.6). This was despite a moderate but significant decrease in body weight in male DEL82 HOM compared to WT on a HFD. In summary, although a HFD did cause reduced glucose tolerance in all groups tested, DEL82 animals did not present with an altered glucose tolerance phenotype, independent of sex, diet or age.

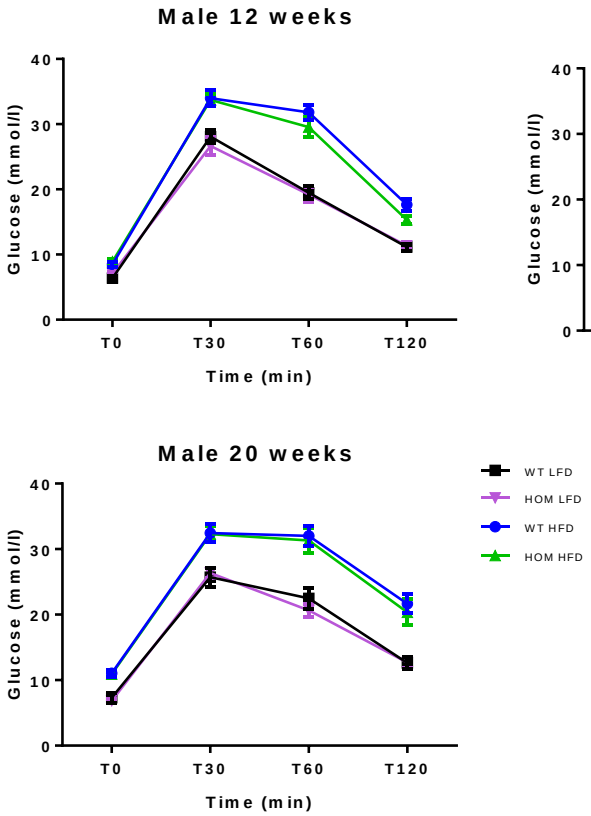


Fig. 4.6 | DEL82 does not alter glucose tolerance in intraperitoneal glucose tolerance tests. Mice were fasted overnight and intraperitoneal glucose tolerance tests (IPGTTs) were performed the next morning. N numbers for each group are: male HFD WT (n=17), HOM (n=17); male LFD WT (n=13), HOM (n=15); female HFD WT (n=17), HOM (n=14); LFD WT (n=11), HOM (n=12). Data was analysed using repeated measures two-way ANOVA, Bonferroni's multiple comparisons test, data are expressed as mean ± SEM.

4.3.5 Adipose Depots Weights

We established that only male DEL82 animals on a HFD present with a body weight and fat mass phenotype. Therefore, subsequent data collection aimed to further characterise the effect of DEL82 in this group. Therefore, more in-depth phenotyping was performed predominantly in DEL82 males on a HFD. Following completion of the metabolic phenotyping pipeline *in vivo*, male DEL82 mice on a HFD were sacrificed and tissues isolated for downstream analysis.

Additionally, fat depot weights were measured for iWAT and gWAT depots. As expected, there was no difference in iWAT, gWAT and BAT depot weights from female DEL82 HOM and WT animals (Fig. 4.7A). Unexpectedly, although overall fat mass was reduced, male mice homozygous for DEL82 showed a significant increase (P=0.002) in gWAT weights with an average of 1.3-fold (equivalent to 0.4g) increased gWAT fat mass compared to WT (Fig. 4.7A). iWAT fat mass was not significantly different in both groups. The difference in gonadal fat mass between DEL82 HOM and controls persisted when normalised to overall body weight (P=0.004) or fat mass (P=0.004) (Fig. 4.7B). Given the differential fat mass accumulation in DEL82 males, the iWAT/gWAT ratio is significantly lower for the mutant animals compared to WT animals (Fig. 4.7C; P=0.01), suggesting depot differences in fat mass accumulation during the development of obesity in these mice.

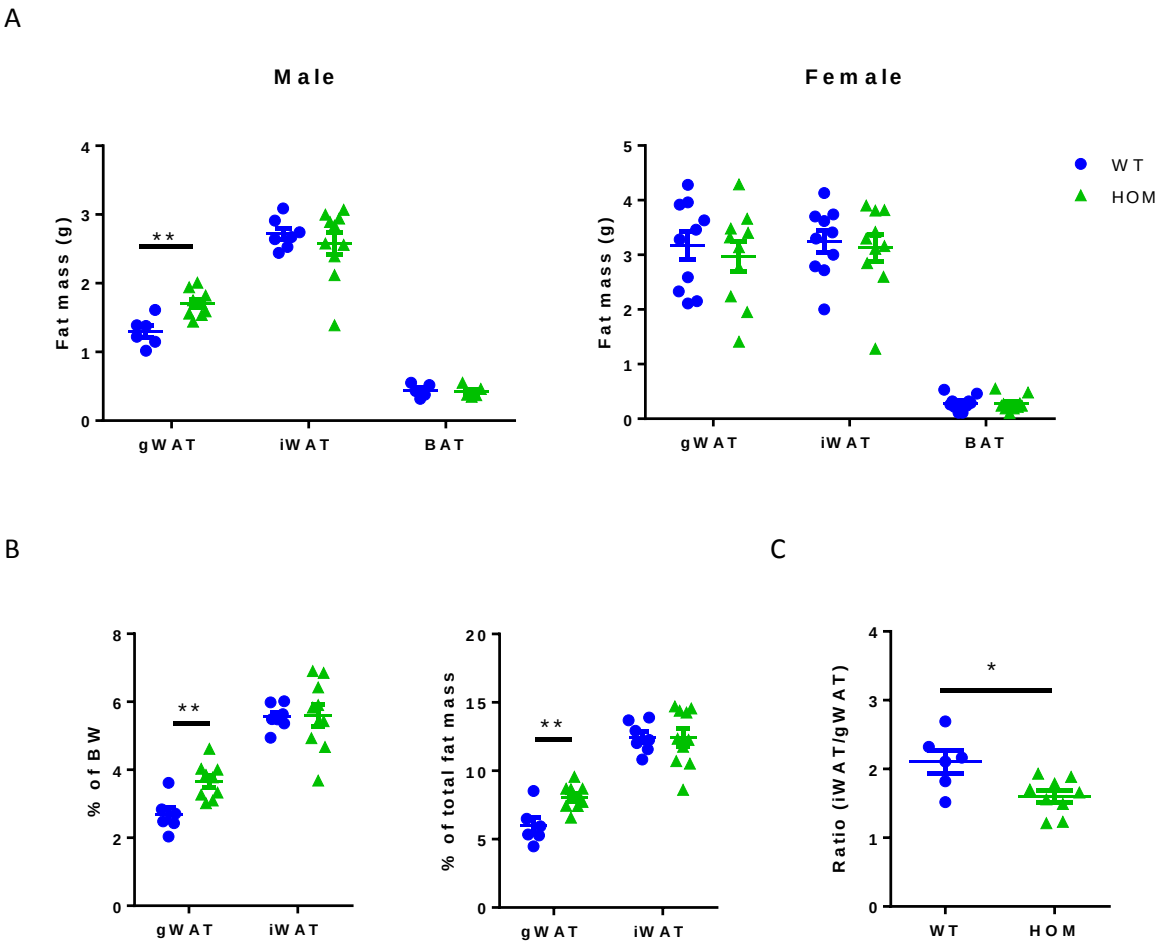


Fig. 4.7 | DEL82 causes differential fat accumulation in male mice on a HFD.
Fat pads were excised from male and female DEL82 HOM and WT mice on HFD at 6 months of age. (A) Total fat mass weight of gWAT, iWAT and BAT for male and female. (B) Male fat pad weights for gWAT and iWAT were further normalised to body weight (BW) and total fat mass. (C) Ratio of male iWAT/gWAT. Statistical significance was determined using multiple Student's t-tests, data is shown as individual data points and mean ± SEM with *p<0.05 **p<0.01.

4.3.6 Histology

To determine whether DEL82 had an effect on dermal WAT thickness, two skin samples, one from the ventral (abdominal) area (adWAT) and one from the dorsal area (behind the lower limbs) (sdWAT) were taken. All samples were excised from male DEL82 HOM and WT mice on HFD at 6 months of age for histological processing. Skin samples were fixed in formaldehyde before a standard H&E protocol was performed for visualisation (see Material and Methods).

The superficial adipose layer was significantly reduced in DEL82 in both anatomical locations assessed, although the difference was more striking in sdWAT (Fig. 4.9). Quantitative analysis revealed that there was a 35% and 32% reduction in adWAT (P=0.046) and sdWAT (P=0.005) thickness, respectively (Fig. 4.9B). As dermal WAT (dWAT) represents a large proportion of overall fat mass in mice (Kasza *et al.*, 2016; Reeder *et al*., 2016), the reduction in dWAT could account for the overall decreased fat mass seen in DEL82 HOM males on a HFD.

Current work aims to investigate whether DEL82 has an effect on adipose tissue morphology (e.g. adipocyte size and/or number) in iWAT and gWAT using histological analysis.

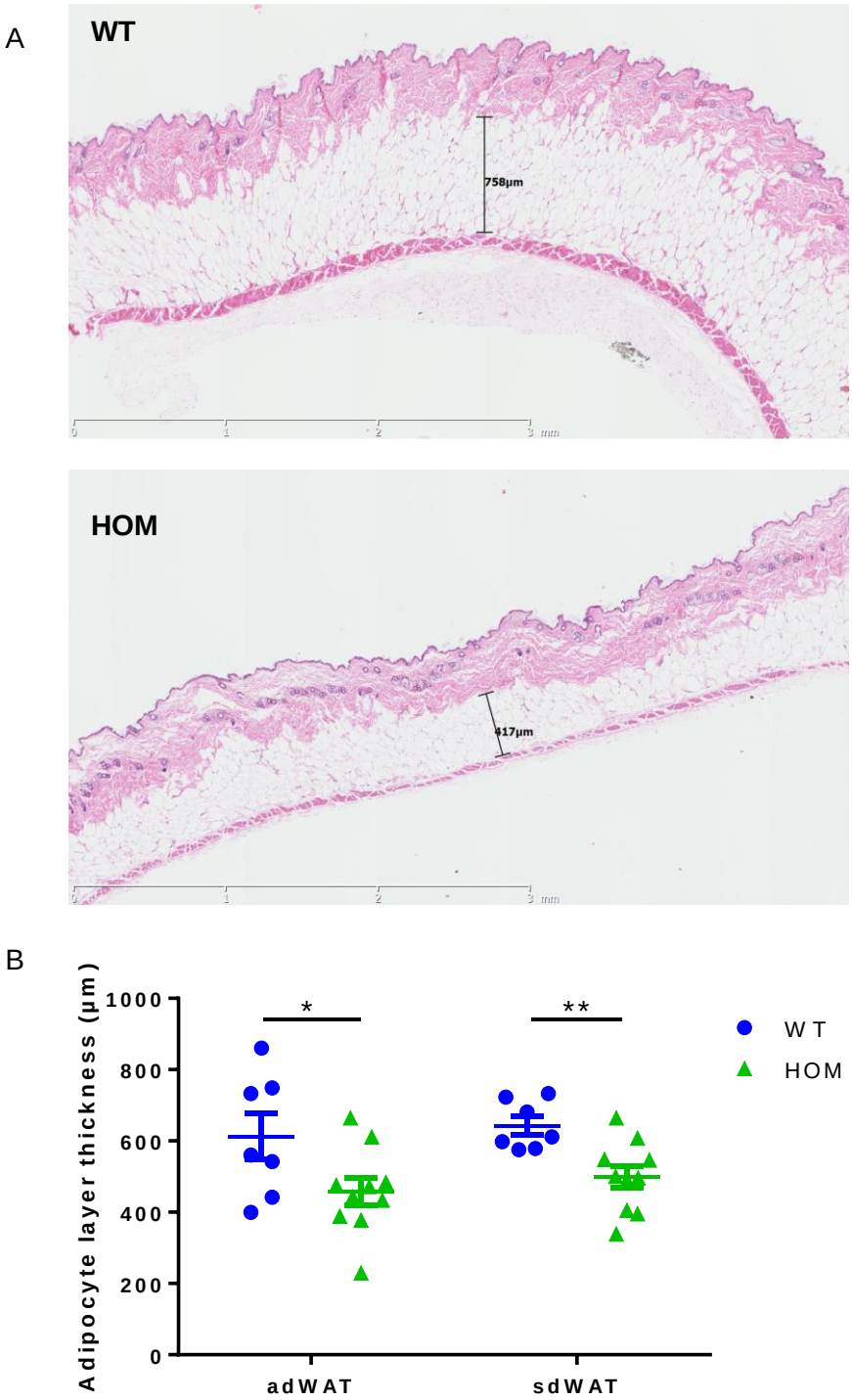


Fig. 4.8 | DEL82 causes decreased size in superficial subcutaneous fat layer in male mice on a HFD.

Skin was excised from male DEL82 HOM and WT mice on HFD at 6 months of age for histological processing. (A) Representative images of the histological appearance of skin samples from WT and DEL82 HOM dermal WAT. (B) Relative adipocyte layer (dWAT) thickness for skin samples obtained from ventral/abdominal area (adWAT) and the dorsal area/back (sdWAT) from WT (n=7) and DEL82 HOM (n=10); dWAT thickness was calculated from histological assay. Statistical significance was determined using multiple Student's t-tests, data is shown as individual data points and mean ± SEM with *P<0.05 **P<0.01.

4.3.7 Clinical Chemistry in Blood Plasma

To determine whether DEL82 males on a HFD have global changes in metabolism, circulating metabolites were analysed in blood plasma. Terminal blood samples were collected at 6 months of age from WT and DEL82 HOM animals in the non-fasted state. Plasma was isolated from whole blood samples and several metabolites measured, namely total cholesterol, LDL (low density lipoprotein), HDL (high density lipoprotein), circulating free fatty acids, triglycerides, glucose and glycerol (Fig. 4.10). In plasma taken from DEL82 HOM males, total cholesterol and LDL levels were significantly reduced (P=0.008 and P=0.034, respectively) compared to WT plasma. There was also a trend for decreased HDL levels in DEL82 HOM, albeit not statistically significant. No statistically significant differences were observed for glucose, glycerol, triglycerides, and free fatty acid concentrations between the genotypes.

In obese individuals, visceral fat area is inversely correlated to LDL, HDL and total cholesterol levels (Hoenig *et al.*, 2011). Consistent with this, DEL82 HOM males had increased visceral fat mass compared to WT littermates, but decreased LDL, HDL and total cholesterol levels, suggesting decreased synthesis and/or increased absorption of cholesterol in DEL82 HOM animals.

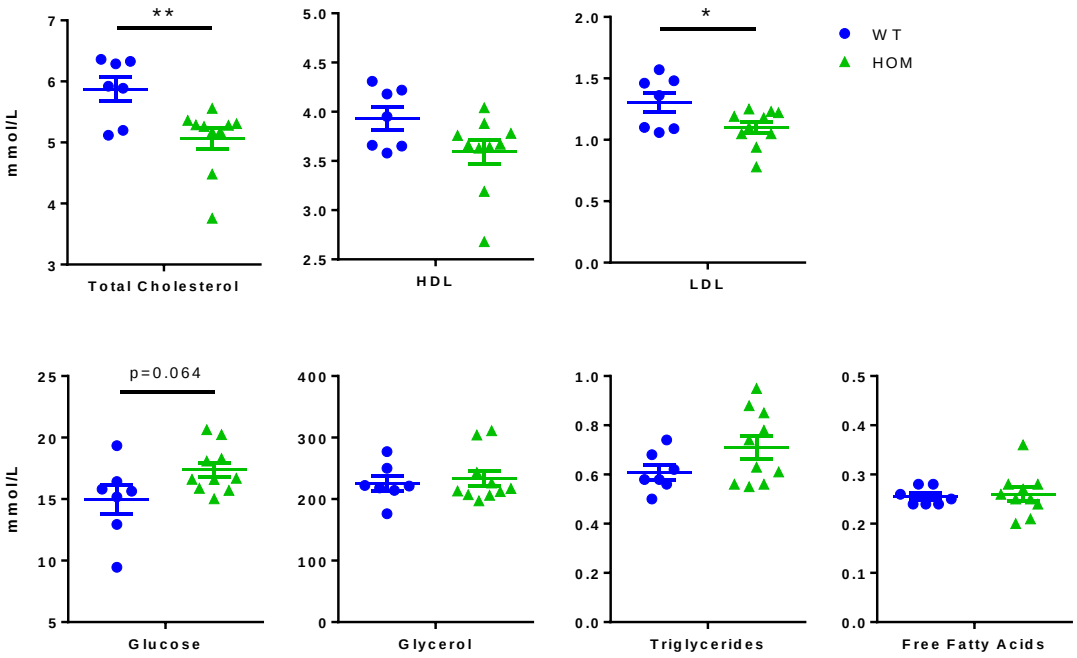


Fig. 4.9 | DEL82 causes altered blood metabolites.
Blood was taken in a non-fasted state. Retro-orbital bleed was performed and plasma was isolated at 6 months of age. Plasma metabolites were measured and data is shown as individual data points and mean ± SEM. Data was analysed using multiple t-tests, corrected for multiple comparisons using the Sidak-Bonferroni method: * P<0.05, ** P<0.01. HDL = high-density lipoprotein; LDL = low-density lipoprotein.

4.5 Discussion

Data presented in this chapter elucidate the phenotypic consequences of DEL82 on the organismal level. My results show that DEL82 causes differential fat accumulation and ~4% reduced overall body weight in DEL82 which is due to ~5% decreased fat mass, with unaltered lean mass.

In many ways, the phenotype resembles key characteristics of human GWAS. For example, the effect of DEL82 on whole body weight is subtle, yet significant. Furthermore, the subtle resistance to accumulate as much weight as WT littermates in DEL82 HOM males is only apparent when challenged with a HFD. Indeed, the effects of common variants identified through GWAS are characteristically dependent on an environmental trigger, and in the case of obesity, such a trigger is an obesogenic environment.

In human, variants in *FTO* were first identified to be associated with T2D (Frayling *et al.*, 2007), but when adjusted for BMI, this association disappeared, suggesting that the effect of variants in *FTO* on T2D risk were mediated via their association with overall obesity. Although mice on a HFD had striking glucose intolerance when compared to animals that received a LFD, there was no effect of genotype on diabetic phenotype as measured by the means of an IPGTT. This is likely explained by the subtle difference of fat mass and body weight between WTs and HOMs that is not sufficient of having an effect on glycaemic traits.

Unexpectedly, there was no difference in iWAT mass between genotypes, making it clear that the decrease in fat mass in DEL82 HOM males on a HFD is not explained by a decrease in iWAT depot size. And although overall body fat mass was reduced in DEL82 homozygous animals, gonadal adipose tissue mass was significantly increased. Therefore, neither of those tissues can explain overall reduced fat mass in DEL82 HOM males observed on a HFD.

Importantly, gWAT and iWAT mass make up only a small percentage of overall fat mass (6% and 13%, respectively). Additionally, there is evidence for a non-equal proportion of visceral

to subcutaneous fat mass in mice in the literature. Whole-body micro-CT imaging studies have shown that under baseline conditions, C57BL/6J mice have approximately 2-fold higher subcutaneous compared to visceral depot volume (Assini *et al.*, 2015). In contrast, MRI-derived measurements of WAT volumes revealed that subcutaneous WAT volume was approximately equal to visceral WAT volume (Kasza *et al.*, 2016). We also know that a HFD increases the volume of both adipose depots, but has a larger effect on visceral than subcutaneous fat mass expansion in male mice (Assini *et al.*, 2015; Wang *et al.*, 2013). Therefore, the net effect of a mutation that potentially causes less subcutaneous WAT expansion and increased visceral WAT expansion in mice could likely result in overall decreased fat mass for a period of time. With a prolonged HFD exposure, however, the difference in depot expansion would presumably disappear. More measurements aiming to quantify different adipose depots (such as MRI measurements) would be necessary to answer this question and to accurately determine the spatio-temporal pattern of HFD-induced WAT expansion in DEL82 animals.

In recent years, evidence has accumulated that suggests that adipocytes other than in major depots are important in physiology and pathophysiology (Hepler, Vishvanath, & Gupta, 2017; Shook *et al.*, 2016). Dermal WAT (dWAT) is a relatively unexplored adipose depot and recent evidence suggests that dWAT thickness increases in response to cold exposure and may regulate thermoregulation (Kasza *et al.*, 2014). I found that the superficial adipose layer in skin is significantly reduced in DEL82 HOM males following a HFD. Dermal adipose tissue makes up a large proportion of the overall fat mass in mice as measurement of fat tissue volumes using MRI showed that dWAT volume accounts for approximately 50% of total WAT volume in mice (Kasza *et al.*, 2016). Therefore, the reduction seen in dWAT could account for the overall decreased fat mass in DEL82 HOM males on a HFD. Even more strikingly, Kasza *et al.* (2016) also found that dWAT thickness mirrors the magnitude of weight gain in male mice on a HFD, suggesting the kinetics of WAT accumulation in dWAT parallels overall body weight increase (Kasza *et al.*, 2016).

The variability in dWAT thickness observed within a genotype is somewhat expected, as it is known that thickness of dWAT varies with a number of factors, such as hair growth and body site (back/dorsal or belly/ventral skin). In mice, skin patches undergo folliculogenesis approximately one week out of three, which results in at least a 2-fold increased thickness of dWAT (Kasza *et al.*, 2016). In future studies, it would be interesting to investigate which target genes might be responsible for the effect of DEL82 on dWAT thickness.

Ongoing experiments include analysis of energy expenditure and food intake using the CLAMS as well as WAT histology with the aim to determine whether DEL82 results in hypertrophic and/or hyperplastic WAT growth.

In summary, deletion of the mouse orthologous region around human rs1421085 results in a metabolic phenotype in mice that includes reduced body weight and fat mass gain, reduced dWAT thickness as well as altered blood metabolites. The described effect is subtle, specific to males, and only apparent in mice that have been challenged with a diet high in fat. All these descriptions are consistent with what we know about the features of most common variants from human GWAS and ultimately what we understand about the context-specificity of gene regulatory networks.

5 DEL82 mechanism

5.1 Introduction

In Chapter 4, I described the effect of DEL82 on the organismal-level and revealed that DEL82 results in an adipose-related phenotype in male mice on a HFD. In this chapter, I will assess the regulatory impact of DEL82 and characterise its cellular consequences in male adipose tissue and primary pre-adipocytes from both gWAT and iWAT.

By examining large domains of long-range 3D chromatin interactions using Hi-C technology, it is possible to predict putative target genes of regulatory elements as previously described (Dixon *et al.*, 2012). From Hi-C data and resulting maps of TADs, it was previously shown that *FTO* intronic elements are within TAD boundaries that encompass 8 putative candidate genes (Claussnitzer *et al.*, 2015), including *RPGRIP1L, FTO, IRX3, IRX5,* and *IRX6*.

Previous studies have predicted a number of candidate causal genes and tissues based on expression QTLs, including *IRX3* in pancreas (Ragvin *et al.*, 2010), brain (Smemo *et al.*, 2014), and adipocytes (Claussnitzer *et al.*, 2015), *IRX5* in adipocytes (Claussnitzer *et al.*, 2015), *RBL2* in lymphocytes (Jowett *et al.*, 2010) and *RPGRIP1L* and *FTO* in iPSC-derived neurons (Stratigopoulos *et al.*, 2016).

Given that the development of obesity is primarily attributed to altered processes in the hypothalamus and adipose tissue (Berry *et al.*, 2014) and given the evidence in humans for *FTO* risk alleles to alter appetite (Karra *et al.,* 2013), I focused identification of effector transcripts in DEL82 mice on brain and adipose. The objectives of this chapter are:

1. To identify the gene(s) and cell type(s)/tissue(s) responsible for the phenotype described in Chapter 4
2. To understand the mechanism by which DEL82 affects adipose tissue function and/or development
3. To study the mechanism of the depot-specific effect of DEL82 described in Chapter 4

5.2 Results

5.2.1 DEL82 is within a regulatory element for *Irx3* and *Irx5*

First I set out to identify the relevant cell type and effector transcript that could explain the decreased body weight and fat mass phenotype observed in DEL82 HOM males described in Chapter 4. To identify the targets of the regulatory element deleted in DEL82, gene expression of 5 putative target genes (*Rpgrip1l, Fto, Irx3, Irx5*, and *Irx6*) within the 2Mb *Fto* locus were evaluated by means of qPCR analysis. Obesogenic WAT expansion is mainly regulated by hypothalamic signals as well as WAT-intrinsic mechanisms (Berry *et al.*, 2014). Therefore, I chose to assess target gene expression in iWAT- and gWAT-derived pre-adipocytes as well as hypothalamus. Additionally, these are the cell types and tissues that have previously been implicated in *FTO* risk variant dissection (Claussnitzer *et al.*, 2015; Smemo *et al.*, 2014; Stratigopoulos *et al.*, 2016). Tissues were isolated from male and female WT and DEL82 HOM animals at 6-8 weeks of age. Mature WAT was digested using collagenase treatment and pre-adipocytes isolated by centrifugation. Adipocyte precursors were then cultured and RNA isolated before (D0) and two days after treatment with adipogenic cocktail (D2).

I found that DEL82 caused significantly decreased *Irx3* and *Irx5* expression in male adipocyte progenitors isolated from iWAT (Fig. 5.1, 5.2). This decreased expression of *Irx3* and *Irx5* in male iWAT-derived pre-adipocytes was maintained during early differentiation (Fig. 5.1), as adipocytes that have been treated with adipogenic cocktail for 2 days showed comparable alteration of target genes (Fig. 5.1).

In males, iWAT-derived pre-adipocytes homozygous for DEL82 had on average 0.65-fold (P=0.0014) and 0.71-fold (P=0.0114) reduction in *Irx3* and *Irx5*, respectively, whereas iWAT-derived pre-adipocytes heterozygous for DEL82 had 0.81-fold (P=0.0348) and 0.79-fold (P=0.021) reduction in *Irx3* and *Irx5* mRNA expression, respectively (Fig. 5.4). These data suggest a dose-dependent effect of DEL82 on target gene expression in male adipocytes.

As expected, no effect of heterozygous DEL82 on target gene expression was observed in female pre-adipocytes (Fig. 5.4).

In gWAT-derived pre-adipocytes of DEL82 HOM males, *Irx3* expression was ~1.4-fold and ~1.5-fold upregulated at D0 and D2, respectively, without having a detectable effect on *Irx5* expression (Fig. 5.3). No gene expression changes were observed in female iWAT- or gWAT-derived pre-adipocytes at any time-point (Fig. 5.3).

DEL82 had no detectable effect on *Irx3* and *Irx5* expression in the hypothalamus of male or female animals (Fig. 5.1), suggesting a cell-type specific effect of DEL82 on gene expression. *Fto, Rpgrip1l* and *Irx6* expression were unaffected in any cell type examined (Fig. 5.1, 5.2, 5.3). Taken together, out of the 6 potential target genes measured in 3 relevant tissue types, only *Irx3* and *Irx5* were altered in pre-adipocytes.

Interestingly, expression analysis of mature gWAT and iWAT depots from male DEL82 HOM and control mice revealed that *Irx3* and *Irx5* were not altered in mature WAT (Fig. 5.3), indicating that target gene expression changes and regulatory activity are restricted to early differentiating pre-adipocytes. This is consistent with human data, where *FTO* risk allele carriers have altered *IRX3* and *IRX5* expression specifically in pre-adipocytes (Claussnitzer *et al.*, 2015).

Further, the difference in direction of target gene alteration between gWAT and iWAT suggests that the regulatory element in *Fto* intron 1 might have both enhancing and repressing activity depending on adipose depot. For future studies, it would be interesting investigate target gene expression in pre-adipocytes from additional adipose depots, such as mesenteric and pericardial adipose depots. Additionally, whether adipose tissue-derived stem cells isolated from older animals and animals that have been challenged with a HFD show similar alterations in target gene expression is unaddressed.

Whether any of the identified effector transcripts are altered on the protein level has not been tested in this thesis.

The non-detectable changes of target gene expression in DEL82 HOM females are consistent with the absence of an *in vivo* phenotype, suggesting a regulatory effect of DEL82 that is either specific to, or much stronger in males. This is very surprising, given there is currently no evidence published for a sex-specific effect of *FTO* risk alleles in humans. More experiments are necessary to explain the sex-specific effect of DEL82 in mice.

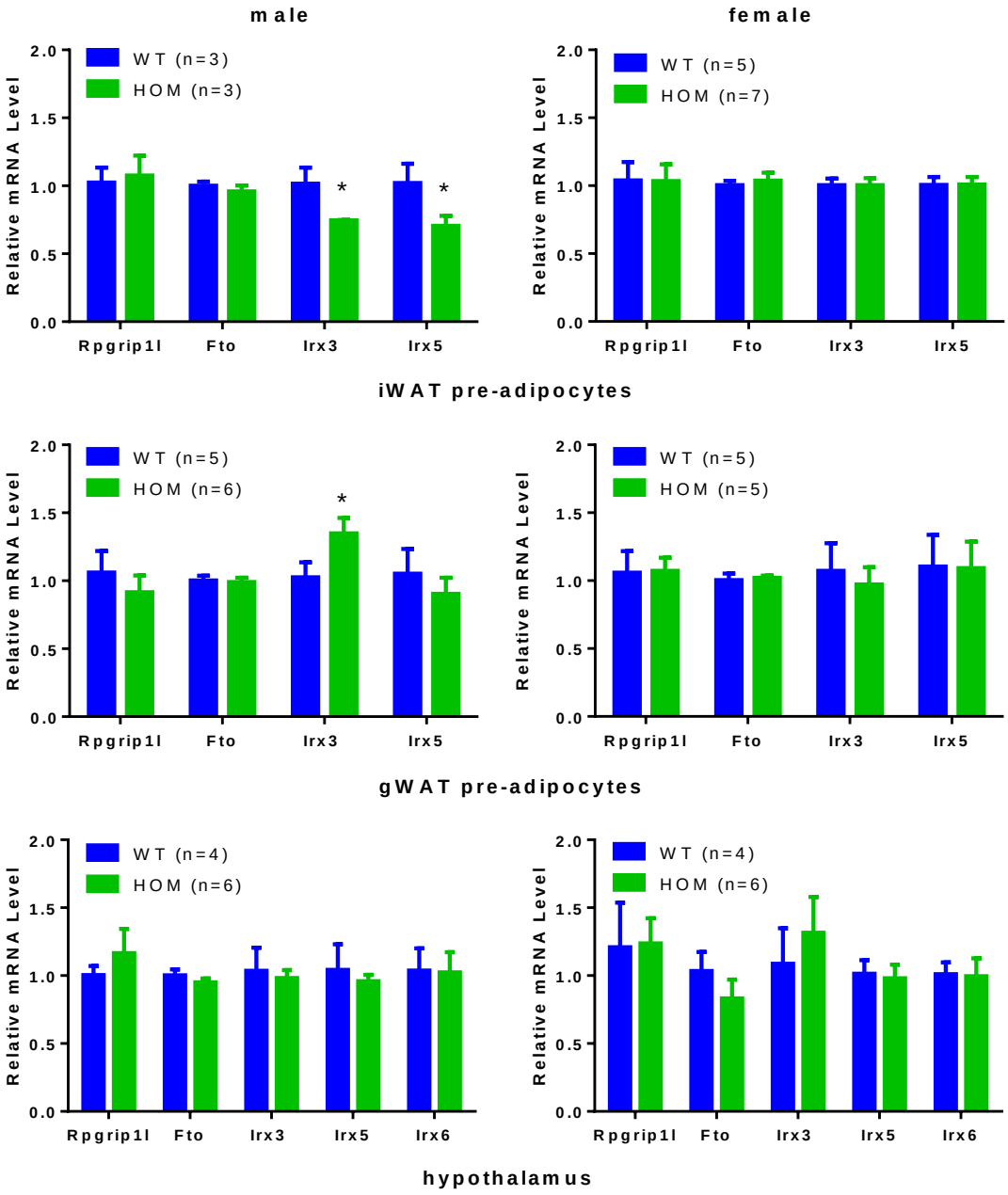


Fig. 5.1 | DEL82 regulates *Irx3* and *Irx5* in male adipocyte progenitors.
Pre-adipocytes from iWAT (top) and gWAT (middle) were isolated and hypothalamus (bottom) was excised from male and female DEL82 WT and HOM mice at 6-8 weeks of age. Gene expression of candidate target genes *Rpgrip1l, Fto, Irx3, Irx5* and *Irx6* was measured by the means of qPCR (normalised to *Canx*). Irx6 expression was too low to be detected in adipocytes. N numbers are indicated in the figure. Statistical significance was assessed using Student's t-test; data are expressed as mean ± SEM *P<0.05.

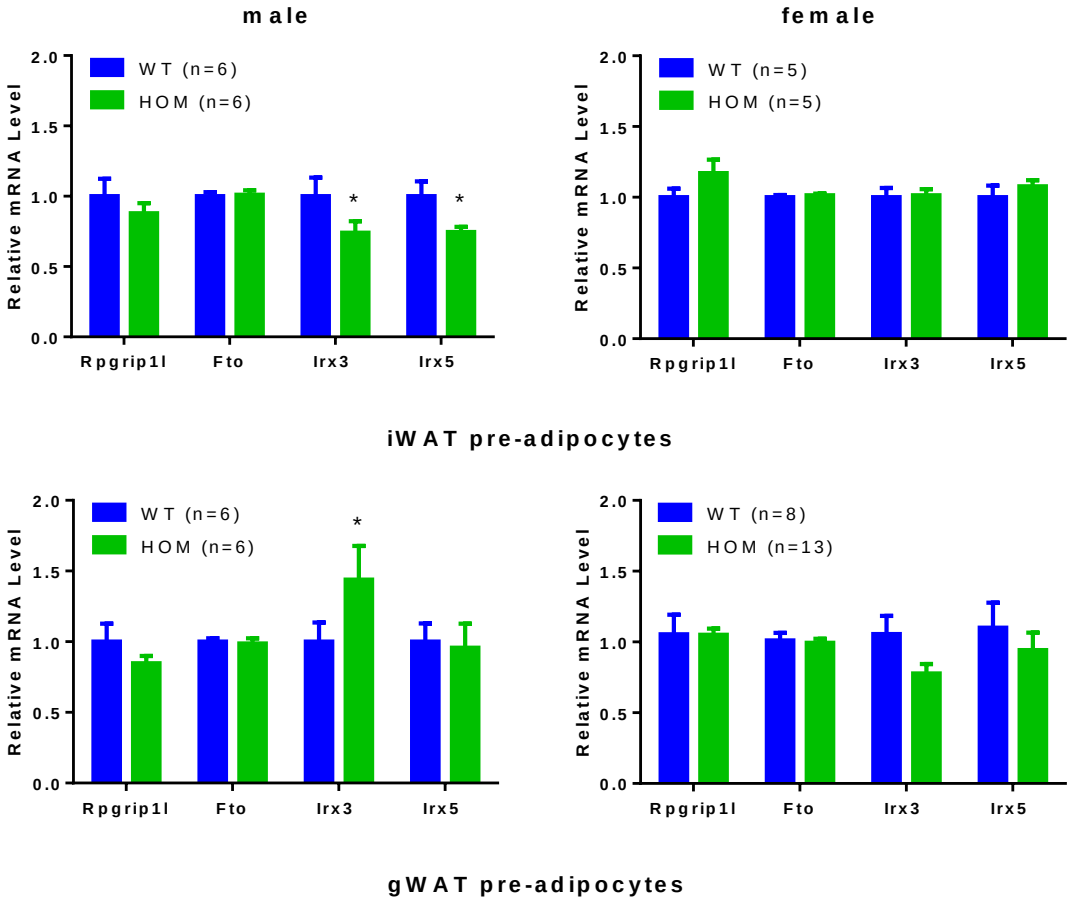


Fig. 5.2 | DEL82 regulation of *Irx3* and *Irx5* in male adipocyte progenitors is maintained in early differentiation.
Pre-adipocytes from iWAT (top) and gWAT (bottom) were isolated from male and female DEL82 WT and HOM mice at 6-8 weeks of age. Pre-adipocytes were cultured and differentiated for 2 days. Gene expression of candidate target genes *Rpgrip1l*, *Fto*, *Irx3* and *Irx5* was measured by the means of qPCR (normalised to *Canx*). N numbers are indicated in the figure. Statistical significance was assessed using Student's t-test; data are expressed as mean ± SEM *P<0.05.

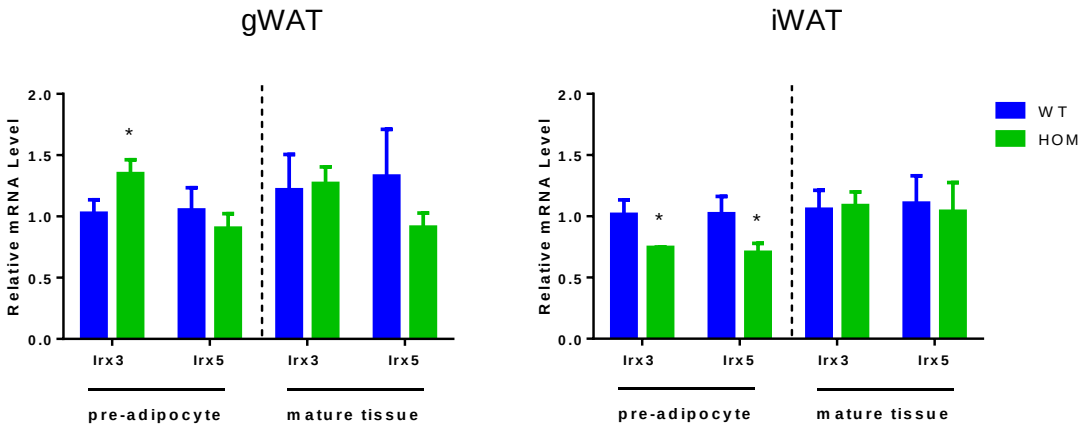


Fig. 5.3 | DEL82 regulatory disruption does not alter *Irx3* and *Irx5* in mature adipose tissue.
Graph shows mRNA expression for *Irx3* and *Irx5* in adipocyte progenitors and mature tissue isolated from gWAT (left) and iWAT (right) of male DEL82 WT (n=5-7) and HOM (n=6-9) animals. Pre-adipocytes were isolated from 6-8 week old mice, and mature tissues were excised from 6 months old mice. Gene expression was measured using qPCR (normalized to *Canx*). Data for pre-adipocytes has previously been used in Fig 5.1. Statistical significance was assessed using Student's t-test; data are expressed as mean ± SEM *P<0.05.

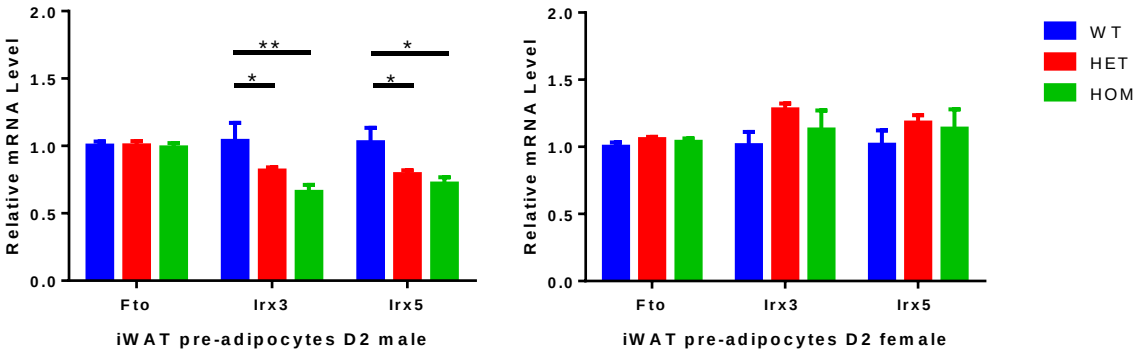
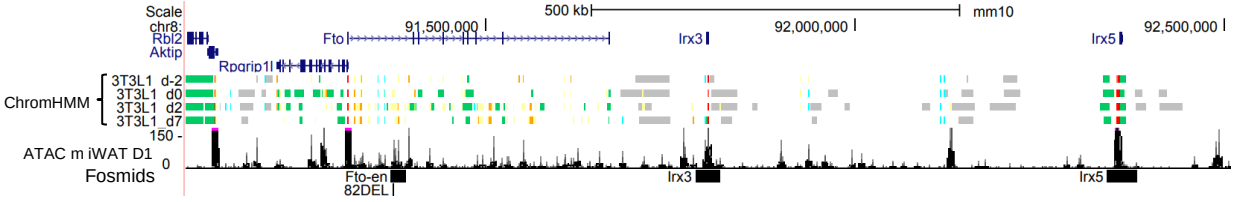


Fig. 5.4 | DEL82 regulatory disruption of *Irx3* and *Irx5* is dose-dependent.
Graph shows mRNA expression for *Fto, Irx3* and *Irx5* in adipocyte progenitors isolated from iWAT of male DEL82 WT (n=6), HET (n=8) and HOM (n=4) and female DEL82 WT (n=4), HET (n=9) and HOM (n=5) animals. Pre-adipocytes were isolated from 6-8 week old mice. Gene expression was measured using qPCR (normalized to *Canx*). Data for WT and HOM pre-adipocytes has previously been used in Fig 5.2. Statistical significance was assessed using two-way ANOVA with Bonferroni's multiple comparisons test; data are expressed as mean ± SEM *P<0.05, **P<0.01.

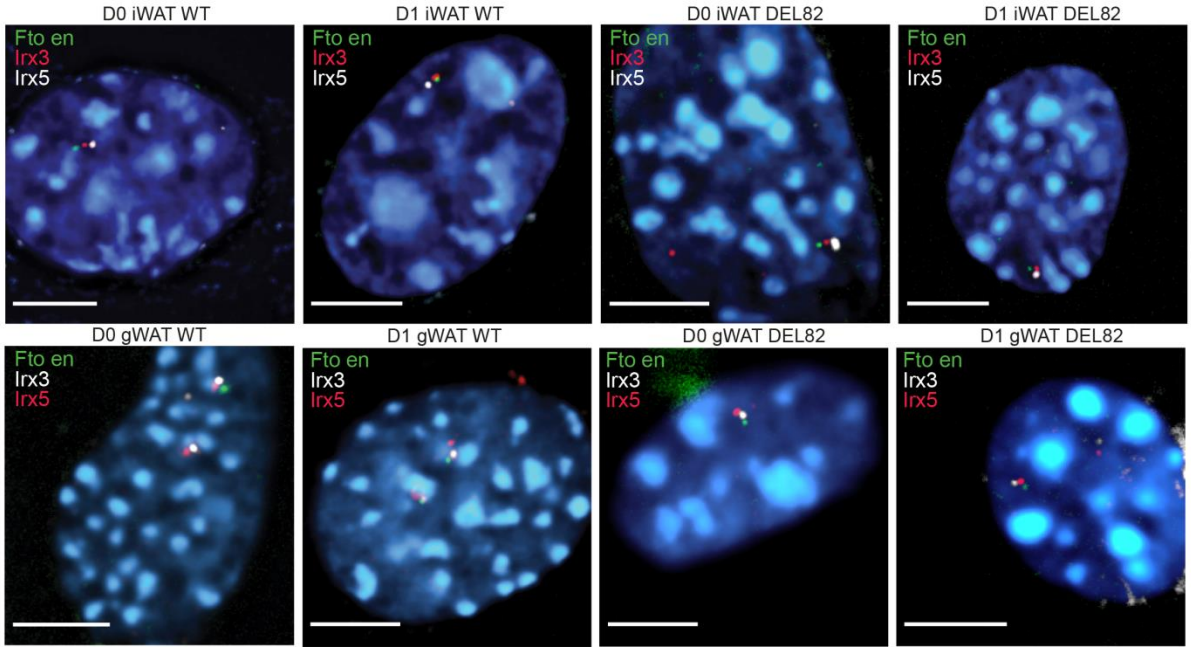
Chromatin looping is a popular model describing how very long-range regulatory elements are in contact and can communicate with their target gene promoter (Benabdallah & Bickmore, 2015); however, the link between loop formation and gene activation is currently unclear. There is evidence for both, stable enhancer-target gene contacts that are present even when the target gene is not active (Ghavi-Helm *et al.*, 2014; Montavon *et al.*, 2011), as well as spatially and temporally restricted enhancer-gene looping that only occurs when a target gene is activated. To assess the proximity between the *Fto* enhancer element (Fto-en), *Irx3* and *Irx5*, we performed 3D Fluorescence in situ hybridisation (FISH). Following fixation and permeabilisation of pre-adipocytes, slides were sent to Iain Williamson (Wendy Bickmore group) at The MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, where hybridization and imaging was performed.

FISH is a technique that is based on the hybridisation of fluorescently labelled nucleic acid probes to complementary sequences in genomic DNA. This allows visualization of the proximity of individual genomic regions in the nucleus. Here, we use 3D-FISH and conventional wide-field deconvolution microscopy, with the objective of determining whether DEL82 results in an altered DNA interaction landscape in pre-adipocytes from gWAT and iWAT that could explain DEL82-induced *Irx3* and *Irx5* mRNA expression changes. Therefore, FISH was performed in iWAT- and gWAT-derived primary pre-adipocytes of male WT (n=3) and DEL82 HOM (n=3). Based on my previous finding that i) early differentiation triggers chromatin to remodel and to become accessible at *Fto* enhancer elements and ii) that target gene expression at D2 of differentiating iWAT pre-adipocytes was statistically more significant than at D0, I decided to prepare cells at two time-points: undifferentiated (D0) and 24h post adipogenic induction (D1). Fosmids were designed to hybridise to Fto-en (enhancer element that overlaps rs1421085), *Irx3* and *Irx5* (Fig. 5.5A).

A


Scale
chr8:
500 kb
mm10
91,500,000
92,000,000
92,500,000
Rbl2
Aktip
Fto
Irx3
Irx5
Rpgrip1l
ChromHMM
3T3L1 d-2
3T3L1 d0
3T3L1 d2
3T3L1 d7
150
ATAC m iWAT D1
0
Fosmids
Fto-en
82DEL
Irx3
Irx5

B


D0 iWAT WT
D1 iWAT WT
D0 iWAT DEL82
D1 iWAT DEL82
D0 gWAT WT
D1 gWAT WT
D0 gWAT DEL82
D1 gWAT DEL82
Fto en
Irx3
Irx5

C

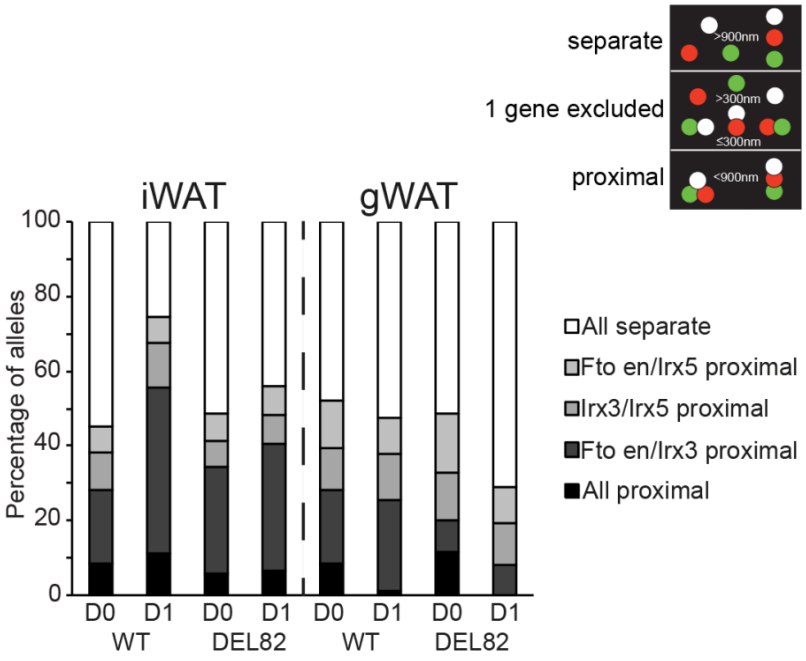

separate
1 gene excluded
proximal
>900nm
>300nm
≤300nm
<900nm
iWAT
gWAT
Percentage of alleles
100
80
60
40
20
0
D0
D1
D0
D1
D0
D1
D0
D1
WT
DEL82
WT
DEL82
All separate
Fto en/Irx5 proximal
Irx3/Irx5 proximal
Fto en/Irx3 proximal
All proximal

Fig. 5.5 | FISH in iWAT- and gWAT-derived pre-adipocytes for Fto-en*, Irx3* and *Irx5*.
Primary pre-adipocytes from WT (n=3) and DEL82 HOM (n=3) animals were isolated, cultured and either fixed undifferentiated (D0) or one day post adipogenic stimulation (D1). Pre-adipocyte nuclei were then incubated with fluorescent probes for Fto-en, Irx3 and Irx5. Probe distances were determined for 62-90 nuclei per experimental group. More details about the experimental procedures and the probe location can be found in the Materials and Methods. (A) Location of genes over a 1.6 Mb mouse genomic locus containing *Irx3 and Irx5*, with the location of adipocyte-specific regulatory elements shown below. Tracks 5-8 show chromatin annotations generated using ChromHMM (Ernst & Kellis, 2012) from histone modifications generated in 3T3-L1 pre-adipocytes (Mikkelsen *et al.*, 2010). Yellow indicates active enhancers. Track 9 shows open chromatin map (ATAC-seq) in iWAT pre-adipocytes at day 1 of adipogenic induction. The bottom track shows the positions to which the fosmids used for FISH hybridise as well as the location of the 82DEL mutation. (B) Representative images showing distance between probes for Fto-en, Irx3 and Irx5 in adipocyte nuclei. Scale bars: 5µm. (C) Percentages of interprobe distances. In situ hybridisation and measurements of distances between probes was performed by Iain Williamson.

5.2.1.1 FISH in iWAT

At both temporal stages (D0 and D1), co-localisation frequencies for the fosmid probes that hybridise to the combination Fto en/Irx3 were higher than for Fto-en/Irx5 or Irx3/Irx5 (Fig. 5.5B), suggesting that in the nucleus of iWAT pre-adipocytes, the genomic region between Fto-en and Irx3 is folded into a compact chromatin domain, which is at its most compact at D1 of adipogenic differentiation. The spatial distances between Fto-en and Irx3 were significantly shorter in D1 WT cells compared to D0 WT cells (P=0.001, Fisher's exact test) (Fig. 5.6), indicating that adipogeneic induction is accompanied by co-localisation of Fto-en and Irx3 in iWAT pre-adipocytes. The proportion of co-localised (<200 nm apart) probes for the *Fto* enhancer and *Irx3* was nearly 30% in D1 WT pre-adipocytes (Fig. 5.6), consistent with active gene-enhancer co-localisation during long-range regulation. This data provide evidence for the *Irx3* promoter to form functional connections with the *Fto* regulatory element in early differentiating adipocyte precursors and that sequence variants in *Fto* intron 1 have the potential to affect *Irx3* regulation.

In DEL82 HOM cells I observed significantly less co-localisation of Fto-en and Irx3 probes compared to WT cells at D1 of differentiation (P<0.05) (Fig. 5.7), suggesting that DEL82 causes a failure of *Irx3* to interact with Fto-en upon induction of adipogenesis. This is consistent with reduced levels of *Irx3* mRNA expression in DEL82 HOM pre-adipocytes compared to WT.

The distances between the other probe pairs, Irx3/Irx5 and Fto-en/Irx5 were not significantly different across the time-points (Fig. 5.7), albeit there was a trend for the distance between *Irx3* and *Irx5* to be reduced in DEL82 HOM at D1 compared to WT (P=0.057; Mann Whitney U test) suggesting that DEL82 disrupts interactions of these genes. This leads to a proposed model of a 3D regulatory landscape at the *Fto* locus in which Fto-en is in contact with *Irx3*, which in turn might be in contact with *Irx5*, and DEL82 disrupts these interactions.

In summary, there is a significantly increased co-localisation of Fto-en with Irx3 but not Irx5 in D1 WT cells in iWAT. These data are consistent with decreased *Irx3* expression in DEL82 HOM iWAT primary adipocytes. However, this data cannot explain the reduced *Irx5* expression observed in DEL82 HOM iWAT pre-adipocytes, suggesting that (given the FISH experiment had enough power) the effect of DEL82 on *Irx5* expression might not primarily be driven by direct binding events between Fto-en and *Irx5*, but potentially by deceased interaction between *Irx3* and *Irx5* or perhaps a secondary, downstream *trans*-regulatory effect.

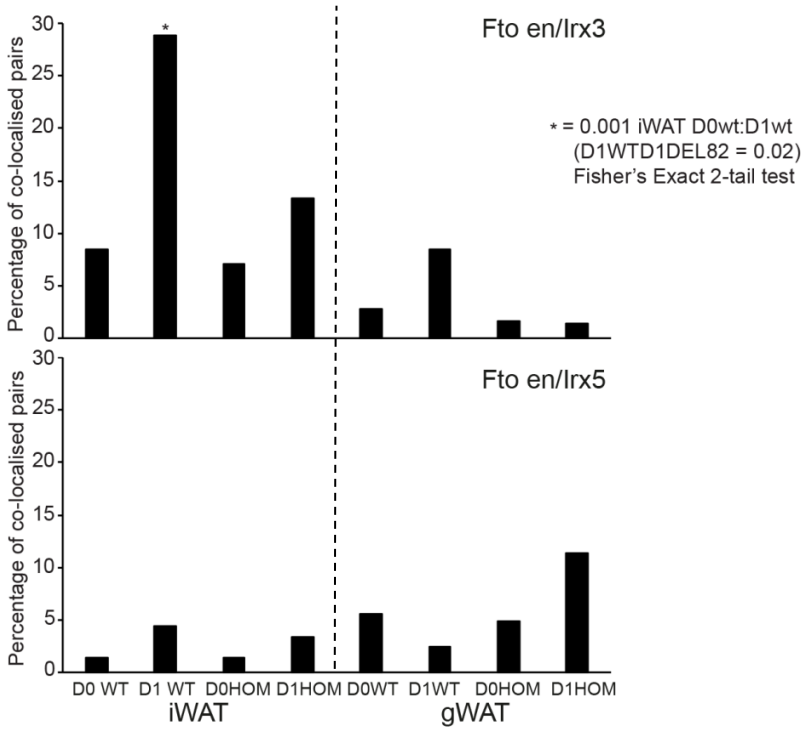


Fig. 5.6 | Proximity between probes for Fto-en/Irx3 and Fto-en/Irx5 in DEL82 and WT pre-adipocytes.
Pre-adipocytes were isolated from gWAT and iWAT and isolated before (D0) and 1 day after adipogenic induction (D1). Statistical significance between data sets was determined using Fisher's Exact 2-tail tests, *P<0.05. In situ hybridisation and measurements of distances between probes was performed by Iain Williamson.

5.2.1.2 FISH in gWAT

In gWAT-derived pre-adipocytes, Fto-en and Irx3 probes were noticeably further apart than in iWAT samples (Fig. 5.6; Fig. 5.7), suggesting that the proximity between these two DNA elements is greater in iWAT than gWAT pre-adipocytes. Additionally, the spatial distances between Fto-en/Irx3, Irx3/Irx5 and Fto-en/Irx5 were greater in DEL82 HOM pre-adipocytes compared to WT (Fig. 5.7), suggesting that DEL82 disrupts interactions between Fto-en and its target genes *Irx3* and *Irx5*. These differences were significant for all probe combinations (P=0.0017 Fto-en/Irx3; P=0.043 Irx3/Irx5; P=0.0065 Fto-en/Irx5) and, intriguingly, were specific to D0 of adipogenesis as differences in co-localisation frequencies between WT and DEL82 were no longer detected at the later stage (D1) (Fig. 5.7). These data indicate a dynamic chromatin conformation at the Fto/IrxB regulatory circuitry in gWAT.

From the spatial relationships of the three genomic loci across the Fto-IrxB region, we can infer a 3D conformation of the *Fto* regulatory domain in which Fto-en is in contact with both Irx3 and Irx5 in gWAT-derived pre-adipocytes. Given the stronger effect of DEL82 disruption between Fto-en/Irx3 and Fto/Irx5 than Irx3/Irx5, I suggest a model in which Fto-en acts as an anchor point between Irx3 and Irx5.

Together this data show that DEL82 disturbs promoter-regulator interactions at D1 and D0 in iWAT and gWAT pre-adipocytes, respectively. *Irx3* has the strongest association to the *Fto* enhancer element, whereas *Irx5* is further distant to Fto-en.

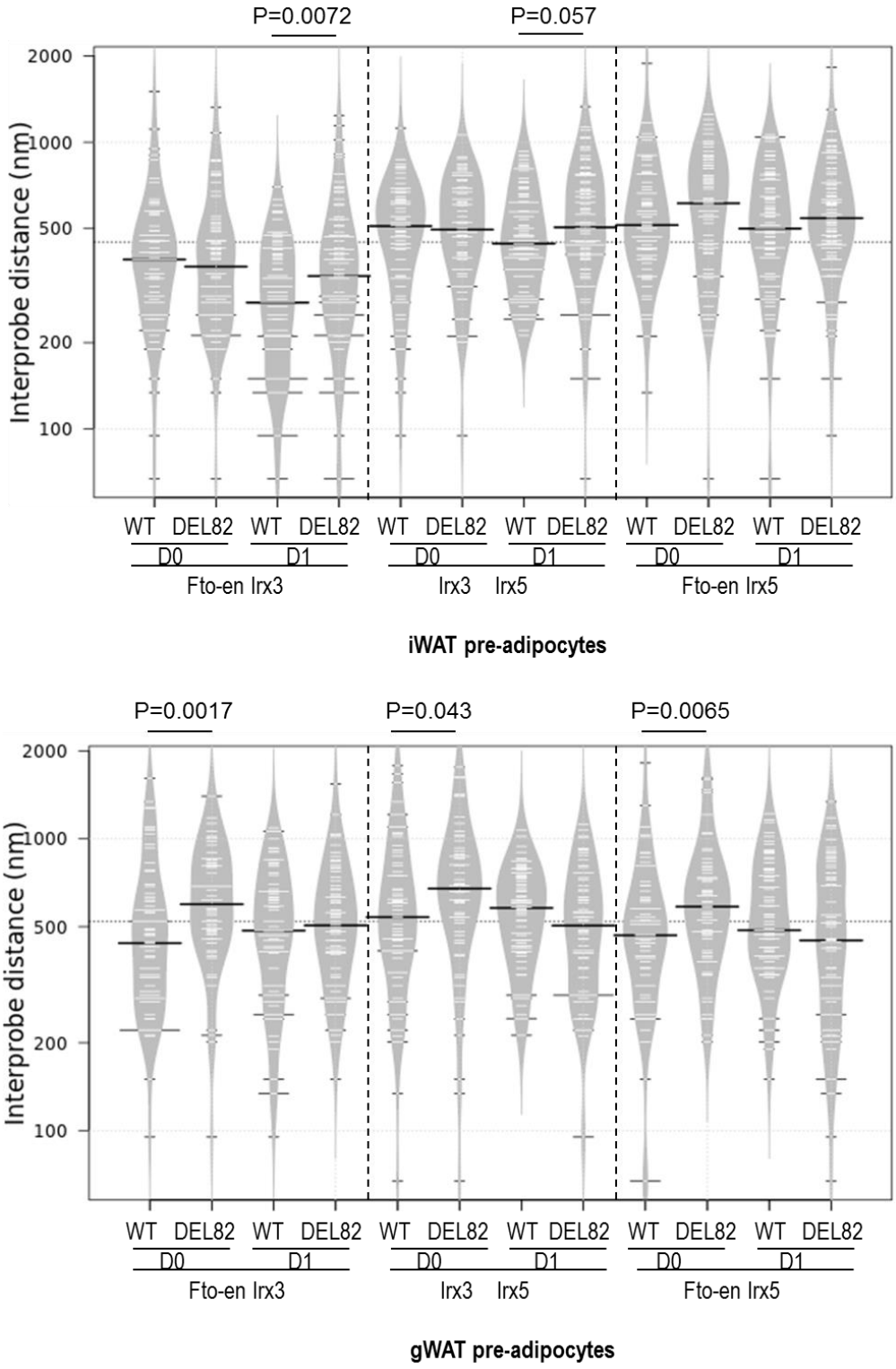


Fig. 5.7 | DEL82 interrupts co-localisation of interporbes for *Irx3* and a putative regulatory element in *Fto*.
Data represents quantitative analysis of interprobe distances between different probe combinations (Fto-en/Irx3; Irx3/Irx5; Fto-en/Irx5) in iWAT pre-adipocytes (top) and gWAT pre-adipocytes (bottom) before (D0) and 24h post induction of adipogenesis (D1). Box plots were generated using the online plotting tool BoxPlotR (http://boxplot.bio.ed.ac.uk/). Black lines show the medians; white lines represent individual data points; polygons represent the estimated density of the data. The statistical significance between data sets was examined Mann Whitney U tests. In situ hybridisation and measurements of distances between probes was performed by Iain Williamson.

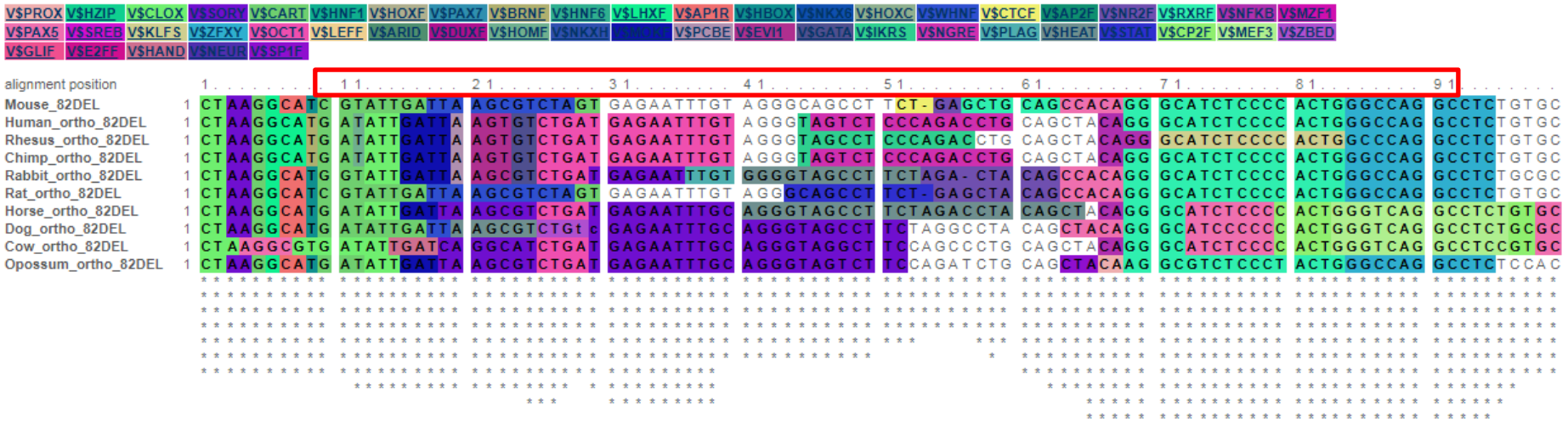
5.2.2 DEL82 sequence contains 2 highly conserved TF modules

The nature of the DEL82 mutation does not exclude the possibility that target gene manipulation in male pre-adipocytes could be due to the absence of motifs downstream of the mouse orthologous region around rs1421085. Additionally, deleting 82 nucleotides potentially created a new motif at this region that could bind TFs.

Based on motif module analysis, the 82 nucleotides deleted in DEL82 are predicted to harbour a large number of motifs for several families of TFs in different species (Fig. 5.8A), suggesting that this sequence acts as an active regulatory region. Conserved TFBS patterns in orthologous regions were assessed across 10 species and revealed that DEL82 harbours two regions of cross-species functionally conserved motif modules (a module is defined as a set of binding site motifs, whose order and distance range is conserved across species) (Fig. 5.8B). The family of TFs predicted to bind to these regions can be found in Fig. 5.8 and include HNF1, HOXF, PAX7, BRNF, HNF6, LHXF, NKX6, NFKB, MZF1, SREB, KLFS, RXRF.

The amount of predicted TFs at these regions is not surprising given that TFs frequently cluster at specific sites, termed TF hotspots (Rasmus Siersbæk *et al.*, 2014). Future experimental dissection is necessary to validate whether the effect of DEL82 on target gene expression is due to the creation or loss of TFBSs, and further, which TFs are directly implicated in this effect.

A


V$PROX V$HZIP V$CLOX V$SORY V$CART V$HNF1 V$HOXF V$PAX7 V$BRNF V$HNF6 V$LHXF V$AP1R V$HBOX V$NKX6 V$HOXC V$WHNF V$CTCF V$AP2F V$NR2F V$RXRF V$NFKB V$MZF1
V$PAX5 V$SREB V$KLFS V$ZFXY V$OCT1 V$LEFF V$ARID V$DUXF V$HOMF V$NKXH V$PCBE V$EVI1 V$GATA V$IKRS V$NGRE V$PLAG V$HEAT V$STAT V$CP2F V$MEF3 V$ZBED
V$GLIF V$E2FF V$HAND V$NEUR V$SP1F
alignment position
Mouse_82DEL
Human_ortho_82DEL
Rhesus_ortho_82DEL
Chimp_ortho_82DEL
Rabbit_ortho_82DEL
Rat_ortho_82DEL
Horse_ortho_82DEL
Dog_ortho_82DEL
Cow_ortho_82DEL
Opossum_ortho_82DEL

B


V$CLOX V$SORY V$CART V$HNF1 V$HOXF V$PAX7 V$BRNF V$HNF6 V$LHXF V$NKX6 V$HOXC V$NFKB V$MZF1 V$SREB V$KLFS V$RXRF
alignment position
Mouse_82DEL
Human_ortho_82DEL
Rhesus_ortho_82DEL
Chimp_ortho_82DEL
Rabbit_ortho_82DEL
Rat_ortho_82DEL
Horse_ortho_82DEL
Dog_ortho_82DEL
Cow_ortho_82DEL
Opossum_ortho_82DEL

Fig. 5.8 | DEL82 harbours several conserved motifs.
Position of DEL82 +4/-8nt was scanned for TFBSs using Genomatix (Munich, Germany) and conserved TFBS patterns in orthologous regions was assessed across 10 species. (A) Shows TFBSs sites that occur in at least one species. (B) shows TFBSs present in at least 80% of named species. Two regulatory modules within DEL82 show high phylogenic conservation. Location of DEL82 is indicated with a red box.

5.2.3.1 Mitochondrial copy number

To assess the effect of DEL82 on WAT function, adipose tissues were excised from male mice at 6 months of age that had been on a HFD since weaning. Adipose tissues from DEL82 HOM had a significantly increased number of mitochondria (Fig. 5.9) with 10.8-fold (P=0.014) and 6.6-fold (P=0.002) increase in gWAT and iWAT, respectively, compared to WT animals. This is particulary interesting given that DEL82 had opposing effects on target gene expression in gWAT compared to iWAT. Despite this, the effect of increased *Irx3* expression in gWAT and decreased *Irx3* and *Irx5* expression in iWAT on mitochondrial copy number is comparable in both depots.

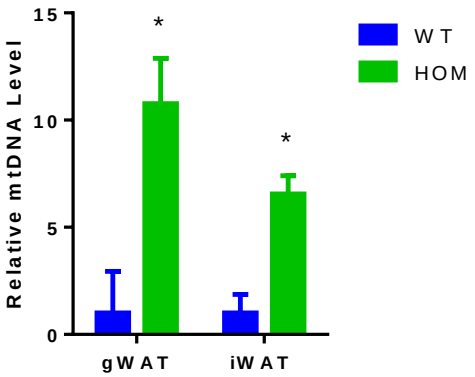


Fig. 5.9 | Adipose tissue from DEL82 HOM mice have increased mitochondrial copy number. DNA was isolated from mature adipose depots and mitochondrial DNA level measured using qPCR for a probe detecting mtDNA, namely *mtND1* (normalized to genomic DNA *Gapdh*). WT n=7 and HOM n=9. Statistical significance was assessed using Student's t-test; data are expressed as mean ± SEM *P<0.05.

5.2.3.2 Gene expression in mature WAT depots

I next quantified the effect that DEL82 had on thermogenic capacity in mature iWAT and gWAT isolated from HOM and WT male mice by performing gene expression analysis. The fat depots were isolated at 6 months of age following exposure to HFD since weaning. Gene expression changes in DEL82 fat pads were profiled for a number of commonly used markers for thermogensis and mitochondrial respiration (*Adrb3, Pgc1a, Ucp1, Cox71, Cox8b, Prdm16, Dio2, Elovl3*).

DEL82 had a significant effect on β-adrenergic receptor 3 (*Adrb3*) expression in gWAT ($P<0.0001$), with DEL82 homozygotes having ~3-fold increased levels compared to WT animals (Fig. 5.10). In contrast, in iWAT of these mice no difference in *Adrb3* levels was detected. Instead, *Elovl3* was more than 2-fold increased in DEL82 HOM iWAT compared to WT ($P=0.0003$) and there was a trend ($P=0.2$), albeit not statistically significant, for increased *Dio2* levels in DEL82 HOM compared to WT (Fig. 5.10). Although mitochondrial copy number was significantly increased in WAT of both depots, the mitochondrial biogenesis master regulator *Pgc1a* was unaltered in both depots (Fig. 5.10). I hpothesize that stimulation of the tissue and/or measuring gene expression in mature WAT-derived pre-adipocytes would have yielded insight into whether *Pgc1a* has been up-regulated at an earlier stage. What I capture by profiling mature adipose tissue are the consequences of altered gene regulatory mechanisms that occurred throughout tissue development.

Subsequently, I assessed adipogenic gene expression in these mature WAT depots. In gWAT from DEL82, I observed a trend for an elevated profile of adipogenic and mature adipocyte markers (*Pparg, Cebpa, Fabp4, Plin1, Fasn*), however, only *Fasn* showed statistically significant upregulation (~1.5-fold) in DEL82 HOM compared to WT gWAT ($P=0.043$) (Fig. 5.11). *Fasn* is an enzyme that catalyzes the first committed step in *de novo* lipogenesis. Although lipogenesis is believed to only be a minor contributor to whole body lipid stores in a modern human consuming a typical diet high in fat, evidence suggests that *de novo* lipogenesis acts as a metabolic signal transmitter autonomous of the generation of

lipid stores (Lodhi *et al.*, 2011). Additionaly, SNPs in *FASN* were identified through GWAS to be associated with WHR, BMI-related traits and also *FASN* transcript levels (Schleinitz *et al.*, 2014).

Similarly to gWAT, *Fasn* was also signifcanty upregulated (~1.3-fold) in iWAT of DEL82 HOM males (P=0.025) (Fig. 5.11). Intriguingly, although gWAT mass was increased in DEL82 HOM males (consistent with increased *Fasn* expression), overall iWAT mass was unaltered in DEL82 HOM males, despite upregulation of *Fasn*. Notably, gene expression of *Fasn* is lower in gWAT than iWAT (Lee *et al.*, 2017) and therefore increasing it moderately might have a more drastic effect in gWAT on cell function.

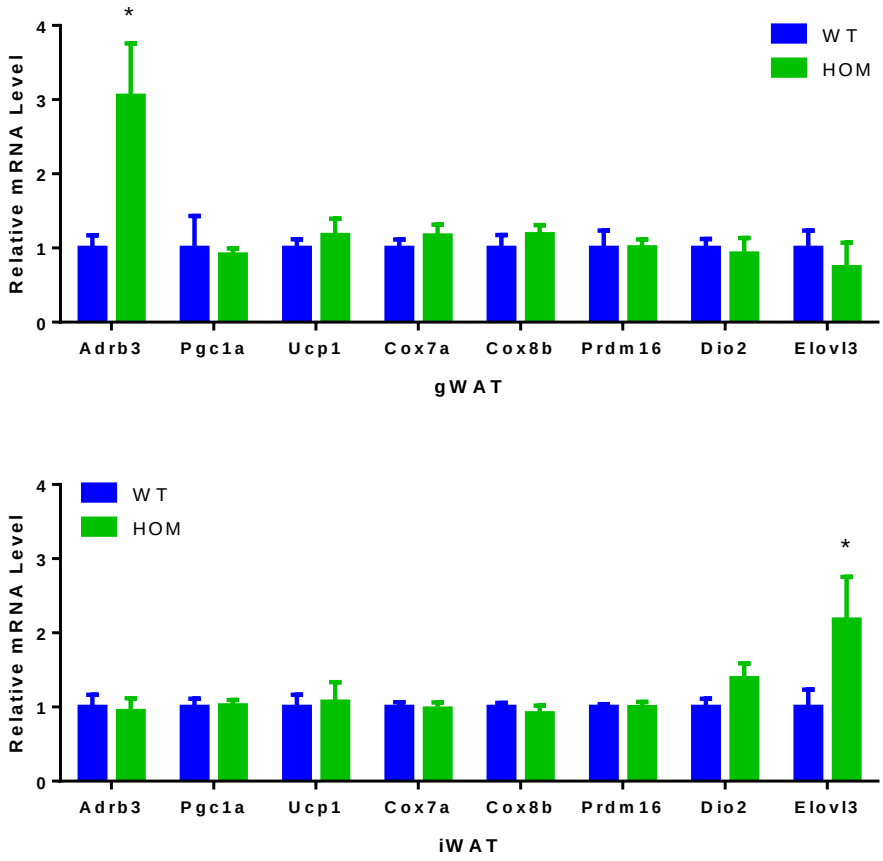


Fig. 5.10 | Gene expression of thermogenesis and mitochondrial function markers in DEL82 and WT adipose tissues.
Mature gWAT (top) and iWAT (bottom) depots were excised from male DEL82 WT (n=7) and HOM (n=9) animals at 20-22 weeks of age following HFD treament since weaning. Gene expression changes were evaluated by means of qPCR (normalised to *Canx*). Statistical significance was assessed using Student's t-test; data are expressed as mean ± SEM *P<0.05.

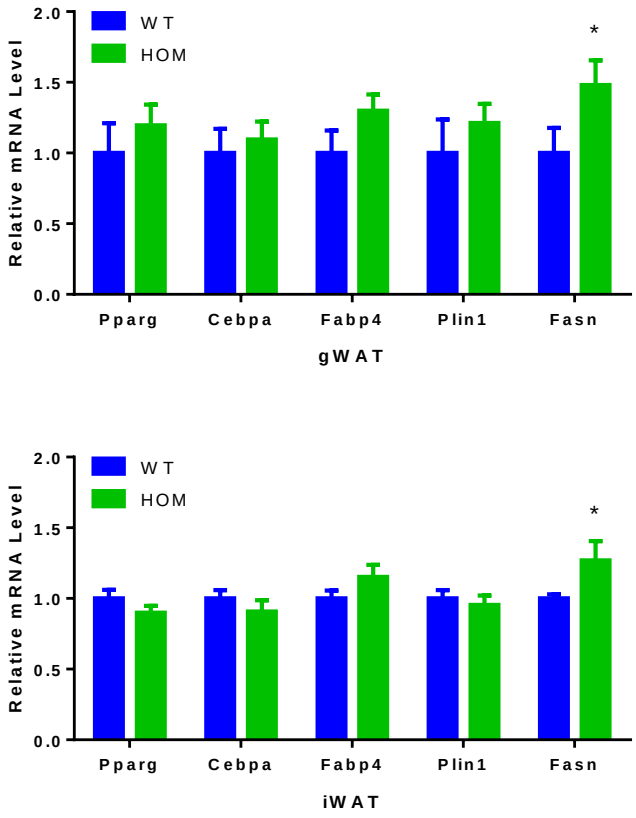


Fig. 5.11 | Adipogenic gene expression changes in DEL82.
Mature gWAT (top) and iWAT (bottom) depots were excised from male DEL82 WT (n=7) and HOM (n=9) animals at 6 months of age following HFD treament since weaning. Gene expression changes were evaluated by means of qPCR (normalised to *Canx*). Statistical significance was assessed using Student's t-test; data are expressed as mean ± SEM *P<0.05.

5.2.3.3 OCR of in vitro differentiated pre-adipocytes

To assess whether gene expression changes in *Adrb3* in DEL82 gWAT had functional consequences on cellular oxygen consumption rate, pre-adipocytes from WT and DEL82 HOM male mice were isolated to perform a mitochondrial stress test using the Seahorse Bioflux XF24 Analyser. The Seahorse XF24 analyser can be used to measure OCR in live cultured cells in real-time. Measurements were taken under basal conditions and following the administration of mitochondrial OXPHOS inhibitors oligomycin (complex 5 inhibitor), FCCP (mitochondrial uncoupler) and Antimycin / Rotenone (complex 3 and complex 1 inhibitor, respectively). This allows the generation of a bioenergetics profile of the cells, including basal mitochondrial OXPHOS function, mitochondrial proton leak, ATP production, maximal respiration and spare respiratory capacity. Pre-adipocytes were allowed to differentiate in the assay plates and subsequent experiments were performed either under basal conditions or after overnight isoproteronol stimulation. Isoproteronol is a β-adrenergic receptor agonist that is widely used to induce browning *in vitro*.

Under basal conditions, DEL82 did not affect OCR profiles in gWAT- or iWAT-derived pre-adipocytes (Fig. 5.12; Fig. 5.13). This is consistent with the absence of gene expression changes in unstimulated DEL82 pre-adipocytes (Appendix Fig. S3). However, β-adrenergic stimulation resulted in significantly increased OCR and proton leak (Fig. 5.13) in DEL82 adipocytes isolated from gWAT, indicative of uncoupled respiration. Increased β-adrenergic responsiveness in gWAT-derived pre-adipcytes is consistent with increased *Adrb3* expression in mature tissue (Fig. 5.10). In contrast, β-adrenergic stimulation in differentiated pre-adipocytes isolated from DEL82 iWAT had no effect on OCR (Fig. 5.12).

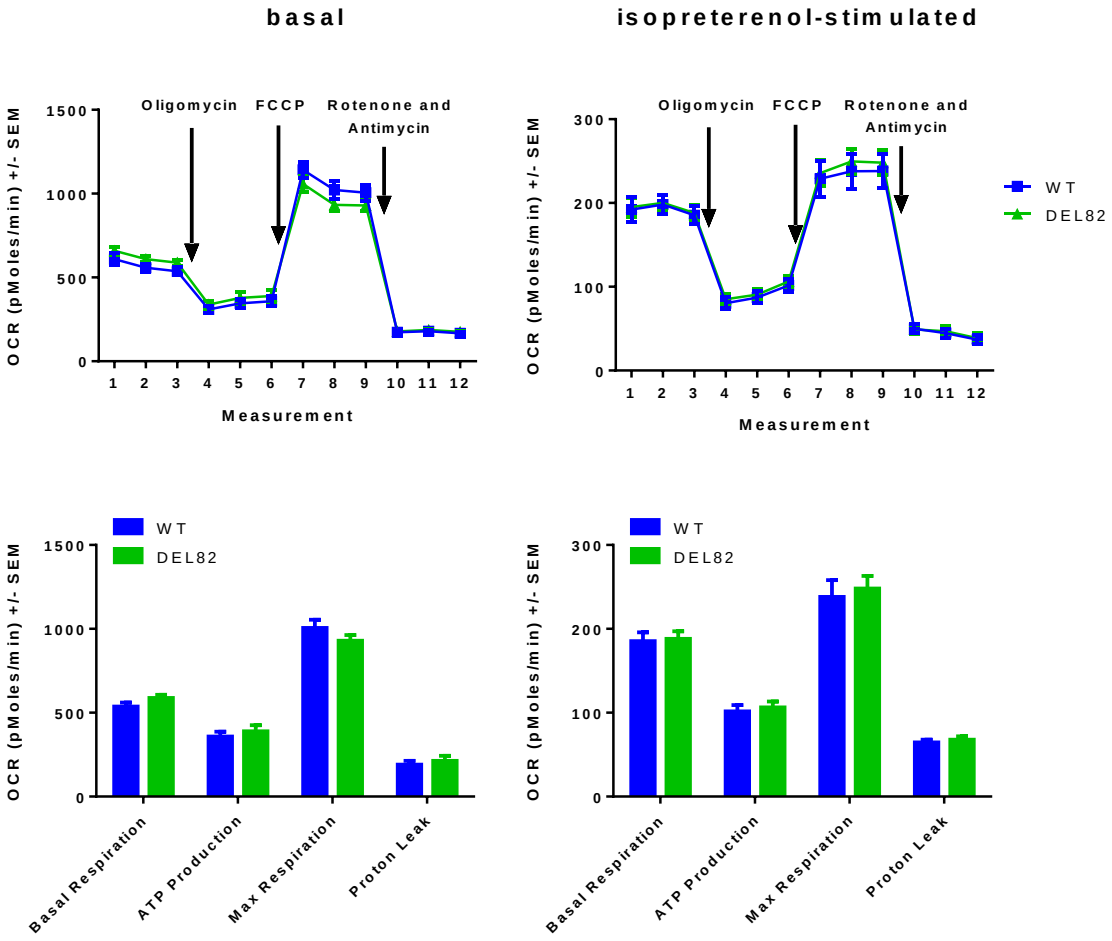


Fig. 5.12 | Seahorse in iWAT under basal conditions and isoproterenol-stimulated.
Assay was performed in iWAT-derived pre-adipocytes at terminal differentiation with 10 technical replicates per sample. Data shown is representative of 3 repeated experiments per condition obtaining comparable results. Data are expressed as mean ± SEM.

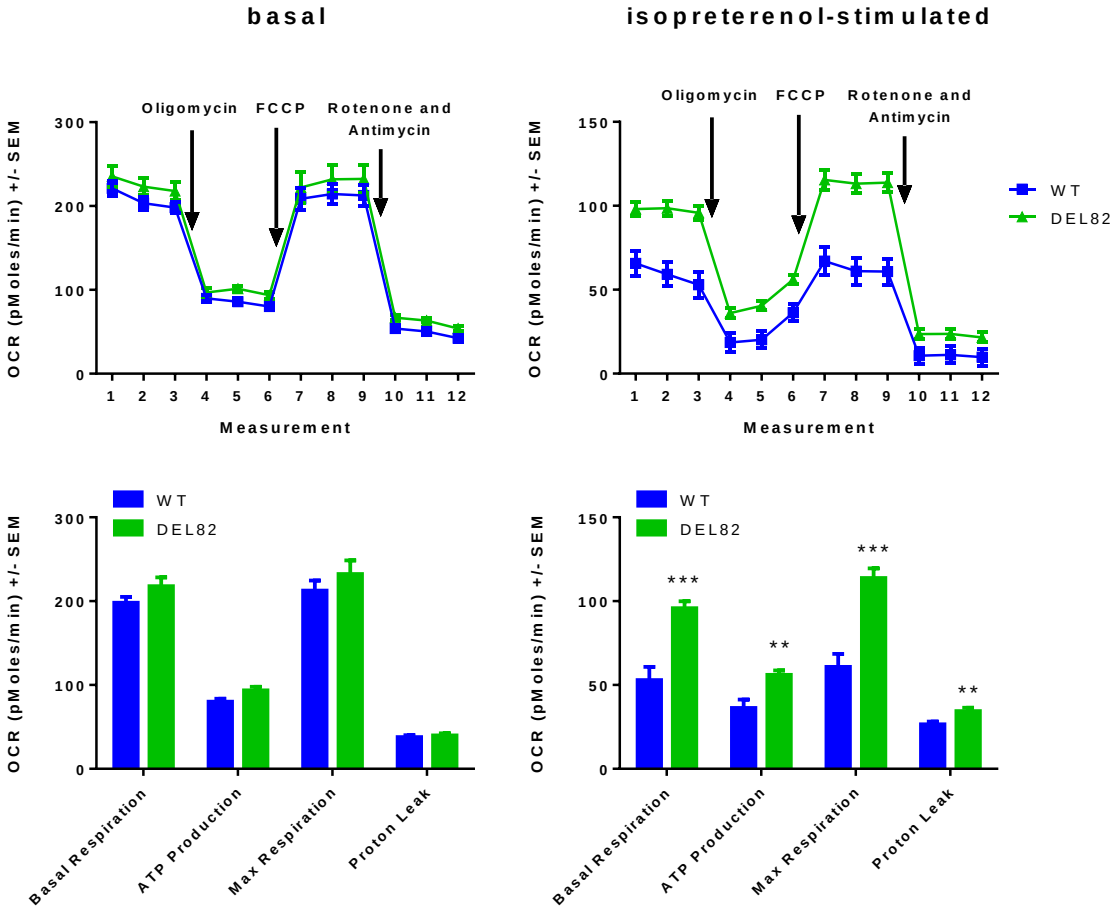


Fig. 5.13 | Seahorse in gWAT under basal conditions and isoproterenol-stimulated.
Assay was performed in gWAT-derived pre-adipocytes at terminal differentiation with 10 technical replicates per sample. Data presented is representative of 3 repeated experiments for basal and 2 for stimulated condition obtaining comparable results. Statistical significance was assessed using Student's t-test; data are expressed as mean ± SEM *P<0.05, **P<0.01, ***P<0.001.

It is worth pointing out that for the OCR experiment, pre-adipocytes were isolated from young animals (in absence of a phenotype) whereas for the gene expression assays in 5.2.4.2, mature WAT depots were derived from 6-months old animals that had been on a HFD and that presented with a body fat mass phenotype in DEL82. These differences in experimental design need to be considered when comparing results.

In summary, DEL82 had an effect on β-adrenergic stimulated OCR in gWAT-derived, but not iWAT-derived, pre-adipocytes, suggesting that the genetic effect of DEL82 is depot-specific and dependent on environmental triggers.

5.3 Discussion

Data presented in this chapter elucidate parts of the likely mechanism by which DEL82 causes altered body weight and fat mass regulation in male mice. By combining a number of genetics approaches together with 3D FISH imaging we identified the DEL82 target genes in pre-adipocytes. My results show that DEL82 causes dysregulation of *Irx3* and *Irx5* specifically in adipocyte precursors (an effect that is complex in nature), and that the cellular consequences of altered *Irx3* and *Irx5* expression in pre-adipocytes has lasting effects on mature tissue function.

In several ways, the findings presented in this chapter indicate a deep conservation of *cis*-regulatory networks between human and mouse at the *FTO* locus. Human pre-adipocytes isolated from subcutaneous WAT of rs1421085 risk allele carriers have increased *IRX3* and *IRX5* expression and this effect is restricted to an early developmental stage of the cell (Claussnitzer *et al.*, 2015). I hypothesized that deletion of this *FTO* intronic enhancer element around human rs1421085 in mouse would disrupt enhancer activity and cause reduced *Irx3* and *Irx5* expression specifically in pre-adipocytes. Both target genes and target cell type identified in human *FTO* risk allele carriers were recapitulated in my mouse model, in the expected direction. However, my DEL82 mutant displayed additional levels of

regulatory specificity. Alteration of target genes was only detectable in males, the number of identified target genes was different for the subcutaneous and visceral adipose depots studied, and the direction of target gene alteration was in the opposite direction in pre-adipocytes isolated from iWAT and gWAT. Although the similarities between human and mouse data point towards a regulatory conserved network, it will be important to validate the new discoveries made in this mouse model in human cells in the future, to establish how deep context-specificity is conserved between the species. Irrespectively of this, the data presented in this chapter clearly emphasizes the need to study adipocyte gene regulation not only in the right cell type and developmental stage, but also in both sexes and in different adipose tissue depots.

It has previously been hypothesized that variant rs1421085 potentially disrupts interaction between the *Fto* enhancer element and the target promoters for *IRX3* and *IRX5* (Claussnitzer *et al.*, 2015; Smemo *et al.*, 2014). Using FISH, my results show that in pre-adipocytes, DEL82 interrupts the binding of *Irx3* to its regulator in *Fto* intron 1, resulting in altered *Irx3* expression in early differentiating pre-adipocytes. Co-localisation of the *Fto* enhancer and *Irx3* (less than 200nm apart) was nearly 30% in D1 WT pre-adipocytes, similar levels as those observed for Shh-ZRS and Hoxd13-GCR in mouse distal posterior limb buds (Williamson *et al*., 2016). This provides clear evidence for *Irx3* to form functional connections with the *Fto* regulatory element in early adipocyte precursors and that sequence variants in *Fto* intron 1 have a direct effect on *Irx3* regulation. *Irx5,* on the other hand, was more distant to *Fto* intron 1, and DEL82 selectively interrupted interactions between Fto-en and *Irx5* in gWAT-derived pre-adipocytes, but not iWAT ones, suggesting cell type specific regulator-target interaction.

As illustrated in a report by Smemo *et al.* (2014) ChIA-PET using polymerase 2 in MCF7 (human breast adenocarcinoma) cells shows interactions between obesity associated regions in *FTO* intron 1 and *IRX5*. In contrast, Hi-C data from ENCODE in human ES cells and fetal lung did not show such an interaction, emphasizing the tissue-specificity of

genomic organization and the need to study the relevant cell type. Although FISH clearly showed that the *FTO* obesity-associated interval in mouse co-localises with *Irx3* in early pre-adipocytes, there are certainly more experiments required to unravel the link between regulatory elements within *Fto* and *Irx5* expression in pre-adipocytes.

Importantly, altered *Irx3* and *Irx5* expression in DEL82 pre-adipocytes has functional consequences in their development with persistent implications in mature tissue function. In my model, deleting a regulatory element of 82 nucleotides results in a ~30% alteration of *Irx3* and *Irx5* expression, resembling fine-tuning of gene expression. Surprisingly, although target gene direction and number are distinct in gWAT and iWAT, the effect of increased *Irx3* expression in gWAT and decreased *Irx3* and *Irx5* expression in iWAT results in similar cellular consequences, such as increased number of mitochondria. At the same time, both tissues show clear differences, such as in expression levels of *Adrb3* and responsiveness to β-adrenergic stimulation. Although mitochondrial copy number is significantly upregulated in mature adipose tissues of both tissues, *Irx3* and *Irx5* are no longer dysregulated. This suggests that the effect of altered *Irx3* and *Irx5* expression in pre-adipocytes during early differentiation has lasting effects on adipose tissue function (Fig. 5.14). This is consistent with human data, where it was shown that the effect of rs1421085 on *IRX3* and *IRX5* activity was restricted to early adipogenesis, however, the effects on tissue function persisted in mature adipocytes (Claussnitzer *et al.*, 2015).

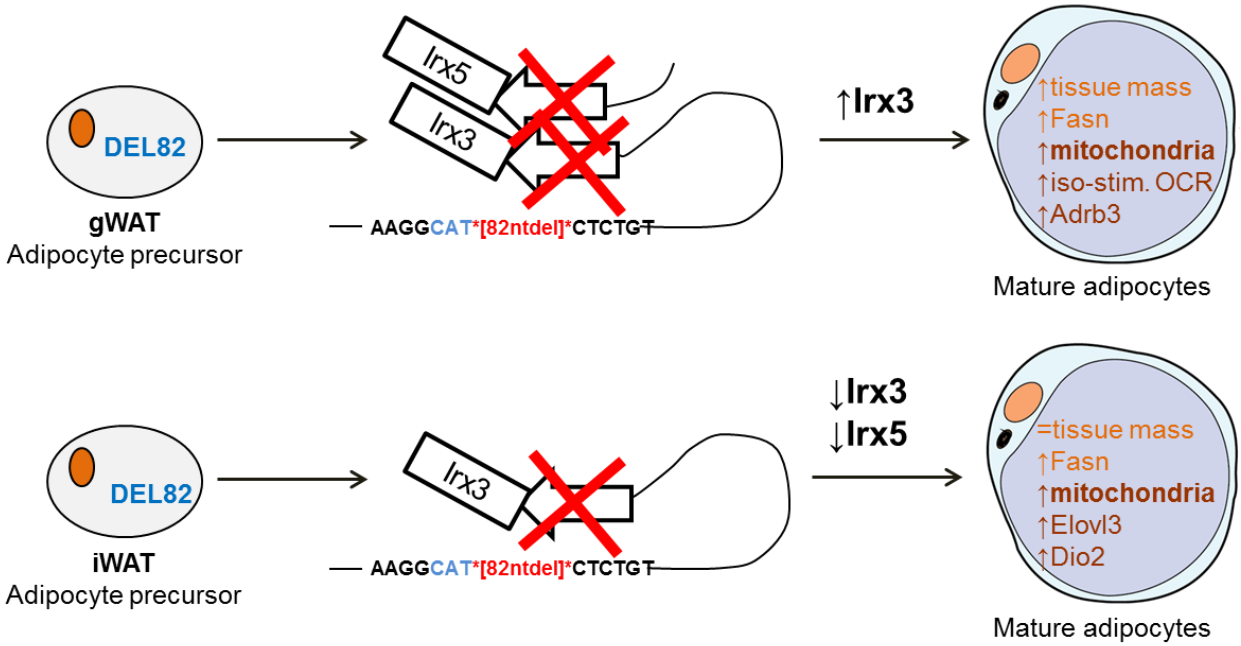


Fig. 5.14 | Mechanistic model summarizing DEL82 consequences on chromatin interactions and cellular function in adipocyte precursors from gWAT and iWAT.
In gWAT, DEL82 disrupts interactions between the *Fto* regulatory element and its targets *Irx3* and *Irx5*, resulting in increased expression of *Irx3* in pre-adipocytes. The consequences of upregulation of *Irx3* in adipocyte precursors are complex and involve increased tissue mass, possibly due to increased *Fasn* expression. At the same time, cells differentiate into more brown-like cells measured by increased amount of mitochondria and increased expression of *Adrb3* expression. This in turn results in pre-adipocytes to be responsive to β-adrenergic stimulation and consequently increased oxygen consumption and proton leak. In iWAT, DEL82 results in decreased *Irx3* and *Irx5* expression. Decreased *Irx3* expression in iWAT-derived pre-adipocytes is due to decreased binding of *Irx3* to the *Fto* regulatory element as shown by FISH. The effect of DEL82 on *Irx5* is currently not known, but is likely downstream of decreased *Irx3* expression. Reduced *Irx3* and *Irx5* expression will lead to cells developing into mature adipocytes that have an increased amount of mitochondria, increased gene expression of browning markers such as *Elovl3* and *Dio2*, but also increased *Fasn* expression.

Increased *Irx3* expression in pre-adipocytes of gWAT corresponds to increased tissue weight in later life as shown in Chapter 4. However, decreased *Irx3* and *Irx5* expression in pre-adipocytes isolated from iWAT is not translating into a significant decrease in tissue volume. Current histological analysis of iWAT will address the question whether – although there is no overall difference in fat mass - DEL82 results in altered adipocyte size and/or number. The reasons for the absence of an iWAT tissue weight difference between DEL82 and WT could be due to several reasons, including the possibility that *Irx3* and *Irx5* have antagonistic effects on adipogenesis (further discussed in Chapter 6), that the particular time the measurement was taken is not representative of effects that might happen earlier/later. In the future, it would be interesting to sample more depots of both visceral and subcutaneous origin to pin-point the precise tissue-specific mechanisms that seem to underlie the DEL82 *in vivo* phenotype in males.

The limiting step in the OCR assays was the number of assays that could be performed on one day with one machine, making it necessary to stagger experiments over several days. However, as all experimental conditions were present on the same assay plate on the same day, these differences did not interfer with the intepretation of genotype-mediated differences on OCR observed in DEL82 following β-adrenergic stimulation. Additionally, every experiment was confirmed with 2-3 additional experiments, as stated in the figure legends.

Noteworthy, the strongest genetic effect in humans on cellular energy expenditure was stimulus-induced (β-adrenergic) thermoregulation (Claussnitzer *et al.*, 2015), however, regarding whole-body energy expenditure there are currently no human studies addressing the effect of *FTO* variants following β-adrenergic stimulants (e.g., by cold or exercise). Consistent with human cellular studies, β-adrenergic stimulation was also critical for eliciting a phenotype in mouse, as seen in the Seahorse analysis.

In summary, DEL82 is within a regulatory element for *Irx3* and *Irx5* that is active in pre-adipocytes of male mice. In iWAT- and gWAT-derived pre-adipocytes, DEL82 causes decreased *Irx3*/*Irx5* and increased *Irx3* expression, respectively, suggesting a depot-specific effect of *FTO* regulation. The cellular consequences of DEL82 are complex and involve increased amount of mitochondria in both adipose depots studied, albeit distinct expression pattern of mitochondrial and adipogenic genes. These findings go beyond what is currently known about the mechanistic underpinnings of human *FTO* risk allele carriers and suggest an even more complex *cis*-regulatory landscape in *FTO* intron 1 that varies for sex, adipose depot, and developmental stage.

6 *IRX3* and *IRX5* mechanisms

6.1 Introduction

IRX3 and IRX5 are transcription factors that are largely uncharacterised in adipose tissue biology and only have recently gained attention in metabolism following reports by Smemo *et al.* (2014) and Claussnitzer *et al.* (2015). Although *Irx3* has been studied in knockout animal models before, with a striking effect on adipose tissue where lack of *Irx3* resulted in adipocyte browning and reduced overall fat mass (Smemo *et al.*, 2014), the authors concluded that the effects on adipose tissue were driven by the hypothalamus. In the case of human adipocyte studies, Claussnitzer *et al.* (2015) showed that risk allele carriers have altered *IRX3* and *IRX5* expression, and they further characterized downstream consequences of the net effect of alteration of both target genes. Similar to Smemo *et al.* (2014), this study mainly focused on the effect of *IRX3* and *IRX5* on adipocyte browning. However, the functions of these genes have not been addressed individually in adipocytes yet and even more importantly, very little is known about the adipocyte-intrinsic mechanisms of *IRX3* and *IRX5*. In this Chapter, I aimed to address the functional roles of *IRX3/Irx3* and *IRX5/Irx5* in adipocyte function and development. I was particularly interested in outlining possible similarities and differences in *IRX3/Irx3* and *IRX5/Irx5* function.

6.2 Results

6.2.1 Gene expression of *IRX3* and *IRX5*

First, I set out to compare the patterns of *IRX3* and *IRX5* mRNA expression across tissues. Using RNA-seq data from the GTEx Portal (https://www.gtexportal.org/home/) revealed that *IRX3* and *IRX5* are widely expressed across the 53 tissues for which data is available (Ardlie *et al.*, 2015). Both genes are comparably expressed in females and males (Fig. 6.1). The expression pattern of *IRX3* and *IRX5* are similar, suggesting that they are co-regulated. To further test this observation, I downloaded the RNA-seq data from GTEx and plotted the

expression (RPKM) of *IRX3* against *IRX5* across 53 tissues. Indeed, there was significant correlation between *IRX3* and *IRX5* expression, with a coefficient R^2 = 0.9 (Fig. 6.2). When comparing *IRX3* expression with that of other genes in the locus, it became clear that *IRX3* was not co-expressed with its upstream neighbours *FTO* and *RPGRIP1L,* but showed some correlation with *IRX6* expression (R^2 = 0.49), a gene downstream of *IRX5*. This suggests that *IRX* family members at this locus, namely *IRX3, IRX5* and *IRX6*, share their regulatory machinery.

From the GTEx RNA-seq data, it was apparent that *IRX3* and *IRX5* are expressed at higher levels in subcutaneous adipose tissue compared to visceral adipose tissue (Fig. 6.1). Using qPCR, I assessed gene expression of *Fto, Irx3* and *Irx5* in iWAT- and gWAT-derived primary pre-adipocytes. Comparing RNA levels between these two depots revealed that consistent with human RNA-seq data, *Irx3* and *Irx5* were expressed much higher in pre-adipocytes originating from iWAT compared to gWAT (Fig. 6.3). This suggests that the differential gene expression in WAT depots is conserved in human and mouse.

To study the expression pattern of *IRX3* and *IRX5* during adipocyte development, I next visualised the openness of both gene promoters using ATAC-seq data in hWAT during 4 time-points of differentiation. Chromatin openness is based on ATAC tag counts. Promoter regions for *IRX3* and *IRX5* show increasing accessibility during adipocyte differentiation. Promoter annotations from ChromHMM are also stretched at later time-points of differentiating hASC (Fig. 7.4). Increasing accessibility of *IRX3* and *IRX5* gene promoters with progressing adipocyte maturation suggest that *IRX3* and *IRX5* transcription increases as differentiation progresses.

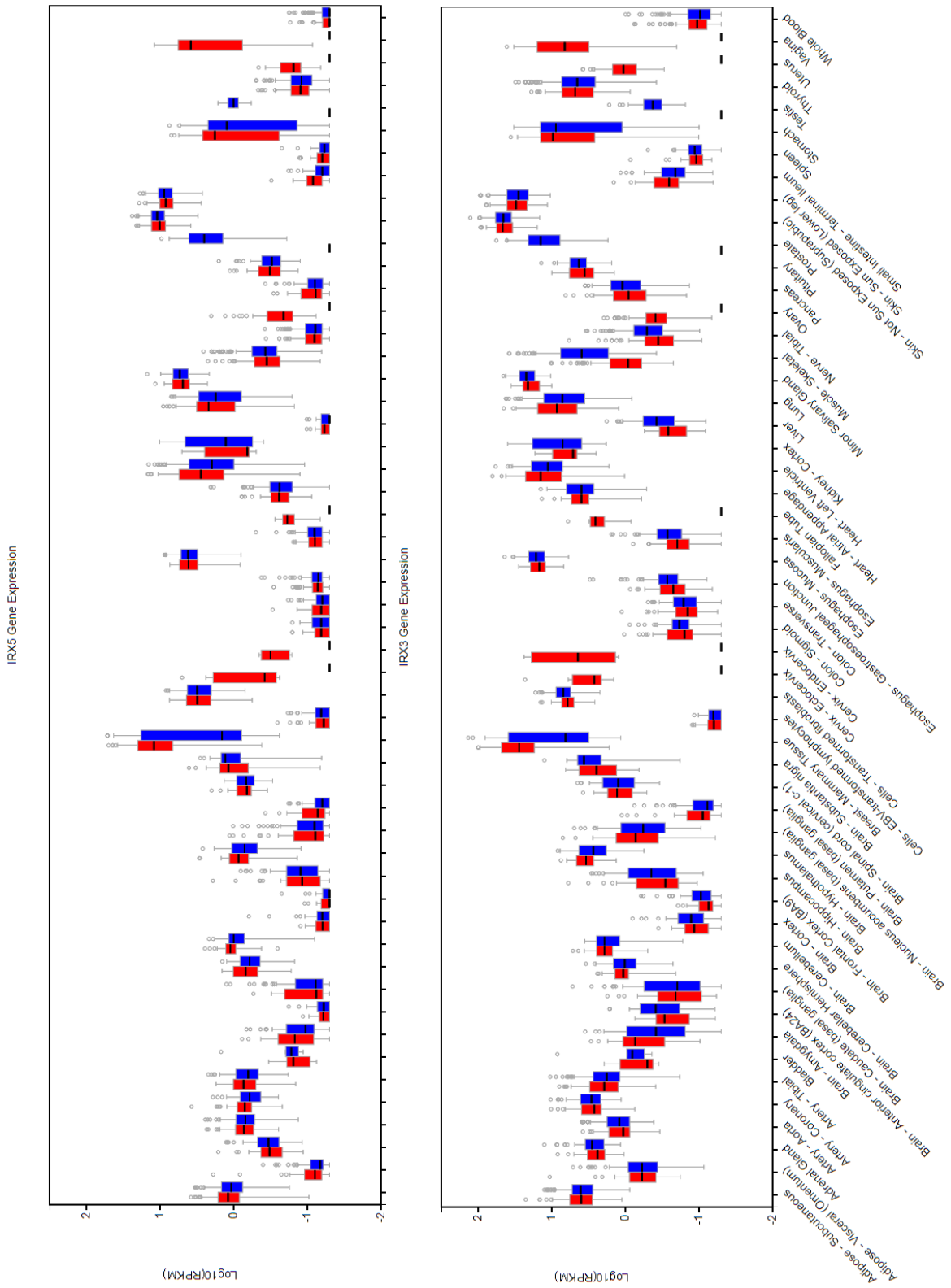


Fig. 6.1 | *IRX3* and *IRX5* are widely expressed across tissues.
Gene expression levels of *IRX3* and *IRX5* across human tissues in females (red) and males (blue). RPKM refers to Reads Per Kilobase per Million mapped reads. Figure exported from GTEx (Ardlie *et al.*, 2015).

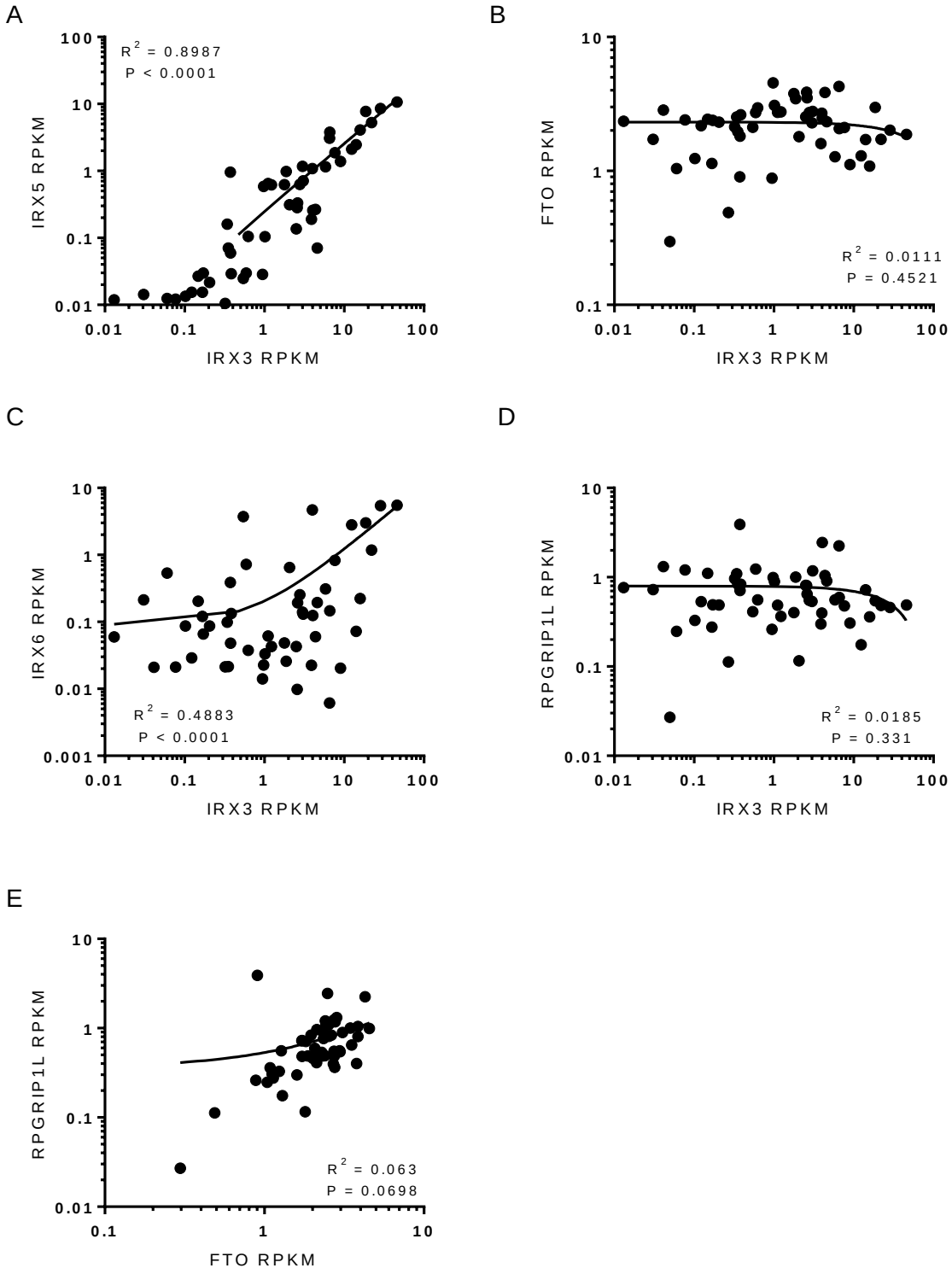


Fig. 6.2 | ***IRX3, IRX5* and *IRX6* are co-regulated across different tissues**.
Gene expression data was downloaded from GTEx (Ardlie *et al.*, 2015) for 53 human tissues and linear regression was analysed for expression correlation for (A) *IRX3* and *IRX5*, (B) *IRX3* and *FTO*, (C) *IRX3* and *IRX6*, (D) *IRX3* and *RPGRIP1L*, and (E) *FTO* and *RPGRIP1L*. RPKM refers to Reads Per Kilobase per Million mapped reads.

To confirm this, I studied the expression pattern of *Irx3* and *Irx5* during adipocyte development. I isolated the SVF which contains adipocyte precursors, from mouse iWAT and extracted RNA at 5 time-points (D0, D1, D2, D4, D8) during adipogenic differentiation. As a positive control, I measured expression of *Plin1*, a gene that is known to be upregulated during adipogenesis, across the 5 time-points. *Plin1* expression showed the expected pattern and gradual increase during adipogenesis, significantly upregulated in terminal differentiated adipocytes, at day 8 of adipogenic treatment (Fig. 6.5). Gene expression profiling of *Irx3* and *Irx5* revealed that both of these genes were also upregulated during adipogenesis (Fig. 6.5). *Irx3* expression gradually increased as differentiation progresses and after 8 days of adipogenic stimulation, *Irx3* is 10-fold higher than at day 0 (Fig. 6.5). *Irx5* showed a slightly different profile, increasing during differentiation but peaking at D4 of differentiation, with over 30-fold increased mRNA levels (Fig. 6.5).

Differential gene expression during the course of adipocyte differentiation points towards a regulatory role for *IRX3*/*Irx3* and *IRX5*/*Irx5* in adipose tissue development.

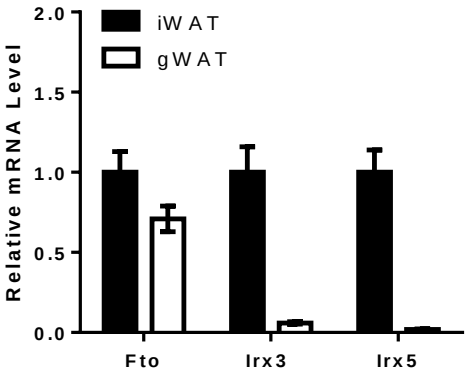
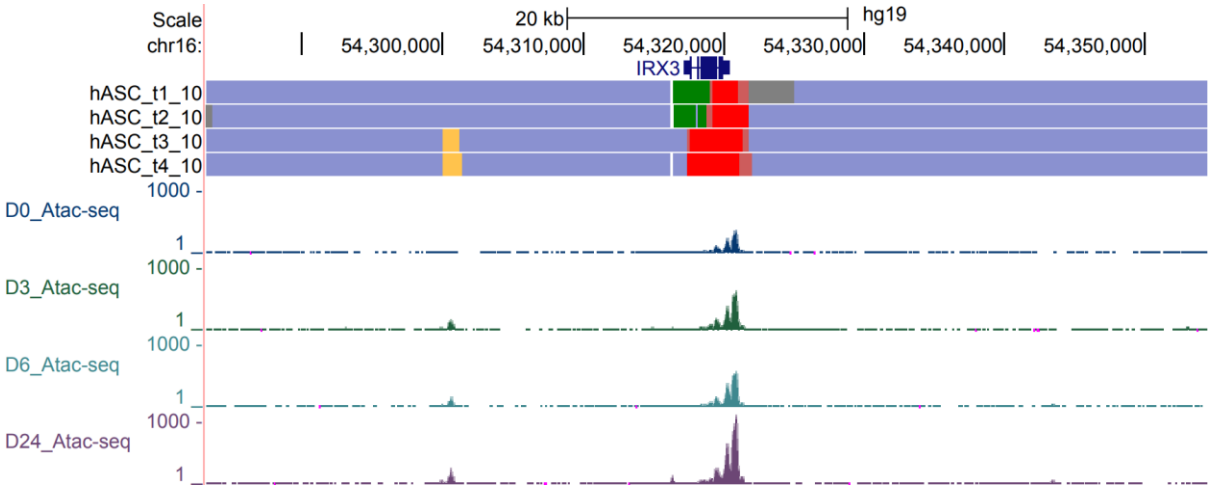


Fig. 6.3 | Differential gene expression in iWAT- and gWAT-derived pre-adipocytes.
Pre-adipocytes from iWAT (n=3) and gWAT (n=1) were isolated from mice at 6-8 weeks of age. Gene expression of *Fto, Irx3* and *Irx5* was measured by the means of qPCR (normalised to *18S*). Data are shown as mean ± SEM.

A



B

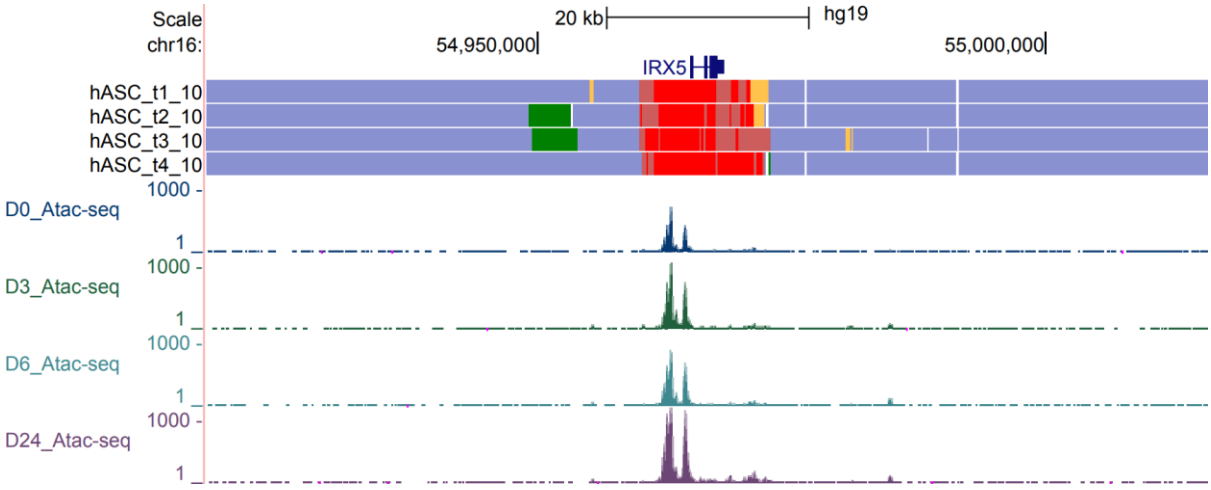


Fig. 6.4 | ATAC-seq site analysis during hWAT adipogenesis at *IRX3* and *IRX5* promoters. Promoter regions for *IRX3* (A) and *IRX5* (B) show increasing accessibility during adipocyte differentiation. ATAC-seq data and chromatin state annotations are visualized using the UCSC Genome Browser. Chromatin state annotations using ChromHMM from human adipocytes (hASC) during 4 time-points during differentiation (top to bottom: day-2, day0, day3, day9) (Mikkelsen *et al.*, 2010) are shown in the first 4 tracks in each image. The bottom 4 tracks show ATAC-seq histograms from hWAT adipocytes (Xue *et al.*, 2015) during 4 time-points during differentiation (top to bottom: day0, day3, day6, day24). All histograms are shown on the same scale and high values were truncated as necessary.

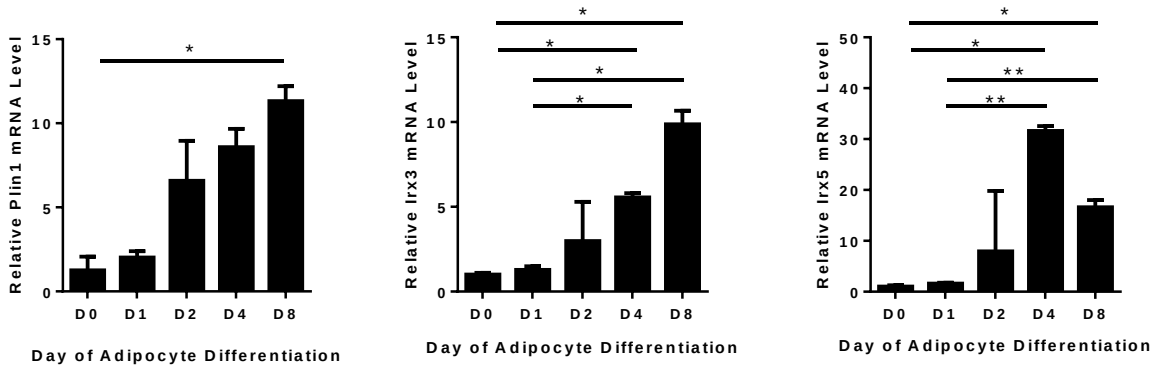


Fig. 6.5 | *Irx3* and *Irx5* expression increases during adipogenesis.
IWAT-derived primary pre-adipocytes were isolated from female mice at 6-8 weeks of age. Gene expression of candidate target genes *Fto, Irx3* and *Irx5* was measured by the means of qPCR (normalised to geometric mean of *Canx* and *Ubc*). Statistical significance was assessed using One-way ANOVA with multiple comparisons, Bonferroni correction; data are shown as mean ± SEM; *P<0.05; **P<0.01.

6.2.2 *Irx3* and adipogenesis

To determine whether *Irx3* has an effect on adipocyte development and browning, 3T3-L1 pre-adipocytes were treated with siRNA for *Irx3* or control RNA for 2 days before differentiation was induced. Cells treated with siIrx3 and control vector (siCON) were assessed for differentiation and thermogenic capacity by the means of gene expression changes of relevant markers and mitochondrial respiration.

First, mitochondrial stress tests were performed using the Seahorse XF24 analyser to determine whether siIrx3 treatment in pre-adipocytes affects mitochondrial OXPHOS function. 3T3-L1 cells were treated with siIrx3 or siCON before differentiation for 7 days. Assays were performed before differentiation (A) and after differentiation (B) and statistical analysis for basal respiration, ATP production, maximal respiration, and proton leak (C) was performed for the third measurement of each stage of the mitochondrial stress test.

As expected, in undifferentiated pre-adipocytes mitochondrial oxidative metabolism was marginal (Fig. 6.6) as oxygen consumption rate did not decrease following addition of OXPHOS inhibitors. This is consistent with the literature, where it has been shown that energy production in pre-adipocytes relies mainly on glycolytic ATP production (Keuper *et al.*, 2014) and that during differentiation, this is replaced by mitochondrial oxidative metabolism. This increase in mitochondrial respiration during adipogenesis is accompanied by the coinciding upregulation of *PGC1a* and *CytC* (Keuper *et al.*, 2014). No difference was observed in the bioenergetics profile between siIrx3- and siCON-treated cells (Fig. 6.6), suggesting that in pre-adipocytes that have not undergone differentiation yet, alteration of *Irx3* does not affect oxidative metabolism, mainly because the cells rely on glycolytic ATP production at this stage.

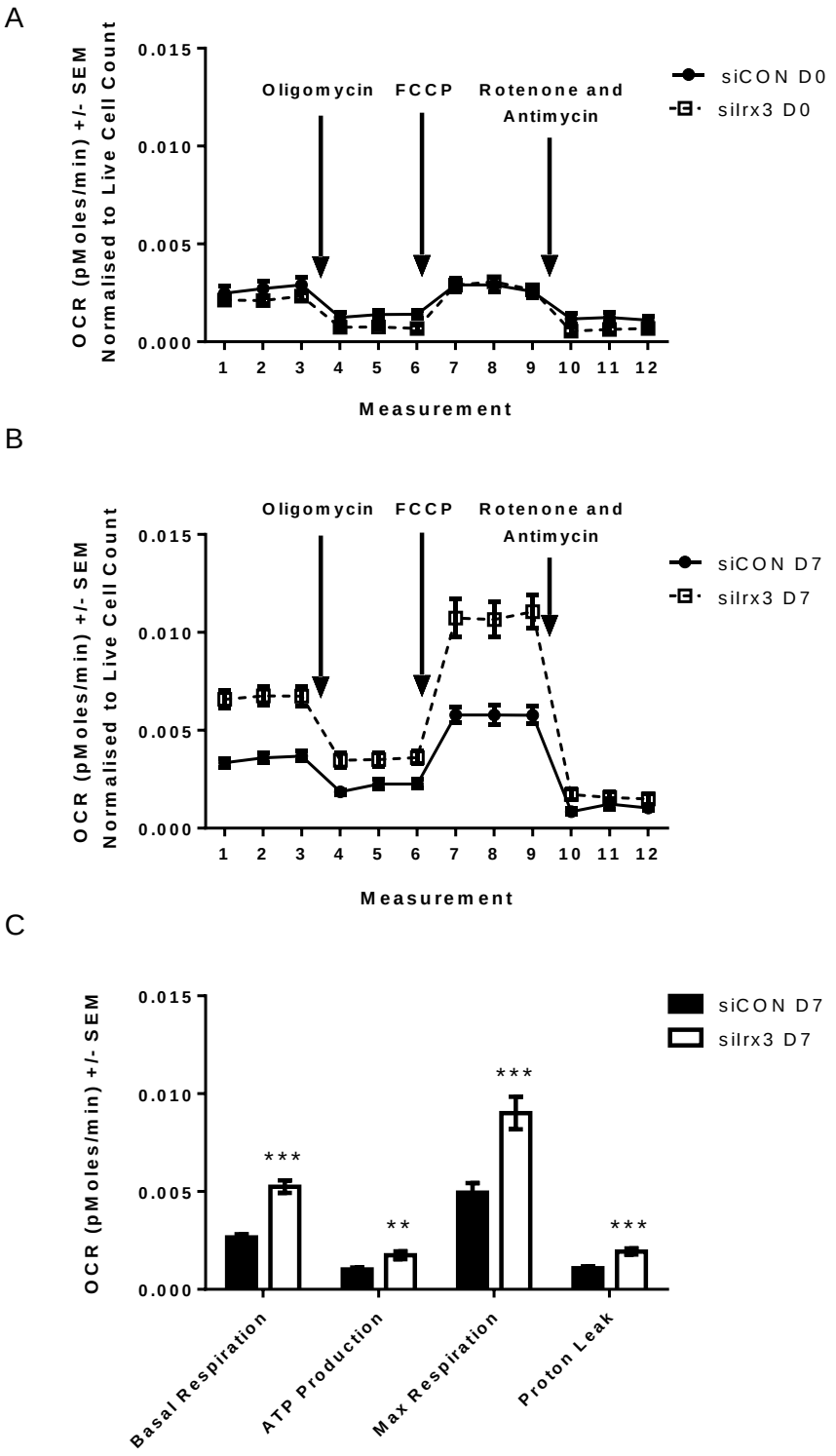


Fig. 6.6 | *Irx3* modulates oxygen consumption rate in 3T3-L1 pre-adipocytes.
Oxygen consumption rate (OCR) was measured at baseline and after oligomycin, FCCP and Rotenone / Antimycin treatment using the Seahorse XF24 analyser. 3T3-L1 cells were treated with siIrx3 or control vector (siCON) before differentiated for 7 days. Assays were performed in 3T3-L1 before differentiation (A) and after differentiation (B) and statistical analysis for basal respiration, ATP production, maximal respiration, and proton leak (C) were performed for the third measurement of each stage of the mitochondrial stress test (see material and Methods for more details). Experiments

were performed with 10 technical replicates and further confirmed in an independent experiment obtaining comparable results. Results were normalized to live cell count. Statistical significance was assessed using Student's t-test; data are shown as mean ± SEM *P<0.05, **P<0.01, ***P<0.001.

In differentiated 3T3-L1 pre-adipocytes, addition of FCCP caused an increased in OCR, and this increase was significantly larger for cells that had been treated with siIrx3 compared to siCON (Fig. 6.6). In addition to an increase in maximal respiration, siIrx3 treated adipocytes show significantly increased basal respiration, ATP production and proton leak, at day 8 of differentiation (Fig. 6.6), all of which is consistent with a more brown-like adipocyte. As 3T3-L1 pre-adipocytes have a hyper-methylation of the *Ucp1* promoter (Alexander Pfeifer, oral communication), the increase in proton leak is unlikely explained by increased *Ucp1* expression, but must be due to *Ucp1*-independent mechanisms.

We hypothesised that the increase in mitochondrial function observed in siIrx3-treated pre-adipocytes was due to up-regulation in mitochondrial biogenesis. To test this hypothesis, gene expression of key markers of mitochondrial biogenesis were analysed next. *Pgc1a* is often referred to as the master regulator of mitochondrial biogenesis. *Pgc1a* was significantly upregulated in siIrx3-treated 3T3-L1 pre-adipocytes at day 8 of differentiation compared to control-transfected cells (Fig. 6.7). Furthermore, several genes involved in the ETC and mitochondrial respiration were increased in siIrx3-treated cells as well, namely *Dio2, Cox7a, Cox8b, Elovl3* (Fig. 6.7), suggesting that Irx3 is a repressor of the mitochondrial program in differentiating pre-adipocytes.

Next, I tested gene expression of genes involved in adipocyte differentiation and function. *Pparg* and *Cebpa* are often referred to as the master regulators of adipocyte differentiation and both genes were significantly upregulated at day 8 of differentiation in siIrx3-treated 3T3-L1 pre-adipocytes compared to controls (Fig. 6.7). *Fabp4, Plin1, Fasn* and *Lipe* were also upregulated following siIrx3 treatment (Fig. 6.7), suggesting that *Irx3* acts as an inhibitor for both, adipocyte differentiation and lipid uptake in mature adipocytes. This is consistent with my hypothesis that reduction of *Irx3* in pre-adipocytes results in these cells to differentiate into mature adipocytes that have increased thermogenic capacity.

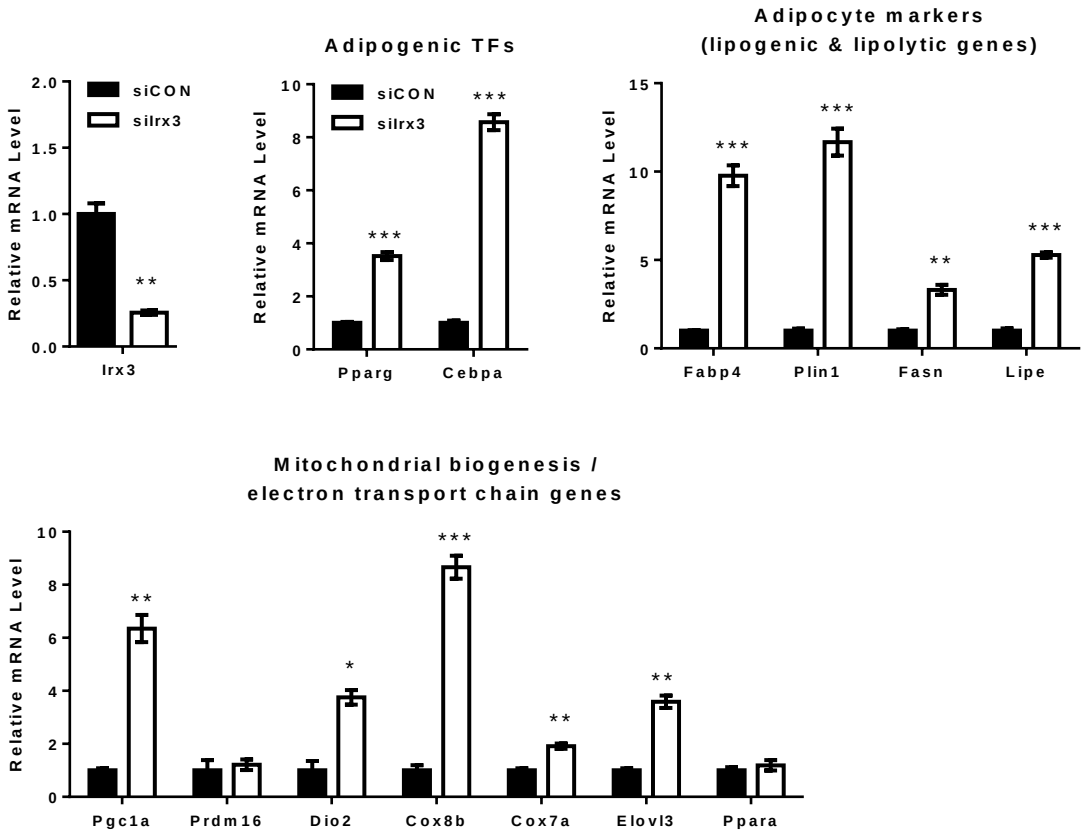


Fig. 6.7 | *Irx3* plays a role in adipocyte development.
siIrx3 treatment results in increased mRNA expression of adipogenic, lipogenic/lipolytic and mitochondrial genes in 3T3-L1 at day 5 of adipogenic differentiation. Experiments were performed in technical triplicates. Statistical significance was assessed via independent Student's t-test. Data are shown as mean ± SEM with *P<0.05; **P<0.01; ***P<0.001.

6.2.3 IRX3 ChIP-seq

To identify the targets of the transcription factor IRX3, we performed ChIP-seq in early differentiating mouse primary pre-adipocytes of both visceral and subcutaneous white adipocyte depots. Adipocyte precursors were isolated from iWAT and gWAT (female B6N), cultured, and crosslinked one day before (D-1) and one day after adipogenic induction (D1), as illustated on the schematic of the experimental setup in Fig. 6.9. Pre-adipocytes were prepared and chromatin fixation was performed in-house and shipped to Active Motif (California, US) were all subsequent steps were executed (including informatics service comprising of QC and peak calling; Fig.6.10). Analysis and pooling of replicates was further assisted by Li Wang from the Broad Institute of Harvard/MIT. The IRX3 antibody used for the IP was abcam ChIP Grade ab25703 (lot GR198517-1) and was tested for specificity via Western Blotting of protein that was obtained from 3T3-L1 pre-adipocytes that have been treated with siIrx3 or siCON (Fig. 6.8).

I identified 951 and 3688 peaks in iWAT D-1 and D1 samples, respectively (Table 6.1). Notably, induction of adipogenesis resulted in 3.9-fold increased number of peaks in D1 samples compared to D-1. This is consistent with my ATAC-seq data, where I have shown that induction of adipogenesis is accompanied by reorganisation of the chromatin landscape and the dynamic opening of many sites. It has previously been shown that this increase in accessibility initiated by induction of the adipogenic program is critical for cell differentiation, as many of the critical regulatory sites of adipgenesis become occupied by TFs (Siersbaek *et al.*, 2012; R. Siersbæk & Mandrup, 2011; Rasmus Siersbæk *et al.*, 2014). Within the first 4 hours post adipogenic stimulation, the chromatin landscape undergoes dramatic remodelling and analysis of global pattern of chromatin accessibility revealed a dramatic increase in the number of DNase I hypersensitive sites (DHS) in response to adipogenic stimuli during differentiation (Siersbaek *et al.,* 2012).

Differential peak location enrichment analysis using GREAT (McLean *et al.*, 2010) between D-1 and D1 in iWAT samples showed that the strongest association is at “Strongly down-

regulated at 8-96h during differentiation of 3T3-L1 cells (fibroblast) into adipocytes" (Table 6.2), further providing evidence for IRX3 to act as a transcriptional repressor of adipogenesis.

To further exploit this data in finding the possible mechansisms by which IRX3 regulates adipogenesis, I used the PANTHER pathway analysis tool (Mi *et al.*, 2017; Thomas *et al.*, 2003) for the lists of identified genes. In gonadal-derived pre-adipocytes, I identified >2000 peaks in proximal promoter regions (≤ 1kb). A statitical overrepresentation test revealed that mitochondrial genes – and particularly mitochondrial ribosomal genes - were significantly overrepresented (P=2.2E-11) in this data set (Table 6.3; Fig. 6.11). I next set out to understand the functional relationship of the expressed genes using STRING (Szklarczyk *et al.*, 2017). STRING is a webtool that integrated biologicaly meaningful physical and functional associations between proteins, with the aim to guide system-wide understanding of cellular function. Using STRING, the identified mitochondrial genes clustered into three main groups: mitochondrial ribosomal machinery, mitochondrial complex I assembly genes and mitochondrial membrane transporters (Fig. 6.12). Fig. 6.13 shows representative histograms for selected complex 1 assembly genes derived from IRX3 ChIP-seq D1. A list of all mitochondrial genes identified through IRX3 ChIP-seq in gWAT and iWAT at D1 of adipogenesis can be found in the Appendix, Table S1 and Table S2, respectively. I propose that IRX3 affects early adipocyte development by modulation of mitochondrial genes that are involved in OXPHOS, protein transport and processing in the mitochondrion. Manipulation of these processes will ultimatly affect mitochondrial respiratory capacity and differentiation capacity of the cell.

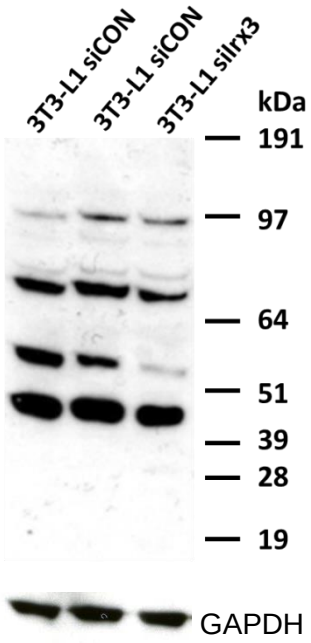


Fig. 6.8 | Western Blot using ChIP-grade antibody against IRX3.
IRX3 antibody used was abcam ab25703 (lot GR198517-1). IRX3 is predicted to be 52kDa. 3T3-L1 pre-adipocytes were treated with siCON (lanes 1-2) or siIrx3 (lane 3) before protein was isolated and Western Blot performed. Bottom panel shows GAPDH on same blot (antibody used was anti-GAPDH antibody (ab8245)).

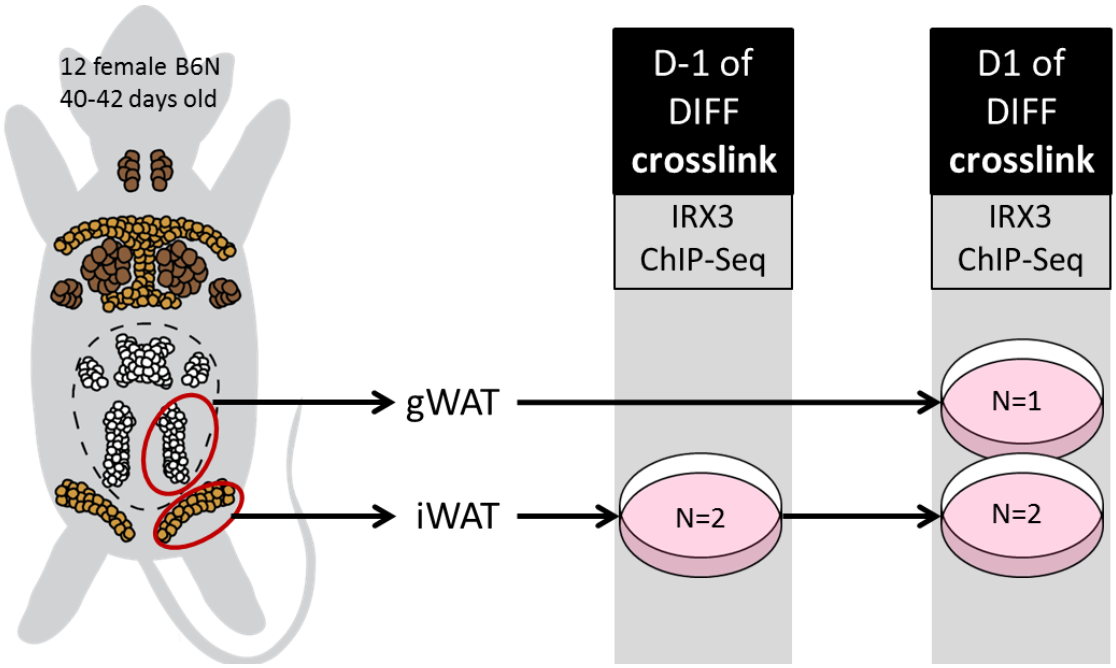
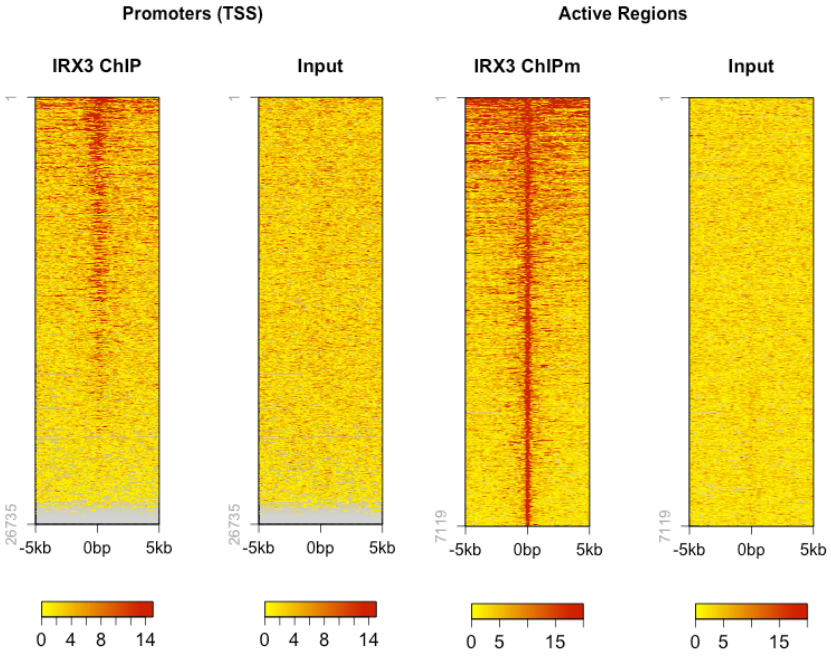


Fig. 6.9 | Experimental Setup.
ChIP-seq for IRX3 was performed on iWAT- and gWAT-derived primary pre-adipocytes isolated from B6N female mice one day before (D-1) and 24h post induction of adipogenesis (D1) using a differentiation (DIFF) cocktail containing insulin, IBMX and Dexa.

A



B

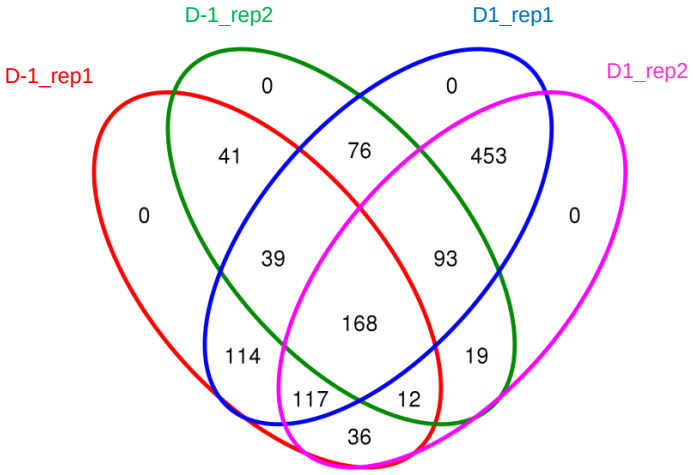


Fig. 6.10 | ChIP-seq QC.
(A) Heatmaps representing the location of peaks. Tag distributions across target regions such as Active Regions (= all peak regions), transcription start sites (TSS) or gene bodies are determined and presented as heatmaps (values in x-axis, regions in y-axis). (B) VENN diagram illustrating the overlap between peaks for all four samples (two replicates and two time-points) from iWAT ChIP-seq experiment. The diagram shows the number of Active Regions for the various categories/overlap sets.

Table 6.1 | Number of Peaks in IRX3 ChIP-seq iWAT data sets.
Reference list Mus musculus 22320, Input: IRX3 ChIP-seq iWAT D-1 and D1. Data analysed by Li Wang, Broad Institute.

	Peaks D-1	Peaks D1	Diff Peaks up_reg (FDR0.05/FDR0.01)	Diff Peaks down_reg (FDR0.05/FDR0.01)
SPP+IDR (optimal)	951	3688	108/29	70/47

Table 6.2 | Gene ontology analysis for differential peaks identified in IRX3 ChIP-seq for iWAT.
GREAT tool (McLean *et al.*, 2010) was used to assess pathways enriched for differential peaks identified at D-1 and D1. Input: IRX3 ChIP-seq iWAT D-1 and D1. Data analysed by Li Wang, Broad Institute.

Term Name	Binom Rank	Binom Raw P-Value	Binom Fold Enrichment
Strongly down-regulated at 8-96h during differentiation of 3T3-L1 cells (fibroblast) into adipocytes.	1	2.8e-8	7.5

Table 6.3 | Statistical Overrepresentation Test using IRX3 ChIP-seq data.
PANTHER (Mi *et al.*, 2017; Thomas *et al.*, 2003) was used to identify statistically overrepresented cellular components in gWAT ChIP-seq for IRX3. Reference list Mus musculus 22320. Bonferroni correction was used for multiple testing. GO cellular component complete refer to mitochondrion (GO:0005739); mitochondrial part (GO:0044429); mitochondrial matrix (GO:0005759); mitochondrial ribosome (GO:0005761).

GO cellular component complete	Fold Enrichment	+/-	P value
gWAT D1 IRX3 ChIP			
mitochondrion	1.67	+	2.2E-11
mitochondrial part	1.75	+	0.000082
mitochondrial matrix	2.21	+	0.0132
mitochondrial ribosome	3.62	+	0.00167
iWAT D1 IRX3 ChIP			
Mitochondrion	1.86	+	8.61E-12
mitochondrial part	2.04	+	3.91E-09
mitochondrial matrix	2.69	+	1.44E-08
mitochondrial ribosome	4.27	+	0.00125
iWAT D-1 IRX3 ChIP			
Mitochondrion	1.81	+	0.0024
mitochondrial part	2.04	+	0.0108

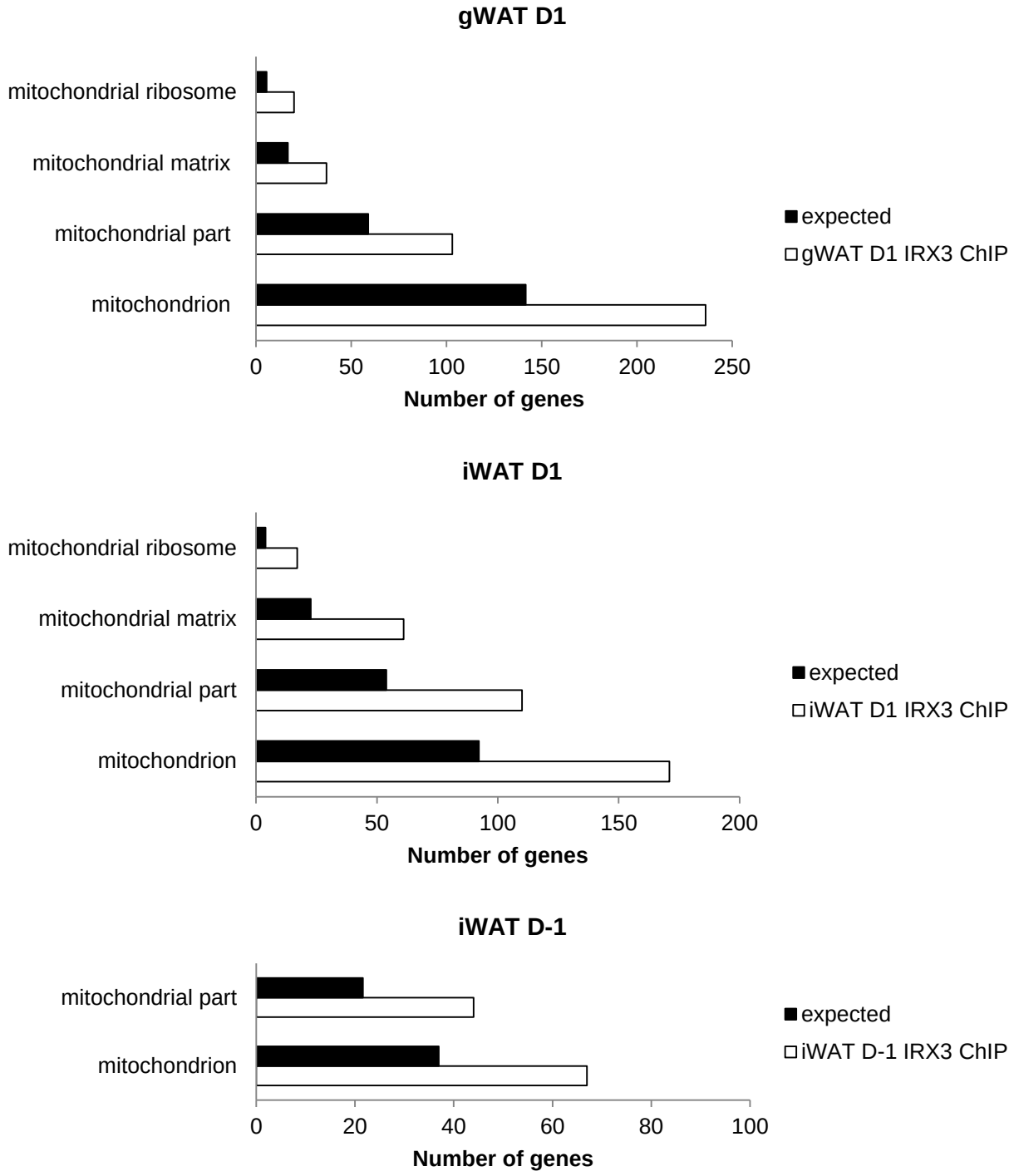


Fig. 6.11 | Statistical Overrepresentation Test.
PANTHER (Mi *et al.*, 2017; Thomas *et al.*, 2003) was used to identify statistically overrepresented cellular components in gWAT and iWAT ChIP-seq data sets for IRX3. Reference list Mus musculus 22320. Bonferroni correction was used for multiple testing.

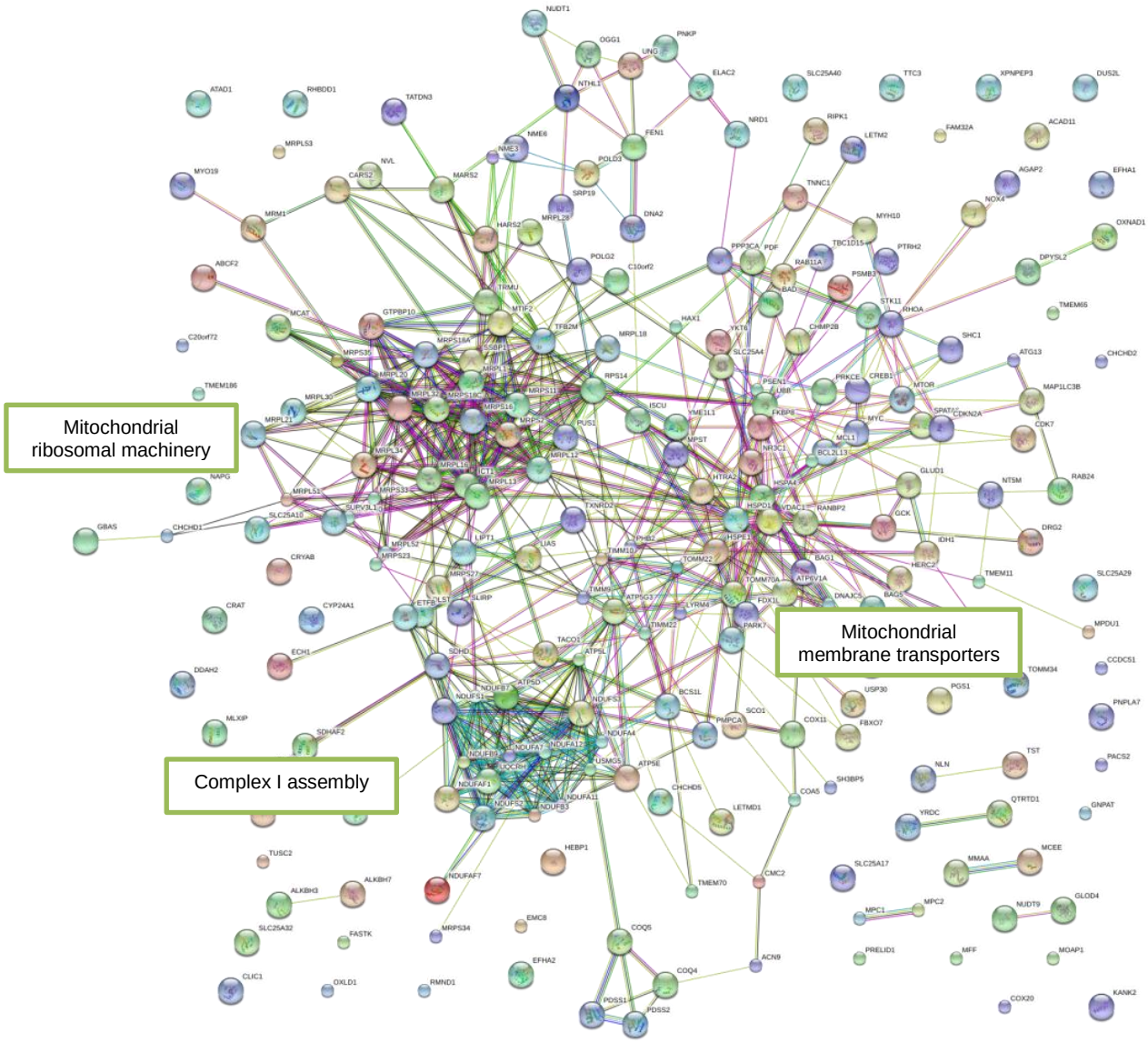
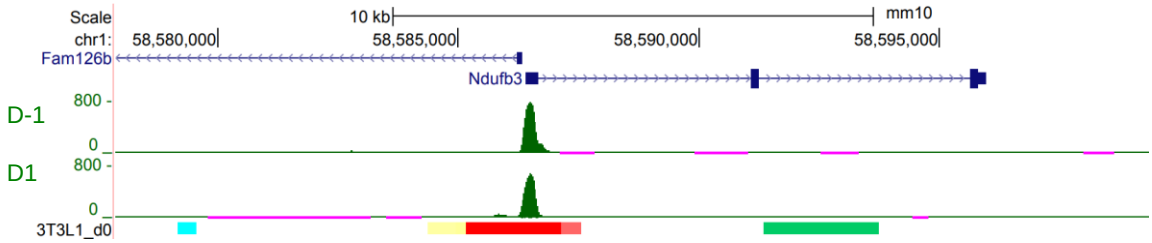
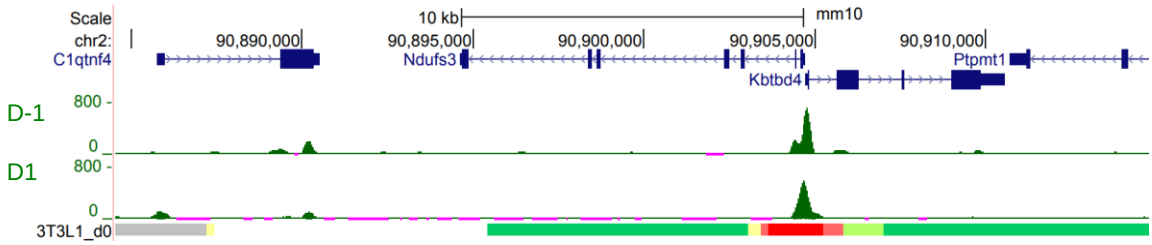


Fig. 6.12 | STRING analysis of mitochondrial genes identified by IRX3 ChIP-seq.
Reference list Homo sapiens. Input: 236 mitochondrion genes identified by PANTHER in gWAT ChIP-seq data set. STRING resource was used at http://string-db.org/ (Szklarczyk *et al.*, 2017).

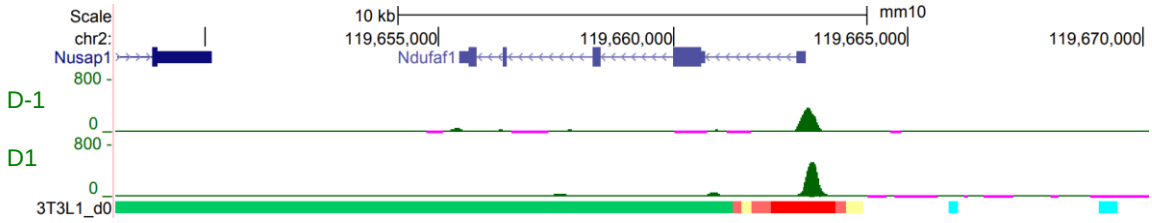
A *Ndufb3*



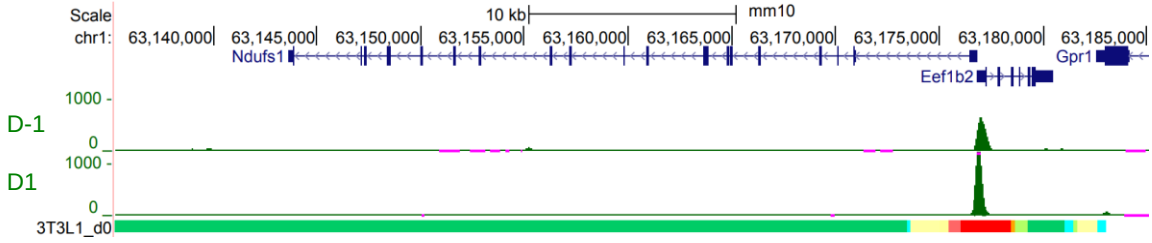
B *Ndufs3*



C *Ndufaf1*



D *Ndufs1*



E *Ndufa12*

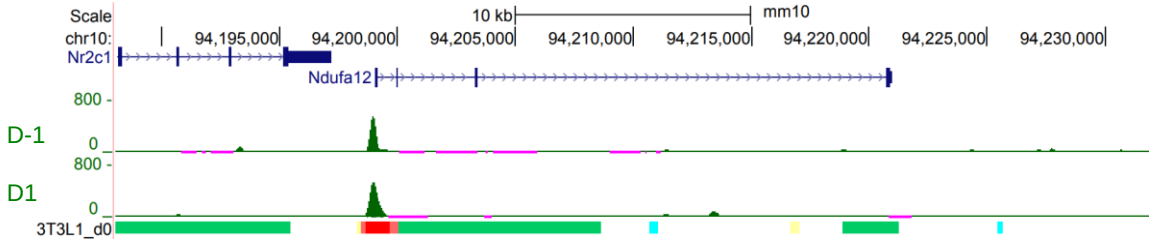


Fig. 6.13 | Representative IRX3 ChIP-seq profiles for selected OXPHOS complex 1 genes.
To assess the quality of the generated ChIP-seq data, histograms were visualized over gene promoters of identified peaks using the UCSC Genome Browser. Chromatin state annotations using ChromHMM from mouse pre-adipocytes (undifferentiated) (Mikkelsen *et al.*, 2010) are shown in the bottom tracks in each image. Tracks one and two show ChIP-seq histograms from iWAT pre-adipocytes before (D-1) and 24h after induction of adipogenesis (D1), respectively. All histograms within each browser screenshot are shown on the same scale and high values were truncated as necessary. ChIP-seq peaks overlap histone modifications for promoters (red annotation) and show characteristic peak pattern. Input controls did not show any peaks (not shown). Histograms are shown for (A) *Ndufb3* (NADH dehydrogenase (ubiquinone) 1 beta subcomplex subunit 3), (B) *Ndufs3* (NADH dehydrogenase (ubiquinone) iron-sulfur protein 3, mitochondrial), (C) *Ndufaf1* (Complex I intermediate-associated protein 30, mitochondrial), (D) *Ndufs1* (NADH-ubiquinone oxidoreductase 75 kDa subunit, mitochondrial), and (E) *Ndufa12* (NADH dehydrogenase (ubiquinone) 1 alpha subcomplex subunit 12).

6.2.4 *Irx5* and adipogenesis

To study the effect of *Irx5* on adipocyte development, 3T3-L1 pre-adipocytes were treated with silencing RNA for *Irx5* (siIrx5) and control RNA (siCON) for two days before differentiation was induced. Two days of siIrx5 treatment resulted in significant decrease in *Irx5* expression (Fig. 6.14), with 90% reduction of *Irx5* mRNA levels in siIrx5 treated 3T3-L1 compared to siCON treated cells. At day 8 of differentiation, siIrx5-treated cells accumulated fewer lipids than siCON-treated cells, as shown by decreased Oil-Red O staining (Fig. 6.14). This was consistent with significantly reduced mRNA levels of *Pparg, Cebpa* and *Fabp4* in *Irx5*-silenced 3T3-L1 (Fig. 6.14). Furthermore, siIrx5 treatment had an effect on mitochondrial gene expression, with *Pgc1a* and *Dio2* being reduced to 50% compared to control-treated cells. *Prdm16* and *Elvol3*, on the other hand, were significantly upregulated (~2.5-fold and ~3-fold, respectively) in siIrx5 3T3-L1 cells, suggesting that *Irx5* has effects that are agonistic as well as antagonistic towards a browning program in pre-adipocytes. In conclusion, 90% reduction of *Irx5* in 3T3-L1 pre-adipocytes results in markedly reduced differentiation capacity of these cells.

To confirm these findings, the *Irx5* silencing experiment was repeated in primary pre-adipocytes isolated from iWAT of male B6N mice (Fig. 6.15). Consistent with previous results, siIrx5-treated adipocytes differentiated into mature adipocytes that had reduced mRNA levels of *Pparg, Cebpa, Fabp4* and *Pgc1a* compared to cells treated with control RNA. Additionally, gene expression of the β-adrenergic receptor *Adrb3* was significantly reduced, suggesting that *Irx5* increases the responsiveness of pre-adipocytes to adrenergic simulation and consequently increases browning-capacity of those cells. In contrast to results obtained in 3T3-L1 adipocytes, *Dio2* was upregulated in siIrx5 iWAT-derived primary adipocytes, indicating that there are differences in *Irx5*-mediated processes in these two models.

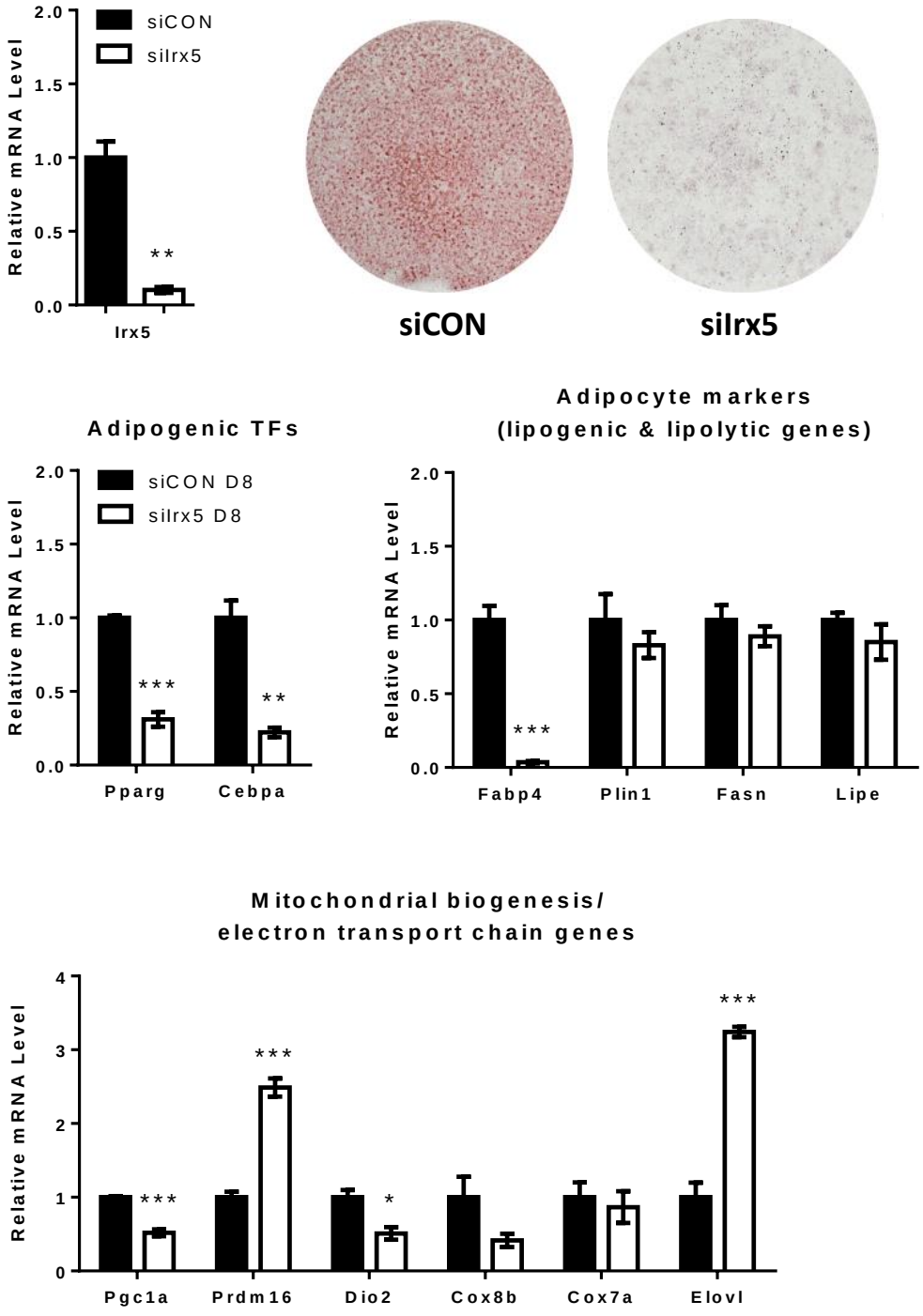


Fig. 6.14 | *Irx5* affects adipocyte development.
siIrx5 treatment results in decreased lipid accumulation at day 8 of DIFF as shown by Oil-Red O staining in 3T3-L1 (A). This is consistent with decreased mRNA expression (normalised to *Gapdh*) of adipogenic, lipogenic and lipolytic genes in 3T3-L1 (B) as well as cultured iWAT isolated from male B6N (C). Experiments were performed in technical triplicates. Statistical significance was assessed via independent Student's t-test. Data are shown as mean ± SEM with *P<0.05; **P<0.01; ***P<0.001.

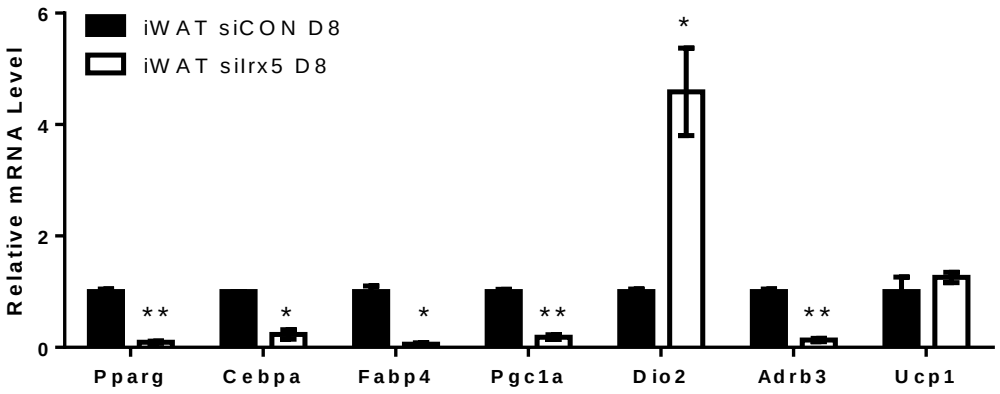


Fig. 6.15 | *Irx5* regulates adipogenic and browning markers in iWAT.
Pre-adipocytes from iWAT were isolated from mice at 6-8 weeks of age and differentiated for 8 days before RNA was extracted. Gene expression of adipogenic and mitochondrial markers was measured by the means of qPCR (normalised to *18S*). Experiments were performed on technical duplicates. Statistical significance was assessed using Student's t-test; data are shown as mean ± SEM with *P<0.05; **P<0.01.

6.3 Discussion

Very little is currently known about the adipocyte-intrinsic functions of *Irx3* and *Irx5*. To study the role of *Irx3* and *Irx5* in adipocyte development, I treated 3T3-L1 and primary adipocyte precursors isolated from iWAT of B6N with siRNA for *Irx3* and *Irx5* and differentiated these for 7-8 days. siIrx5 treatment in 3T3-L1 as well as primary pre-adipocytes from male iWAT inhibits development of mature adipocytes drastically. In contrast, siIrx3 appears to have the opposite effect, suggesting that *Irx3* and *Irx5* have opposite effect on adipogenesis when silenced in 3T3-L1. I hypothesise that although *Irx3* and *Irx5* are members of the same gene family and co-regulated, they have distinct functions. It is also interesting to hypothesise that due to differential gene expression between different WAT depots, alteration of these genes in *FTO* risk allele carriers has different effects in different adipose tissues.

Following siIrx3 treatment, adipocyte precursors develop into brown-like adipocytes, as measured by elevated expression of genes involved in mitochondrial biogenesis. This is consistent with results by Claussnitzer *et al.* (2015) who showed that decreased expression of *IRX3* and *IRX5* increased brown-like development of human adipocyte precursors. Additionally, I show here that siIrx3 treatment results in stimulation of adipogenesis as shown by increased expression of adipocyte markers 8 days post induction of differentiation in 3T3-L1.

As previously mentioned, energy production in pre-adipocytes relies mainly on glycolytic ATP production (Keuper *et al.*, 2014) and when pre-adipocytes commit do become adipocytes, mitochondrial oxidative metabolism is induced. Early upregulation of *Pgc1a* and genes involved in oxidative phosphorylation (OXPHOS) is crucial for differentiation. The role of *Irx3* in decreasing complex 1 (CI) assembly genes as well as decreasing mitochondrial membrane transporters and therefore reducing the import of nuclear-transcribed proteins into the mitochondrion will have direct impact on cellular differentiation and function. Obesity was found to be associated with the downregulation of genes that are involved in OXPHOS in white adipose tissue, with Complex 1 (CI) and Complex 4 (CIV) being the main respiratory

chain complexes affected (B. Fischer *et al.*, 2015). Comparing single respiratory chain complex components for complexes I–V showed that there is a significant reduction in the protein amounts for CI (NDUFB8) and CIV (MT-CO2) in subcutaneous WAT from female obese individuals compared to non-obese (B. Fischer *et al.*, 2015). These changes translated into reduced OXPHOS in mitochondria from adipocytes of females with higher BMI. The authors concluded that mitochondrial dysfunction in WAT plays a key role in the pathogenesis of obesity and associated comorbidities (B. Fischer *et al.*, 2015). It has also been hypothesized that obese individuals have fewer beige adipocytes and therefore are primed to gain weight on high-fat diets and that targeting WAT browning therapeutically presents an interesting avenue to pursue (reviewed in Harms & Seale, 2013).

Together, multiple lines of evidence support the idea that IRX3 acts as a transcriptional repressor in early differentiating pre-adipocytes and decreased levels of *Irx3* in pre-adipocytes (as it is the case for *FTO* protective allele carriers) causes pre-adipocytes to develop into more brown-like adipocytes with increased thermogenic capacity and increased proton leak (Fig. 7.16).

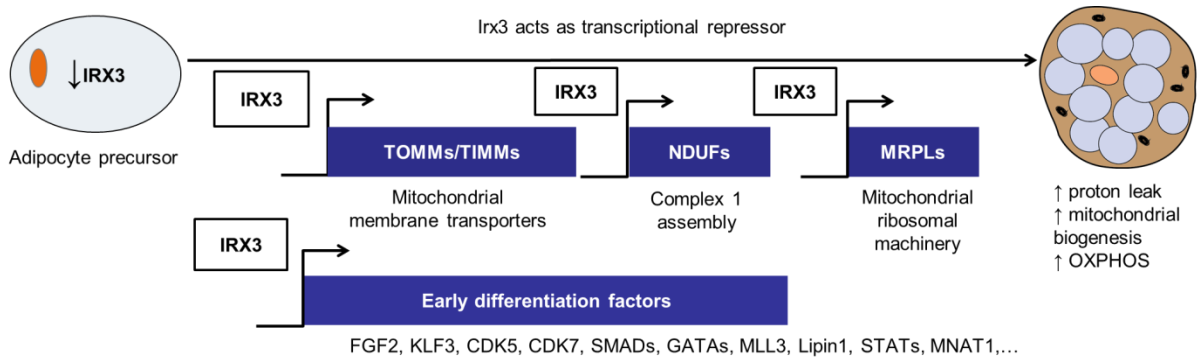


Fig. 6.16 | *Irx3* acts as a transcriptional repressor during early adipogenic differentiation. IRX3 is a transcription factor that inhibits transcription of mitochondrial membrane transporters (TOMMs/TIMMs), Complex I assembly genes (NDUFs), genes involved in the mitochondrial ribosomal machinery (MRPLs) as well as a number of early differentiation factors, including *Klf3* and *Gata* family members. Consequently, reduced *Irx3* levels in pre-adipocytes (as it is the case for *FTO* protective allele carriers) results in cells to differentiate into more brown-like adipocytes with characteristic increase in proton leak, mitochondrial biogenesis and OXPHOS.

7 Phenotype and mechanism of FTO-RS1421085

7.1 Introduction

In the past decade, human and mouse studies have pointed towards a number of functional variants and target genes in the *FTO* locus and mechanistic studies of selected target genes have proven a role for many of these genes in the development of obesity. However, the field has reported conflicting results on the number and direction of effector transcripts and even conflicting results on the function of those genes. For example, global *Fto* knockout animals have reduced fat mass (together with a severe developmental phenotype) (Church *et al.*, 2009), whereas adult onset *Fto* knockout results in increased fat mass (McMurray *et al.*, 2013), and more recently, global hypomorphism of *Fto* has been linked to hyperphagic obesity (Stratigopoulos *et al.*, 2016). These reports emphasize the need to study animal mutants that have target genes expression changes at the right level, the right direction and at the right developmental stage.

Additionally, there is currently evidence that the CC risk genotype for rs1421085 results in increased expression of *IRX3* and *IRX5* (Claussnitzer *et al.*, 2015) as well as reduced expression of *FTO* and *RPGRIP1L* (Stratigopoulos *et al.*, 2016). Both studies further characterise the cellular and organismal consequences of altered gene expression of respective genes and link them to an obesity phenotype.

Together those findings suggest that the BMI-association at the *FTO* locus is founded on a complex regulatory network that very likely cannot be explained by one tissue or target gene in isolation, but needs to be studied as a systemic regulatory circuitry that will affect several organs and several mechanism at several developmental stages that, together, are responsible for the large effect size of the most prominent obesity locus. The creation of knock-in mouse models with high-risk *FTO* alleles should facilitate the determination of the contribution of these gene variants to obesity (assuming conservation – of course).

At the mouse orthologous location of rs1421085, WT mice of strains B6N, C57BL/6J (B6J) and C3H/HeH (C3H) harbour the TT genotype, that is associated with non-risk in human. To assess whether a single nucleotide change to the human CC risk genotype has an effect on target gene expression, body weight and composition, we utilised CRISPR/Cas9-mediated genome editing to create a mouse model that harbours the T-to-C conversion on a B6N background, termed FTO-RS1421085. In this Chapter I aimed to evaluate the degree of functional conservation at this position on target gene expression changes between human and mouse and to decipher the consequences of editing rs1421085 on body composition in mice.

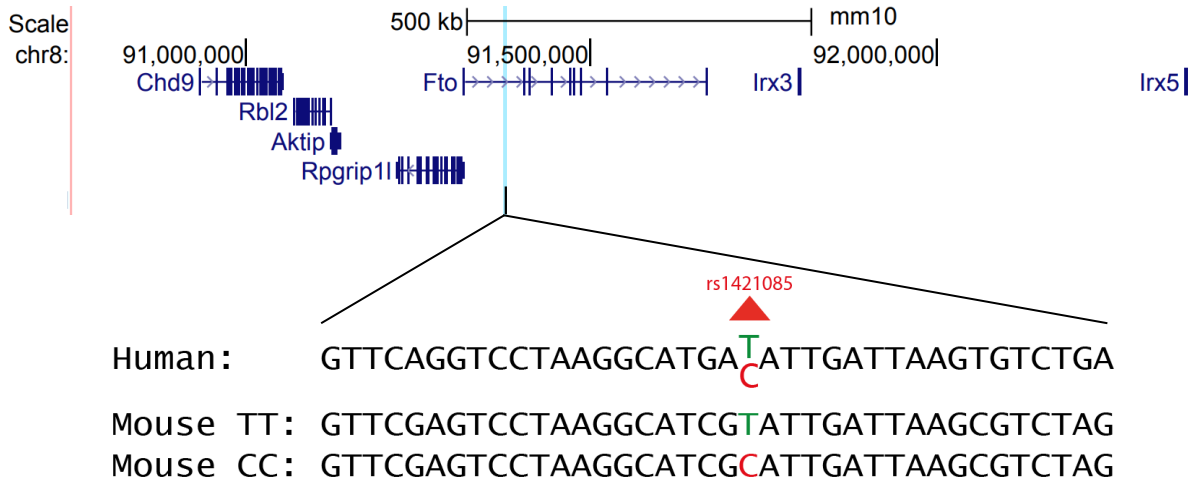
7.2 Results

7.2.1 Target gene expression in FTO-RS1421085

First, I set out to assess whether a single nucleotide change from T-to-C (mimicking the human C risk allele for rs1421085; Fig. 7.1A) has a comparable effect on *Irx3* and *Irx5* expression that was observed in human risk allele carriers (Claussnitzer *et al.*, 2015), where pre-adipocytes with the risk allele isolated from female subcutaneous WAT showed ~2-fold increase in *IRX3* and *IRX5* expression at day 2 of adipogenic differentiation.

Primary pre-adipocytes isolated from iWAT and gWAT depots of male and female mice were isolated, cultured and differentiated for two days using an adipogenic cocktail before RNA was isolated and gene expression measured using qPCR. Pre-adipocytes from female iWAT harbouring the CC genotype had significantly increased levels of *Irx3* and *Irx5* compared to the WT TT genotype (Fig. 7.1B). Specifically, *Irx3* and *Irx5* mRNA levels were 2-fold (P=0.025) and 1.7-fold (P=0.03) increased, respectively, in iWAT of CC carriers compared to TT. This effect size is comparable to *IRX3* and *IRX5* upregulation in risk allele carriers of human subcutaneous-derived pre-adipocytes that has been reported in humans (Claussnitzer *et al.*, 2015).

A



B

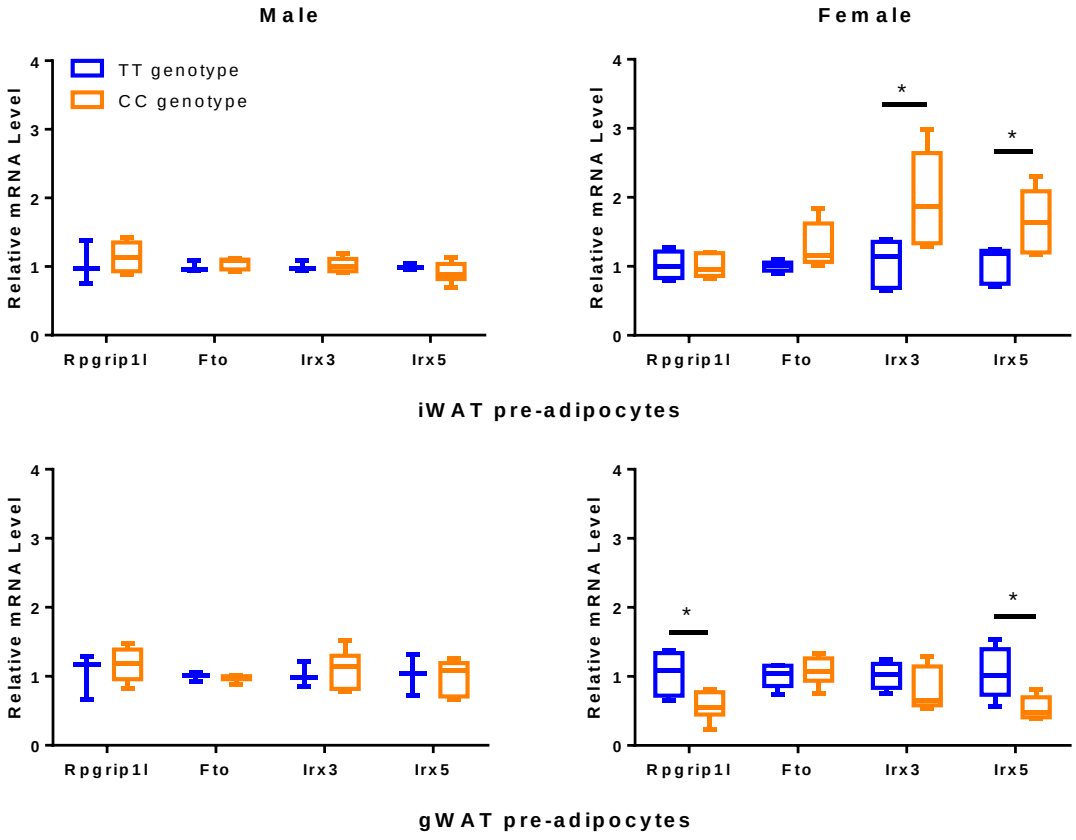


Fig. 7.1 | T-to-C editing at the orthologue site of human rs1421085 in mouse alters gene expression of target transcripts in females.
(A) Schematic illustrating human and mouse sequence conservation at rs1421085 and risk (red) and non-risk (black). (B) Primary pre-adipocytes were isolated from male (left) with TT genotype (n=3) and CC genotype (n=7) and from female (right) with TT genotype (n=5) and CC genotype (n=6) from both iWAT (top) and gWAT (bottom) depots. Cells were treated with adipogenic cocktail for two days before RNA was isolated. Gene expression changes in T-to-C conversion of putative target genes were measured using qPCR (normalised to *Canx*). Horizontal line within each box represents the median, the top and bottom of each box represent the interquartile range and the bars indicate the range. Statistical significance was assessed using Student's t-test; *P<0.05.

Unexpectedly, in gWAT-derived pre-adipocytes of females carrying the CC genotype, *Rpgrip1l* and *Irx5* expression were downregulated 0.52-fold (P=0.016) and 0.53-fold (P=0.021), respectively (Fig. 7.1). There is currently no equivalent data in humans that could be used for comparison. In male animals, the CC genotype had no detectable effect on expression of candidate target genes *Rpgrip1l, Fto, Irx3,* and *Irx5* in iWAT and gWAT pre-adipocytes (Fig. 7.1B).

To further establish the temporal context dependency, target gene expression was additionally measured in at day 0 and day 8 of differentiation in female-derived pre-adipocytes. For comparative visualization, day 2 data that was previously shown in Fig. 7.1 was also included in Fig. 7.2. In gWAT-derived primary pre-adipocytes from CC carriers, *Rpgrip1l* and *Irx5* were downregulated specifically at day 2 of differentiation (Fig. 7.2B,H), whereas *Fto* was downregulated specifically in terminal differentiated cells (P=0.012) at day 8 (Fig. 7.2D). In iWAT-derived pre-adipocytes from CC carriers, *Irx3* was specifically upregulated at day 2 of adipogenesis (Fig. 7.2E), whereas *Irx5* was upregulated at day 2 (Fig. 7.2G). Interestingly, none of the measured genes was significantly different between TT and CC carriers in undifferentiated cells at day 0, although trends for decreased *Irx3* and *Irx5* expression in gWAT were obvious. This suggests that regulation of target genes was initiated by induction of the adipogenic program.

In summary, T-to-C editing in mouse to mimic the human risk variant at rs1421085 results in alteration of four target genes, which show high dependency for developmental time-point, sex and cell type. These data suggest that functional conservation at this location is indeed at the nucleotide level in addition to being highly conserved at the module level as previously described (Claussnitzer *et al.*, 2014). Moreover, using this mouse model I also observe a previously undescribed sex- and depot-specific effect of this variant, a finding that requires validation in human.

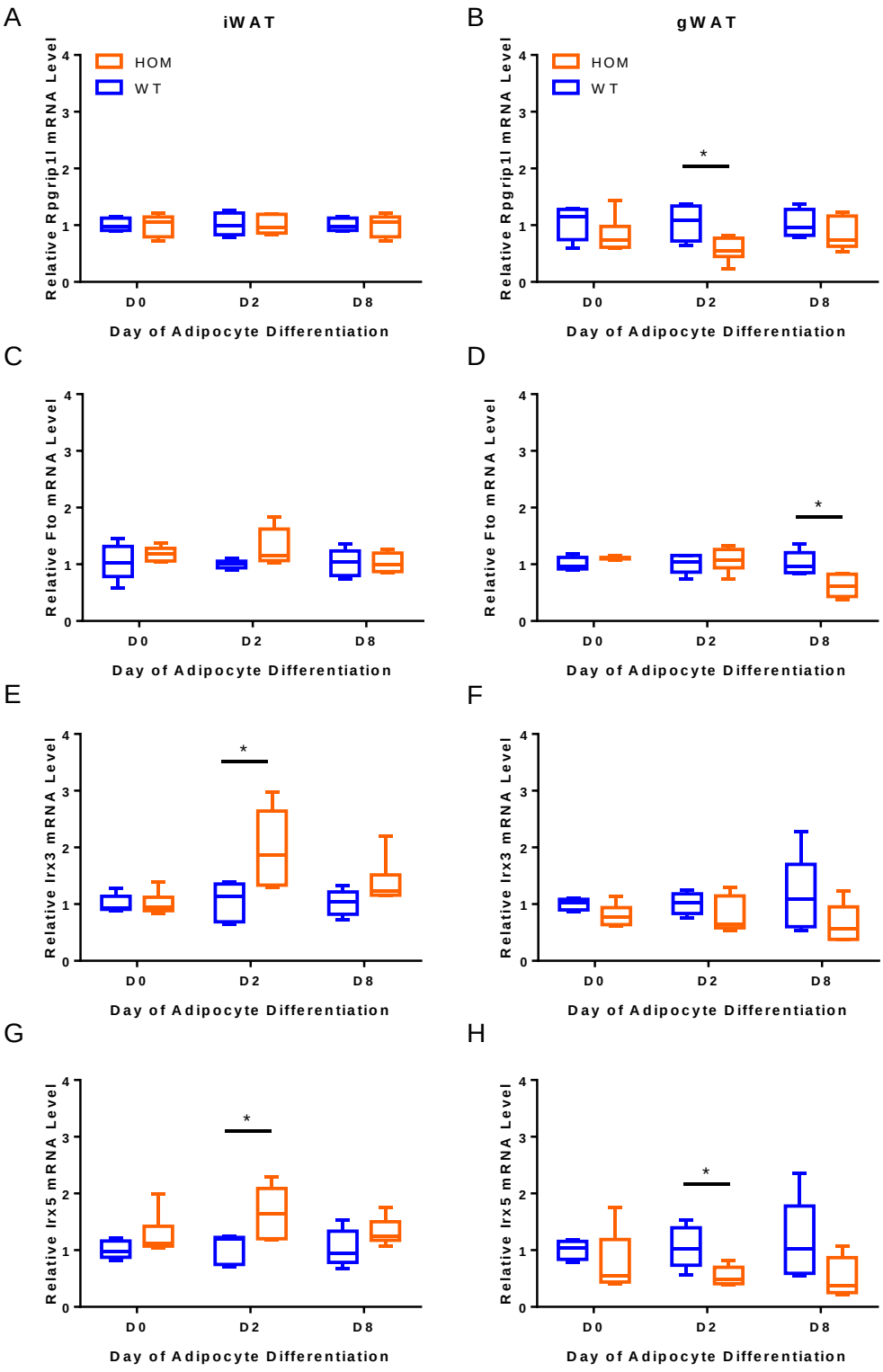


Fig. 7.2 | T-to-C conversion in female mice results in a depot-specific regulatory disruption of *Rpgrip1l, Fto, Irx3* and *Irx5* in pre-adipocytes during differentiation.
Graph shows mRNA expression for *Rpgrip1l* (A,B), *Fto* (C,D), *Irx3* (E,F) and *Irx5* (G,H) in adipocyte progenitors isolated from iWAT (left) and gWAT (right) of female FTO-RS1421085 animals harbouring the TT genotype (n=4-5) and CC genotype (n=5-6). Pre-adipocytes were isolated from 6-8 week old mice and differentiated for eight days. RNA was isolated before (D0) two days post adipogenic induction (D2) and in terminally differentiated adipocytes (D8). Gene expression was

measured using qPCR (normalized to *Canx*). Data for D2 has previously been used in Fig.7.1. Horizontal line within each box represents the median, the top and bottom of each box represent the interquartile range and the bars indicate the range. Statistical significance was assessed using Student's t-test; *P<0.05

7.2.2 Identification of conserved TFBSs affected by rs1421085

To identify the upstream regulator of rs1421085 in mouse, I next examined disrupted motifs in conserved TFBSs that overlap rs1421085. The T-to-C substitution at rs1421085 disrupts conserved motifs for a number of regulators, with highest scoring ones being ARID/ ARID5A, ESRR/ ESRRA and HBOX/ VAX (Table 7.1). Motifs for each of the predicted TFBSs affected by the SNP can be found in the appendix (Appendix Table S.3). Given the sex-specific nature of the risk allele in mouse, estrogen related receptor alpha (ESRRA) is particularly interesting. However, *Esrra* gene expression does not differ between female and male pre-adipocytes in unstimulated conditions (Appendix Fig. S.4). It will be necessary in the future to experimentally validate the effect of this variant on TF binding by e.g. using electrophoresis-mobility-shift-assays (EMSAs). Also, there is currently no sex-difference published for *FTO* risk variants in humans.

Table 7.1 | Rs1421085 alters a number of conserved TFBSs.

Seq. name	SNP pos.	allele
rs1421085	501	C → T

lost/ new	Family/ matrix	Further Information	Opt. thresh.	Start pos.	End pos.	Strand	Core sim.	Matrix sim.
lost	RXRF/ THRB.02	Thyroid hormone receptor, beta (ER5 - everted repeat, spacer 5)	0.75	482	506	-	0.81	0.751
	STEM/ OCT3_4.01	POU domain, class 5, transcription factor 1	0.81	488	506	+	1	0.813
	ESRR/ ESRRA.01	Estrogen-related receptor alpha	0.87	494	516	-	0.77	0.877
	HBOX/ VAX2.01	Ventral anterior homeobox 2	0.85	498	516	+	1	0.856
new	ARID/ ARID5A.01	AT rich interactive domain 5A (MRF1-like)	0.87	495	515	+	1	0.884
	HNF1/ HNF1.01	Hepatic nuclear factor 1	0.8	496	512	-	0.767	0.806
	LEFF/ LEF1.04	TCF/LEF-1 (secondary DNA binding preference)	0.84	496	512	-	1	0.841
	BRNF/ BRN3.03	POU class 4 homeobox 3 (POU4F3), BRN3C	0.83	497	515	+	0.753	0.842
	HNF1/ HNF1.02	Hepatic nuclear factor 1	0.77	498	514	+	0.785	0.79

SNPInspector analysis was run on rs1421085 in the centre of a 1001 bp sequence. TFBS were analysed in a matrix for vertebrates. The optimized matrix threshold and matrix similarity refer to the optimized value defined in a way that a minimum number of matches is found in non-regulatory test sequences (i.e. with this matrix similarity the number of false positive matches is minimized). The "core sequence" of a matrix is defined as the (usually 4) highest conserved positions of the matrix. The maximum core similarity of 1.0 is only reached when the highest conserved bases of a matrix match exactly in the sequence. Only matches that contain the "core sequence" of the matrix with a score higher than the core similarity are listed in the output. Matrix similarity - A perfect match to the matrix gets a score of 1.00 (each sequence position corresponds to the highest conserved nucleotide at that position in the matrix), a "good" match to the matrix usually has a similarity of > 0.80. Mismatches in highly conserved positions of the matrix decrease the matrix similarity more than mismatches in less conserved regions.

7.2.3 RS1421085 body composition

Next, to assess whether T-to-C conversion at rs1421085 is sufficient in having an effect on the organismal level, TT and CC carriers were put on a HFD at weaning and body composition was measured every two weeks using Echo-MRI (performed by Liz Bentley). My results show that at 14-weeks of age, males with the CC genotype have significantly reduced fat mass and a trend for overall decreased body weight compared to TT animals (Fig. 7.3). Males harbouring the TC genotype show an intermediate phenotype on body weight and fat mass at 14-weeks of age and have significantly reduced fat mass at 12-weeks of age (Fig. 7.3). This is surprising, given that in humans the CC risk genotype is associated with increased BMI. No body composition phenotype was observed in female mice (Fig. 8.3).

7.2.4 DEL10 around the rs1421085 risk variant in mouse

As a second off-target side-product of our CRISPR/Cas9 guide, I gained a mouse mutant that carries a deletion of 10 nucleotides, just downstream of the nucleotide of interest at rs1421085 in mouse (Fig. 7.4A). Homozygous animals for this allele have been termed DEL10. I hypothesized that this mutation has no effect on any target genes in primary pre-adipocytes. As previously, cells were isolated from 6-week old mice before differentiation was induced for 2 days. Gene expression of putative target genes was assessed using qPCR.

DEL10 was not sufficient to alter *Rpgrip1l, Fto, Irx3* or *Irx5* expression in pre-adipocytes isolated from male and female iWAT and gWAT (Fig. 7.4B). These results suggest that the effect of DEL82 and FTO-RS1421085 on *Rpgrip1l, Fto, Irx3* and *Irx5* expression is independent on the upstream genomic context of rs1421085 in mouse.

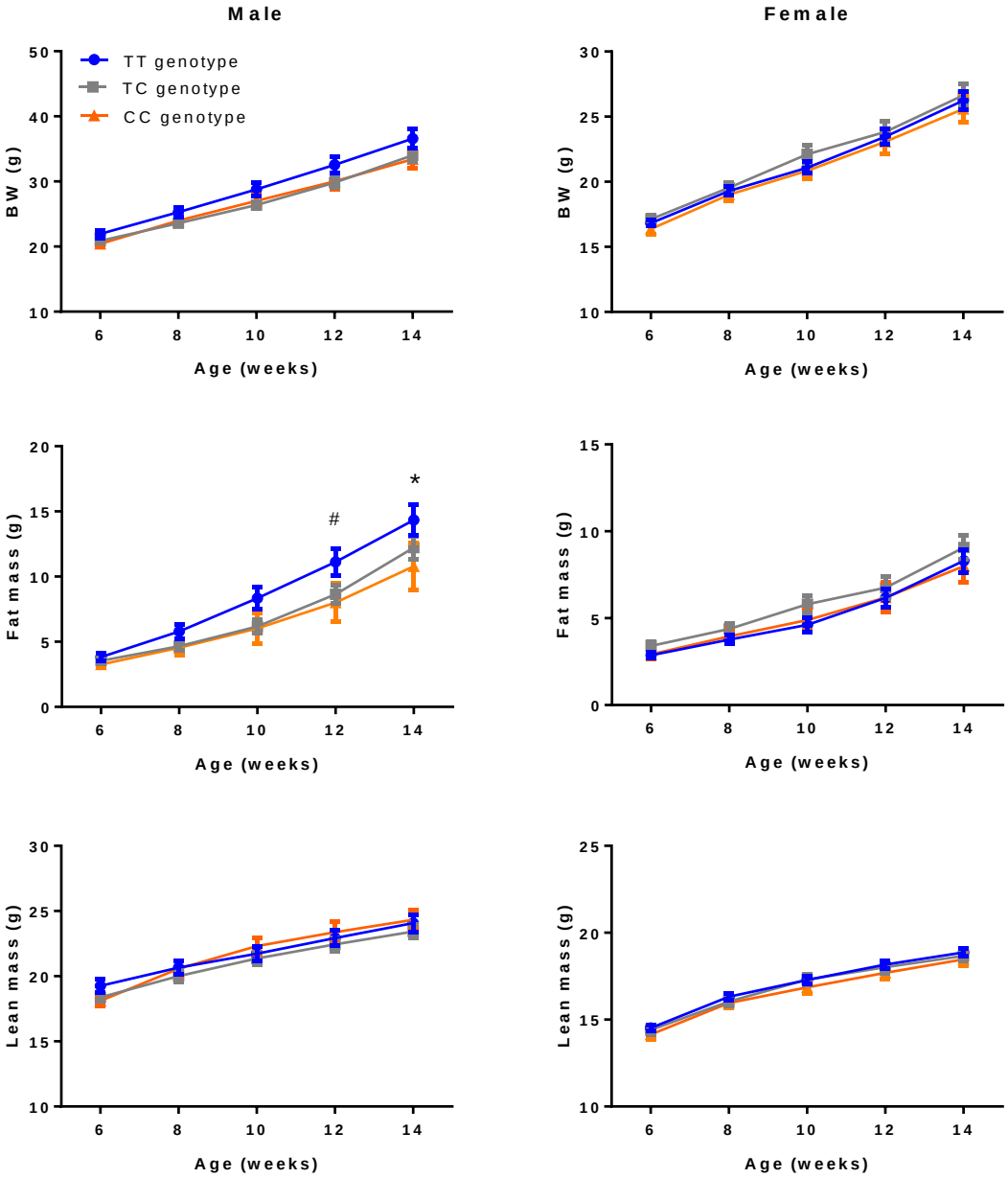


Fig. 7.3 | FTO-RS1421085 T-to-C causes reduced body weight and fat mass in males on a HFD. Male (left) and female (right) mice with TT (WT), TC and CC genotypes were fed a high-fat diet (HFD) and body composition assessed using Echo-MRI. (A) Body weight, (B) Fat mass and (C) Lean mass in male and female FTO-RS1421085 mice. N numbers for each group are: male TT (n=16), TC (n=23), CC (n=8); female TT (n=22), TC (n=18), CC (n=13). Statistical significance was determined using repeated measures two-way ANOVA, with Bonferroni's multiple comparisons test. Data are expressed as mean ± SEM with $^{\#}P<0.05$ between TT and TC genotypes and *P<0.05 between TT and CC genotypes.

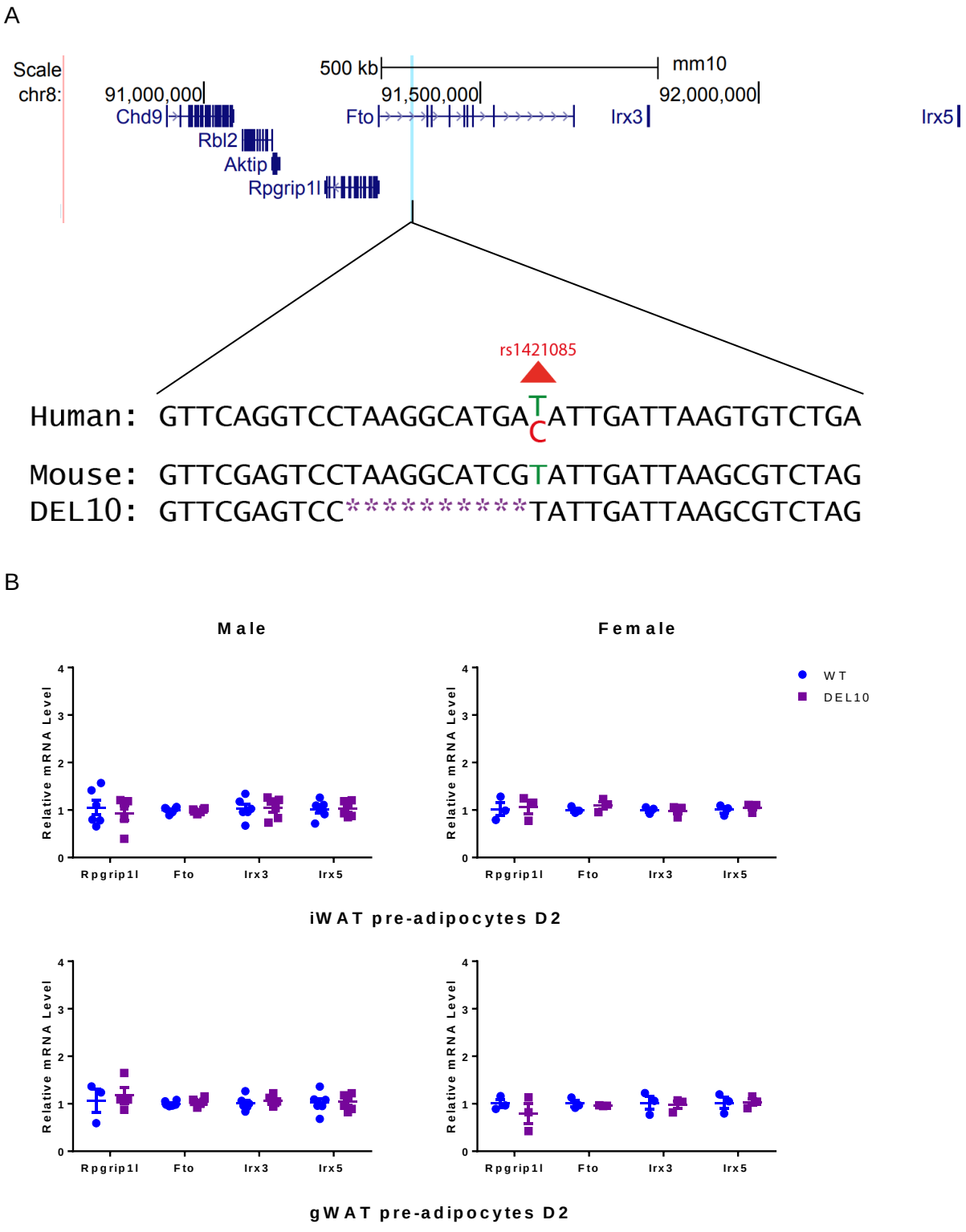


Fig. 7.4 | DEL10 does not alter gene expression of putative target transcripts.
(A) Schematic illustrating mouse sequence at rs1421085 and its motif (blue) and location of 10 nucleotide deletion (purple). (B) Primary pre-adipocytes were isolated from male WT (n=3-6) and DEL10 HOM (n=4-6) from both iWAT (left) and gWAT (right) depots. (C) Female primary pre-adipocytes were isolated from WT (n=3) and DEL10 HOM (n=3) from both iWAT (left) and gWAT (right) depots. Cells were treated with adipogenic cocktail for two days before RNA was isolated. Gene expression of putative target genes was measured using qPCR (normalised to *Canx*). Data are expressed as individual data points including a horizontal line for mean ± SEM.

7.3 Discussion

Given the high conservation scores between human and mouse at the FTO locus, I hypothesized that a single-nucleotide change from T-to-C in mice at rs1421085 that mimics the human BMI risk allele would result in similar target gene manipulation in both species. My results show a striking similarity to the direction, effect size as well as spatial-temporal effect of target gene alterations observed in human risk allele carriers. Available data in human subcutaneous pre-adipocytes from female risk allele carriers shows selective upregulation of *IRX3* and *IRX5* by a factor of approximately 2 and 2.5, respectively (Claussnitzer *et al.*, 2015). My data in corresponding mouse pre-adipocytes (matched for sex and depot) revealed genotype-associated expression changes that suggest similar long-range (1.2-Mb) genetic control between human and mouse, with approximately 2-fold increase in *Irx3* and *Irx5* expression. It would be interesting for future studies to address the question whether there is a dose-dependent effect of the risk allele on target gene expression. Furthermore, my results point towards a previously unknown mechanism of the mouse orthologous region around rs1421085, namely that this element acts as a repressive sequence in gWAT, rather than an enhancer, a finding that requires validation in human. In essence, my data suggests that rs1421085 is within an element that regulates at least 4 target genes (*Rpgrip1l, Fto, Irx3* and *Irx5*) in pre-adipocytes. It is not clear, however, whether all mentioned genes are direct targets of the enhancer around rs1421085 or whether some genes are altered through downstream *trans*-regulatory pathways.

Although global knock-out and overexpression models have clearly highlighted a role for *Fto* in the development of adipocytes, where decreased *Fto* specifically during early adipogenesis results in decreased lipid uptake and fat mass (Merkestein *et al.*, 2015; Tews, Fischer-Posovszky, & Wabitsch, 2011; Zhang *et al.*, 2015; Zhao *et al.*, 2014), my data shows that pre-adipocytes of mice harbouring the human risk allele have 50% reduced *Fto* expression in differentiated gWAT pre-adipocytes. My previous work (Merkestein *et al.*, 2015) revealed that altering *Fto* two days after induction of differentiation has no effect on

adipogenic differentiation, indicating that the stimulatory effect of increased *Fto* on adipogenesis is restricted to early differentiation. Given that *Fto* is not altered in the CC risk genotype of early differentiating pre-adipocytes, *Fto*-mediated adipogenesis is unlikely the mechanism that explains the BMI-association at this locus. Interestingly, although global FTO-KO in mice has a severe developmental phenotype including decreased fat mass (Church *et al.*, 2009; Hess & Brüning, 2014), adult-onset FTO-KO shows the reverse phenotype, including increased fat mass and specifically increased gWAT weight and adipocyte size (McMurray *et al.*, 2013). Therefore I hypothesise that reduced *Fto* expression in mature gonadal adipocytes could result in increased gWAT mass and adipocyte size. Future studies will aim to test this hypothesis.

Given the sex-specific effect observed in FTO-RS1421085 on target gene expression, ESRRA appears to be particularly interesting as a potential upstream regulator. ESRRA is a transcriptional regulator that binds as a monomer to the consensus extended half-site TnAAGGTCA (Badis *et al.*, 2009). Not much is known about ESRRA in adipose tissue. It has been implicated in adipogenesis and adipocyte browning (Akter *et al.*, 2008; Cheung *et al.*, 2013; Delhon *et al.*, 2009; Gantner *et al.*, 2016; Ju *et al.*, 2009) and shows organ and sex-specific expression pattern (in contrast to ubiquitous expression of estrogen receptors). More experiments are necessary to validate which regulators are functional in FTO-RS1421085 mouse pre-adipocytes and to compare these to human.

Two obvious questions arising from this work are why is there no phenotype in the female animals (despite target gene alterations that match human data) and why does the obesity-risk allele results in reduced fat mass gain in male animals (despite no target gene manipulation in adipocytes)? Based on my data and the high regulatory conservation between human and mouse, it can be speculated that rs1421085 is not in isolation responsible for the BMI-signal and that additional SNPs in the risk haplotype are functional as well as causal, which may explain the large effect size of *FTO*. Additionally, adipocytes are most likely not the only effector tissue and rs1421085 must be functional in other tissues.

The development of obesity has been described as a consequence of a positive energy balance due to overconsumption of food and decreased energy expenditure in form of physical activity (Spiegelman & Flier, 2001). It is well established that energy homeostasis is regulated by neuronal signaling within the hypothalamus (Sternson & Atasoy, 2014), specifically via the paraventricular nucleus and the leptin-melanocortin pathway involving neuropeptide Y/agouti-related protein (NPY/AgRP) and pro-opiomelanocortin (POMC) circuits (Yazdi *et al.*, 2015). Dysregulation of this signaling and consequently the imbalance between food intake and energy expenditure leads to obesity. Future studies are currently underway to assess (i) potential effector transcript expression in other tissues (mainly relevant areas in the brain), (ii) the nature of the sex-specific effect, and (iii) whether there is an effect of rs1421085 risk alleles on effector transcript expression in pre-adipocytes of male human risk allele carriers. On the organismal level, FTO-RS1421085 animals will further undergo a metabolic phenotyping pipeline which will include CLAMS and IPGTT, measurements of oxygen consumption, food intake and glucose tolerance.

In summary, a single-nucleotide change in B6N at rs1421085 from T-to-C recapitulates available human data (Claussnitzer *et al.*, 2015) with striking similarity in target gene expression changes and the effect size of those genes. But I also unravel additional mechanisms, such as a tightly orchestrated regulation of several target genes in pre-adipocytes by a variant in the *Fto* locus and an element that acts simultaneously as an enhancer and repressor, depending on cell type, all of which requires validation in human.

8 Discussion and Future Perspective

Obesity association at the *FTO* locus is unusual in some ways; risk variants have a high frequency (44%) and a relatively high effect size (~1.5kg per risk allele), suggesting positive selection of the risk haplotype. Human and animal studies have pointed towards several functional variants, candidate effector transcripts (based on both proximity to the association signal and functional candidates) as well as plausible mechanisms in several tissues through which disease risk could be inferred at this locus. Nevertheless, the vast growing field of *FTO* locus dissection has also yielded a large amount of conflicting and inconclusive findings, leaving obesity risk at this locus still to be fully explained.

There is certainly a need for new strategies to decipher human risk-variants associated with diseases. Furthermore, there is currently a lack of evidence in the literature regarding the utility of mouse models in the functional dissection of *cis*-regulatory landscapes, particularly for human disease-risk associations. I have tried to address these gaps in knowledge creating mouse mutants that manipulate risk sequences at the *FTO* locus. The questions to be asked of the phenotype of such a mouse would be: Can we recapitulate risk circuitry discovered in human? Can we determine its contribution to obesity? The work presented in this thesis elucidates a potential mechanism for a regulatory element that overlaps a human risk variant for obesity and further pinpoints towards a depot- and sex-specific effect in mice. The results presented allow us to draw conclusions on *cis*-regulatory conservation between human and mouse at human risk sites and to judge feasibility of using mouse models in non-coding variant risk dissection.

One of the aims of this thesis was to create a mouse that harbours the human risk variant rs1421085 and to use this model to functionally dissect the mechanism of this regulatory element. Using CRISPR/Cas9-mediated genome editing, I was able to create the desired point-mutation at the mouse orthologous region at human rs1421085, but also generated off-

target deletions around this site, namely a 82 nucleotide deletion at the desired motif that stretches mainly downstream of the SNP and a 10 nucleotide deletion that is located upstream of the motif of interest, termed DEL82 and DEL10, respectively. The primary objectives of these models were to determine target genes, and affected tissues and consequently to dissect the mechanism of the human disease-associated sequence variant.

8.1 Comparing *FTO/Fto* regulation in human and mouse

A key question of this thesis was: to what extent do the cellular and organismal consequences of DEL82 and FTO-RS1421085 resemble the human obesity association at the regulatory, cellular and organismal level? This chapter aims to summarize and discuss the key findings of this thesis and compares them to relevant published findings in human. Table 8.1 summarizes key findings in human and compares them to data generated in the mouse lines DEL82 and FTO-RS1421085.

8.2 Target genes

Based on human findings, I hypothesized that removing the enhancer element for *Irx3* and *Irx5* in mouse would result in decreased expression of those genes in the relevant cell type (assuming that we are not creating a new and active TFBS that would alter target gene expression). As expected, I found that in mice, removing this regulatory element causes a reduction of *Irx3* and *Irx5* in pre-adipocytes of subcutaneous origin (iWAT). Unexpectedly, however, this was not the case in pre-adipocytes from gWAT. In gWAT-derived pre-adipocytes, *Irx3* expression was upregulated while *Irx5* was unchanged.

Given the high epigenomic conservation at the *FTO* locus, I further hypothesized that creating a mouse that harbours the human risk variant at rs1421085 would show comparable target gene regulation as observed in human. On a global scale, the *cis*-regulatory landscape is the major site of cross-species divergence (Denas *et al.*, 2015). My findings, however, indicate that the regulatory circuitry at the *FTO* locus is conserved at the nucleotide level between human and mouse. In both species, T-to-C conversion results in

the same target gene manipulation in female subcutaneous pre-adipocytes, with remarkable similarities in direction and effect size. We discovered an additional mechanism, where SNP-knockin in mouse results in decreased levels of *Rpgrip1l, Irx5* and *Fto* in gWAT-derived pre-adipocytes. Consistent in both mouse models (DEL82 and FTO-RS1421085), the element around rs1421085 acts as a transcriptional repressor in visceral gWAT, rather than an enhancer. It is unclear, however, if all target gene alterations are caused directly through regulatory perturbations or via secondary downstream *trans*-regulatory mechanisms. In the case of *Irx3* and *Irx5*, my FISH data certainly supports the notion of direct interactions between *Fto* regulatory regions and those genes. Given the possibility that all of the effector transcript changes observed in FTO-RS1421085 are mediated via direct interactions of *Fto* intronic regulatory elements and target gene promoters (that span over kilobases away), it is intriguing to postulate that regions that are associated through TADs will likely interact at some point in some contexts and that risk variants will more often than not affect multiple genes within these domains.

Decreased expression of *Rpgrip1l* and *Fto* for the rs1421085 risk allele in mouse is in line with findings by Stratigopoulos *et al.* (2016), who showed that the rs1421085 risk allele decreases *RPGRIP1L* and *FTO* expression in human iPSC-derived neurons (Stratigopoulos *et al.*, 2016). It will be necessary in the future to measure possible target genes in additional tissues (different adipose depots, brain and other organs) and from mice at different ages as well as following different environmental stimuli, to complete the picture of regulatory complexity at this locus. More experiments are also necessary to investigate which target genes are altered by direct binding events between regulators within *FTO* and gene promoters.

8.3 Target developmental stage

Gene regulation is often restricted to specific developmental windows. In adipose tissue, transcriptional regulation is tightly orchestrated during adipogenesis and just after pre-adipocytes are triggered to commit and develop into mature cells, the chromatin remodels

and thousands of regulatory sites become accessible for TF binding (Siersbaek *et al.*, 2012; Siersbæk & Mandrup, 2011). Using FISH revealed that the *FTO* obesity-associated interval orthologue in mouse co-localizes with *IRX3* in pre-adipocytes from iWAT, however, this effect was only seen once differentiation has been triggered. For gWAT-derived pre-adipocytes, DEL82-mediated disruption of chromatin contact between *Fto* regulatory elements and *Irx3/Irx5* was restricted to undifferentiated adipocyte precursors. Gene expression data and ATAC-seq profiles during adipocyte development in human and mouse cells has further confirmed this observation and together, these data sets support the idea that rs1421085 is within a conserved regulatory element where activity is strongest in early differentiating pre-adipocytes. This is consistent with human studies, where the *FTO* risk allele is linked to an effect specifically in early developing adipocytes (Claussnitzer *et al.*, 2015; Landgraf *et al.,* 2016). Target gene expression changes peak in significance once differentiation is triggered and are subsequently lost in mature adipose tissue in both human (Claussnitzer *et al.*, 2015) and mouse.

Interestingly, although *Rpgrip1l, Irx3* and *Irx5* were clearly dysregulated at early time-points in FTO-RS1421085 risk allele mice, *Fto* expression was downregulated in differentiated adipocytes. This is somewhat surprising given the lack of an eQTL in human visceral adipose tissue for *FTO* risk variants. I could speculate that this effect might be specific to species, depot, developmental stage, age at isolation and/or due to sample size and power. There are clearly more experiments needed in the future to unravel this further.

8.4 Target tissue/cell type

Using gene expression analysis has clearly identified pre-adipocytes as a target cell type for the mouse orthologous region around rs1421085. I have also shown that DEL82 does not affect target genes in the hypothalamus. To make the claim of an adipocyte-specific effect of FTO-RS1421085 and DEL82, however, it would be necessary in the future to expand target gene testing to additional tissues, including BAT, pancreas, muscle, heart, liver and additional regions of the brain. Additionally, the lack of altered target genes in pre-adipocytes

of male RS1421085 animals in the presence of a body weight phenotype in these mice suggests that this variant has additional target tissues, with functional effects on body weight.

8.5 Regulator

Claussnitzer *et al.* (2015) have previously identified ARID5B as an upstream regulator of rs1421085-mediated disease risk. Using predictive tools to identify conserved TFBSs that are altered by rs1421085 revealed that there are a number of TFs that are lost or created with T-to-C conversion. Although I have not investigated possible TFs bound to the orthologous region in mouse, it would be of great interest and importance to perform additional studies aiming to identify TFs - and more specifically sex-specific and hormone sensitive TFs - bound to the mouse sequence and establish whether these regulators are common in human and mouse. Through doing so it may be possible to shed light on the sex-specific nature of target gene manipulation that I observed in mouse. For DEL82, it will be necessary in the future to establish whether the phenotypic effect observed in mouse is due to the loss of several TFBSs, the gain of a new motif, or a combination of both.

Table 8.1 | Comparison between human *FTO* risk phenotype in adipocytes, FTO-RS1421085 mouse and DEL82 mouse.

		Human **rs1421085 C risk allele**	**Mouse** **Rs1421085 T→C "risk allele"**	**Mouse** **Motif deletion DEL82**
***Cis*-regulatory landscape**				
Target gene(s) & direction of effect	♀subcut	Increased expression of *IRX3* and *IRX5*[1] Effect size: 2-fold	Increased expression of *Irx3* and *Irx5* Effect size: ~2.5-fold	No effect
	♀visceral	?	Decreased expression of *Rpgrip1l, Fto* and *Irx5* Effect size: ~0.5-fold	No effect
	♂subcut	?	No effect	Decreased expression of *Irx3* and *Irx5* Effect size: ~0.7-fold Decreased proximity between *Fto* regulatory element and *Irx3*
	♂visceral	?	No effect	Increased expression of *Irx3* Effect size: ~1.5-fold Decreased proximity between *Fto* regulatory element and *Irx3* and *Irx5*
	Brain	iPSC-derived neurons: *AKTIP, RPGRIP1L* and *FTO*[7]	Hypothalamus: ?	Hypothalamus: No effect
Target tissue / cell type		Adipocytes[1] Based on elongated enhancer signature in adipocytes Brain[7] Based on gene expression changes in iPSC-derived neurons	Adipocytes Based on gene expression changes in adipose	Adipocytes Based on gene expression changes in adipose and hypothalamus
Target developmental stage		Early development	Early development and terminally differentiated adipocytes	Early development
Upstream regulator		ARID5B[1] in adipose CUX1[7] in brain	?	?

	Predicted: RXRF, OCT3, ESRRA, HBOX, ARID/ARID5A, HNF1, LEFF, BRNF		
Cellular and organismal consequences			
Body weight / BMI	Increased in both sexes[2]	Decreased in males (CC) on HFD	Decreased in DEL82 males on HFD
Body fat mass	Increased in both sexes[3]	Decreased in males (CC) on HFD	Decreased in DEL82 males on HFD
Mitochondrial copy number in adipose tissue	Down in risk genotype[1]	?	Up in DEL82 male iWAT and gWAT
Adipose tissue histology	Increased adipocyte diameter[1,4]	?	Dermal WAT thickness decreased in DEL82 males
OCR	Down in risk genoype, unable to respond to β-adrenergic stimulation[1]	?	Up in gWAT pre-adipocytes upon β-adrenergic stimulation. No difference in iWAT-derived cells
***In vivo* OCR / energy expenditure**	Controversial[5,6]	?	?
Food intake / appetite	Increased[5,6]	?	?

[1](Claussnitzer *et al.*, 2015); [2](Frayling *et al.*, 2007); [3](Wåhlén, Sjölin & Hoffstedt, 2008); [4](Fox *et al.*, 2012); [5](Speakman, 2015); [6](Hakanen *et al.*, 2009); [7](Stratigopoulos *et al.*, 2016); ? = No data

8.6 Body weight and fat mass

8.6.1 DEL82 phenotype

I hypothesized that removal of mouse orthologous region around rs1421085 would result in decreased body weight and fat mass in mice on a HFD. Indeed, DEL82 males on a HFD have a subtle reduction in overall body weight and fat mass when challenged with a HFD. More surprisingly, I also show that DEL82 causes differential fat accumulation. While there is no difference in iWAT weight between DEL82 and WT males, gWAT weight is significantly increased in DEL82 animals on a HFD.

Notably, the overall *in vivo* consequences of DEL82 on body weight and fat mass are (as expected) more subtle than those reported for global and adipose-specific *Irx3* or *Fto* manipulations (Church *et al.*, 2010; Claussnitzer *et al.*, 2015; Merkestein *et al.*, 2015; Smemo *et al.*, 2014). In germline knockout models, IRX3-KO resulted in striking metabolic and developmental phenotypes in mouse as discussed in the introduction. However, in my model, deleting a regulatory element of 82 nucleotides resulted in a ~30% reduction of *Irx3* and *Irx5* in pre-adipocytes, resembling gene expression changes that are more consistent with regulatory fine-tuning. The moderate phenotype together with the HFD-dependency of the body weight effect is very much in line with what we know about GWAS. Common variants implicated in common diseases by GWAS have characteristically a low to modest effect size (Bush & Moore, 2012). This is somewhat expected, given that most of these variants fall into gene regulatory regions. Additionally, in the case of *FTO*, nutrition has the strongest environmental effect on obesity-risk at this locus (Harbron *et al.*, 2014; Young *et al.*, 2016). Using genetic, anthropometric and lifestyle variables collected as part of the UK Biobank, Young *et al.* (2016) assessed gene-by-environment interactions and how they modify the effect of *FTO* variants on BMI. They report significant interactions between rs1421085 and a number of lifestyle and environmental factors, including alcohol

consumption and mean sleep duration, with overall diet having the strongest effect on modifying *FTO* risk (Young *et al.*, 2016).

In the human population, individuals heterozygous and homozygous for the *FTO* risk allele are on average ~1.5 and ~3 kg heavier, respectively, compared to individuals homozygous for the protective allele (Frayling *et al.*, 2007). This implies a dose-dependent effect of the *FTO* risk allele on BMI. This is consistent with my DEL82 mouse phenotype, where mice heterozygous for the DEL82 deletion have an intermittent phenotype on body weight and fat mass. In addition, DEL82 has a dose-dependent effect on target gene expression, where iWAT-derived pre-adipocytes heterozygous for DEL82 have intermediate reduction in *Irx3* and *Irx5* when compared to DEL82 homozygous.

8.6.2 FTO-RS1421085 phenotype

Similar to DEL82, the single nucleotide mutant FTO-RS1421085 presents with a body composition phenotype on a HFD in males that is dose-dependent. At 14-weeks of age, male animals with CC genotype display on average 12% less body fat compared to TT genotype. One puzzle emerging from this study is the direction of the body composition phenotype. In humans the *FTO* risk alleles have an increasing effect on BMI (Frayling *et al.*, 2007), whereas the risk allele for rs1421085 in mouse has a decreasing effect on overall body weight and fat mass in male animals. Given the opposite effect of the C allele at rs1421085 on body composition in human and mouse suggests that rs1421085 is certainly functional, but may not be solely causal for the BMI association in human. Additional variants might play a role at this locus. Alternatively, although the gene regulatory circuitry shows remarkable conservation between human and mouse, there is the possibility that the functional consequences of altered *RPGRIP1L/Rpgrip1l*, *IRX3/Irx3* and *IRX5/Irx5* expression are not the same in both species.

There is substantial evidence within the literature documenting the importance of mouse background strains (Andolina *et al.*, 2015; Dominguez-Salazar *et al*., 2004; Fontaine &

Davis, 2016; Mekada *et al.*, 2009). Mouse strain has drastic effects on phenotypes for many of the target genes investigated in this thesis. For example, from my preliminary data and through communication with others in the field (Marcelo Nobrega, oral communication) we know that the effect of IRX3-KO in mice depends heavily on the background strain. Although IRX3-KO mice on a mixed background show a striking developmental and metabolic phenotype (Smemo *et al.*, 2014), mice on a B6 background are protected from any effects of loss of *Irx3*. Hence it is, perhaps, not surprising that at the organismal level, altered *Irx3* expression in my model (which is B6N) might not reliably result in a phenotype. Although there are advantages to using mice on a homogenous background including defined genomes, relatively stable phenotypes, and comprehensive information that is freely available from published studies and databases, there might be arguments for using heterogeneous populations for studying human genome-wide signals which would resemble the diversity in genetic make-up in which the signal was discovered in in the first place. Strain-specific differences in study outcomes and the need for models that reflect the genetic diversity of human populations has led to the generation of recombinant inbred, collaborative cross, and diversity outbred strains, all of which resemble genetically diverse mouse populations (Kõks *et al.*, 2016).

An additional level of complexity comes with the fact that it is extremely difficult to account for all the different developmental and environmental factors that contribute to disease risk in human and mouse, such as questions on when to start a HFD (which will affect adipose tissue development), what a dietary challenge should be constituted of and when to harvest appropriate tissues for downstream analysis. Additionally, the litter size and microbial environment have recently been described to mediate a bi-stable epigenetic switch that foreshadows obesity (Dalgaard *et al.*, 2016). B6 and C3H mice have very distinct kinetics of body weight and fat mass accumulation (unpublished). For future studies, it would be interesting to assess the effects of DEL82 and RS1421085 on a different mouse background

strain, such as C3H, which responds drastically to a HFD challenge with an extreme obesity phenotype.

Finally, the presence of a body weight and fat mass phenotype in male FTO-RS1421085 CC allele carriers with the absence of target gene alteration in pre-adipocytes suggest that i) pre-adipocytes are not the only cell type affected by the risk allele, ii) there might be additional target transcripts other than the ones investigated or ii) there is an off-target mutation that accounts for the phenotype in males.

8.7 Adipose tissue phenotype

The adipose tissue data presented in this thesis are predominantly derived from gWAT and iWAT, two adipose depots that are easily excised and that are commonly used as representative for visceral and subcutaneous WAT, respectively. Although the exact anatomic location of gonadal and inguinal fat pads in mice does not precisely match human, these mouse depots are believed to have a similar impact on whole-body metabolism as human visceral (Foster *et al.*, 2011) and gluteofemoral WAT (Tran *et al.*, 2008), respectively. Together, iWAT and gWAT account for approximately 20% of total fat mass in 6 months old male B6N mice on a HFD. Approximately half of all adipose tissue is located in subcutaneous depots in dermal WAT (Reeder *et al.*, 2016), making dWAT a major site of fat deposition. DEL82 had an effect on all tested adipose depot sites, by regulating depot weight, gene expression, and/or thickness, albeit with very distinct phenotypes between the depots. Other metabolically important depots, such as mesenteric, renal, or pericardial fat pads that represent visceral WAT depots have not been studied in this model yet. In the future, it would be interesting to extend the study design presented in this thesis and to assess DEL82 and FTO-RS1421085 regulatory consequences in all of the depots mentioned, and to determine target genes and cellular consequences in a systematic and depot-specific manner. This would likely give insights into the relationships between different WAT depots, both on regulatory and functional aspects. Findings presented in this thesis

highlight the necessity of performing adipose characterisation in several depots when assessing a complex phenotype such as obesity, something that human studies have so far fallen short of (except in fat distribution and WHR studies) likely due to the difficulty of obtaining matching depots in one individual. My results further stress the importance to view adipose tissue not as one organ, but many, with distinct developmental, regulatory and functional properties.

8.8 Sex-specific effect

An obvious question arising from the work presented in this thesis is concerning the absence of a detectable effect in female DEL82 mice on body weight and target gene expression. The absence of a female phenotype is especially surprising given the fact that FTO-RS1421085 as well as human data clearly show that pre-adipocytes isolated from female risk-allele carriers have altered *IRX3* and *IRX5* expression and that the *FTO* GWAS association with BMI has not been reported to be sex-specific.

Although I lack sufficient data to explain the sex-specific effect of DEL82 and FTO-RS1421085, I can exclude several possibilities and hypothesize about possible reasons. Given that ATAC-seq histograms in mouse pre-adipocytes show accessible chromatin regions in both depots and both female and male samples at the relevant location proposes that it is not the presence or absence of the regulatory element within *Fto* itself that is responsible for the sex-specific effect of DEL82. However, I have not examined the regulators bound to this sequence, and there could potentially be different levels or cofactors of critical TFs in females and males that could result in a sex-specific effect of motif occupancy.

It is worth noting that we and others have observed that the variability of HFD-induced body weight gain and fat mass gain is considerably larger in females, making subtle phenotypes harder to detect in female mice. Additionally, B6N animals have a distinct pattern of fat accumulation when challenged with a HFD, and as previously mentioned, using a different

background strain that is more prone to fat accumulation might reveal whether the absence of a female phenotype is due to the mouse strain used in this study. Most likely, however, the difference between the sexes is due to hormonal differences between females and males and the resulting difference in cellular phenotypes and transcriptional profiles. More work is necessary to unravel the sex-specific effect described in this thesis.

8.9 Cellular mechanism

In human pre-adipocytes that carry the *FTO* protective allele, relatively lower expression of *IRX3* and *IRX5* results in pre-adipocytes to develop into more brown-like adipocytes as measured via increased thermogenic capacity, increased mitochondrial copy number and gene expression changes that are consistent with a browning phenotype (Claussnitzer *et al.*, 2015). Importantly, β-adrenergic stimulation was critical for the effect of the *FTO* protective allele on adipocyte browning *in vitro*. In contrast, the *FTO* risk allele shifts pre-adipocytes towards a more “white-like” program. On an organismal level, decreased *Irx3* expression is also associated with adipose tissue browning in mice (Smemo *et al.*, 2014).

Similarly, in my DEL82 model, decreased *Irx3* and *Irx5* expression in iWAT-derived pre-adipocytes results in increased mitochondrial copy number and increased expression of *Elovl3. Elovl3* encodes an enzyme that is involved in lipid accumulation and metabolic activity in brown adipocytes during the early phase of the tissue recruitment and plays a role in lipid storage and in resistance to diet-induced obesity (Westerberg *et al.*, 2006). Surprisingly, however, increased *Irx3* expression in gWAT-derived pre-adipocytes mimics the effect of decreases *Irx3* and *Irx5* expression in iWAT-derived pre-adipocytes on mitochondrial copy number in mature WAT, suggesting that it is not the direction of *Irx3*, but the intrinsic properties of the adipose tissue that mediate the effect of *Irx3* alteration on mitochondrial number.

Notably, I have shown that, although gene regulation was restricted to early pre-adipocytes, the consequences of altered *Irx3* and *Irx5* on adipocyte function was maintained in mature

WAT. Overall, mature WAT depots of DEL82 males present with a distinct set of gene expression changes of genes involved in lipid storage and synthesis.

Both, IRX3 and IRX5, are largely uncharacterised in adipose tissue. Chapters 3 and 4 of this thesis focused on the function of these genes in adipocytes. Given the complex nature of the *Fto* regulatory circuitry, designing experiments to study *Irx3* and *Irx5* function have to address several aspects, including sex and depot origin of the cell model as well as direction, level and number of target gene manipulation (i.e. double-KO or single gene KO). Using the well characterised mouse pre-adipocyte cell line 3T3-L1 to perform silencing of *Irx3* and *Irx5* revealed that both genes act in an antagonistic fashion and while *Irx3* inhibits adipogenesis and OCR, *Irx5* seems to promote it. In contrast, pre-adipocytes isolated from DEL82 males show no detectable phenotypic alterations in unstimulated conditions (Table 8.2). Indeed, β-adrenergic stimulation is necessary for the effect of DEL82 on OCR in gWAT-derived primary pre-adipocytes.

Notably, *Irx3* and *Irx5* are expressed higher in subcutaneous than visceral fat. Expression levels of both genes are very low in gWAT. Nevertheless, 1.4-fold increase of *Irx3* in pre-adipocytes from gWAT leads to a significantly larger mature fat pad in male mice on a HFD *in vivo*. Therefore, low mRNA expression level is an insufficient criterion for excluding a potential target gene.

Table 8.2 | Comparison between effect of altered Irx3 and Irx5 expression in different models.

	Target gene direction	**Tissue (HFD-stimulated)**	***In vitro* differentiated**
DEL82 male gWAT	↑Irx3 – 140% OE	↑ tissue mass ↑ mitochondria ↑ Ardb3 = Elovl3 ↑ Fasn	↑isoproteronol-stimulated OCR
DEL82 male iWAT	↓Irx3 – 30% KD ↓Irx5 – 30% KD	= tissue mass ↑ mitochondria = Adrb3 ↑ Elovl3 ↑ Fasn	= adipogenesis = OCR = mitochondria = Pgc1a = mitochondrial gene expression
3T3-L1 siIrx3	↓Irx3 – 70% KD		↑adipogenesis ↑OCR ↑ mitochondrial gene expression
3T3-L1 siIrx5	↓Irx5 – 90% KD		↓adipogenesis ↓lipogenesis ↓Pgc1a ↑Elovl3, Prdm16

The mouse as a tool in dissecting human non-coding GWA signals

Although the similarities between human and mouse data point towards a regulatory conserved network, it will be important to validate the new discoveries made in this mouse model in human cells in the future to establish how deeply context-specificity is conserved between the species and in to evaluate the feasibility of using the mouse for future non-coding GWA variant dissection at other loci.

Using the mouse to study non-coding regulatory elements that have direct disease-relevance in human could have several benefits. Most importantly, by using mice it is possible to design and perform experiments precisely around the developmental and adult stages of interest and to measure a wealth of parameters in the same animal, giving insight into *in vivo* and *in vitro* biology in the same organism. As mentioned before, the data presented in this thesis clearly emphasize the need to study adipocyte gene regulation not only in the right cell type and developmental stage, but also in both sexes and in different adipose tissue depots. Particularly important for adipose studies, all major depots are easily excisable from the same animal. The mouse is also a well-controllable system and can be challenged with diet or exercise for example, both of which have implications in obesity.

The findings described in this thesis have further implications on current practice in mouse model engineering. Traditionally, genetic manipulation in model organisms as well as *in vitro* models involves the insertion or deletion of large cassettes in non-coding and coding regions of the genome, mainly around genes of interest. However, it is likely that many of the non-coding regions targeted with selection cassettes disrupt the regulatory landscapes of the underlying sequence. My results show that a single-nucleotide change in a non-coding region can result in a significant decrease in body weight in male mice. This is interesting in the light of the IMPC project or ENU-induced mutagenesis screens, the latter being a strategy where thousands of single-nucleotide changes are introduced and may likely be co-inherited along with the described neighbouring coding mutations. It has recently been

shown that for 81 mouse lines with ENU-induced mutations and a variety of phenotypes, whole-genome-sequencing only identified 44 coding mutations that are believed to underlie the observed phenotype (Potter *et al.*, 2016) which leaves almost half of all observed phenotypes unexplained. It is interesting to speculate that in a number of cases, this might be explained by non-coding regulatory disruptions. These are important considerations when designing and interpreting mouse mutants.

Interestingly, DEL82 and FTO-RS1421085 show a number of differences, not only in the direction of phenotype effects (as expected), but also the presence or absence of certain features (Table 8.1). These differences in DEL82 and FTO-RS1421085 indicate that deletion of an element might not necessarily resemble the complex effects of a single nucleotide variant; deleting one or a set of TFBSs has different phenotypic outcomes than a single-nucleotide change which, for example, alters affinity of TFs or creates a new motif.

8.10 Impact

My results show that a single nucleotide change from T-to-C at rs1421085 in mouse selectively alters *Irx3* and *Irx5* in pre-adipocytes of female subcutaneous WAT. To my knowledge, this is the first mouse model of a non-coding GWAS signal that results in gene expression changes that models human risk allele carriers in that they appear to (i) be cell type-specific; (ii) alter all target genes; (iii) alter target genes at a relevant level and direction; and (iv) alter target genes at the relevant time of development. My results show a striking similarity to the spatial-temporal effect of target gene alterations observed in human risk allele carriers. More importantly, my results also point towards previously unknown mechanisms of the mouse orthologous region at rs1421085, namely that this element i) acts as a repressive sequence in visceral WAT-derived pre-adipocytes, rather than an enhancer; ii) has at least four target genes in pre-adipocytes (*Rpgrip1l, Fto, Irx3* and *Irx5*); and iii) has a clear sex-specific effect; all of which are findings that require validation in human.

8.11 Conclusions

The key findings of the work presented in this thesis are that mouse models of human risk variants i) recapitulated human risk-allele carriers on the cellular level suggesting a high conservation of human and mouse regulation at the *FTO/Fto* regulatory element; and ii) allowed us to discover novel aspects of regulatory complexity at the *FTO/Fto* locus. Additionally, DEL82 and FTO-RS1421085 resemble many features of human GWAS: the effects are subtle, triggered by environmental stimuli (β-adrenergic and dietary) and are specific to sex and tissue.

Targeted genome editing using CRISPR/Cas9 (Ran *et al.*, 2013) allowed us to test the phenotypic effect of altering the predicted causal nucleotide rs1421085 in its endogenous genomic context and in isolation from the other obesity-associated genetic variants in the same haplotype. Importantly, SNP-knockin in mouse is in principle a powerful approach to test whether an identfied variant has an effect on the organismal level. My data indicates that although rs1421085 is causal for the change in *Irx3* and *Irx5*, it might not solely be responsible for the BMI association. These data implicate the importance of a risk haplotype (rather than individual risk variant) and from the data presented here and by others, it seems likely that several variants are at play that cumulatively account for the large effect size of *FTO*-mediated risk on BMI. Therefore, an important goal for the future will be to establish whether the overall BMI phenotype in humans is due to one or a combination of variants that act in different tissues and/or at different times.

The level of regulatory complexity discovered using these models further stresses the importance of developing new experimental approaches to unravel GWAS variants. It can be argued that based on the data presented in this thesis it becomes clear that it is not sufficient to infer target genes based on variant proximity and function in a relevant cell type. Several functional candidates are present at the *FTO* locus. Instead, gene regulation across the interval and chromatin biology should be a focus when studying complex trait signals, such

as obesity; for obesity signals, important considerations are the choice of adipose depot and brain region, developmental time-points and environmental stimulants (β-adrenergic, dietary, etc.), which all influence disease risk. Importantly, although I propose that cellular and murine models can be powerful tools when used to dissect the mechanisms underlying the association of genetic variants with human phenotypes, these models are not a substitute for the detailed phenotyping of the species in which the association was discovered.

Together, data presented in this thesis point towards new levels of complexity at the *FTO/Fto* locus, where obesity-risk-associated regions might not only be specific to tissue and developmental stage, but also sex and depot. Finally, my data indicates that not only can a regulatory region have several target genes, the number and direction of the effect on those genes can be different in different cell types and developmental stages. Whether these additional discoveries made in the mouse have relevance in human should be priority in future follow-up studies.

References

Aida, T., Imahashi, R., & Tanaka, K. (2014). Translating human genetics into mouse: The impact of ultra-rapid *in vivo* genome editing. *Development, Growth & Differentiation*, *56*(1), 34–45. https://doi.org/10.1111/dgd.12101

Akter, M. H., Yamaguchi, T., Hirose, F., & Osumi, T. (2008). Perilipin, a critical regulator of fat storage and breakdown, is a target gene of estrogen receptor-related receptor α. *Biochemical and Biophysical Research Communications*, *368*(3), 563–568. https://doi.org/10.1016/j.bbrc.2008.01.102

Andolina, D., Puglisi-Allegra, S., & Ventura, R. (2015). Strain-dependent differences in corticolimbic processing of aversive or rewarding stimuli. *Frontiers in Systems Neuroscience*, *8*, 207. https://doi.org/10.3389/fnsys.2014.00207

Appleton, S. L., Seaborn, C. J., Visvanathan, R., Hill, C. L., Gill, T. K., Taylor, A. W., & Adams, R. J. (2013). Diabetes and Cardiovascular Disease Outcomes in the Metabolically Healthy Obese Phenotype: A cohort study. *Diabetes Care*, *36*(8), 2388–2394. https://doi.org/10.2337/dc12-1971

Ardlie, K. G., Deluca, D. S., Segre, A. V., Sullivan, T. J., Young, T. R., Gelfand, E. T., … Dermitzakis, E. T. (2015). The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. *Science*, *348*(6235), 648–660. https://doi.org/10.1126/science.1262110

Arner, E., Westermark, P. O., Spalding, K. L., Britton, T., Ryden, M., Frisen, J., … Arner, P. (2010). Adipocyte Turnover: Relevance to Human Adipose Tissue Morphology. *Diabetes*, *59*(1), 105–109. https://doi.org/10.2337/db09-0942

Arnone, M. I., & Davidson, E. H. (1997). The hardwiring of development: organization and function of genomic regulatory systems. *Development (Cambridge, England)*, *124*(10), 1851–64. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9169833

Arrighi, N., Moratal, C., Clément, N., Giorgetti-Peraldi, S., Peraldi, P., Loubat, A., … Dechesne, C. A. (2015). Characterization of adipocytes derived from fibro/adipogenic progenitors resident in human skeletal muscle. *Cell Death & Disease*, *6*(4), e1733–e1733. https://doi.org/10.1038/cddis.2015.79

Assini, J. M., Mulvihill, E. E., Burke, A. C., Sutherland, B. G., Telford, D. E., Chhoker, S. S., … Huff, M. W. (2015). Naringenin Prevents Obesity, Hepatic Steatosis, and Glucose Intolerance in Male Mice Independent of Fibroblast Growth Factor 21. *Endocrinology*, *156*(6), 2087–2102. https://doi.org/10.1210/en.2014-2003

Badis, G., Berger, M. F., Philippakis, A. A., Talukder, S., Gehrke, A. R., Jaeger, S. A., … Bulyk, M. L. (2009). Diversity and Complexity in DNA Recognition by Transcription Factors. *Science*, *324*(5935), 1720–1723. https://doi.org/10.1126/science.1162327

Baker, K., & Beales, P. L. (2009). Making sense of cilia in disease: The human ciliopathies. *American Journal of Medical Genetics Part C: Seminars in Medical Genetics*, *151C*(4), 281–295. https://doi.org/10.1002/ajmg.c.30231

Bannister, A. J., & Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Cell Research*, *21*(3), 381–395. https://doi.org/10.1038/cr.2011.22

Beer, N. L., Tribble, N. D., McCulloch, L. J., Roos, C., Johnson, P. R. V., Orho-Melander, M., & Gloyn, A. L. (2009). The P446L variant in GCKR associated with fasting plasma glucose and triglyceride levels exerts its effect through increased glucokinase activity in liver. *Human Molecular Genetics*, *18*(21), 4081–4088. https://doi.org/10.1093/hmg/ddp357

Ben-Haim, M. S., Moshitch-Moshkovitz, S., & Rechavi, G. (2015). FTO: linking m6A demethylation to

adipogenesis. *Cell Research*, *25*(1), 3–4. https://doi.org/10.1038/cr.2014.162

Benabdallah, N. S., & Bickmore, W. A. (2015). Regulatory Domains and Their Mechanisms. *Cold Spring Harbor Symposia on Quantitative Biology*, *80*, 45–51. https://doi.org/10.1101/sqb.2015.80.027268

Berry, R., Jeffery, E., & Rodeheffer, M. S. (2014). Weighing in on Adipocyte Precursors. *Cell Metabolism*, *19*(1), 8–20. https://doi.org/10.1016/j.cmet.2013.10.003

Berulava, T., & Horsthemke, B. (2010). The obesity-associated SNPs in intron 1 of the FTO gene affect primary transcript levels. *European Journal of Human Genetics : EJHG*, *18*(9), 1054–6. https://doi.org/10.1038/ejhg.2010.71

Bille, D. S., Chabanova, E., Gamborg, M., Fonvig, C. E., Nielsen, T. R. H., Thisted, E., … Holm, J.-C. (2012). Liver fat content investigated by magnetic resonance spectroscopy in obese children and youths included in multidisciplinary treatment. *Clinical Obesity*, *2*(1–2), 41–49. https://doi.org/10.1111/j.1758-8111.2012.00038.x

Billon, N., Iannarelli, P., Monteiro, M. C., Glavieux-Pardanaud, C., Richardson, W. D., Kessaris, N., … Dupin, E. (2007). The generation of adipocytes by the neural crest. *Development*, *134*(12), 2283–2292. https://doi.org/10.1242/dev.002642

Boissel, S., Reish, O., Proulx, K., Kawagoe-Takaki, H., Sedgwick, B., Yeo, G. S. H., … Colleaux, L. (2009). Loss-of-function mutation in the dioxygenase-encoding FTO gene causes severe growth retardation and multiple malformations. *American Journal of Human Genetics*, *85*(1), 106–11. https://doi.org/10.1016/j.ajhg.2009.06.002

Bosse, A., Zülch, A., Becker, M. B., Torres, M., Gómez-Skarmeta, J. L., Modolell, J., & Gruss, P. (1997). Identification of the vertebrate Iroquois homeobox gene family with overlapping expression during early development of the nervous system. *Mechanisms of Development*, *69*(1–2), 169–81. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9486539

Boyle, A. P., Hong, E. L., Hariharan, M., Cheng, Y., Schaub, M. A., Kasowski, M., … Snyder, M. (2012). Annotation of functional variation in personal genomes using RegulomeDB. *Genome Research*, *22*(9), 1790–7. https://doi.org/10.1101/gr.137323.112

Breschi, A., Djebali, S., Gillis, J., Pervouchine, D. D., Dobin, A., Davis, C. A., … Guigó, R. (2016). Gene-specific patterns of expression variation across organs and species. *Genome Biology*, *17*(1), 151. https://doi.org/10.1186/s13059-016-1008-y

Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y., & Greenleaf, W. J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nature Methods*, *10*(12), 1213–8. https://doi.org/10.1038/nmeth.2688

Buenrostro, J. D., Wu, B., Chang, H. Y., & Greenleaf, W. J. (2015). ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Current Protocols in Molecular Biology*, *109*, 21.29.1-9. https://doi.org/10.1002/0471142727.mb2129s109

Bush, W. S., & Moore, J. H. (2012). Chapter 11: Genome-Wide Association Studies. *PLoS Computational Biology*, *8*(12), e1002822. https://doi.org/10.1371/journal.pcbi.1002822

Cantile, M., Procino, A., D'Armiento, M., Cindolo, L., & Cillo, C. (2003). HOX gene network is involved in the transcriptional regulation of in vivo human adipogenesis. *Journal of Cellular Physiology*, *194*(2), 225–36. https://doi.org/10.1002/jcp.10210

Carithers, L. J., Ardlie, K., Barcus, M., Branton, P. A., Britton, A., Buia, S. A., … Moore, H. M. (2015). A Novel Approach to High-Quality Postmortem Tissue Procurement: The GTEx Project. *Biopreservation and Biobanking*, *13*(5), 311–319. https://doi.org/10.1089/bio.2015.0032

Carrat, G. R., Hu, M., Nguyen-Tu, M.-S., Chabosseau, P., Gaulton, K. J., van de Bunt, M., … Rutter,

G. A. (2017). Decreased STARD10 Expression Is Associated with Defective Insulin Secretion in Humans and Mice. *The American Journal of Human Genetics*, *100*(2), 238–256. https://doi.org/10.1016/j.ajhg.2017.01.011

Chau, Y.-Y., Bandiera, R., Serrels, A., Mart?nez-Estrada, O. M., Qing, W., Lee, M., … Hastie, N. (2014). Visceral and subcutaneous fat have different origins and evidence supports a mesothelial source. *Nature Cell Biology*, *16*(4), 367–375. https://doi.org/10.1038/ncb2922

Chen, X., Luo, Y., Jia, G., Liu, G., Zhao, H., & Huang, Z. (2017). FTO Promotes Adipogenesis through Inhibition of the Wnt/β-catenin Signaling Pathway in Porcine Intramuscular Preadipocytes. *Animal Biotechnology*, 1–7. https://doi.org/10.1080/10495398.2016.1273835

Chen, X., Zhou, B., Luo, Y., Huang, Z., Jia, G., Liu, G., & Zhao, H. (2016). Tissue Distribution of Porcine FTO and Its Effect on Porcine Intramuscular Preadipocytes Proliferation and Differentiation. *PloS One*, *11*(3), e0151056. https://doi.org/10.1371/journal.pone.0151056

Cheng, Y., Ma, Z., Kim, B.-H., Wu, W., Cayting, P., Boyle, A. P., … Snyder, M. P. (2014). Principles of regulatory information conservation between mouse and human. *Nature*, *515*(7527), 371–375. https://doi.org/10.1038/nature13985

Cheung, N. K. M., Cheung, A. C. K., Ye, R. R., Ge, W., Giesy, J. P., & Au, D. W. T. (2013). Expression profile of oestrogen receptors and oestrogen-related receptors is organ specific and sex dependent: the Japanese medaka *Oryzias latipes* model. *Journal of Fish Biology*, *83*(2), 295–310. https://doi.org/10.1111/jfb.12164

Christoffels, V. M., Keijser, A. G. M., Houweling, A. C., Clout, D. E. W., & Moorman, A. F. M. (2000). Patterning the Embryonic Heart: Identification of Five Mouse Iroquois Homeobox Genes in the Developing Heart. *Developmental Biology*, *224*(2), 263–274. https://doi.org/10.1006/dbio.2000.9801

Chu, A. Y., Deng, X., Fisher, V. A., Drong, A., Zhang, Y., Feitosa, M. F., … Fox, C. S. (2017). Multiethnic genome-wide meta-analysis of ectopic fat depots identifies loci associated with adipocyte development and differentiation. *Nature Genetics*, *49*(1), 125–130. https://doi.org/10.1038/ng.3738

Church, C. D., Berry, R., & Rodeheffer, M. S. (2014). Isolation and study of adipocyte precursors. *Methods in Enzymology*, *537*, 31–46. https://doi.org/10.1016/B978-0-12-411619-1.00003-3

Church, C., Lee, S., Bagg, E. A. L., McTaggart, J. S., Deacon, R., Gerken, T., … Cox, R. D. (2009). A mouse model for the metabolic effects of the human fat mass and obesity associated FTO gene. *PLoS Genetics*, *5*(8), e1000599. https://doi.org/10.1371/journal.pgen.1000599

Church, C., Moir, L., McMurray, F., Girard, C., Banks, G. T., Teboul, L., … Cox, R. D. (2010). Overexpression of Fto leads to increased food intake and results in obesity. *Nature Genetics*, *42*(12), 1086–92. https://doi.org/10.1038/ng.713

Claussnitzer, M., Dankel, S. N., Kim, K.-H., Quon, G., Meuleman, W., Haugen, C., … Kellis, M. (2015). FTO Obesity Variant Circuitry and Adipocyte Browning in Humans. *The New England Journal of Medicine*, *373*(10), 895–907. https://doi.org/10.1056/NEJMoa1502214

Claussnitzer, M., Dankel, S. N., Klocke, B., Grallert, H., Glunk, V., Berulava, T., … Laumen, H. (2014). Leveraging cross-species transcription factor binding site patterns: From diabetes risk loci to disease mechanisms. *Cell*, *156*(1–2), 343–358. https://doi.org/10.1016/j.cell.2013.10.058

Cohen, D. R., Cheng, C. W., Cheng, S. H., & Hui, C. C. (2000). Expression of two novel mouse Iroquois homeobox genes during neurogenesis. *Mechanisms of Development*, *91*(1–2), 317–21. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10704856

Cohen, M., Syme, C., Deforest, M., Wells, G., Detzler, G., Cheng, H.-L., … Hamilton, J. (2014). Ectopic fat in youth: the contribution of hepatic and pancreatic fat to metabolic disturbances. *Obesity (Silver Spring, Md.)*, *22*(5), 1280–6. https://doi.org/10.1002/oby.20674

Coutinho, T., Goel, K., Corr?a de S?, D., Kragelund, C., Kanaya, A. M., Zeller, M., … Lopez-Jimenez, F. (2011). Central Obesity and Survival in Subjects With Coronary Artery Disease. *Journal of the American College of Cardiology*, *57*(19), 1877–1886. https://doi.org/10.1016/j.jacc.2010.11.058

Dahlman, I., Rydén, M., Brodin, D., Grallert, H., Strawbridge, R. J., & Arner, P. (2016). Numerous Genes in Loci Associated With Body Fat Distribution Are Linked to Adipose Function. *Diabetes*, *65*(2), 433–7. https://doi.org/10.2337/db15-0828

Dalgaard, K., Landgraf, K., Heyne, S., Lempradl, A., Longinotto, J., Gossens, K., … Pospisilik, J. A. (2016). Trim28 Haploinsufficiency Triggers Bi-stable Epigenetic Obesity. *Cell*, *164*(3), 353–64. https://doi.org/10.1016/j.cell.2015.12.025

Dankel, S. N., Fadnes, D. J., Stavrum, A. K., Stansberg, C., Holdhus, R., Hoang, T., … Mellgren, G. (2010). Switch from stress response to homeobox transcription factors in adipose tissue after profound fat loss. *PLoS ONE*, *5*(6). https://doi.org/10.1371/journal.pone.0011033

Daoud, H., Zhang, D., McMurray, F., Yu, A., Luco, S. M., Vanstone, J., … Armour, C. M. (2016). Identification of a pathogenic *FTO* mutation by next-generation sequencing in a newborn with growth retardation and developmental delay. *Journal of Medical Genetics*, *53*(3), 200–207. https://doi.org/10.1136/jmedgenet-2015-103399

Degner, J. F., Pai, A. A., Pique-Regi, R., Veyrieras, J.-B., Gaffney, D. J., Pickrell, J. K., … Pritchard, J. K. (2012). DNase I sensitivity QTLs are a major determinant of human expression variation. *Nature*, *482*(7385), 390–4. https://doi.org/10.1038/nature10808

Dekker, J., Marti-Renom, M. A., & Mirny, L. A. (2013). Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. *Nature Reviews Genetics*, *14*(6), 390–403. https://doi.org/10.1038/nrg3454

Delhon, I., Gutzwiller, S., Morvan, F., Rangwala, S., Wyder, L., Evans, G., … Fournier, B. (2009). Absence of Estrogen Receptor-Related-α Increases Osteoblastic Differentiation and Cancellous Bone Mineral Density. *Endocrinology*, *150*(10), 4463–4472. https://doi.org/10.1210/en.2009-0121

Denas, O., Sandstrom, R., Cheng, Y., Beal, K., Herrero, J., Hardison, R. C., & Taylor, J. (2015). Genome-wide comparative analysis reveals human-mouse regulatory landscape and evolution. *BMC Genomics*, *16*(1), 87. https://doi.org/10.1186/s12864-015-1245-6

DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium, Dia. G. R. A. M. (DIAGRAM), Asian Genetic Epidemiology Network Type 2 Diabetes (AGEN-T2D) Consortium, A. G. E. N. T. 2 D. (AGEN-T., South Asian Type 2 Diabetes (SAT2D) Consortium, S. A. T. 2 D. (SAT2D), Mexican American Type 2 Diabetes (MAT2D) Consortium, M. A. T. 2 D. (MAT2D), Type 2 Diabetes Genetic Exploration by Nex-generation sequencing in muylti-Ethnic Samples (T2D-GENES) Consortium, T. 2 D. G. E. by N. sequencing in multi-E. S. (T2D-G., Mahajan, A., … Morris, A. P. (2014). Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. *Nature Genetics*, *46*(3), 234–44. https://doi.org/10.1038/ng.2897

Dickel, D. E., Barozzi, I., Zhu, Y., Fukuda-Yuzawa, Y., Osterwalder, M., Mannion, B. J., … Pennacchio, L. A. (2016). Genome-wide compendium and functional assessment of in vivo heart enhancers. *Nature Communications*, *7*, 12923. https://doi.org/10.1038/ncomms12923

Dimas, A. S., Lagou, V., Barker, A., Knowles, J. W., Magi, R., Hivert, M.-F., … MAGIC Investigators. (2014). Impact of Type 2 Diabetes Susceptibility Variants on Quantitative Glycemic Traits Reveals Mechanistic Heterogeneity. *Diabetes*, *63*(6), 2158–2171. https://doi.org/10.2337/db13-0949

Dixon, J. R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., … Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*, *485*(7398), 376–80. https://doi.org/10.1038/nature11082

Dominguez-Salazar, E., Bateman, H. L., & Rissman, E. F. (2004). Background matters: the effects of estrogen receptor alpha gene disruption on male sexual behavior are modified by background strain. *Hormones and Behavior*, *46*(4), 482–90. https://doi.org/10.1016/j.yhbeh.2004.05.006

Emerging Risk Factors Collaboration, Wormser, D., Kaptoge, S., Di Angelantonio, E., Wood, A. M., Pennells, L., … Danesh, J. (2011). Separate and combined associations of body-mass index and abdominal adiposity with cardiovascular disease: collaborative analysis of 58 prospective studies. *Lancet (London, England)*, *377*(9771), 1085–95. https://doi.org/10.1016/S0140-6736(11)60105-0

ENCODE Project Consortium, T. E. P. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature*, *489*(7414), 57–74. https://doi.org/10.1038/nature11247

Ernst, J., & Kellis, M. (2012). ChromHMM: automating chromatin-state discovery and characterization. *Nature Methods*, *9*(3), 215–216. https://doi.org/10.1038/nmeth.1906

Ernst, J., Kheradpour, P., Mikkelsen, T. S., Shoresh, N., Ward, L. D., Epstein, C. B., … Bernstein, B. E. (2011). Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature*, *473*(7345), 43–49. https://doi.org/10.1038/nature09906

Ferrara, A. (2007). Increasing prevalence of gestational diabetes mellitus: a public health perspective. *Diabetes Care*, *30 Suppl 2*(Supplement 2), S141-6. https://doi.org/10.2337/dc07-s206

Fischer, B., Schöttl, T., Schempp, C., Fromme, T., Hauner, H., Klingenspor, M., & Skurk, T. (2015). Inverse relationship between body mass index and mitochondrial oxidative phosphorylation capacity in human subcutaneous adipocytes. *American Journal of Physiology. Endocrinology and Metabolism*, *309*(4), E380-7. https://doi.org/10.1152/ajpendo.00524.2014

Fischer, J., Koch, L., Emmerling, C., Vierkotten, J., Peters, T., Brüning, J. C., & Rüther, U. (2009). Inactivation of the Fto gene protects from obesity. *Nature*, *458*(7240), 894–898. https://doi.org/10.1038/nature07848

Flannick, J., & Florez, J. C. (2016). Type 2 diabetes: genetic data sharing to advance complex disease research. *Nature Reviews Genetics*, *17*(9), 535–549. https://doi.org/10.1038/nrg.2016.56

Fontaine, D. A., & Davis, D. B. (2016). Attention to Background Strain Is Essential for Metabolic Research: C57BL/6 and the International Knockout Mouse Consortium. *Diabetes*, *65*(1), 25–33. https://doi.org/10.2337/db15-0982

Fonvig, C. E., Bille, D. S., Chabanova, E., Nielsen, T. R. H., Thomsen, H. S., & Holm, J.-C. (2012). Muscle fat content and abdominal adipose tissue distribution investigated by magnetic resonance spectroscopy and imaging in obese children and youths. *Pediatric Reports*, *4*(1), e11. https://doi.org/10.4081/pr.2012.e11

Foster, M. T., Shi, H., Seeley, R. J., & Woods, S. C. (2011). Removal of intra-abdominal visceral adipose tissue improves glucose tolerance in rats: Role of hepatic triglyceride storage. *Physiology & Behavior*, *104*(5), 845–854. https://doi.org/10.1016/j.physbeh.2011.04.064

Fox, C. S., Liu, Y., White, C. C., Feitosa, M., Smith, A. V, Heard-Costa, N., … Borecki, I. B. (2012). Genome-wide association for abdominal subcutaneous and visceral adipose reveals a novel locus for visceral fat in women. *PLoS Genetics*, *8*(5), e1002695. https://doi.org/10.1371/journal.pgen.1002695

Frayling, T. M., Timpson, N. J., Weedon, M. N., Zeggini, E., Freathy, R. M., Lindgren, C. M., … McCarthy, M. I. (2007). A Common Variant in the FTO Gene Is Associated with Body Mass Index and Predisposes to Childhood and Adult Obesity. *Science*, *316*(5826), 889–894. https://doi.org/10.1126/science.1141634

Fuchsberger, C., Flannick, J., Teslovich, T. M., Mahajan, A., Agarwala, V., Gaulton, K. J., … McCarthy, M. I. (2016). The genetic architecture of type 2 diabetes. *Nature*, *536*(7614), 41–47.

https://doi.org/10.1038/nature18642

Gaborit, N., Sakuma, R., Wylie, J. N., Kim, K.-H., Zhang, S.-S., Hui, C.-C., & Bruneau, B. G. (2012). Cooperative and antagonistic roles for Irx3 and Irx5 in cardiac morphogenesis and postnatal physiology. *Development*, *139*(21), 4007–4019. https://doi.org/10.1242/dev.081703

Gantner, M. L., Hazen, B. C., Eury, E., Brown, E. L., & Kralli, A. (2016). Complementary Roles of Estrogen-Related Receptors in Brown Adipocyte Thermogenic Function. *Endocrinology*, *157*(12), 4770–4781. https://doi.org/10.1210/en.2016-1767

Gao, X., Shin, Y.-H., Li, M., Wang, F., Tong, Q., & Zhang, P. (2010). The Fat Mass and Obesity Associated Gene FTO Functions in the Brain to Regulate Postnatal Growth in Mice. *PLoS ONE*, *5*(11), e14005. https://doi.org/10.1371/journal.pone.0014005

Gardiner, W. J. K., & Teboul, L. (2009). Overexpression transgenesis in mouse: pronuclear injection. *Methods in Molecular Biology (Clifton, N.J.)*, *561*, 111–26. https://doi.org/10.1007/978-1-60327-019-9_8

Gaulton, K. J. (2017). Mechanisms of Type 2 Diabetes Risk Loci. *Current Diabetes Reports*, *17*(9), 72. https://doi.org/10.1007/s11892-017-0908-x

Gaulton, K. J., Ferreira, T., Lee, Y., Raimondo, A., Mägi, R., Reschen, M. E., … DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium. (2015). Genetic fine mapping and genomic annotation defines causal mechanisms at type 2 diabetes susceptibility loci. *Nature Genetics*, *47*(12), 1415–25. https://doi.org/10.1038/ng.3437

Gesta, S., Bezy, O., Mori, M. A., Macotela, Y., Lee, K. Y., & Kahn, C. R. (2011). Mesodermal developmental gene Tbx15 impairs adipocyte differentiation and mitochondrial respiration. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(7), 2771–6. https://doi.org/10.1073/pnas.1019704108

Gesta, S., Blüher, M., Yamamoto, Y., Norris, A. W., Berndt, J., Kralisch, S., … Kahn, C. R. (2006). Evidence for a role of developmental genes in the origin of obesity and body fat distribution. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(17), 6676–81. https://doi.org/10.1073/pnas.0601752103

Gesta, S., Tseng, Y. H., & Kahn, C. R. (2007). Developmental Origin of Fat: Tracking Obesity to Its Source. *Cell*, *131*(2), 242–256. https://doi.org/10.1016/j.cell.2007.10.004

Ghavi-Helm, Y., Klein, F. A., Pakozdi, T., Ciglar, L., Noordermeer, D., Huber, W., & Furlong, E. E. M. (2014). Enhancer loops appear stable during development and are associated with paused polymerase. *Nature*, *512*(7512), 96. https://doi.org/10.1038/nature13417

Gilad, Y., & Mizrahi-Man, O. (2015). A reanalysis of mouse ENCODE comparative gene expression data. *F1000Research*, *4*, 121. https://doi.org/10.12688/f1000research.6536.1

Giordano, A., Frontini, A., & Cinti, S. (2016). Convertible visceral fat as a therapeutic target to curb obesity. *Nature Reviews. Drug Discovery*, *15*(6), 405–424. https://doi.org/10.1038/nrd.2016.31

Gómez-Marín, C., Tena, J. J., Acemel, R. D., López-Mayorga, M., Naranjo, S., de la Calle-Mustienes, E., … Gómez-Skarmeta, J. L. (2015). Evolutionary comparison reveals that diverging CTCF sites are signatures of ancestral topological associating domains borders. *Proc Natl Acad Sci U S A*, *112*(24), 201505463. https://doi.org/10.1073/pnas.1505463112

Grove, K. L., Fried, S. K., Greenberg, A. S., Xiao, X. Q., & Clegg, D. J. (2010). A microarray analysis of sexual dimorphism of adipose tissues in high-fat-diet-induced obese mice. *International Journal of Obesity*, *34*(6), 989–1000. https://doi.org/10.1038/ijo.2010.12

Gulati, P., Avezov, E., Ma, M., Antrobus, R., Lehner, P., O'Rahilly, S., & Yeo, G. S. H. (2014). Fat mass and obesity-related (FTO) shuttles between the nucleus and cytoplasm. *Bioscience Reports*, *34*(5), 621–628. https://doi.org/10.1042/BSR20140111

Gulati, P., Cheung, M. K., Antrobus, R., Church, C. D., Harding, H. P., Tung, Y.-C. L., … Yeo, G. S. H. (2013). Role for the obesity-related FTO gene in the cellular sensing of amino acids. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(7), 2557–62. https://doi.org/10.1073/pnas.1222796110

Hakanen, M., Raitakari, O. T., Lehtimäki, T., Peltonen, N., Pahkala, K., Sillanmäki, L., … Rönnemaa, T. (2009). FTO Genotype Is Associated with Body Mass Index after the Age of Seven Years But Not with Energy Intake or Leisure-Time Physical Activity. *The Journal of Clinical Endocrinology & Metabolism*, *94*(4), 1281–1287. https://doi.org/10.1210/jc.2008-1199

Harbron, J., Van der Merwe, L., Zaahl, M. G., Kotze, M. J., & Senekal, M. (2014). Fat mass and obesity-associated (FTO) gene polymorphisms are associated with physical activity, food intake, eating behaviors, psychological health, and modeled change in body mass index in overweight/obese caucasian adults. *Nutrients*, *6*(8), 3130–3152. https://doi.org/10.3390/nu6083130

Harms, M., & Seale, P. (2013). Brown and beige fat: development, function and therapeutic potential. *Nat Med*, *19*(10), 1252–1263. https://doi.org/10.1038/nm.3361

Health Effects of Overweight and Obesity in 195 Countries over 25 Years. (2017). *New England Journal of Medicine*, NEJMoa1614362. https://doi.org/10.1056/NEJMoa1614362

Heid, I. M., Jackson, A. U., Randall, J. C., Winkler, T. W., Qi, L., Steinthorsdottir, V., … Lindgren, C. M. (2010). Meta-analysis identifies 13 new loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. *Nature Genetics*, *42*(11), 949–60. https://doi.org/10.1038/ng.685

Heitmann, B. L., & Lissner, L. (2011). Hip Hip Hurrah! Hip size inversely related to heart disease and total mortality. *Obesity Reviews : An Official Journal of the International Association for the Study of Obesity*, *12*(6), 478–81. https://doi.org/10.1111/j.1467-789X.2010.00794.x

Hepler, C., Vishvanath, L., & Gupta, R. K. (2017). Sorting out adipocyte precursors and their role in physiology and disease. *Genes & Development*, *31*(2), 127–140. https://doi.org/10.1101/gad.293704.116

Hess, M. E., & Brüning, J. C. (2014). The fat mass and obesity-associated (FTO) gene: Obesity and beyond? *Biochimica et Biophysica Acta - Molecular Basis of Disease*, *1842*(10), 2039–2047. https://doi.org/10.1016/j.bbadis.2014.01.017

Hoenig, M. R., Cowin, G., Buckley, R., McHenery, C., & Coulthard, A. (2011). Low density lipoprotein cholesterol is inversely correlated with abdominal visceral fat area: a magnetic resonance imaging study. *Lipids in Health and Disease*, *10*, 12. https://doi.org/10.1186/1476-511X-10-12

Houweling, A. C., Dildrop, R., Peters, T., Mummenhoff, J., Moorman, A. F., Rüther, U., & Christoffels, V. M. (2001). Gene and cluster-specific expression of the Iroquois family members during mouse development. *Mechanisms of Development*, *107*(1–2), 169–74. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11520674

Hughes, J. R., Roberts, N., McGowan, S., Hay, D., Giannoulatou, E., Lynch, M., … Higgs, D. R. (2014). Analysis of hundreds of cis-regulatory landscapes at high resolution in a single, high-throughput experiment., *46*(2), 205–12. https://doi.org/10.1038/ng.2871

Hurst, L. D., Pál, C., & Lercher, M. J. (2004). The evolutionary dynamics of eukaryotic gene order. *Nature Reviews Genetics*, *5*(4), 299–310. https://doi.org/10.1038/nrg1319

Jeffery, E., Church, C. D., Holtrup, B., Colman, L., & Rodeheffer, M. S. (2015). Rapid depot-specific activation of adipocyte precursor cells at the onset of obesity. *Nature Cell Biology*, *17*(4), 376–385. https://doi.org/10.1038/ncb3122

Jeffery, E., Wing, A., Holtrup, B., Sebo, Z., Kaplan, J. L., Saavedra-Peña, R., … Rodeheffer, M. S. (2016). The Adipose Tissue Microenvironment Regulates Depot-Specific Adipogenesis in

Obesity. *Cell Metabolism*, *24*(1), 142–50. https://doi.org/10.1016/j.cmet.2016.05.012

Jorgensen, J. S., & Gao, L. (2005). Irx3 is differentially up-regulated in female gonads during sex determination. *Gene Expression Patterns : GEP*, *5*(6), 756–62. https://doi.org/10.1016/j.modgep.2005.04.011

Jowett, J. B. M., Curran, J. E., Johnson, M. P., Carless, M. A., Göring, H. H. H., Dyer, T. D., … Blangero, J. (2010). Genetic variation at the FTO locus influences RBL2 gene expression. *Diabetes*, *59*(3), 726–32. https://doi.org/10.2337/db09-1277

Ju, D., He, J., Zheng, X., & Yang, G. (2009). [Cloning, expression of the porcine estrogen-related receptor alpha gene and its effect on lipid accumulation in mature adipocytes]. *Sheng Wu Gong Cheng Xue Bao = Chinese Journal of Biotechnology*, *25*(11), 1627–32. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/20222459

Kamimae-Lanning, A. N., Krasnow, S. M., Goloviznina, N. A., Zhu, X., Roth-Carter, Q. R., Levasseur, P. R., … Marks, D. L. (2015). Maternal high-fat diet and obesity compromise fetal hematopoiesis. *Molecular Metabolism*, *4*(1), 25–38. https://doi.org/10.1016/j.molmet.2014.11.001

Kanneganti, T.-D., & Dixit, V. D. (2012). Immunological complications of obesity. *Nature Immunology*, *13*(8), 707–712. https://doi.org/10.1038/ni.2343

Karastergiou, K., Fried, S. K., Xie, H., Lee, M.-J., Divoux, A., Rosencrantz, M. A., … Smith, S. R. (2013). Distinct Developmental Signatures of Human Abdominal and Gluteal Subcutaneous Adipose Tissue Depots. *The Journal of Clinical Endocrinology & Metabolism*, *98*(1), 362–371. https://doi.org/10.1210/jc.2012-2953

Karra, E., O'Daly, O. G., Choudhury, A. I., Yousseif, A., Millership, S., Neary, M. T., … Batterham, R. L. (2013). A link between FTO, ghrelin, and impaired brain food-cue responsivity. *Journal of Clinical Investigation*, *123*(8), 3539–3551. https://doi.org/10.1172/JCI44403

Kasowski, M., Grubert, F., Heffelfinger, C., Hariharan, M., Asabere, A., Waszak, S. M., … Snyder, M. (2010). Variation in transcription factor binding among humans. *Science (New York, N.Y.)*, *328*(5975), 232–5. https://doi.org/10.1126/science.1183621

Kasza, I., Hernando, D., Roldán-Alzate, A., Alexander, C. M., Reeder, S. B., & Alexander, C. (2016). Thermogenic profiling using magnetic resonance imaging of dermal and other adipose tissues. *JCI Insight*, *1*(13), 35–45. https://doi.org/10.1172/jci.insight.87146

Kasza, I., Suh, Y., Wollny, D., Clark, R. J., Roopra, A., Colman, R. J., … Alexander, C. M. (2014). Syndecan-1 is required to maintain intradermal fat and prevent cold stress. *PLoS Genetics*, *10*(8), e1004514. https://doi.org/10.1371/journal.pgen.1004514

Kent, W. J., Sugnet, C. W., Furey, T. S., Roskin, K. M., Pringle, T. H., Zahler, A. M., & Haussler, D. (2002). The human genome browser at UCSC. *Genome Research*, *12*(6), 996–1006. https://doi.org/10.1101/gr.229102. Article published online before print in May 2002

Keuper, M., Jastroch, M., Yi, C.-X., Fischer-Posovszky, P., Wabitsch, M., Tschöp, M. H., & Hofmann, S. M. (2014). Spare mitochondrial respiratory capacity permits human adipocytes to maintain ATP homeostasis under hypoglycemic conditions. *FASEB Journal : Official Publication of the Federation of American Societies for Experimental Biology*, *28*(2), 761–70. https://doi.org/10.1096/fj.13-238725

Kim, S.-N., Jung, Y.-S., Kwon, H.-J., Seong, J. K., Granneman, J. G., & Lee, Y.-H. (2016). Sex differences in sympathetic innervation and browning of white adipose tissue of mice. *Biology of Sex Differences*, *7*, 67. https://doi.org/10.1186/s13293-016-0121-7

Kim, S., Lun, M., Wang, M., Senyo, S., Guillermier, C., Patwari, P., & Steinhauser, M. (2014). Loss of White Adipose Hyperplastic Potential Is Associated with Enhanced Susceptibility to Insulin Resistance. *Cell Metabolism*, *20*(6), 1049–1058. https://doi.org/10.1016/j.cmet.2014.10.010

Knight, J. (2014). Approaches for establishing the function of regulatory genetic variants involved in disease. *Genome Medicine*, *6*(10), 92. https://doi.org/10.1186/s13073-014-0092-4

Kõks, S., Dogan, S., Tuna, B. G., González-Navarro, H., Potter, P., & Vandenbroucke, R. E. (2016). Mouse models of ageing and their relevance to disease. *Mechanisms of Ageing and Development*, *160*, 41–53. https://doi.org/10.1016/j.mad.2016.10.001

Kong, D., Tong, Q., Ye, C., Koda, S., Fuller, P. M., Krashes, M. J., … Lowell, B. B. (2012). GABAergic RIP-Cre Neurons in the Arcuate Nucleus Selectively Regulate Energy Expenditure. *Cell*, *151*(3), 645–657. https://doi.org/10.1016/j.cell.2012.09.020

Krishnaswami, S. R., Grindberg, R. V, Novotny, M., Venepally, P., Lacar, B., Bhutani, K., … Lasken, R. S. (2016). Using single nuclei for RNA-seq to capture the transcriptome of postmortem neurons. *Nature Protocols*, *11*(3), 499–524. https://doi.org/10.1038/nprot.2016.015

Krueger, K. C., Costa, M. J., Du, H., & Feldman, B. J. (2014). Characterization of Cre Recombinase Activity for In Vivo Targeting of Adipocyte Precursor Cells. *Stem Cell Reports*, *3*(6), 1147–1158. https://doi.org/10.1016/j.stemcr.2014.10.009

Kulzer, J. R., Stitzel, M. L., Morken, M. A., Huyghe, J. R., Fuchsberger, C., Kuusisto, J., … Mohlke, K. L. (2014). A Common Functional Regulatory Variant at a Type 2 Diabetes Locus Upregulates ARAP1 Expression in the Pancreatic Beta Cell. *The American Journal of Human Genetics*, *94*(2), 186–197. https://doi.org/10.1016/j.ajhg.2013.12.011

Landgraf, K., Scholz, M., Kovacs, P., Kiess, W., & Körner, A. (2016). FTO Obesity Risk Variants Are Linked to Adipocyte IRX3 Expression and BMI of Children - Relevance of FTO Variants to Defend Body Weight in Lean Children? *PloS One*, *11*(8), e0161739. https://doi.org/10.1371/journal.pone.0161739

Lecoutre, S., Deracinois, B., Laborie, C., Eberlé, D., Guinez, C., Panchenko, P. E., … Breton, C. (2016). Depot- and sex-specific effects of maternal obesity in offspring's adipose tissue. *The Journal of Endocrinology*, *230*(1), 39–53. https://doi.org/10.1530/JOE-16-0037

Lee, B.-C., & Lee, J. (2014). Cellular and molecular players in adipose tissue inflammation in the development of obesity-induced insulin resistance. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, *1842*(3), 446–462. https://doi.org/10.1016/j.bbadis.2013.05.017

Lee, J. M., & Sonnhammer, E. L. L. (2003). Genomic gene clustering analysis of pathways in eukaryotes. *Genome Research*, *13*(5), 875–82. https://doi.org/10.1101/gr.737703

Lee, Y.-H., Kim, S.-N., Kwon, H.-J., & Granneman, J. G. (2017). Metabolic heterogeneity of activated beige/brite adipocytes in inguinal adipose tissue. *Scientific Reports*, *7*(January), 39794. https://doi.org/10.1038/srep39794

Li, G., Ruan, X., Auerbach, R. K., Sandhu, K. S., Zheng, M., Wang, P., … Ruan, Y. (2012). Extensive Promoter-Centered Chromatin Interactions Provide a Topological Basis for Transcription Regulation. *Cell*, *148*(1–2), 84–98. https://doi.org/10.1016/j.cell.2011.12.014

Liu, C.-T., Buchkovich, M. L., Winkler, T. W., Heid, I. M., African Ancestry Anthropometry Genetics Consortium, I. B., GIANT Consortium, C. S., … Adrienne Cupples, L. (2014). Multi-ethnic fine-mapping of 14 central adiposity loci. *Human Molecular Genetics*, *23*(17), 4738–44. https://doi.org/10.1093/hmg/ddu183

Liu, L., Zhang, M., Xia, Z., Xu, P., Chen, L., & Xu, T. (2011). Caenorhabditis elegans ciliary protein NPHP-8, the homologue of human RPGRIP1L, is required for ciliogenesis and chemosensation. *Biochemical and Biophysical Research Communications*, *410*(3), 626–631. https://doi.org/10.1016/j.bbrc.2011.06.041

Locke, A. E., Kahali, B., Berndt, S. I., Justice, A. E., Pers, T. H., Day, F. R., … Speliotes, E. K. (2015). Genetic studies of body mass index yield new insights for obesity biology. *Nature*, *518*(7538), 197–206. https://doi.org/10.1038/nature14177

Lodhi, I. J., Wei, X., & Semenkovich, C. F. (2011). Lipoexpediency: de novo lipogenesis as a metabolic signal transmitter. *Trends in Endocrinology & Metabolism*, *22*(1), 1–8. https://doi.org/10.1016/j.tem.2010.09.002

Manolopoulos, K. N., Karpe, F., & Frayn, K. N. (2010). Gluteofemoral body fat as a determinant of metabolic health. *International Journal of Obesity*, *34*(6), 949–959. https://doi.org/10.1038/ijo.2009.286

McCulloch, L. J., van de Bunt, M., Braun, M., Frayn, K. N., Clark, A., & Gloyn, A. L. (2011). GLUT2 (SLC2A2) is not the principal glucose transporter in human pancreatic beta cells: implications for understanding genetic association signals at this locus. *Molecular Genetics and Metabolism*, *104*(4), 648–53. https://doi.org/10.1016/j.ymgme.2011.08.026

McLean, C. Y., Bristor, D., Hiller, M., Clarke, S. L., Schaar, B. T., Lowe, C. B., … Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. *Nature Biotechnology*, *28*(5). https://doi.org/10.1038/nbt.1630

McMurray, F., Church, C. D., Larder, R., Nicholson, G., Wells, S., Teboul, L., … Cox, R. D. (2013). Adult onset global loss of the fto gene alters body composition and metabolism in the mouse. *PLoS Genetics*, *9*(1), e1003166. https://doi.org/10.1371/journal.pgen.1003166

McTaggart, J. S., Lee, S., Iberl, M., Church, C., Cox, R. D., & Ashcroft, F. M. (2011). FTO is expressed in neurones throughout the brain and its expression is unaltered by fasting. *PloS One*, *6*(11), e27968. https://doi.org/10.1371/journal.pone.0027968

Mekada, K., Abe, K., Murakami, A., Nakamura, S., Nakata, H., Moriwaki, K., … Yoshiki, A. (2009). Genetic Differences among C57BL/6 Substrains. *Exp. Anim*, *58*(2), 141–149. https://doi.org/10.1538/expanim.58.141

Merkestein, M., Laber, S., McMurray, F., Andrew, D., Sachse, G., Sanderson, J., … Cox, R. D. (2015). FTO influences adipogenesis by regulating mitotic clonal expansion. *Nature Communications*, *6*. https://doi.org/10.1038/ncomms7792

Meyre, D., Proulx, K., Kawagoe-Takaki, H., Vatin, V., Gutierrez-Aguilar, R., Lyon, D., … Yeo, G. S. (2009). Prevalence of loss of function {FTO} mutations in lean and obese individuals. *Diabetes*, *59*(January), 311–318. https://doi.org/db09-0703 [pii] 10.2337/db09-0703

Mi, H., Huang, X., Muruganujan, A., Tang, H., Mills, C., Kang, D., & Thomas, P. D. (2017). PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. *Nucleic Acids Research*, *45*(D1), D183–D189. https://doi.org/10.1093/nar/gkw1138

Mianné, J., Codner, G. F., Caulder, A., Fell, R., Hutchison, M., King, R., … Teboul, L. (2017). Analysing the outcome of CRISPR-aided genome editing in embryos: Screening, genotyping and quality control. *Methods (San Diego, Calif.)*. https://doi.org/10.1016/j.ymeth.2017.03.016

Michailidou, Z., Morton, N. M., Moreno Navarrete, J. M., West, C. C., Stewart, K. J., Fernández-Real, J. M., … Ratcliffe, P. J. (2015). Adipocyte Pseudohypoxia Suppresses Lipolysis and Facilitates Benign Adipose Tissue Expansion. *Diabetes*, *64*(3), 733–745. https://doi.org/10.2337/db14-0233

Mikkelsen, T. S., Xu, Z., Zhang, X., Wang, L., Gimble, J. M., Lander, E. S., & Rosen, E. D. (2010). Comparative Epigenomic Analysis of Murine and Human Adipogenesis. *Cell*, *143*(1), 156–169. https://doi.org/10.1016/j.cell.2010.09.006

Misra, A., Garg, A., Abate, N., Peshock, R. M., Stray-Gundersen, J., & Grundy, S. M. (1997). Relationship of anterior and posterior subcutaneous abdominal fat to insulin sensitivity in nondiabetic men. *Obesity Research*, *5*(2), 93–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9112243

Montavon, T., Soshnikova, N., Mascrez, B., Joye, E., Thevenet, L., Splinter, E., … Duboule, D. (2011). A regulatory archipelago controls Hox genes transcription in digits. *Cell*, *147*(5), 1132–

45. https://doi.org/10.1016/j.cell.2011.10.023

Morris, A. P., Voight, B. F., Teslovich, T. M., Ferreira, T., Segrè, A. V, Steinthorsdottir, V., … DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium. (2012). Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes., *44*(9). https://doi.org/10.1038/ng.2383

Multhaup, M. L., Seldin, M. M., Jaffe, A. E., Lei, X., Kirchner, H., Mondal, P., … Feinberg, A. P. (2015). Mouse-human experimental epigenetic analysis unmasks dietary targets and genetic liability for diabetic phenotypes. *Cell Metabolism*, *21*(1), 138–49. https://doi.org/10.1016/j.cmet.2014.12.014

Nazare, J.-A., Smith, J. D., Borel, A.-L., Haffner, S. M., Balkau, B., Ross, R., … Després, J.-P. (2012). Ethnic influences on the relations between abdominal subcutaneous and visceral adiposity, liver fat, and cardiometabolic risk profile: the International Study of Prediction of Intra-Abdominal Adiposity and Its Relationship With Cardiometabolic Risk/Intra-Abdominal Adiposity. *The American Journal of Clinical Nutrition*, *96*(4), 714–26. https://doi.org/10.3945/ajcn.112.035758

Ng, M., Fleming, T., Robinson, M., Thomson, B., Graetz, N., Margono, C., … Gakidou, E. (2014). Global, regional, and national prevalence of overweight and obesity in children and adults during 1980?2013: a systematic analysis for the Global Burden of Disease Study 2013. *The Lancet*, *384*(9945), 766–781. https://doi.org/10.1016/S0140-6736(14)60460-8

Nicholas, L. M., Rattanatray, L., MacLaughlin, S. M., Ozanne, S. E., Kleemann, D. O., Walker, S. K., … McMillen, I. C. (2013). Differential effects of maternal obesity and weight loss in the periconceptional period on the epigenetic regulation of hepatic insulin-signaling pathways in the offspring. *FASEB Journal : Official Publication of the Federation of American Societies for Experimental Biology*, *27*(9), 3786–96. https://doi.org/10.1096/fj.13-227918

O'Rahilly, S., & Farooqi, I. S. (2000). *The Genetics of Obesity in Humans*. *Endotext*. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/25905292

Oben, J. A., Mouralidarane, A., Samuelsson, A.-M., Matthews, P. J., Morgan, M. L., McKee, C., … Taylor, P. D. (2010). Maternal obesity during pregnancy and lactation programs the development of offspring non-alcoholic fatty liver disease in mice. *Journal of Hepatology*, *52*(6), 913–20. https://doi.org/10.1016/j.jhep.2009.12.042

Palmer, B. F., & Clegg, D. J. (2015). The sexual dimorphism of obesity. *Molecular and Cellular Endocrinology*, *402*, 113–9. https://doi.org/10.1016/j.mce.2014.11.029

Pasquali, L., Gaulton, K. J., Rodríguez-Seguí, S. A., Mularoni, L., Miguel-Escalada, I., Akerman, I., … Ferrer, J. (2014). Pancreatic islet enhancer clusters enriched in type 2 diabetes risk-associated variants. *Nature Genetics*, *46*(2), 136–43. https://doi.org/10.1038/ng.2870

Pellegrinelli, V., Carobbio, S., & Vidal-Puig, A. (2016). Adipose tissue plasticity: how fat depots respond differently to pathophysiological cues. *Diabetologia*, *59*(6), 1075–1088. https://doi.org/10.1007/s00125-016-3933-4

Penfold, N. C., & Ozanne, S. E. (2015). Developmental programming by maternal obesity in 2015: Outcomes, mechanisms, and potential interventions. *Hormones and Behavior*, *76*, 143–52. https://doi.org/10.1016/j.yhbeh.2015.06.015

Pennacchio, L. A., Ahituv, N., Moses, A. M., Prabhakar, S., Nobrega, M. A., Shoukry, M., … Rubin, E. M. (2006). In vivo enhancer analysis of human conserved non-coding sequences. *Nature*, *444*(7118), 499–502. https://doi.org/10.1038/nature05295

Pers, T. H., Karjalainen, J. M., Chan, Y., Westra, H.-J., Wood, A. R., Yang, J., … Franke, L. (2015). Biological interpretation of genome-wide association studies using predicted gene functions. *Nature Communications*, *6*, 5890. https://doi.org/10.1038/ncomms6890

Phillips, L. K., & Prins, J. B. (2008). The link between abdominal obesity and the metabolic syndrome.

Current Hypertension Reports, *10*(2), 156–64. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18474184

Pischon, T., Boeing, H., Hoffmann, K., Bergmann, M., Schulze, M. B., Overvad, K., … Riboli, E. (2008). General and abdominal adiposity and risk of death in Europe. *The New England Journal of Medicine*, *359*(20), 2105–20. https://doi.org/10.1056/NEJMoa0801891

Pomerantz, M. M., Ahmadiyeh, N., Jia, L., Herman, P., Verzi, M. P., Doddapaneni, H., … Freedman, M. L. (2009). The 8q24 cancer risk variant rs6983267 shows long-range interaction with MYC in colorectal cancer. *Nature Genetics*, *41*(8), 882–4. https://doi.org/10.1038/ng.403

Potter, P. K., Bowl, M. R., Jeyarajan, P., Wisby, L., Blease, A., Goldsworthy, M. E., … Brown, S. D. M. (2016). Novel gene function revealed by mouse mutagenesis screens for models of age-related disease. *Nature Communications*, *7*, 12444. https://doi.org/10.1038/ncomms12444

Prentice, A. M. (2006). The emerging epidemic of obesity in developing countries. *International Journal of Epidemiology*, *35*(1), 93–9. https://doi.org/10.1093/ije/dyi272

Pruim, R. J., Welch, R. P., Sanna, S., Teslovich, T. M., Chines, P. S., Gliedt, T. P., … Willer, C. J. (2010). LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics*, *26*(18), 2336–2337. https://doi.org/10.1093/bioinformatics/btq419

Ragvin, A., Moro, E., Fredman, D., Navratilova, P., Drivenes, Ø., Engström, P. G., … Becker, T. S. (2010). Long-range gene regulation links genomic type 2 diabetes and obesity risk regions to HHEX, SOX4, and IRX3. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(2), 775–80. https://doi.org/10.1073/pnas.0911591107

Ramachandrappa, S., Raimondo, A., Cali, A. M. G., Keogh, J. M., Henning, E., Saeed, S., … Farooqi, I. S. (2013). Rare variants in single-minded 1 (SIM1) are associated with severe obesity. *Journal of Clinical Investigation*, *123*(7), 3042–3050. https://doi.org/10.1172/JCI68016

Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., & Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nature Protocols*, *8*(11), 2281–2308. https://doi.org/10.1038/nprot.2013.143

Reeder, S., Reeder, S., Wen, Z., Pelc, N., Klos, K., & Alexander, C. (2016). Thermogenic profiling using magnetic resonance imaging of dermal and other adipose tissues. *Magn Reson Med*, *51*(1), 35–45. https://doi.org/10.1172/JCI.INSIGHT.87146

Reggiani, L., Raciti, D., Airik, R., Kispert, A., & Brändli, A. W. (2007). The prepattern transcription factor Irx3 directs nephron segment identity. *Genes & Development*, *21*(18), 2358–70. https://doi.org/10.1101/gad.450707

Roadmap Epigenomics Consortium, A., Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., … Kellis, M. (2015). Integrative analysis of 111 reference human epigenomes. *Nature*, *518*(7539). https://doi.org/10.1038/nature14248

Ronkainen, J., Huusko, T. J., Soininen, R., Mondini, E., Cinti, F., Mäkelä, K. a, … Salonurmi, T. (2015). Fat mass- and obesity-associated gene Fto affects the dietary response in mouse white adipose tissue. *Scientific Reports*, *5*, 9233. https://doi.org/10.1038/srep09233

Ronkainen, J., Mondini, E., Cinti, F., Cinti, S., Sebért, S., Savolainen, M. J., & Salonurmi, T. (2016). Fto-Deficiency Affects the Gene and MicroRNA Expression Involved in Brown Adipogenesis and Browning of White Adipose Tissue in Mice. *International Journal of Molecular Sciences*, *17*(11). https://doi.org/10.3390/ijms17111851

Sanchez-Gurmaches, J., & Guertin, D. A. (2014). Adipocytes arise from multiple lineages that are heterogeneously and dynamically distributed. *Nature Communications*, *5*. https://doi.org/10.1038/ncomms5099

Schleinitz, D., Böttcher, Y., Blüher, M., & Kovacs, P. (2014). The genetics of fat distribution.

Diabetologia, *57*(7), 1276–1286. https://doi.org/10.1007/s00125-014-3214-z

Schwessinger, R., Suciu, M. C., McGowan, S. J., Telenius, J., Taylor, S., Higgs, D. R., & Hughes, J. R. (2017). Sasquatch: predicting the impact of regulatory SNPs on transcription factor binding from cell- and tissue-specific DNase footprints. *Genome Research*, *27*(10), 1730–1742. https://doi.org/10.1101/gr.220202.117

Scuteri, A., Sanna, S., Chen, W.-M., Uda, M., Albai, G., Strait, J., … Abecasis, G. R. (2007). Genome-Wide Association Scan Shows Genetic Variants in the FTO Gene Are Associated with Obesity-Related Traits. *PLoS Genetics*, *3*(7), e115. https://doi.org/10.1371/journal.pgen.0030115

Shen, L., Thompson, P. M., Potkin, S. G., Bertram, L., Farrer, L. A., Foroud, T. M., … Saykin, A. J. (2014). Genetic analysis of quantitative phenotypes in AD and MCI: Imaging, cognition and biomarkers. *Brain Imaging and Behavior*, *8*(2), 183–207. https://doi.org/10.1007/s11682-013-9262-z

Shook, B., Rivera Gonzalez, G., Ebmeier, S., Grisotti, G., Zwick, R., & Horsley, V. (2016). The Role of Adipocytes in Tissue Regeneration and Stem Cell Niches. *Annual Review of Cell and Developmental Biology*, *32*, 609–631. https://doi.org/10.1146/annurev-cellbio-111315-125426

Shungin, D., Winkler, T. W., Croteau-Chonka, D. C., Ferreira, T., Locke, A. E., Mägi, R., … Mohlke, K. L. (2015). New genetic loci link adipose and insulin biology to body fat distribution. *Nature*, *518*(7538), 187–196. https://doi.org/10.1038/nature14132

Siersbaek, M. S., Loft, A., Aagaard, M. M., Nielsen, R., Schmidt, S. F., Petrovic, N., … Mandrup, S. (2012). Genome-Wide Profiling of Peroxisome Proliferator-Activated Receptor in Primary Epididymal, Inguinal, and Brown Adipocytes Reveals Depot-Selective Binding Correlated with Gene Expression. *Molecular and Cellular Biology*, *32*(17), 3452–3463. https://doi.org/10.1128/MCB.00526-12

Siersbæk, R., & Mandrup, S. (2011). Transcriptional networks controlling adipocyte differentiation. *Cold Spring Harbor Symposia on Quantitative Biology*, *76*, 247–255. https://doi.org/10.1101/sqb.2011.76.010512

Siersbæk, R., Rabiee, A., Nielsen, R., Sidoli, S., Traynor, S., Loft, A., … Mandrup, S. (2014). Transcription factor cooperativity in early adipogenic hotspots and super-enhancers. *Cell Reports*, *7*(5), 1443–55. https://doi.org/10.1016/j.celrep.2014.04.042

Small, K. S., Hedman, Å. K., Grundberg, E., Nica, A. C., Thorleifsson, G., Kong, A., … MuTHER Consortium. (2011). Identification of an imprinted master trans regulator at the KLF14 locus related to multiple metabolic phenotypes. *Nature Genetics*, *43*(6), 561–564. https://doi.org/10.1038/ng.833

Smemo, S., Tena, J. J., Kim, K.-H., Gamazon, E. R., Sakabe, N. J., Gómez-Marín, C., … Nóbrega, M. a. (2014). Obesity-associated variants within FTO form long-range functional connections with IRX3. *Nature*, *507*(7492), 371–5. https://doi.org/10.1038/nature13138

Snijder, M. B., Dekker, J. M., Visser, M., Bouter, L. M., Stehouwer, C. D. A., Kostense, P. J., … Seidell, J. C. (2003). Associations of hip and thigh circumferences independent of waist circumference with the incidence of type 2 diabetes: the Hoorn Study. *The American Journal of Clinical Nutrition*, *77*(5), 1192–7. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12716671

Spalding, K. L., Arner, E., Westermark, P. O., Bernard, S., Buchholz, B. A., Bergmann, O., … Arner, P. (2008). Dynamics of fat cell turnover in humans. *Nature*, *453*(7196), 783–7. https://doi.org/10.1038/nature06902

Speakman, J. R. (2015). The “Fat Mass and Obesity Related” (FTO) gene: Mechanisms of Impact on Obesity and Energy Balance. *Current Obesity Reports*, *4*(1), 73–91. https://doi.org/10.1007/s13679-015-0143-1

Spiegelman, B. M., & Flier, J. S. (2001). Obesity and the regulation of energy balance. *Cell*, *104*(4),

531–43. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11239410

Stergachis, A. B., Neph, S., Sandstrom, R., Haugen, E., Reynolds, A. P., Zhang, M., … Stamatoyannopoulos, J. A. (2014). Conservation of trans-acting circuitry during mammalian regulatory evolution. *Nature*, *515*(7527), 365–370. https://doi.org/10.1038/nature13972

Sternson, S. M., & Atasoy, D. (2014). Agouti-Related Protein Neuron Circuits That Regulate Appetite. *Neuroendocrinology*, *100*(2–3), 95–102. https://doi.org/10.1159/000369072

Stratigopoulos, G., Burnett, L. C., Rausch, R., Gill, R., Penn, D. B., Skowronski, A. A., … Leibel, R. L. (2016). Hypomorphism of Fto and Rpgrip1l causes obesity in mice. *Journal of Clinical Investigation*, *126*(5), 1897–1910. https://doi.org/10.1172/JCI85526

Stratigopoulos, G., LeDuc, C. A., Cremona, M. L., Chung, W. K., & Leibel, R. L. (2011). Cut-like homeobox 1 (CUX1) regulates expression of the fat mass and obesity-associated and retinitis pigmentosa GTPase regulator-interacting protein-1-like (RPGRIP1L) genes and coordinates leptin receptor signaling. *The Journal of Biological Chemistry*, *286*(3), 2155–70. https://doi.org/10.1074/jbc.M110.188482

Stratigopoulos, G., Martin Carli, J. F., O'Day, D. R., Wang, L., LeDuc, C. A., Lanzano, P., … Leibel, R. L. (2014). Hypomorphism for RPGRIP1L, a Ciliary Gene Vicinal to the FTO Locus, Causes Increased Adiposity in Mice. *Cell Metabolism*, *19*(5), 767–779. https://doi.org/10.1016/j.cmet.2014.04.009

Stratigopoulos, G., Padilla, S. L., LeDuc, C. A., Watson, E., Hattersley, A. T., McCarthy, M. I., … Leibel, R. L. (2008). Regulation of Fto/Ftm gene expression in mice and humans. *American Journal of Physiology. Regulatory, Integrative and Comparative Physiology*, *294*(4), R1185-96. https://doi.org/10.1152/ajpregu.00839.2007

Szklarczyk, D., Morris, J. H., Cook, H., Kuhn, M., Wyder, S., Simonovic, M., … von Mering, C. (2017). The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible. *Nucleic Acids Research*, *45*(D1), D362–D368. https://doi.org/10.1093/nar/gkw937

Tchoukalova, Y. D., Koutsari, C., Votruba, S. B., Tchkonia, T., Giorgadze, N., Thomou, T., … Jensen, M. D. (2010). Sex- and Depot-Dependent Differences in Adipogenesis in Normal-Weight Humans. *Obesity*, *18*(10), 1875–1880. https://doi.org/10.1038/oby.2010.56

Tews, D., Fischer-Posovszky, P., Fromme, T., Klingenspor, M., Fischer, J., Rüther, U., … Wabitsch, M. (2013). FTO deficiency induces UCP-1 expression and mitochondrial uncoupling in adipocytes. *Endocrinology*, *154*(9), 3141–3151. https://doi.org/10.1210/en.2012-1873

Tews, D., Fischer-Posovszky, P., & Wabitsch, M. (2011). Regulation of FTO and FTM expression during human preadipocyte differentiation. *Hormone and Metabolic Research*, *43*(1), 17–21. https://doi.org/10.1055/s-0030-1265130

Thomas, P. D., Campbell, M. J., Kejariwal, A., Mi, H., Karlak, B., Daverman, R., … Narechania, A. (2003). PANTHER: a library of protein families and subfamilies indexed by function. *Genome Research*, *13*(9), 2129–41. https://doi.org/10.1101/gr.772403

Thomsen, S. K., Ceroni, A., van de Bunt, M., Burrows, C., Barrett, A., Scharfmann, R., … Gloyn, A. L. (2016). Systematic Functional Characterization of Candidate Causal Genes for Type 2 Diabetes Risk Variants. *Diabetes*, *65*(12), 3805–3811. https://doi.org/10.2337/db16-0361

Tran, T. T., & Kahn, C. R. (2010). Transplantation of adipose tissue and stem cells: role in metabolism and disease. *Nature Reviews Endocrinology*, *6*(4), 195–213. https://doi.org/10.1038/nrendo.2010.20

Tran, T. T., Yamamoto, Y., Gesta, S., Kahn, C. R., Leibel, R. L., Ferrante, A. W., … al., et. (2008). Beneficial Effects of Subcutaneous Fat Transplantation on Metabolism. *Cell Metabolism*, *7*(5), 410–420. https://doi.org/10.1016/j.cmet.2008.04.004

Tung, Y. C. L., Ayuso, E., Shan, X., Bosch, F., O'Rahilly, S., Coll, A. P., & Yeo, G. S. H. (2010). Hypothalamic-specific manipulation of Fto, the ortholog of the human obesity gene FTO, affects food intake in rats. *PLoS ONE*, *5*(1). https://doi.org/10.1371/journal.pone.0008771

Uezumi, A., Fukada, S., Yamamoto, N., Ikemoto-Uezumi, M., Nakatani, M., Morita, M., … Tsuchida, K. (2014). Identification and characterization of PDGFRα+ mesenchymal progenitors in human skeletal muscle. *Cell Death & Disease*, *5*(4), e1186. https://doi.org/10.1038/cddis.2014.161

van de Bunt, M., Manning Fox, J. E., Dai, X., Barrett, A., Grey, C., Li, L., … Gloyn, A. L. (2015). Transcript Expression Data from Human Islets Links Regulatory Signals from Genome-Wide Association Studies for Type 2 Diabetes and Glycemic Traits to Their Downstream Effectors. *PLOS Genetics*, *11*(12), e1005694. https://doi.org/10.1371/journal.pgen.1005694

Vierstra, J., Rynes, E., Sandstrom, R., Zhang, M., Canfield, T., Hansen, R. S., … Stamatoyannopoulos, J. A. (2014). Mouse regulatory DNA landscapes reveal global principles of cis-regulatory evolution. *Science*, *346*(6212), 1007–1012. https://doi.org/10.1126/science.1246426

Visel, A., Minovitsky, S., Dubchak, I., & Pennacchio, L. A. (2007). VISTA Enhancer Browser--a database of tissue-specific human enhancers. *Nucleic Acids Research*, *35*(Database issue), D88-92. https://doi.org/10.1093/nar/gkl822

Visel, A., Taher, L., Girgis, H., May, D., Golonzhka, O., Hoch, R., … Rubenstein, J. R. (2013). A High-Resolution Enhancer Atlas of the Developing Telencephalon. *Cell*, *152*(4), 895–908. https://doi.org/10.1016/j.cell.2012.12.041

Visscher, P. M., Wray, N. R., Zhang, Q., Sklar, P., McCarthy, M. I., Brown, M. A., & Yang, J. (2017). 10 Years of GWAS Discovery: Biology, Function, and Translation. *The American Journal of Human Genetics*, *101*(1), 5–22. https://doi.org/10.1016/j.ajhg.2017.06.005

Visser, M., Kayser, M., & Palstra, R.-J. (2012). HERC2 rs12913832 modulates human pigmentation by attenuating chromatin-loop formation between a long-range enhancer and the OCA2 promoter. *Genome Research*, *22*(3), 446–55. https://doi.org/10.1101/gr.128652.111

Voight, B. F., Scott, L. J., Steinthorsdottir, V., Morris, A. P., Dina, C., Welch, R. P., … GIANT Consortium. (2010). Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nature Genetics*, *42*(7), 579–89. https://doi.org/10.1038/ng.609

Wåhlén, K., Sjölin, E., & Hoffstedt, J. (2008). The common rs9939609 gene variant of the fat mass- and obesity-associated gene FTO is related to fat cell lipolysis. *Journal of Lipid Research*, *49*(3), 607–11. https://doi.org/10.1194/jlr.M700448-JLR200

Wajchenberg, B. L., Giannella-Neto, D., da Silva, M. E., & Santos, R. F. (2002). Depot-specific hormonal characteristics of subcutaneous and visceral adipose tissue and their relation to the metabolic syndrome. *Hormone and Metabolic Research = Hormon- Und Stoffwechselforschung = Hormones et Metabolisme*, *34*(11–12), 616–21. https://doi.org/10.1055/s-2002-38256

Wang, C., Liu, W., Nie, Y., Qaher, M., Horton, H. E., Yue, F., … Kuang, S. (2017). Loss of MyoD Promotes Fate Transdifferentiation of Myoblasts Into Brown Adipocytes. *EBioMedicine*, *16*, 212–223. https://doi.org/10.1016/j.ebiom.2017.01.015

Wang, P., Yang, F.-J., Du, H., Guan, Y.-F., Xu, T.-Y., Xu, X.-W., … Miao, C.-Y. (n.d.). Involvement of leptin receptor long isoform (LepRb)-STAT3 signaling pathway in brain fat mass- and obesity-associated (FTO) downregulation during energy restriction. *Molecular Medicine (Cambridge, Mass.)*, *17*(5–6), 523–32. https://doi.org/10.2119/molmed.2010.00134

Wang, Q. A., Tao, C., Gupta, R. K., & Scherer, P. E. (2013). Tracking adipogenesis during white adipose tissue development, expansion and regeneration. *Nature Medicine*, *19*(10), 1338–44. https://doi.org/10.1038/nm.3324

Wang, X., Zhu, L., Chen, J., & Wang, Y. (2015). mRNA m^6A methylation downregulates adipogenesis

in porcine adipocytes. *Biochemical and Biophysical Research Communications*, *459*(2), 201–7. https://doi.org/10.1016/j.bbrc.2015.02.048

Wang, Y., Zhang, B., Zhang, L., An, L., Xu, J., Li, D., … Yue, F. (2017). The 3D Genome Browser: a web-based browser for visualizing 3D genome organization and long-range chromatin interactions. *bioRxiv*. Retrieved from http://www.biorxiv.org/content/early/2017/02/27/112268

Ward, L. D., & Kellis, M. (2012). HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Research*, *40*(D1), D930–D934. https://doi.org/10.1093/nar/gkr917

Wei, F.-Y., Suzuki, T., Watanabe, S., Kimura, S., Kaitsuka, T., Fujimura, A., … Tomizawa, K. (2011). Deficit of tRNALys modification by Cdkal1 causes the development of type 2 diabetes in mice. *Journal of Clinical Investigation*, *121*(9), 3598–3608. https://doi.org/10.1172/JCI58056

Wernstedt?Asterholm, I., Tao, C., Morley, T., Wang, Q., Delgado-Lopez, F., Wang, Z., & Scherer, P. (2014). Adipocyte Inflammation Is Essential for Healthy Adipose Tissue Expansion and Remodeling. *Cell Metabolism*, *20*(1), 103–118. https://doi.org/10.1016/j.cmet.2014.05.005

Westerberg, R., Månsson, J.-E., Golozoubova, V., Shabalina, I. G., Backlund, E. C., Tvrdik, P., … Jacobsson, A. (2006). ELOVL3 Is an Important Component for Early Onset of Lipid Recruitment in Brown Adipose Tissue. *Journal of Biological Chemistry*, *281*(8), 4958–4968. https://doi.org/10.1074/jbc.M511588200

Weyer, C., Foley, J. E., Bogardus, C., Tataranni, P. A., & Pratley, R. E. (2000). Enlarged subcutaneous abdominal adipocyte size, but not obesity itself, predicts Type II diabetes independent of insulin resistance. *Diabetologia*, *43*(12), 1498–1506. https://doi.org/10.1007/s001250051560

Williams, C. L., Li, C., Kida, K., Inglis, P. N., Mohan, S., Semenec, L., … Leroux, M. R. (2011). MKS and NPHP modules cooperate to establish basal body/transition zone membrane associations and ciliary gate function during ciliogenesis. *The Journal of Cell Biology*, *192*(6), 1023–1041. https://doi.org/10.1083/jcb.201012116

Xin, Y., Kim, J., Okamoto, H., Ni, M., Wei, Y., Adler, C., … Gromada, J. (2016). RNA Sequencing of Single Human Islet Cells Reveals Type 2 Diabetes Genes. *Cell Metabolism*, *24*(4), 608–615. https://doi.org/10.1016/j.cmet.2016.08.018

Xue, R., Lynes, M. D., Dreyfuss, J. M., Shamsi, F., Schulz, T. J., Zhang, H., … Tseng, Y.-H. (2015). Clonal analyses and gene profiling identify genetic biomarkers of the thermogenic potential of human brown and white preadipocytes. *Nature Medicine*, *21*(7), 760–768. https://doi.org/10.1038/nm.3881

Yamamoto, Y., Gesta, S., Lee, K. Y., Tran, T. T., Saadatirad, P., & Kahn, C. R. (2010). Adipose depots possess unique developmental gene signatures. *Obesity (Silver Spring, Md.)*, *18*(5), 872–8. https://doi.org/10.1038/oby.2009.512

Yang, J., Loos, R. J. F., Powell, J. E., Medland, S. E., Speliotes, E. K., Chasman, D. I., … Visscher, P. M. (2012). FTO genotype is associated with phenotypic variability of body mass index. *Nature*, *490*(7419), 267–272. https://doi.org/10.1038/nature11401

Yazdi, F. T., Clee, S. M., & Meyre, D. (2015). Obesity genetics in mouse and human: back and forth, and back again. *PeerJ*, *3*, e856. https://doi.org/10.7717/peerj.856

Young, A. I., Wauthier, F., & Donnelly, P. (2016). Multiple novel gene-by-environment interactions modify the effect of FTO variants on body mass index. *Nature Communications*, *7*, 12724. https://doi.org/10.1038/ncomms12724

Yue, F., Cheng, Y., Breschi, A., Vierstra, J., Wu, W., Ryba, T., … Mouse ENCODE Consortium. (2014). A comparative encyclopedia of DNA elements in the mouse genome. *Nature*, *515*(7527), 355–64. https://doi.org/10.1038/nature13992

Zhang, M., Zhang, Y., Ma, J., Guo, F., Cao, Q., Zhang, Y., … Zhao, R. (2015). The demethylase activity of FTO (Fat mass and obesity associated protein) is required for preadipocyte differentiation. *PLoS ONE*, *10*(7), 1–15. https://doi.org/10.1371/journal.pone.0133788

Zhang, X., Bailey, S. D., & Lupien, M. (2014). Laying a solid foundation for Manhattan – "setting the functional basis for the post-GWAS era." *Trends in Genetics*, *30*(4), 140–149. https://doi.org/10.1016/j.tig.2014.02.006

Zhao, X., Yang, Y., Sun, B.-F., Shi, Y., Yang, X., Xiao, W., … Yang, Y.-G. (2014). FTO-dependent demethylation of N6-methyladenosine regulates mRNA splicing and is required for adipogenesis. *Cell Research*, *24*(12), 1403–1419. https://doi.org/10.1038/cr.2014.151

Zhou, K., Bellenguez, C., Spencer, C. C. A., Bennett, A. J., Coleman, R. L., Tavendale, R., … Pearson, E. R. (2011). Common variants near ATM are associated with glycemic response to metformin in type 2 diabetes. *Nature Genetics*, *43*(2), 117–120. https://doi.org/10.1038/ng.735

Appendix

Fig. S. 1 | Epigenomic annotations.
Figure shows chromatin state model for 12-state model using 6 histone marks from (Mikkelsen *et al.*, 2010) for candidate state annotations (Ernst & Kellis, 2012).

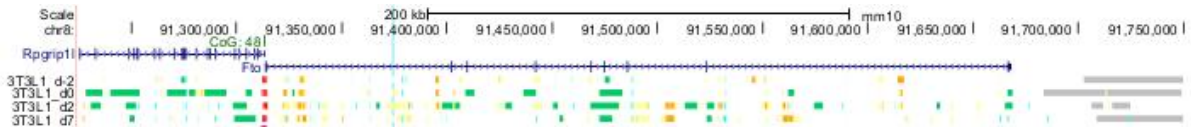


Fig. S. 2 | ChromHMM in 3T3-L1 pre-adipocytes during adipogenesis.
Blue vertical line indicates mouse orthologous region to human rs1421085. ChromHMM (Ernst & Kellis, 2012) was used to annotate adipogenesis in 3T3-L1s (Mikkelsen *et al.*, 2010) at 4 time-points during differentiation corresponding to proliferating cells 2 days before adipogenic induction (d-2), in confluent cells at the induction of differentiation (d0), in early differentiating cells 2 days post induction (d2) and at terminal differentiation at day 7 (d7). Data was visualized using the UCSC Genome Browser (Kent *et al.*, 2002).

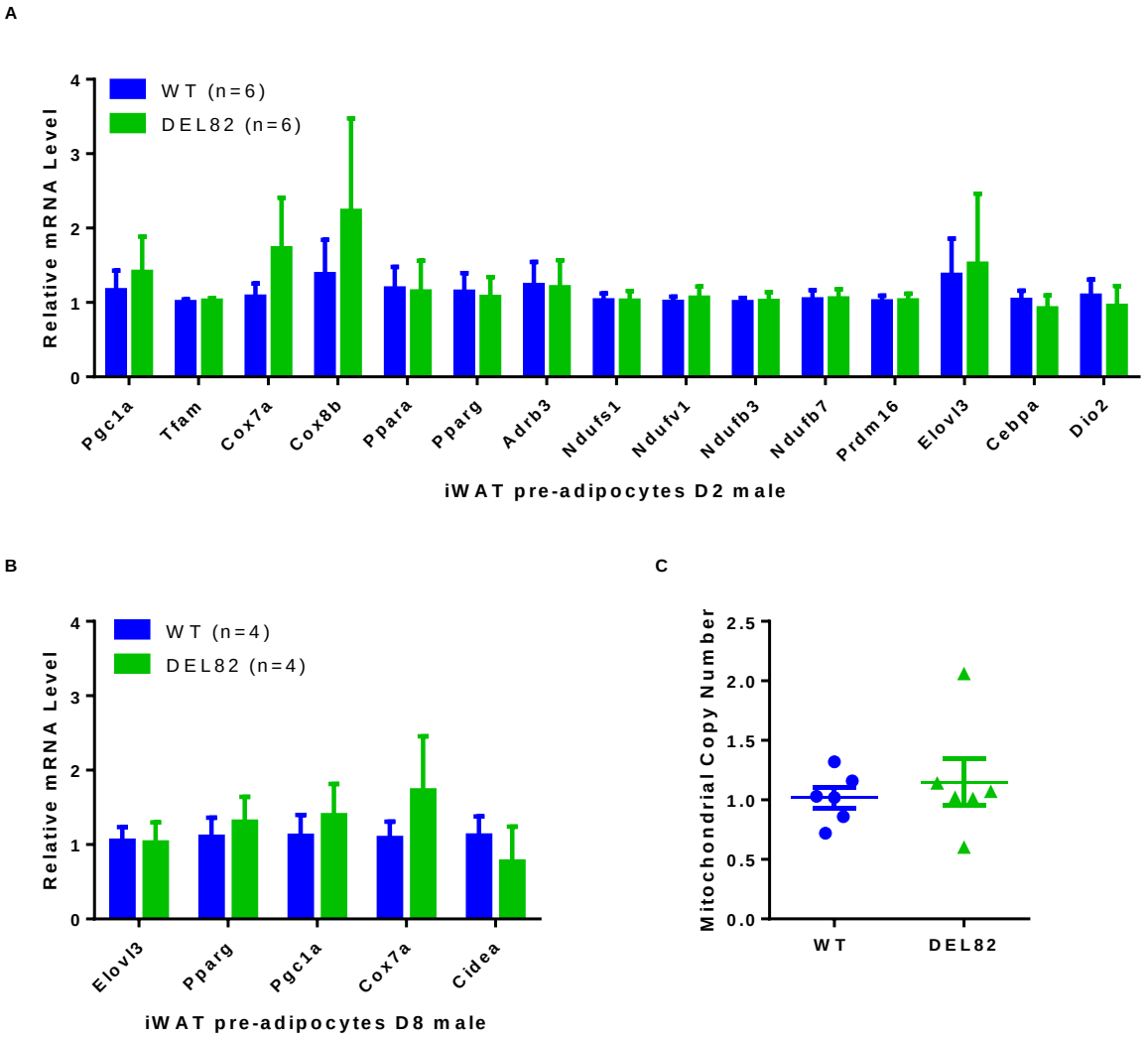


Fig. S. 3 | Gene expression patterns in isolated pre-adipocytes of DEL82 and WT males.
iWAT pre-adipocytes were isolated from males at 6-8 weeks of age and differentiated *in vitro* in absence of β-adrenergic stimulation. Cells were harvested for gene expression analysis at D2 (A) and at terminal differentiation D8 (B) of adipogenic induction. Gene expression was assessed using qPCR (normalised to *Canx*). (C) Mitochondrial content at terminal differentiation pre-adipocytes (D8 of differentiation). N numbers for each group are as indicated in the figure. Statistical significance was assessed using Student's t-test; data are expressed as mean ± SEM.

Table S. 1 | Mitochondrial genes identified by IRX3 ChIP-seq in gWAT.
List of mitochondrial genes that were identified by PANTHER (Mi *et al.*, 2017) to be overrepresented in ChIP-seq data set. Differentiation D1, n=1.

	Gene Name
Abcf2	ATP-binding cassette sub-family F member 2
Acad11	Acyl-CoA dehydrogenase family member 11
Acn9	Protein ACN9 homolog, mitochondrial
Acsf3	Acyl-CoA synthetase family member 3, mitochondrial
Agap2	Arf-GAP with GTPase, ANK repeat and PH domain-containing protein 2
Alkbh3	Alpha-ketoglutarate-dependent dioxygenase alkB homolog 3
Alkbh7	Alpha-ketoglutarate-dependent dioxygenase alkB homolog 7, mitochondrial
Atad1	ATPase family AAA domain-containing protein 1
Atg13	Autophagy-related protein 13
Atp5d	ATP synthase subunit delta, mitochondrial
Atp5e	ATP synthase subunit epsilon, mitochondrial
Atp5g3	ATP synthase F(0) complex subunit C3, mitochondrial
Atp5l	ATP synthase subunit g, mitochondrial
Atp6v1a	V-type proton ATPase catalytic subunit A
Bad	Bcl2 antagonist of cell death
Bag1	BAG family molecular chaperone regulator 1
Bag5	BAG family molecular chaperone regulator 5
Bcl2l13	Bcl-2-like protein 13
Bcs1l	Mitochondrial chaperone BCS1
Cars2	Probable cysteine--tRNA ligase, mitochondrial
Casp8ap2	CASP8-associated protein 2
Ccdc51	Coiled-coil domain-containing protein 51
Cdk7	Cyclin-dependent kinase 7
Cdkn2a	Cyclin-dependent kinase inhibitor 2A, isoform 3
Chchd1	Coiled-coil-helix-coiled-coil-helix domain-containing protein 1
Chchd2	Coiled-coil-helix-coiled-coil-helix domain-containing protein 2, mitochondrial

Chchd5	Coiled-coil-helix-coiled-coil-helix domain-containing protein 5
Chmp2b	Charged multivesicular body protein 2b
Clic1	Chloride intracellular channel protein 1
Cltc	Clathrin heavy chain 1
Cmc2	COX assembly mitochondrial protein 2 homolog
Cmc2	Calcium-binding mitochondrial carrier protein Aralar2
Coa5	Cytochrome c oxidase assembly factor 5
Coq4	Ubiquinone biosynthesis protein COQ4 homolog, mitochondrial
Coq5	2-methoxy-6-polyprenyl-1,4-benzoquinol methylase, mitochondrial
Cox11	Cytochrome c oxidase assembly protein COX11, mitochondrial
Cox20	Cytochrome c oxidase protein 20 homolog
Crat	Carnitine O-acetyltransferase
Creb1	Cyclic AMP-responsive element-binding protein 1
Cryab	Alpha-crystallin B chain
Cstad	CSA-conditional, T cell activation-dependent protein
Cyp24a1	1,25-dihydroxyvitamin D(3) 24-hydroxylase, mitochondrial
Ddah2	N(G),N(G)-dimethylarginine dimethylaminohydrolase 2
Dlst	Dihydrolipoyllysine-residue succinyltransferase component of 2-oxoglutarate dehydrogenase complex, mitochondrial
Dna2	DNA replication ATP-dependent helicase/nuclease DNA2
Dnajc11	DnaJ homolog subfamily C member 11
Dnajc5	DnaJ homolog subfamily C member 5
Dpysl2	Dihydropyrimidinase-related protein 2
Drg2	Developmentally-regulated GTP-binding protein 2
Dus2l	tRNA-dihydrouridine(20) synthase [NAD(P)+]-like
Ech1	Delta(3,5)-Delta(2,4)-dienoyl-CoA isomerase, mitochondrial
Elac2	Zinc phosphodiesterase ELAC protein 2
Emc8	ER membrane protein complex subunit 8
Etfb	Electron transfer flavoprotein subunit beta

Fam32a	Protein FAM32A
Fastk	Fas-activated serine/threonine kinase
Fbxo7	F-box only protein 7
Fdx1l	Adrenodoxin-like protein, mitochondrial
Fen1	Flap endonuclease 1
Fkbp8	Peptidyl-prolyl *cis*-trans isomerase FKBP8
Gbas	Protein NipSnap homolog 2
Gck	Glucokinase
Glod4	Glyoxalase domain-containing protein 4
Glud1	Glutamate dehydrogenase 1, mitochondrial
Gnpat	Dihydroxyacetone phosphate acyltransferase
Gtpbp10	GTP-binding protein 10
Hars2	Probable histidine--tRNA ligase, mitochondrial
Hax1	HCLS1-associated protein X-1
Hebp1	Heme-binding protein 1
Herc2	E3 ubiquitin-protein ligase HERC2
Hspa4	Heat shock 70 kDa protein 4
Hspd1	60 kDa heat shock protein, mitochondrial
Hspe1	10 kDa heat shock protein, mitochondrial
Htra2	Serine protease HTRA2, mitochondrial
Ict1	Peptidyl-tRNA hydrolase ICT1, mitochondrial
Idh1	Isocitrate dehydrogenase [NADP] cytoplasmic
Iscu	Iron-sulfur cluster assembly enzyme ISCU, mitochondrial
Kank2	KN motif and ankyrin repeat domain-containing protein 2
Letm2	LETM1 domain-containing protein LETM2, mitochondrial
Letmd1	LETM1 domain-containing protein 1
Lias	Lipoyl synthase, mitochondrial
Lipt1	Lipoyltransferase 1, mitochondrial

Lyrm4	LYR motif-containing protein 4
Map1lc3b	Microtubule-associated proteins 1A/1B light chain 3B
Mars2	Methionine--tRNA ligase, mitochondrial
Mcat	Mitochondrial carnitine/acylcarnitine carrier protein
Mcat	Malonyl-CoA-acyl carrier protein transacylase, mitochondrial
Mcee	Methylmalonyl-CoA epimerase, mitochondrial
Mcl1	Induced myeloid leukemia cell differentiation protein Mcl-1 homolog
Mff	Mitochondrial fission factor
Mgme1	Mitochondrial genome maintenance exonuclease 1
Micu2	Calcium uptake protein 2, mitochondrial
Micu3	Calcium uptake protein 3, mitochondrial
Mlxip	MLX-interacting protein
Mmaa	Methylmalonic aciduria type A homolog, mitochondrial
Moap1	Modulator of apoptosis 1
Mpc1	Mitochondrial pyruvate carrier 1
Mpc2	Mitochondrial pyruvate carrier 2
Mpdu1	Mannose-P-dolichol utilization defect 1 protein
Mpst	3-mercaptopyruvate sulfurtransferase
Mrm1	rRNA methyltransferase 1, mitochondrial
Mrpl1	39S ribosomal protein L1, mitochondrial
Mrpl12	39S ribosomal protein L12, mitochondrial
Mrpl13	39S ribosomal protein L13, mitochondrial
Mrpl16	39S ribosomal protein L16, mitochondrial
Mrpl18	39S ribosomal protein L18, mitochondrial
Mrpl20	39S ribosomal protein L20, mitochondrial
Mrpl21	39S ribosomal protein L21, mitochondrial
Mrpl28	39S ribosomal protein L28, mitochondrial
Mrpl30	39S ribosomal protein L30, mitochondrial

Mrpl32	39S ribosomal protein L32, mitochondrial
Mrpl34	39S ribosomal protein L34, mitochondrial
Mrpl40	39S ribosomal protein L40, mitochondrial
Mrpl51	39S ribosomal protein L51, mitochondrial
Mrpl52	39S ribosomal protein L52, mitochondrial
Mrpl53	39S ribosomal protein L53, mitochondrial
Mrps11	28S ribosomal protein S11, mitochondrial
Mrps16	28S ribosomal protein S16, mitochondrial
Mrps18a	28S ribosomal protein S18a, mitochondrial
Mrps18c	28S ribosomal protein S18c, mitochondrial
Mrps2	28S ribosomal protein S2, mitochondrial
Mrps23	28S ribosomal protein S23, mitochondrial
Mrps27	28S ribosomal protein S27, mitochondrial
Mrps33	28S ribosomal protein S33, mitochondrial
Mrps34	28S ribosomal protein S34, mitochondrial
Mrps35	28S ribosomal protein S35, mitochondrial
Mtif2	Translation initiation factor IF-2, mitochondrial
Mtor	Serine/threonine-protein kinase mTOR
Myc	Myc proto-oncogene protein
Myg1	UPF0160 protein MYG1, mitochondrial
Myh10	Myosin-10
Myo19	Unconventional myosin-XIX
Napg	Gamma-soluble NSF attachment protein
Ndufa11	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 11
Ndufa12	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 12
Ndufa4	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 4
Ndufa7	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 7
Ndufaf1	Complex I intermediate-associated protein 30, mitochondrial

Ndufaf7	NADH dehydrogenase [ubiquinone] complex I, assembly factor 7
Ndufb3	NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 3
Ndufb7	NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 7
Ndufb9	NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 9
Ndufs1	NADH-ubiquinone oxidoreductase 75 kDa subunit, mitochondrial
Ndufs2	NADH dehydrogenase [ubiquinone] iron-sulfur protein 2, mitochondrial
Ndufs3	NADH dehydrogenase [ubiquinone] iron-sulfur protein 3, mitochondrial
Nln	Neurolysin, mitochondrial
Nme3	Nucleoside diphosphate kinase 3
Nme6	Nucleoside diphosphate kinase 6
Nox4	NADPH oxidase 4
Nr3c1	Glucocorticoid receptor
Nrd1	Nardilysin
Nt5m	5'(3')-deoxyribonucleotidase, mitochondrial
Nthl1	Endonuclease III-like protein 1
Nudt1	7,8-dihydro-8-oxoguanine triphosphatase
Nudt9	ADP-ribose pyrophosphatase, mitochondrial
Nvl	Nuclear valosin-containing protein-like
Ogg1	N-glycosylase/DNA lyase
Oxld1	Oxidoreductase-like domain-containing protein 1
Oxnad1	Oxidoreductase NAD-binding domain-containing protein 1
Pacs2	Phosphofurin acidic cluster sorting protein 2
Park7	Protein DJ-1
Pdf	Protein Pdf
Pdss1	Decaprenyl-diphosphate synthase subunit 1
Pdss2	Decaprenyl-diphosphate synthase subunit 2
Peo1	Twinkle protein, mitochondrial
Pgs1	CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase, mitochondrial

Pgs1	Biglycan
Phb2	Prohibitin-2
Pmpca	Mitochondrial-processing peptidase subunit alpha
Pnkp	Bifunctional polynucleotide phosphatase/kinase
Pnpla7	Patatin-like phospholipase domain-containing protein 7
Pold3	DNA polymerase delta subunit 3
Polg2	DNA polymerase subunit gamma-2, mitochondrial
Ppp3ca	Serine/threonine-protein phosphatase 2B catalytic subunit alpha isoform
Prelid1	PRELI domain-containing protein 1, mitochondrial
Prkce	Protein kinase C epsilon type
Psen1	Presenilin-1
Psmb3	Proteasome subunit beta type-3
Ptrh2	Peptidyl-tRNA hydrolase 2, mitochondrial
Pus1	tRNA pseudouridine synthase A, mitochondrial
Qtrtd1	Queuine tRNA-ribosyltransferase subunit QTRTD1
Rab11a	Ras-related protein Rab-11A
Rab24	Ras-related protein Rab-24
Ranbp2	E3 SUMO-protein ligase RanBP2
Rhbdd1	Rhomboid-related protein 4
Rhoa	Transforming protein RhoA
Ripk1	Receptor-interacting serine/threonine-protein kinase 1
Rmnd1	Required for meiotic nuclear division protein 1 homolog
Rps14	40S ribosomal protein S14
Sco1	Protein SCO1 homolog, mitochondrial
Sdhaf2	Succinate dehydrogenase assembly factor 2, mitochondrial
Sdhd	Succinate dehydrogenase [ubiquinone] cytochrome b small subunit, mitochondrial
Sh3bp5	SH3 domain-binding protein 5
Shc1	SHC-transforming protein 1

Slc25a10	Mitochondrial dicarboxylate carrier
Slc25a11	Mitochondrial 2-oxoglutarate/malate carrier protein
Slc25a17	Peroxisomal membrane protein PMP34
Slc25a29	Mitochondrial carnitine/acylcarnitine carrier protein CACL
Slc25a32	Mitochondrial folate transporter/carrier
Slc25a4	ADP/ATP translocase 1
Slc25a40	Solute carrier family 25 member 40
Slirp	SRA stem-loop-interacting RNA-binding protein, mitochondrial
Spata5	Spermatogenesis-associated protein 5
Srp19	Signal recognition particle 19 kDa protein
Ssbp1	Single-stranded DNA-binding protein, mitochondrial
Stk11	Serine/threonine-protein kinase STK11
Supv3l1	ATP-dependent RNA helicase SUPV3L1, mitochondrial
Taco1	Translational activator of cytochrome c oxidase 1
Tatdn3	Putative deoxyribonuclease TATDN3
Tbc1d15	TBC1 domain family member 15
Tfb2m	Dimethyladenosine transferase 2, mitochondrial
Timm10	Mitochondrial import inner membrane translocase subunit Tim10
Timm22	Mitochondrial import inner membrane translocase subunit Tim22
Timm9	Mitochondrial import inner membrane translocase subunit Tim9
Tmem11	Transmembrane protein 11, mitochondrial
Tmem186	Transmembrane protein 186
Tmem65	Transmembrane protein 65
Tmem70	Transmembrane protein 70, mitochondrial
Tnnc1	Troponin C, slow skeletal and cardiac muscles
Tomm22	Mitochondrial import receptor subunit TOM22 homolog
Tomm34	Mitochondrial import receptor subunit TOM34
Tomm70a	Mitochondrial import receptor subunit TOM70

Trmu	Mitochondrial tRNA-specific 2-thiouridylase 1
Tshz3	Teashirt homolog 3
Tst	Thiosulfate sulfurtransferase
Ttc3	E3 ubiquitin-protein ligase TTC3
Tusc2	Tumor suppressor candidate 2
Txnrd2	Thioredoxin reductase 2, mitochondrial
Ubb	Polyubiquitin-B
Ung	Uracil-DNA glycosylase
Uqcrh	Cytochrome b-c1 complex subunit 6, mitochondrial
Usmg5	Up-regulated during skeletal muscle growth protein 5
Usp30	Ubiquitin carboxyl-terminal hydrolase 30
Vdac1	Voltage-dependent anion-selective channel protein 1
Xpnpep3	Probable Xaa-Pro aminopeptidase 3
Ykt6	Synaptobrevin homolog YKT6
Yme1l1	ATP-dependent zinc metalloprotease YME1L1
Yrdc	YrdC domain-containing protein, mitochondrial
Zfp202	Protein Zfp202

Table S. 2 | Mitochondrial genes identified by IRX3 ChIP-seq in iWAT D1.
List of mitochondrial genes that was identified by PANTHER (Mi *et al.*, 2017) to be overrepresented in IRX3 ChIP-seq data set of iWAT D1, n=2.

	Gene Name
Ict1	Peptidyl-tRNA hydrolase ICT1, mitochondrial
Abhd10	Mycophenolic acid acyl-glucuronide esterase, mitochondrial
Acad11	Acyl-CoA dehydrogenase family member 11
Ak3	GTP:AMP phosphotransferase AK3, mitochondrial
Aldh7a1	Alpha-aminoadipic semialdehyde dehydrogenase
Alkbh3	Alpha-ketoglutarate-dependent dioxygenase alkB homolog 3
Atad1	ATPase family AAA domain-containing protein 1

Atp5j	ATP synthase-coupling factor 6, mitochondrial
C1qbp	Complement component 1 Q subcomponent-binding protein, mitochondrial
Chchd2	Coiled-coil-helix-coiled-coil-helix domain-containing protein 2, mitochondrial
Chchd5	Coiled-coil-helix-coiled-coil-helix domain-containing protein 5
Cltc	Clathrin heavy chain 1
Cox16	Cytochrome c oxidase assembly protein COX16 homolog, mitochondrial
Ddah1	N(G),N(G)-dimethylarginine dimethylaminohydrolase 1
Gfm2	Ribosome-releasing factor 2, mitochondrial
Hemk1	HemK methyltransferase family member 1
Hlcs	Biotin--protein ligase
Mrpl21	39S ribosomal protein L21, mitochondrial
Mrpl30	39S ribosomal protein L30, mitochondrial
Mrpl44	39S ribosomal protein L44, mitochondrial
Mrpl53	39S ribosomal protein L53, mitochondrial
Mrps18c	28S ribosomal protein S18c, mitochondrial
Mrps21	28S ribosomal protein S21, mitochondrial
Mrps27	28S ribosomal protein S27, mitochondrial
Mrps34	28S ribosomal protein S34, mitochondrial
Mtif2	Translation initiation factor IF-2, mitochondrial
Mtx2	Metaxin-2
Mtx3	Metaxin-3
Myh10	Myosin-10
Ndufa12	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 12
Ndufaf1	Complex I intermediate-associated protein 30, mitochondrial
Ndufaf7	NADH dehydrogenase [ubiquinone] complex I, assembly factor 7
Ndufb3	NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 3
Ndufs1	NADH-ubiquinone oxidoreductase 75 kDa subunit, mitochondrial
Ndufs3	NADH dehydrogenase [ubiquinone] iron-sulfur protein 3, mitochondrial

Nfs1	Cysteine desulfurase, mitochondrial
Nme1	Nucleoside diphosphate kinase A
Nol7	Nucleolar protein 7
Nudt8	Nucleoside diphosphate-linked moiety X motif 8, mitochondrial
Nvl	Nuclear valosin-containing protein-like
Park7	Protein DJ-1
Pde12	2',5'-phosphodiesterase 12
Pdf	Peptide deformylase, mitochondrial
Pkm	Pyruvate kinase PKM
Prdx5	Peroxiredoxin-5, mitochondrial
Psen1	Presenilin-1
Ptcd2	Pentatricopeptide repeat-containing protein 2, mitochondrial
Ptrh2	Peptidyl-tRNA hydrolase 2, mitochondrial
Rab11a	Ras-related protein Rab-11A
Rmdn1	Regulator of microtubule dynamics protein 1
Rnf185	E3 ubiquitin-protein ligase RNF185
Romo1	Reactive oxygen species modulator 1
Slc25a25	Calcium-binding mitochondrial carrier protein SCaMC-2
Slc44a1	Choline transporter-like protein 1
Ssbp1	Single-stranded DNA-binding protein, mitochondrial
Stoml2	Stomatin-like protein 2, mitochondrial
Taco1	Translational activator of cytochrome c oxidase 1
Timmdc1	Complex I assembly factor TIMMDC1, mitochondrial
Tmem186	Transmembrane protein 186
Tomm70a	Mitochondrial import receptor subunit TOM70
Trmt10c	Mitochondrial ribonuclease P protein 1
Tufm	Elongation factor Tu, mitochondrial
Tusc2	Tumor suppressor candidate 2

Ubiad1	UbiA prenyltransferase domain-containing protein 1
Usp30	Ubiquitin carboxyl-terminal hydrolase 30
Vars	Valine--tRNA ligase
Vwa8	von Willebrand factor A domain-containing protein 8

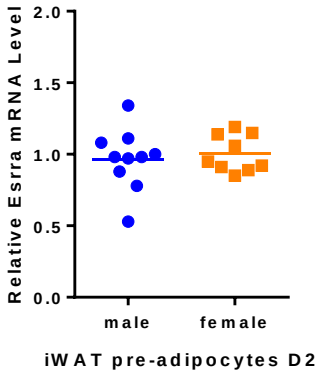
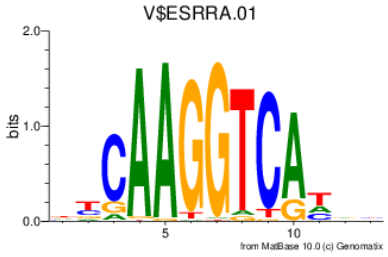
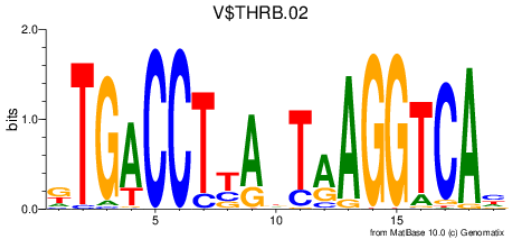
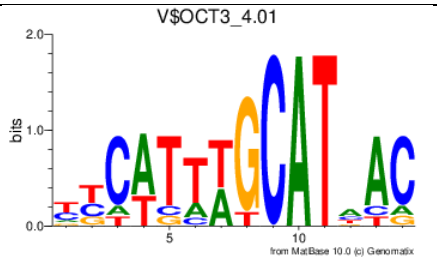
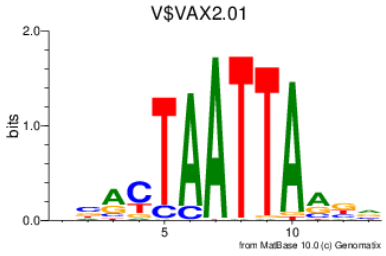
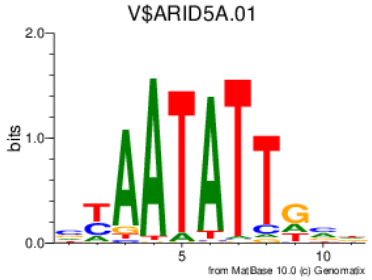
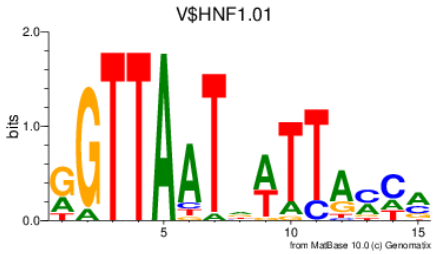
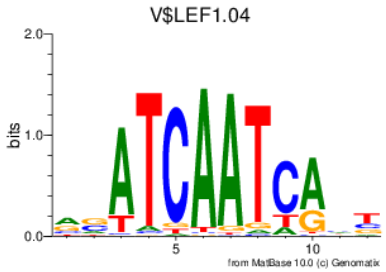
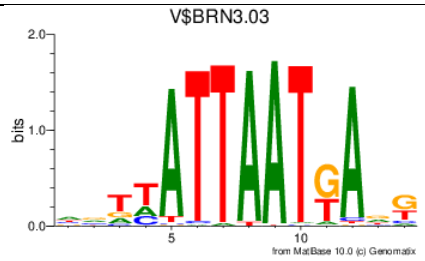
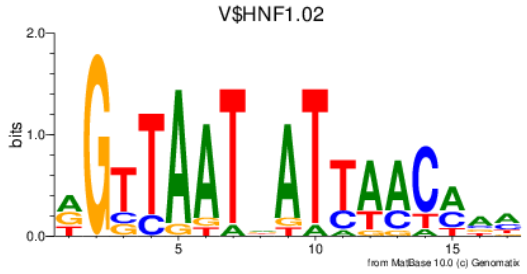


Fig. S. 4 | Esrra mRNA expression is comparable between female and male pre-adipocytes. Primary pre-adipocytes were isolated from male and female iWAT. Cells were treated with adipogenic cocktail for two days before RNA was isolated. Gene expression of *Esrra* was measured using qPCR (normalised to *Canx*). Horizontal line represents mean.

Table S. 3 | SNPInspector revealed several conserved motifs overlapping rs1421085.

	Motif
ESRRA	
THRB	

OCT3	
VAX2	
ARID5A	
HNF1	
LEF1	

BRN3	 V$BRN3.03 bits 2.0 1.0 0.0 5 10 from MatBase 10.0 (c) Genomatix
HNF1	 V$HNF1.02 bits 2.0 1.0 0.0 5 10 15 from MatBase 10.0 (c) Genomatix