

The role and transcriptional regulation of FAM46A in adipocyte differentiation

Présentée le 15 juillet 2021

Faculté des sciences de la vie
Unité du Prof. Deplancke
Programme doctoral en approches moléculaires du vivant

pour l'obtention du grade de Docteur ès Sciences

par

Ann-Kristin HOV

Acceptée sur proposition du jury

Dr C. Canto Alvarez, président du jury
Prof. B. Deplancke, Dr N. Gheldof, directeurs de thèse
Prof. F. Verdeguer, rapporteur
Prof. L. Fajas Coll, rapporteur
Prof. K. Schoonjans, rapporteuse

*All that is gold does not glitter,
Not all those who wander are lost;
The old that is strong does not wither,
Deep roots are not reached by the frost.*

— J.R.R. Tolkien, *The Riddle of Strider*

*To my parents,
And fellow neurodivergents.
We got this.*

Acknowledgements

They say the happiest people are ones who practice gratitude, and I have many people to be grateful for throughout this PhD journey.

First of all, I would like to thank my jury president Dr. Carles Cantó Alvarez and external jury members Prof. Lluís Fajas Coll and Prof. Francisco Verdeguer for reviewing this thesis. A big thank you to Prof. Kristina Schoonjans, not only as a jury member but as a wonderful mentor who gave me advice when I needed it the most. Thank you to my thesis director Prof. Bart Deplancke, who welcomed me into his lab and lab meetings with fruitful discussions and feedback. You were always so kind to include me, so an extra thank you for the lab retreats and laughs over food and drinks. The biggest thank you needs to go to my second thesis director Dr. Nele Gheldof. Thank you for taking a chance on me, for your advice, patience, flexibility and incredible support over the years, especially during this last one. It was not easy for many reasons, and you always tried to find ways make it easier. For that I will always be grateful.

To my collaborators, Dr. Jan Procházka, thank you for our discussions and Goretti, thank you for all your amazing work and for contributing to my PhD. I wish we could have met in person! Many thank you's go to the wonderful team at the Genomics facility at NIHS led by Patrick. Solenn, Coralie, Sylviane, Lorane and Alix you are absolute rockstars for helping me with my work and it would not exist with you.

To my other colleagues at NIHS, it has been a pleasure to know and on occasion work with you. Many of you made the workday so much better with your positive attitudes and chats. Thank you to Jörg and the previous members of the Nutrition and Metabolic Health/Metabolic Phenotyping group for welcoming me when I started; Armand, Jérôme and Christian C., thank you for all the scientific advice and conversations. A big thank you to Christian D. and other CMP members for welcoming me into the group in the last years.

Acknowledgements

Patricia, thank you so much for all your help with my work and listening to me when I needed it. I am so grateful. To the other office ladies, Corinne and Ornella, thank you for all your support, hilarious conversation and for your empathy. I am sorry I cried so much at my desk. Margherita/Margoola, thank you for being you, and everything that comes with it. I am honoured to be your friend. Matt (you know I am crying right now), thank you for our cell culture chats/sings, your continued support and general amazing ways. Angelique and Miriam, thank you for covfefe breaks and silly times. Phil and Chris, for reminding me there's always time for happy hour (cheeeers!). P1 Robin, for letting me play music and hearing your incredible singing voice. Vincenzo, thank you for the coffee breaks from reality. Tanja, for your continued support, even across Switzerland, and traditions that I love.

To my EPFL buddies. Mateusz, thank you for reminding me what hard work (and play) looks like. Amber, thank you for our talks and your knitting enthusiasm. Marjan and Magda, I am beyond grateful to you both for listening and advising me through an incredibly tough time. I would not be here today without you. Arielle, thank you for keeping grounded and for many conversations shared over delicious food.

To my chosen little family of CatCave9/the Catalyst. Thank you for accepting me, for reminding me who I am and for laughs so hard I cried. I can only smile when I think of you all. Daisy, thank you for supporting me through everything, basically. You make it easier to exist (plus you send amazing GIFs). Holly, thank you for always being there, even when we would lose contact for months. Enrico and Ben, for being there, literally until the end. Your support means everything.

Thank you to all the people I met during these past four years who made my experience in Lausanne what it is. I'm sorry our meeting was not impactful enough for me to acknowledge you personally.

Acknowledgements

Last, but not least, with all the gratitude I have, thank you to my incredibly supportive family. Charles, you tried to understand what I was doing and that is all I can ask for. To my wonderful parents, Ketil and Petra. Thank you for letting me pursue what I wanted with no judgement or criticism, for your endless support and wisdom. You have taught me so much, and I will keep these lessons with me always. I would not be who I am without your guidance and example (and coffee machine). Jeg er glad i dere.

- AK

Abstract

Considered one of the top five preventable risks for mortality, obesity is currently a leading global health concern. Available treatments include undergoing a weight-loss regimen. However, successful weight-loss is variable, where the heritability of obesity is considered part of the cause. This patient variability to weight-loss was addressed by Carayol et al. (2017), where protein plasma level changes was associated to genetic variability and changes in body mass index (BMI) under a low-calorie diet (LCD) intervention. One of the most significant association regions was found to regulate downstream gene *FAM46A*; coding for a nucleotidyl-transferase (NTase) protein. Loss of function mutations in *FAM46A* have previously been discovered in multiple human diseases, such as osteogenesis imperfecta (OI). However, there is a lack of understanding of *FAM46A*'s function, especially in the context of adipose tissue and metabolism. This thesis aimed to follow up on the Carayol et al. study by investigating the transcriptional regulation of *FAM46A* and how the associated region is potentially influencing this regulation, as well as how *FAM46A* is involved in adipose tissue function.

When mapping the physical DNA interaction network around the associated region using chromosome conformation capture (3C) in human subcutaneous adipocytes (SGBS cells), no interactions between the *FAM46A* promoter and selected loci of the associated region were found. Interestingly, significant interaction peaks downstream of the *FAM46A* promoter overlapped with regions of open chromatin and were enriched for transcription factor (TF) motifs that are involved in stem cell differentiation and TGF β signalling. SMAD4, part of the TGF β signalling pathways, binds to this region, thus indicating a potential regulator of *FAM46A*.

To address *FAM46A*'s potential role in adipocyte function, *FAM46A* expression was disrupted *in vitro* using differentiating adipocytes. Reduction of *FAM46A* expression in SGBS cells resulted in decreased expression of *PPAR γ* and

Abstract

CEBPA, the master regulators of adipogenesis, complemented by a reduction in lipid accumulation during differentiation. Follow up studies in murine mesenchymal stem cells (MSCs) where *Fam46a* was over-expressed were controversial as this led to a decrease in adipogenic markers (*Adipoq* and *PPAR γ*). Overall, disruption of *FAM46A* expression led to opposing changes in adipogenic potential, indicating the requirement for a more stable cell system to study *FAM46A in vitro*.

To further characterise the role of *Fam46a*, metabolic phenotyping of a *Fam46a* knockout (KO) mouse model under hyper-caloric stress was performed in collaboration with the IMG in Prague. This model revealed a significantly smaller mouse, and increased plasma alkaline phosphatase (ALP), as previously observed. When fed a high-fat diet (HFD), KO mice had reduced overall fat mass, improved glucose clearance and reduced adipocyte size in both subcutaneous and visceral white adipose tissue (subWAT/visWAT) depots. Transcriptomic analysis of these two depots revealed an up-regulation of genes expressed in ribosome related pathways in the subWAT, whereas the visWAT showed the up-regulation of fatty acid metabolism pathways, suggesting that the KO adipose tissue is more metabolically active. Other down-regulated pathways involved stem cell differentiation and BMP/TGF β -signalling.

In conclusion, this study found the possible regulation of the *FAM46A* promoter by TFs involved in adipogenesis regulation, such as SMAD proteins, as well as strengthening its possible functional involvement in the TGF β signalling by enrichment of this pathway in KO studies, as well as stem cell differentiation. This indicates that *FAM46A* may have a general role in regulating stem state and subsequent differentiation.

Keywords: *FAM46A*, adipogenesis, adipose tissue, 3C, gene expression, SGBS, metabolism, transcriptomics.

Résumé

Considérée comme l'un des cinq principaux risques évitables de mortalité, l'obésité est actuellement un problème de santé mondial majeur. Les traitements disponibles comprennent un régime de perte de poids. Cependant, le succès de la perte de poids est variable, l'héritabilité de l'obésité étant considérée comme une partie de la cause. Cette variabilité des patients face à la perte de poids a été abordée par Carayol et al. (2017), où les modifications du taux plasmatique de protéines ont été associées à la variabilité génétique et aux modifications de l'IMC dans le cadre d'un régime hypocalorique. L'une des régions d'association les plus significatives a été trouvée pour réguler le gène en aval FAM46A ; codant pour une protéine nucléotidyl-transférase. Des mutations de perte de fonction dans FAM46A ont déjà été découvertes dans de nombreuses maladies humaines, comme l'ostéogénèse imparfaite. Cependant, il existe un manque de compréhension de la fonction de FAM46A, notamment dans le contexte du tissu adipeux et du métabolisme. Cette thèse avait pour but de faire suite à l'étude de Carayol et al. en étudiant la régulation transcriptionnelle de FAM46A et la façon dont la région associée peut influencer cette régulation, ainsi que la façon dont FAM46A est impliqué dans la fonction du tissu adipeux.

Lors de la cartographie du réseau d'interactions physiques de l'ADN autour de la région associée en utilisant la capture de confirmation du chromosome (3C) dans les adipocytes sous-cutanés humains (cellules SGBS), aucune interaction entre le promoteur de FAM46A et les régions sélectionnées de la région associée n'a été trouvée. Il est intéressant de noter que les pics d'interaction significatifs en aval du promoteur de FAM46A chevauchaient des régions de chromatine ouverte et étaient enrichis en motifs de facteurs de transcription (TF) impliqués dans la différenciation des cellules souches et la signalisation du TGF. SMAD4, qui fait partie des voies de signalisation du TGF, se lie à cette région, indiquant ainsi un régulateur potentiel de FAM46A.

Résumé

Pour étudier le rôle potentiel de FAM46A dans la fonction adipocytaire, l'expression de FAM46A a été perturbée in vitro en utilisant des adipocytes en cours de différenciation. La réduction de l'expression de FAM46A dans les cellules SGBS a entraîné une diminution de l'expression de PPAR γ et CEBPA, les principaux régulateurs de l'adipogenèse, complétée par une réduction de l'accumulation de lipides pendant la différenciation. Des études de suivi sur des cellules souches mésenchymateuses (MSC) murines où Fam46a a été surexprimé ont été controversées car cela a conduit à une diminution des marqueurs adipogéniques (Adipoq et PPAR γ). Dans l'ensemble, la perturbation de l'expression de FAM46A a conduit à des changements opposés dans le potentiel adipogène, indiquant la nécessité d'un système cellulaire plus stable pour étudier FAM46A in vitro.

Pour caractériser davantage le rôle de Fam46a, le phénotypage métabolique d'un modèle de souris knockout (KO) Fam46a sous stress hypercalorique a été réalisé en collaboration avec l'IMG de Prague. Ce modèle a révélé une souris significativement plus petite, et une augmentation de la phosphatase alcaline plasmatique (ALP), comme observé précédemment. Lorsqu'elles sont soumises à un régime riche en graisses (HFD), les souris KO présentent une réduction de la masse grasse globale, une amélioration de la clairance du glucose et une réduction de la taille des adipocytes dans les dépôts de tissu adipeux blanc sous-cutané et viscéral (subWAT/visWAT). L'analyse transcriptomique de ces deux dépôts a révélé une régulation positive des gènes exprimés dans les voies liées aux ribosomes dans le subWAT, tandis que le visWAT a montré une régulation positive des voies du métabolisme des acides gras, ce qui suggère que le tissu adipeux KO est plus actif sur le plan métabolique. D'autres voies régulées à la hausse impliquaient la différenciation des cellules souches et la signalisation BMP/TGF.

En conclusion, cette étude a mis en évidence la régulation possible du promoteur FAM46A par des TFs impliqués dans la régulation de

Résumé

l'adipogenèse, comme les protéines SMAD, ainsi que le renforcement de son implication fonctionnelle possible dans la signalisation du TGF par l'enrichissement de cette voie dans les études KO, ainsi que la différenciation des cellules souches. Cela indique que FAM46A peut avoir un rôle général dans la régulation de l'état de la tige et de la différenciation ultérieure.

Mots clés: FAM46A, adipogenèse, tissu adipeux, 3C, expression génétique, SGBS, métabolisme, transcriptomique.

Table of Contents

Acknowledgements	ii
Abstract	5
Résumé.....	vii
Table of Contents.....	x
General introduction	2
1. Adipose tissue: a plastic endocrine organ	2
1.1. Adipogenesis: from precursor to mature cell	5
1.1.1. Lineage tracing of adipocytes	5
1.1.2. Mesenchymal stem cells: to be fat or to be bone?.....	6
1.1.3. Early adipocyte commitment: pre-adipocytes	8
1.1.4. Terminal differentiation into adipocytes	10
1.1.5. Neutral lipid measurement	11
1.2. Adipose tissue pathology: obesity.....	12
1.2.1. Weight-loss as an obesity treatment.....	12
1.3. FAM46A: a novel adipocyte regulator?.....	14
1.3.1. Role of genetic variation on weight-loss response.....	14
1.3.2. Family with sequence similarity 46	14
1.3.3. Scope of thesis	18
2. Transcriptional regulation of the FAM46A promoter during human adipocyte differentiation.....	20
2.1. Abstract.....	20
2.2. Introduction	22
2.2.1. The 3D genome	22
2.2.2. Topologically associating domains.....	22
2.2.3. Gene expression initiation and the role of chromatin structure.....	23
2.2.4. Chromatin conformation capture.....	24
2.2.5. Research question	25
2.3. Results	27
2.3.1. Optimisation of CaptureC digestion efficiency in SGBS cells	27
2.3.2. CaptureC viewpoint selection and design optimisation.....	30
2.3.3. CaptureC validation	33
2.3.4. Interaction peak calling using peakC	37
2.3.5. Analysis of FAM46A viewpoint for regulatory elements	40
2.3.6. Motif enrichment and pathway analysis of interaction peaks.....	44
2.4. Discussion	51
2.5. Materials and methods	58
2.5.1. Cell culture	58
2.5.2. CaptureC library preparation: 3C libraries	58
2.5.3. CaptureC library preparation: sequence adaptors and capture	60
2.5.4. Sequencing parameters.....	61
2.5.5. Data analysis	61

TOC

3. Functional characterisation of FAM46A in human adipocytes and murine MSCs.....	64
3.1. Abstract.....	64
3.2. Introduction	66
3.2.1. Cellular models for adipogenesis.....	66
3.2.2. Research question	67
3.3. Results	69
3.3.1. FAM46A expression in adipocytes	69
3.3.2. Disruption of FAM46A expression during adipogenesis.....	69
3.3.3. Expression in human adipocyte samples and mouse depots	75
3.3.4. Fam46a expression in murine MSCs.....	77
3.3.5. Disruption of Fam46a expression in murine MSCs.....	80
3.4. Discussion	84
3.5. Materials and methods	87
3.5.1. SGBS cell culture	87
3.5.2. SGBS qPCR	87
3.5.3. Transfection of siRNA and CMV-plasmids	87
3.5.4. Lipid quantification	88
3.5.5. C3H10T1/2 cell culture	88
3.5.6. BM-MSCs cell culture	89
3.5.7. Murine MSCs qPCR.....	89
3.5.8. Lipofectamine transfection of siRNA and CMV-plasmids.....	89
3.5.9. ORO and ALP activity staining	90
4. Functional characterisation of Fam46a in vivo under metabolic challenge.....	92
4.1. Abstract.....	92
4.2. Introduction	94
4.2.1. Mouse models for metabolic studies.....	94
4.2.2. Skeletal dysplasia model is a Fam46a loss of function mutation.....	94
4.2.3. Research question	95
4.3. Results	97
4.3.1. Study design and modifications	97
4.3.2. Body weight and body composition of WT and Fam46a KO mice ...	99
4.3.3. Metabolic phenotyping: IPGTT and TSE.....	102
4.3.4. Plasma biochemistry and histology of BAT and liver	106
4.3.5. subWAT and visWAT: histology and lipid quantification.....	110
4.3.6. Transcriptomics: Differential gene expression between WT and KO ...	113
4.3.7. Pathway analysis: exploratory and specific pathways	119
4.4. Discussion	128
4.5. Materials and methods	137
4.5.1. Animal experiments	137
4.5.2. Generation of Fam46a KO mouse	137
4.5.3. Mouse diets: chow and HFD compositions.....	137
4.5.4. Mouse body weight and composition measurements	138
4.5.5. Indirect calorimetry with TSE Phenomaster.....	138
4.5.6. Intra-peritoneal glucose tolerance test.....	139

4.5.7. Plasma biochemistry measurements	139
4.5.8. Histological preparation and analysis	140
4.5.9. Library preparation for QuantSeq and sequencing	140
4.5.10. Quantseq data filtering and normalisation	141
<i>Final conclusions and outlook</i>	<i>143</i>
<i>References</i>	<i>146</i>
<i>Curriculum vitae</i>	<i>157</i>

Abbreviations and acronyms

Abbreviations

3C	Chromatin conformation capture
4C	Chromosome conformation capture-on-chip
adipo-MSCs	Adipose tissue derived mesenchymal stem cells
ALP	Alkaline phosphatase
ALT	Alanine aminotransferase
ANCOVA	Analysis of covariance
ANOVA	Analysis of variance
APCs	Adipose precursor cells
AST	Aspartate transaminase
ATAC-seq	Assay for Transposase-Accessible Chromatin using sequencing
AUC	Area under the curve
BAT	Brown adipose tissue
BCAA	Branched chain amino acids
BM-MSCs	Bone marrow mesenchymal stem cells
BMI	Body mass index
BMPs	Bone morphogenetic proteins
BW	Body weight
cAMP	Cyclic adenosine monophosphate
CEBPs	CAAT/enhancer-binding proteins
ChIP-seq	Chromatin immunoprecipitation followed by sequencing
CPM	Counts per million reads mapped
Ct	Cycle threshold
DIO	Diet-induced obesity
DNA	Deoxyribonucleic acid
ECM	Extracellular matrix
ENU	N-ethyl-N-nitrosourea
eQTL	Expression quantitative trait loci

Abbreviations

ESCs	Embryonic stem cells
FACS	Fluorescence-activated cell sorting
FAs	Fatty acids
FC	Fold-change
FDR	False discovery rate
GWAS	Genome-wide association studies
H&E	Haematoxylin and eosin
HDL	High-density lipoprotein
HFD	High-fat diet
IPGTT	Intra-peritoneal glucose tolerance test
KO	Knockout
LCD	Low-calorie diet
MIG	Multi-Image Genome
mRNA	Messenger RNA
MSCs	Mesenchymal stem cells
NGS	Next generation sequencing
NTase	Nucleotidyl-transferase
OI	Osteogenesis imperfecta
ORO	Oil red O
PFA	Paraformaldehyde
PPAR γ	Peroxisome proliferator-activated receptor- γ
pQTL	Protein quantitative trait loci
qPCR	Quantitative polymerase chain reaction
QTL	Quantitative trait loci
RER	Respiratory exchange rate
RNA	Ribonucleic acid
RT	Room temperature
scRNA-seq	Single-cell RNA sequencing
SE	Standard error

Abbreviations

SGBS	Simpson Golabi Behmel Syndrome
shRNA	Short hairpin RNA
siNEG	Negative control signal
siRNAs	Small interfering RNAs
SNPs	Single nucleotide polymorphisms
subWAT	Subcutaneous white adipose tissue
SVF	Stromal vascular fraction
TADs	Topologically associating domains
TEAD	TEA domain family
TF	Transcription factor
TFBS	Transcription factor binding sites
TG	Triglyceride
TGF β	Transforming growth factor β
TSS	Transcription start site
visWAT	Visceral white adipose tissue
WAT	White adipose tissue
WHO	World Health Organisation
WT	Wildtype

Chapter 1:

General introduction

Chapter 1

General introduction

1. Adipose tissue: a plastic endocrine organ

Adipose tissue has classically been regarded as the energy storage organ. During caloric excess, neutral lipids are stored in a specialised fat cell, or adipocyte (Birsoy et al. 2013). Reversely, during nutrient deficiency the stored lipids are accessed and broken down by lipolysis, which involves the catabolism of triacylglycerol into non-esterified fatty acids and glycerol that will supply nutrients to other tissues (Lass et al. 2011). While this is one of the core functions of a certain type of adipose tissue, it is now increasingly clear that adipose tissue is far more complex and plays a more central role in overall metabolic homeostasis and energy balance (Choe et al. 2016).

Adipose tissue can be classed into two major groups; white adipose tissue (WAT) and brown adipose tissue (BAT). WAT is known for excess energy storage, insulation and protection of vital organs (Saely et al. 2012). Although able to store lipids, BAT has main roles in non-shivering thermogenesis and energy expenditure (Cannon & Nedergaard 2004). While a WAT adipocyte is characterised by a single unilocular lipid compartment, BAT adipocyte morphology is different with multilocular lipid droplets and an abundance of mitochondria, which contribute to the distinct deep red colour of the organ (Richard & Picard 2011; Grigoraş et al. 2018). BAT can respond to cold exposure by the uncoupling of oxidative phosphorylation in the inner mitochondrial membrane, which produces heat. This process results in high energy expenditure and is essential for maintenance of core body temperature in euthermic mammals (Gaspar et al. 2021). Recently, there has been a particular interest in a BAT-like thermogenic adipocyte known as brite/beige cells (Bartelt & Heeren 2014; Ikeda et al. 2018). Cold exposure, nutrient imbalance and exercise can induce beige adipocyte formation in WAT, either from transdifferentiation of existing white adipocytes or directly from

progenitor cells within the tissue (Paulo & B. Wang 2019). They become distinct from white adipocytes by expressing UCP1, the uncoupling protein responsible for the oxidative phosphorylation that occurs within mitochondria during thermogenesis (Young et al. 1984).

The WAT can be further subcategorised into two depots; subcutaneous white adipose tissue (subWAT) and visceral white adipose tissue (visWAT). Even though similar in appearance and morphology, these depots have their own distinct functions and roles in metabolic regulation (Hwang & Kim 2019). In humans, subWAT is found beneath the skin to insulate and protect against external damage in depots known as the gluteal, abdominal and femoral depots (Figure 1.1.1, blue text) (Bjørndal et al. 2011; Choe et al. 2016). The visWAT, which surround the organs in the intra-abdominal cavity, consist of the perirenal, retroperitoneal, omental and mesenteric depots (Figure 1.1.1, red text) (Choe et al. 2016) and promotes crosstalk between organs due to its proximity (Hwang & Kim 2019). How fat is distributed between these specific depots may determine risk for certain metabolic diseases (Grundy 2015).

Chapter 1

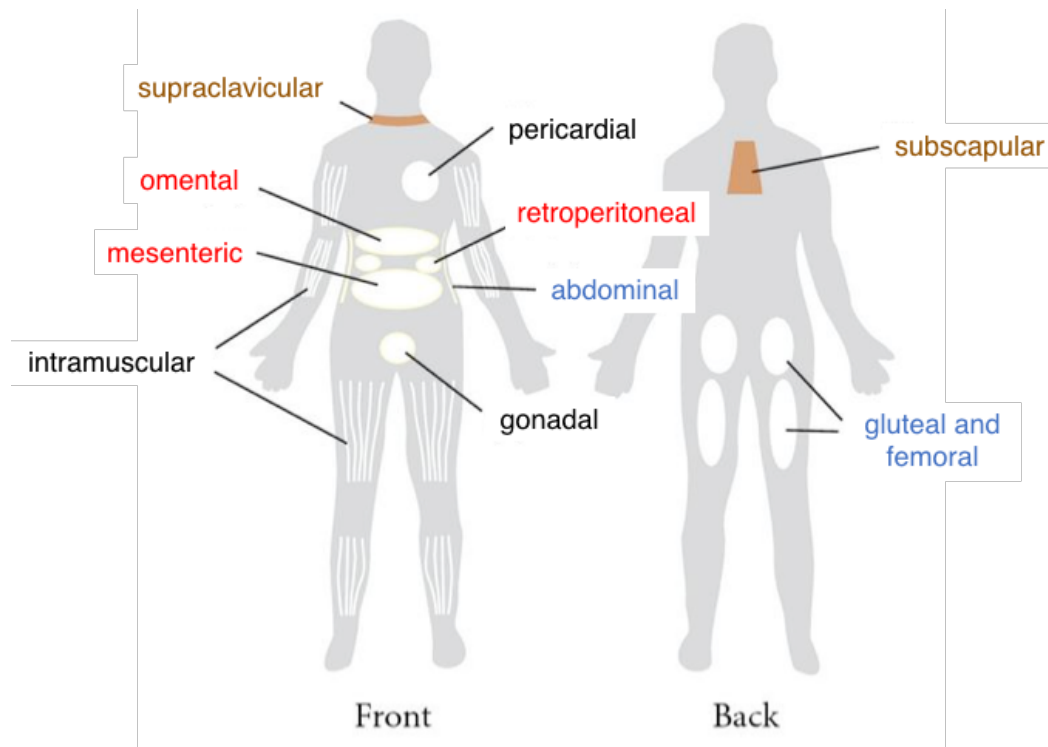


Figure 1.1.1: Overview of main human adipose tissue depots. Figure taken and adapted from Bjørndal et al. (2011).

SubWAT depots are labelled in blue, while visWAT are labelled in red. BAT depots are labelled in brown.

The discovery of leptin and adiponectin in the early 90s (Y. Zhang et al. 1994; Scherer et al. 1995) sparked the consideration of WAT as an endocrine organ. A number of proteins are secreted by the WAT, named adipokines, consisting of cytokines, hormones and metabolites (Rondinone 2006). These adipokines are essential for regulating energy balance by communication with the central nervous system and other nutrient sensing organs. Leptin, for example, suppresses food intake by acting on surface receptors in neurons in the hypothalamus, the satiety centre of the brain, and signals peripheral nerves to induce lipolysis (Fasshauer & Blüher 2015). Adiponectin is abundant in the plasma and has anti-obesity properties by its stimulation of AMP-activated protein kinase (AMPK) to induce lipid oxidation (Yamauchi et al. 2002).

Depending on the nutrient status, the adipose tissue is able to remodel and adapt by changing the number and size of adipocytes. An increase in adipocyte number is known as hyperplasia, while size increase is

hypertrophy (Yamauchi et al. 2002), and dysregulation may lead to metabolic disorders (Haczeyni et al. 2018). Both processes are regulated by external stimuli and certain genetic factors (Spiegelman & Flier 2001). Hyperplasia, and general adipocyte turnover, occurs due to a subpopulation of pre-adipocytes found in the perivascular niche of adipose tissue (Cawthorn et al. 2012). The yearly turnover rate of adipocytes has been estimated to be around 10% (Spalding et al. 2008). Fractionation of the adipose tissue into adipocytes and the stromal vascular fraction (SVF) led to the initial discovery of these stromal cells capable of differentiating into adipocytes, named adipose precursor cells (APCs) (Hollenberg & Vost 1969; Poznanski et al. 1973). The differentiation capacity of APCs has been extensively studied, where differences in APCs sourced from either subWAT or visWAT have different adipogenic properties. visWAT-APCs in general display less proliferation and adipogenic potential than subWAT-APCs (Hauner et al. 1988; Niesler et al. 1998). Understanding how hyperplasia occurs and how it may contribute disease states such as obesity is currently being studied, as it has been observed that the proliferative properties of APCs is increased in obese humans (Roncari et al. 1981).

1.1. Adipogenesis: from precursor to mature cell

1.1.1. Lineage tracing of adipocytes

To understand hyperplasia, it is important to understand where these APCs come from. The developmental origin of adipocytes is still not fully understood. It was previously believed that every adipocyte type emerged from a common mesodermal precursor. However, several studies suggest different origins depending on the adipocyte type. For example, lineage tracing experiments identified adipocyte precursors found in cephalic fat and muscle emerge from the neural crest vs the mesoderm (Billon et al. 2007; Lemos et al. 2012). There is even evidence of depot specific WAT

Chapter 1

precursors, where visWAT precursors are WT1+, supposedly deriving from the lateral plate mesoderm (Chau et al. 2014; Schoettl et al. 2018). The exact origin of subWAT precursors is still unknown. BAT precursor cells originate from the paraxial mesoderm as a mesenchymal precursor shared with skeletal muscle (Sanchez-Gurmaches & Guertin 2014), different from WAT adipocytes. In general, WAT and BAT adipocyte precursors are classically separated by the expression of myogenic factor 5 (Myf5). WAT precursors do not express Myf5, however, recent lineage studies have found that certain WAT precursors express a Myf5 depending on the anatomical location of the WAT depot (Sanchez-Gurmaches & Guertin 2014), which have been suggested to give rise to brite/beige adipocytes within WAT depots. Once developed and precursors are within their respective depots, an orchestrated program of gene expression responding to external cues allows for APCs to differentiate.

1.1.2. Mesenchymal stem cells: to be fat or to be bone?

Mesenchymal stem cells (MSCs) are mesodermal multi-potent cells essential for biological processes such as tissue regeneration and homeostasis. These cell populations are self-renewing and found in the stromal fraction of different tissues, thus explaining their differentiation plasticity. Adipocyte and osteoblast lineages can be traced back to MSCs (T. M. Liu et al. 2007). How an MSC commits to a certain cell type is dependent on a number of critical signalling pathways, such as the transforming growth factor- β (TGF β)/bone morphogenic protein (BMP) signalling pathways (Grafe et al. 2018) (Figure 1.1.2). This pathway determines certain lineages by control of key regulators of differentiation: RUNX2 for osteoblasts, SOX9 for chondrocytes and peroxisome proliferator-activated receptor- γ

(PPAR γ) and CAAT/enhancer-binding proteins (CEBPs) for adipocytes.

TGF β /BMP signalling begins by a pathway specific (TGF β or BMP) ligand binding to transmembrane serine/threonine protein kinase receptors, of which there are two: type I and type II (Y. E. Zhang 2018). This binding activates SMAD-dependent pathways. The SMAD family of proteins can be categorised into receptor regulated (R-SMADs), common partner (co-SMAD) and inhibitory (I-SMADs). R-SMADs SMAD1/5/8 are activated by BMP ligands and SMAD2/3 by TGF ligands. Unphosphorylated R-SMADs remain in the cytoplasm until a ligand prompts their phosphorylation by one of the surface receptors. Phosphorylated R-SMADs (pR-SMADs) form a trimeric complex of two pR-SMADs and one co-SMAD (SMAD4), which then translocates into the nucleus for regulation of specific downstream genes (Hill 2009; Derynck & Budi 2019), such as the aforementioned master regulators. The genes that are regulated depend on which pathway is activated and by which ligands. For adipogenesis, it was primarily thought that the TGF β was inhibitory and BMP induces the maturation of adipocytes, however, this claim has recently come into question (Li & Wu 2020).

Chapter 1

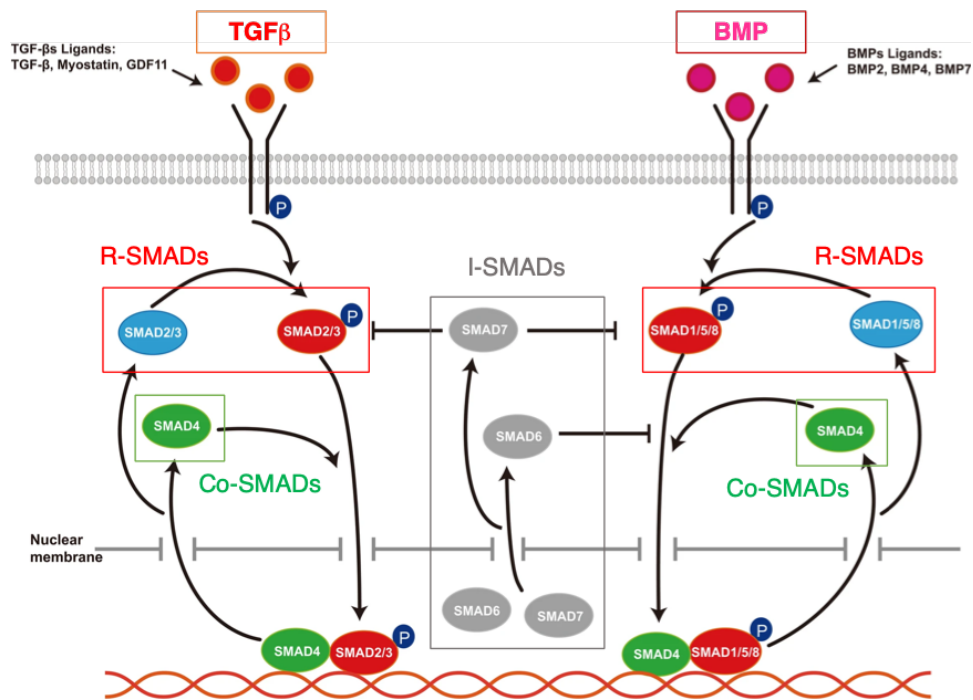


Figure 1.1.2: Overview of TGFβ and BMP signaling. Figure taken and adapted from Li & Wu (2020).

1.1.3. Early adipocyte commitment: pre-adipocytes

The first step of adipocyte differentiation is commitment of a fibroblast-like cell (such as a mesenchymal precursor or another PDGFR α expressing lineage cell) to the adipocyte lineage as pre-adipocyte. This precursor cell undergoes no morphological changes during this step, however, it can no longer commit to any other cell type (such as myoblasts, chondroblasts or osteoblasts) (Cawthorn et al. 2012). Factors that contribute to this first commitment step are several external stimuli such as aforementioned BMPs, where the SMAD4 transcription factor (TF) becomes activated by a SMAD1/5/8 complex (E. A. Wang et al. 1993). This complex translocates to the nucleus to target genes for activation, such as master regulators PPAR γ and CEBPs. Zinc-finger TF ZFP423 is enriched in committed pre-adipocytes and has also been found to have an essential role in this initial step by regulating PPAR γ expression as well as sensitising these pre-adipocytes to pro-adipogenic signals, such as aforementioned BMP

(Gupta et al. 2010). Other TF regulators include ZFP467, which promotes the expression of CEBP α (Quach et al. 2011), and KLF5 which also promotes the expression of PPAR γ and various CEBPs (Oishi et al. 2005) during this first wave of TF expression (Figure 1.1.3, “Commitment”).

Suppression of certain anti-adipogenic pathways is essential for differentiation to continue, such as the Wnt and Hedgehog pathways (Lowe et al. 2011). Over-expression of Wnt1 was shown to inhibit adipogenesis (Ross et al. 2000) as the Wnt pathway blocks the induction of PPAR γ and CEBP α (Christodoulides et al. 2009). Other repressors of adipogenesis include GATA2 and GATA3, TFs that also prevent the expression of PPAR γ (Tong et al. 2000), and PREF1, which targets certain CEBP promoters via SOX9 (Y. Wang & Sul 2009).

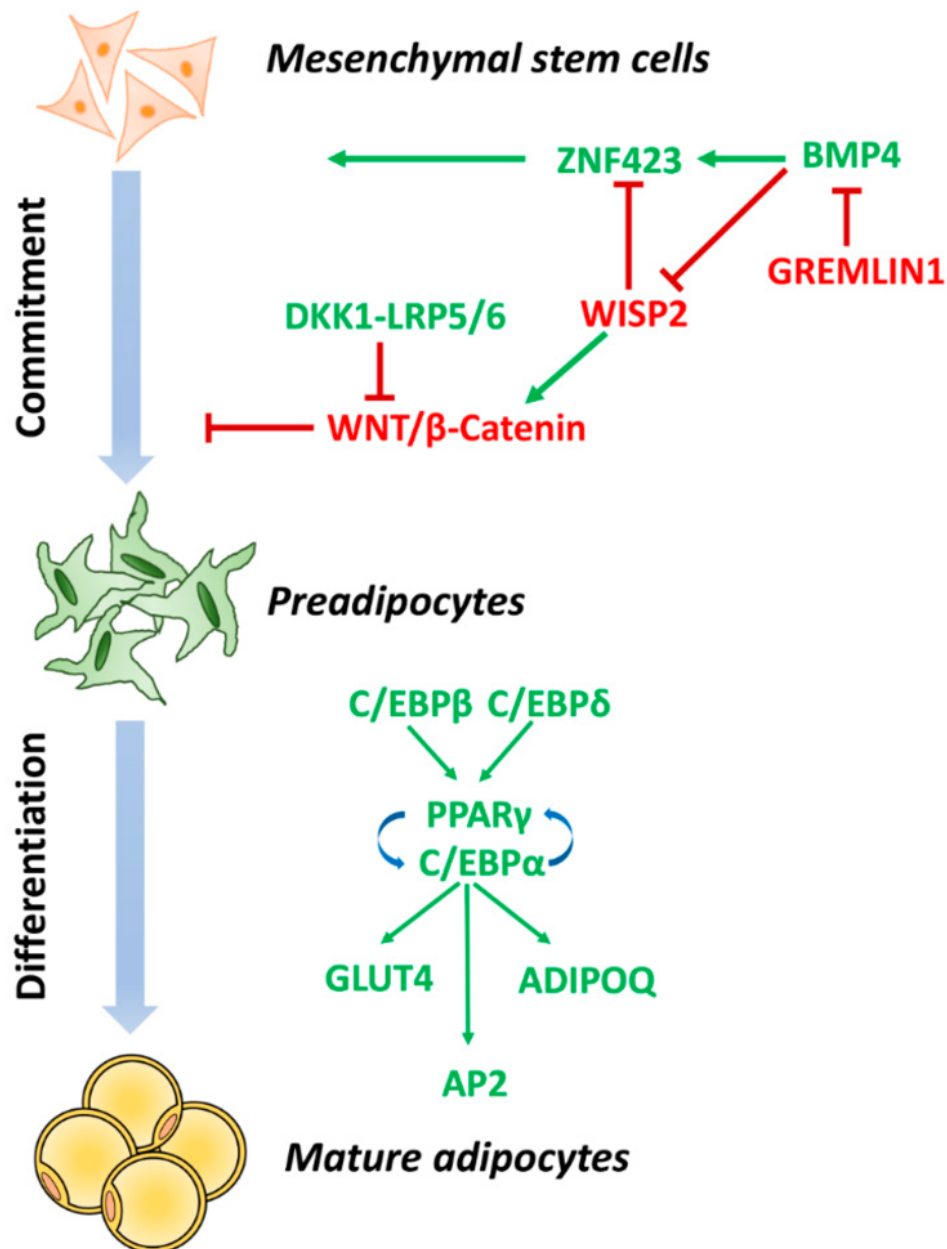


Figure 1.1.3: Overview of molecular mechanisms involved in adipogenesis. Figure taken from Longo et al. (2019)

1.1.4. Terminal differentiation into adipocytes

Multiple TF in the first wave of commitment are involved in regulating the expression of master regulators of adipogenesis, $PPAR\gamma$ and $CEBP\alpha$, whose expression increases in the second wave of TF expression (Figure 1.1.3, “Differentiation”) (Siersbæk & Mandrup 2011). In fact, $PPAR\gamma$ induces expression of $CEBP\alpha$,

which in turn activates PPAR γ for a positive feedback loop (Lowe et al. 2011). At this moment the cell cycle arrest, and differentiation into a mature adipocyte begins. Although the aforementioned description adipogenesis refers to white adipocytes, this “core” gene regulatory network of CEBPs and PPAR γ (second wave) is common to white, brown and beige adipocytes (Carobbio et al. 2019).

1.1.5. Neutral lipid measurement

Once pre-adipocytes have undergone adipogenesis, a mature adipocyte phenotype is achieved: a lipid filled cell. When studying adipogenesis, it may be important to understand how “mature” a cell is and how quickly this state is achieved, especially during disruption of suspected regulators of adipogenesis. This can be measured by the amount of neutral lipids (such as triglycerides (TGs)) accumulated within the cell model of choice during different stages of differentiation. Early stains for lipids included Nile red and oil red O (ORO). Nile red was made by oxidising Nile blue, a fluorescent histochemical stain from the early 1900s that stained neutral lipids pink and acids (such as fatty acids) blue (Smith 1908). Nile red was later introduced in 1985 and subsequently used as a stain for intracellular lipid droplets (Greenspan et al. 1985). Alternatively, ORO staining offered non-fluorescent staining of TGs and cholesteryl oleates, with the addition of quantification of staining by spectrophotometry (Ramirez Zacarias et al. 1992). Several modifications of these methods have been developed, including use on frozen tissue samples (Mehlem et al. 2013). Other fluorescent dyes to stain lipids include boron dipyrromethene 505/515 (BODIPY), a lipophilic dye. With fluorescent microscope image acquisition and specialised programs for quantification of fluorescent signal, these lipid dyes can be used to stain and

Chapter 1

efficiently quantify lipid content within fixed adipocytes to study adipogenesis (Rumin et al. 2015).

1.2. Adipose tissue pathology: obesity

In a modern lifestyle we are consistently exposed to excessive amounts lipids and nutrients. Overconsumption of such nutrient dense foods may lead to the common adipose tissue disease obesity; one of the leading public-health problems in the world. In 2014, the World Health Organisation (WHO) recorded 1.9 billion adults as overweight, of which 600 million were obese (WHO 2020). It is listed as one of the top five global risks for mortality, alongside smoking (WHO 2009). The disease is characterised by an excess of body mass due to an increase of adipose tissue, predominantly in the abdominal region, and a body mass index (BMI) over 30 kg/m² (Bessesen 2008). Patients often suffer from metabolic syndrome, which involves dyslipidaemia (elevated TG, apolipoprotein B and reduced high-density lipoprotein (HDL)) (Grundy 2015) and increased the risk for certain co-morbidities such as cardiovascular disease, type II diabetes and multiple cancers (Haslam & James 2005; Bessesen 2008). These risk factors have been linked to expansion of the visWAT in obese patients due to hypertrophy (Hirsch & Knittle 1970; Choe et al. 2016). It is widely accepted that obesity stems from an imbalance in energy intake and expenditure, leading to the excess energy being stored as TGs in adipose tissue (Schwartz et al. 2017).

1.2.1. Weight-loss as an obesity treatment

The most effective treatment for obesity is to lose the excess weight. One estimate states that 42% of the world population undergoes a weight loss attempt per year (Santos et al. 2017). This is primarily accomplished by reducing calorie intake through

controlled dieting and increasing energy expenditure through exercise (Haslam & James 2005). However, weight regain is extremely common, where it has been found that more than half of the lost weight returns after two years (Anderson et al. 2001). Weight regain has been suggested to be due to a number of mechanisms, such as cell stress, inflammation, changes in adipokine secretion and a reduction in lipolysis (van Baak & Mariman 2019). The extracellular matrix (ECM) plays an essential role in adipocyte size maintenance, creating a limit on the size of the adipocyte during hypertrophy (Khan et al. 2009). Under calorie restriction, adipocytes shrink due the release of fatty acids from lipolysis and the ECM has to remodel to due to the change in space. However, the ECM is not always able to remodel, leading to mechanical stress within the adipose tissue. This stress can lead to impairment in future lipolytic reactions and promotion of precursor differentiation to compensate for the reduced space within the ECM (van Baak & Mariman 2019). These factors can subsequently contribute to weight regain, by refilling the shrunken adipocytes with TGs or increasing the adipocyte number of the tissue, which may also undergo hypertrophy. The most successful diets include a low-calorie diet intervention, which have been shown to decrease the risks for obesity related co-morbidities (Hensrud 2001). However, there is a large amount of variability among individuals in their ability to lose weight, therefore not every diet program will be successful for each patient (Palou & Bonet 2013), leading to the requirement of personalised diets for obese patients (Ferraro et al. 2015).

1.3. FAM46A: a novel adipocyte regulator?

1.3.1. Role of genetic variation on weight-loss response

Weight-loss is a difficult process, however, is becoming more and more necessary in order to tackle the current global epidemic of obesity. Understanding how weight-loss differs on a person by person basis can help obese patients at an individual level. Carayol et al. (2017) addressed this by studying the effect of genetic variation on an individual's plasma protein levels during a weight-loss event. Using data from the DIOGenes intervention trial (Larsen et al. 2010), the plasma levels of 494 obese patients were measured before and after a weight-loss event (low-calorie diet (LCD) intervention). Protein quantitative trait loci (pQTL) associations during the LCD identified single nucleotide polymorphisms (SNPs) regulating plasma protein levels. The most notable association came from two SNPs located in an intergenic region on chromosome 6, between *BCKDHB* and *FAM46A*, associating with leptin. An additional expression quantitative trait loci (eQTL) analysis from subcutaneous adipose tissue biopsies was performed, which discovered that *FAM46A* expression also associated with these variants. Follow-up in vitro assays in human adipocytes showed an inverse relationship between *FAM46A* expression and leptin secretion. This complements the opposite association direction between leptin plasma level changes and *FAM46A* expression found computationally during the LCD, where *FAM46A* is down-regulated. This is the first study to indicate a potential role for *FAM46A* in weight-loss and adipocyte function.

1.3.2. Family with sequence similarity 46

FAM46A/TENT5A is an evolutionarily conserved protein belonging to the terminal nucleotidyl-transferase (NTase) family of proteins (Warkocki et al. 2018). The Family with sequence similarity 46

(FAM46) proteins were initially discovered during an *in silico* screen for proteins containing a NTase fold, and were subsequently classified as non-canonical poly(A) polymerases (Kuchta et al. 2016). This class of enzymes is known to modify RNA by addition of nucleotides to the 3' end, in both the nucleus and the cytosol. FAM46A proteins are found in every known animal genome and have four paralogs in mammals (A-D) that share high similarity in their amino acid sequences. Each protein is comprised of two conserved domains: (i) a NTase domain and (ii) a poly(A) polymerase/20 – 50-oligoadenylate synthetase 1 substrate binding domain (PAP/OAS1 SBD). The NTase domain is the enzymatic domain responsible for the addition of nucleotides, and the predicted catalytic centre is well conserved within the FAM46 proteins, with acidic amino acids being essential to the poly(A) polymerase activity of FAM46C and FAM46D (Mroczek et al. 2017). The PAP/OAS1 SBD in other known NTase enzymes is required for NTP substrate binding and RNA stability (Kuchta et al. 2016; Warkocki et al. 2018). Bioinformatic analysis of all four paralogs predicts that they bind RNA binding proteins, however, in different pathways, allowing for distinct functional roles for each protein (Kuchta et al. 2016).

The FAM46A gene was discovered in 2002 by Lagali et al. when it was known as C6orf37 located on chromosome 6 in the human genome. This region was of interest due to it containing multiple retinal disease loci. This initial discovery found that FAM46A was highly conserved, with homology in other common model organisms. An mRNA interactome screen to identify potential and novel RNA binding proteins in HeLa cells (Castello et al. 2012) classified FAM46A as an RNA binding protein. The expression pattern of FAM46A in humans is not unique to one tissue, with a particularly high expression in the salivary gland (Figure 1.3.1).

Chapter 1

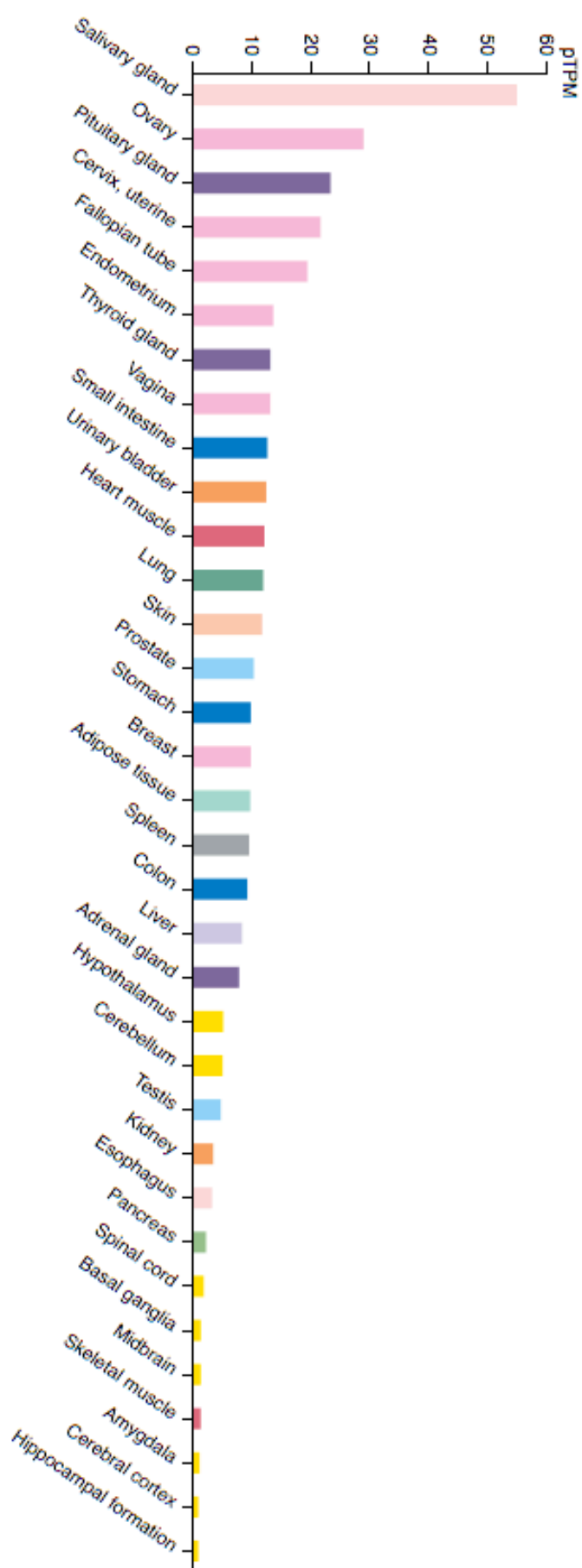


Figure 1.3.1: GTEx dataset of FAM46A expression in human tissues. Plot taken from the Human Protein Atlas Uhlen M et al (2015) Science Data from The GTEx Consortium (2013) Nat. Gen.

Loss of function mutations in FAM46A, differential gene expression and genetic associations within the *FAM46A* gene have been reported in a number of diseases, most notably in osteogenesis imperfecta (OI), where mutations in the NTase domain of FAM46A are found in patients of OI (Doyard et al. 2018; Besio et al. 2019). Skeletal phenotypes have also been observed in a mouse model with a loss of function mutation in the *Fam46a* gene, showing severe skeletal abnormalities, disrupted bone formation and significantly elevated plasma alkaline phosphatase (ALP) (Diener et al. 2016). It has also been suggested that *Fam46a* plays a role in tooth bud formation, as microarray analysis observed a time dependent expression of *Fam46a* in the developing tooth germ in regions containing proliferating cells (Etokebe et al. 2009).

Despite its involvement in multiple diseases, the role and function of FAM46A still remains elusive. A yeast-two hybrid screen for Smad-signalling proteins identified *Fam46a* as a potential interaction partner for SARA, a recruitment protein for Smad2 to the TGF β -II receptor (Colland et al. 2004). A subsequent study in *Xenopus* development observed was *Fam46a* expressed in the pre-placodal ectoderm, and immuno-precipitation assays found that *Fam46a* interacted with Smad-signalling proteins Smad1 and Smad4 (Watanabe et al. 2018). The authors suggest a positive regulation on BMP signalling, where *Fam46a* recruits Smad1 and Smad4 into the nucleus to activate BMP target genes, as well stabilising Smad1 and preventing its ubiquitination.

Other studies have indicated a potential role for FAM46A in metabolic stress response. Mahamoud et al. (2016) found an increase of FAM46A expression in keratinocytes under-going a co-treatment of a glycolysis inhibitor and millimetre waves, while a

Chapter 1

recent publication observed that Fam46a expression was increased in the aorta under a high-fat diet (HFD) (where the mice were insulin resistant). Expression returned to control levels during exercise, indicating that Fam46a may be involved in certain inflammation pathways (S. Liu et al. 2018).

1.3.3. Scope of thesis

This thesis aimed to follow up on the Carayol et al. (2017) study by investigating the transcriptional regulation of FAM46A and how the QTL region is potentially influencing this regulation by observing the physical DNA interaction network around this region using chromosome conformation capture (3C) method CaptureC. FAM46A's potential role in adipocyte function and metabolism was also addressed by disrupting FAM46A expression in differentiating adipocytes and by metabolic phenotyping of a Fam46a knockout (KO) mouse model under hyper-caloric stress (in collaboration with the IMG in Prague). With this study, we hoped to further understand the role and regulation of FAM46A in adipocyte biology and function.

Chapter 2:

Transcriptional regulation of the *FAM46A* promoter during human adipocyte differentiation

2. Transcriptional regulation of the *FAM46A* promoter during human adipocyte differentiation

2.1. Abstract

The current obesity epidemic is one of the most urgent health crises in the developed world, where nearly 40% of adults over 18 are considered overweight or obese (WHO 2020). However, obesity is preventable and can be treated to reduce the risk of co-morbidities, such as cancer and type II diabetes. Current treatment options are aimed at losing the characteristic excess fat mass, where calorie restriction is a popular method (Harvey-Berino 1999). Carayol et al. (2017) investigated the genetic impact on the trajectory of obese patients under-going calorie restriction for weight-loss (DIOgenes (Larsen et al. 2010)). The authors were exploring the effect of patient genotype to the level of plasma markers and if these markers could be used as a predictor of weight-loss outcome. Using pQTL analyses, they identified a cluster of variants, or a “hot spot”, downstream of gene *FAM46A* on chromosome 6. This cluster associated with plasma levels of leptin, a hormone responsible for energy homeostasis. A subsequent eQTL analysis using RNAseq data from subcutaneous adipose tissue samples of these patients revealed that the expression of *FAM46A* also associated with this hot spot region.

Using the 3C method CaptureC to map DNA interaction landscapes, we investigated this hot spot region for its potential regulatory role on *FAM46A* expression. By optimising CaptureC in pre-adipogenic and mature human Simpson Golabi Behmel Syndrome (SGBS) adipocytes, we discovered that the *FAM46A* promoter had stronger and significant interactions with loci upstream of the promoter, rather than the downstream hot spot region, which produced no significant interactions with the *FAM46A* promoter or any other loci of interest. Further investigation of this upstream region highlighted that the significant

interaction peaks were enriched for TF motifs involved in the regulation of cell fate commitment and cell differentiation, including TFs involved in adipogenesis, such as TEADs. Overall, our results suggest that *FAM46A* expression is required to maintain a pre-adipogenic state and this may be dependent on TGF β -BMP pathways.

Chapter 2

2.2. Introduction

2.2.1. The 3D genome

Eukaryotic genomes are highly organised, and this dynamic three-dimensional arrangement allows for an additional layer of complexity to gene regulation and cell identity. The spatial organisation of the genome is hierarchical in size and can be observed from morphologies as large as chromosomes, seen under a microscope, down to a five kilobase pair (kbp) resolution of looping interactions that require more advanced methodologies to be studied (Bonev & Cavalli 2016). Each level of organisation offers a unique purpose, with different structural organisations being seen in different cell types and disease states (Dixon et al. 2012; Krijger & de Laat 2016).

2.2.2. Topologically associating domains

One of the more studied levels of the genome are known as topologically associating domains (TADs). These highly conserved megabase-sized interaction domains were discovered using Hi-C mapping of the genome (Lieberman-Aiden et al. 2009; Dixon et al. 2012). Most regulatory interactions are confined, such as promoter-enhancer contacts, are confined to TADs (Shen et al. 2012), and early Hi-C studies found that CTCF insulator protein is enriched at TAD boundaries (Dixon et al. 2012). In fact, TAD boundaries and replication domain boundaries correlate almost one to one (Pope et al. 2014), and co-regulation of genes have been observed within the same TAD, similar to prokaryotic operons (Soler-Oliva et al. 2017). Removal or disruption of TAD boundaries has led to developmental issues and disease states, primarily cancer (Kaiser & Semple 2017). These findings highlight the importance of genome structure as a functional and regulatory mechanism during replication and gene expression timing, as well as an identifier of cell type and state.

2.2.3. Gene expression initiation and the role of chromatin structure

Understanding how and when gene expression is initiated is one of the most studied questions in the molecular biology field. Gene regulation is a complex blend of multiple factors that can be simplified into three components: regulatory proteins, such as TFs, DNA regulatory elements and chromatin structure (Wasserman & Sandelin 2004) (Figure 2.2.1). To initiate (or repress) gene expression, specific TFs bind to the core promoter of a gene, as well as to a number of *cis*-regulatory elements. These elements include enhancers, silencers and insulators (Spitz & Furlong 2012), which can be identified close to the transcription start site (TSS) of a gene, or at a distal location from the TSS. This distal regulation integrates chromatin structure as a functional element to gene regulation by physically connecting these distal transcription factor binding sites (TFBS) and other co-factors that bind TFs to the promoter of a gene. Interruption of the chromatin structure has been found to be implicated in a number of human diseases, highlighting its importance in molecular regulation (Krijger & de Laat 2016).

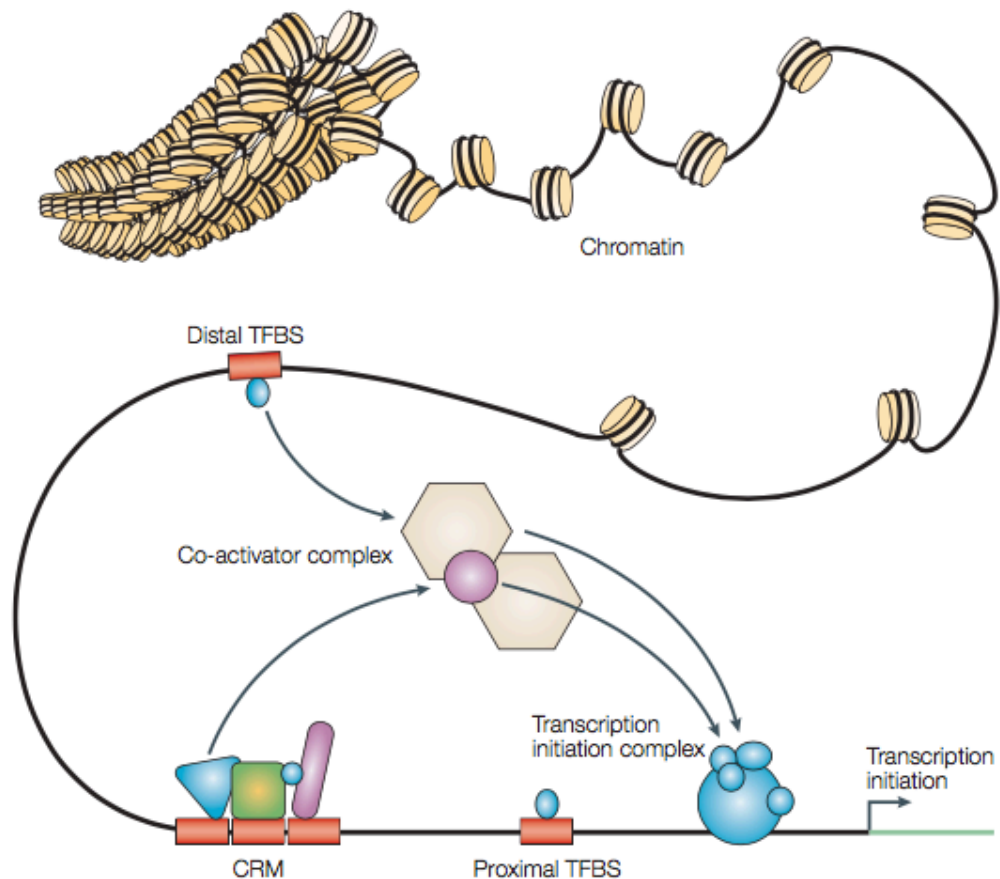


Figure 2.2.1: Overview of gene regulation.

Figure taken from Wasserman & Sandelin (2004) *Nat. Rev. Genet.*

Blue shapes are transcription factors (TFs) that can bind to transcription start sites (TSSs) at gene promoters (green line), to proximal/distal transcription factor binding sites (TFBSs) and cis-regulatory modules (CRMs). All these regulatory elements can come together with other co-factor proteins (beige hexagons and coloured shapes) by chromatin looping to regulate gene expression.

2.2.4. Chromatin conformation capture

As promoter-regulatory element interactions play a key role in gene regulation, 3C techniques, also known as proximity ligation assays, have been developed to measure physical interactions between genomic regions. Since the first method was published in 2002 (Dekker et al. 2002), multiple variations of the technique have been developed to adapt to the user's purpose. Hi-C explores all possible contact frequencies (all-to-all) to identify genome-wide interaction maps, while 3C, and its derivatives, identifies interaction profiles of a specific locus or loci determined by the user (one-to-few or all) (Grob & Cavalli 2018). Hi-C is therefore more appropriate for the identification of higher order structures

such as TADs, while 3C methods are commonly used to understand specific interactions with gene promoters to possibly identify gene regulatory elements. 3C methods ensure a higher resolution of interactions compared to Hi-C, and next generation CaptureC is one of the most high resolution 3C methods currently available (Davies et al. 2016). Here the use of biotinylated DNA oligonucleotides targeting specific regions of interest and a double-capture enrichment step increases the interaction signal compared to similar methods.

As 3C methods require a specific loci of choice, the regions of interest may originate from association studies where SNPs are related with a certain phenotype, also known as genome-wide association studies (GWAS). Using this list of associated variants, 3C can be applied to identify a potential functional role for these variants that may explain the studied phenotype. For example, a GWAS of risk alleles associated with BMI were identified in the non-coding region of the *FTO* gene (Frayling et al. 2007; Dina et al. 2007), however, the functional role for *FTO* in obesity remained controversial. Follow-up 3C studies discovered that these variants were in spatial proximity to a neighbouring gene, *IRX3*, and were rather regulating its expression versus the *FTO* gene itself (Smemo et al. 2014). *IRX3* expression was subsequently found to be a determinant of BMI. This exemplifies the power of 3C techniques to functionally characterise variants found to be associated with a certain phenotype, and to identify novel regulatory elements.

2.2.5. Research question

Connecting to Carayol et al. (2017), who identified a hot-spot region on chromosome 6 that associates with changes in plasma leptin during a low-calorie intervention, we investigated the functional role of this hot spot of variants. When performing a 5 Mb eQTL analysis of the surrounding region in subWAT biopsies from the study participants, only *FAM46A* expression associated with this hot spot region. We were specifically

Chapter 2

interested in understanding if this region contained any regulatory elements controlling *FAM46A* expression as there is no known mechanism of *FAM46A* regulation. It has been found that *FAM46A* interacts with proteins in the TGF β (Colland et al. 2004) and BMP (Watanabe et al. 2018) signalling pathways, which lead to the regulation of certain target genes. Watanabe et al. (2018) propose a mechanism in which *Fam46a* recruits *Smad1* and *Smad4* into the nucleus, however, neither the target genes of this complex, nor regulation of *FAM46A* itself were further explored.

To study the regulation of *FAM46A* expression, we performed CaptureC in human adipocytes (SGBS cells) to map the interaction landscape around the promoter, as well as select loci in the hot spot region. *FAM46A* was differentially expressed in adipose tissue during the weight-loss intervention (Carayol et al. 2017), and during *in vitro* adipocyte differentiation (see Chapter 3), therefore the CaptureC was performed on pre-adipogenic and mature SGBS samples. We believed this would lead to further insights into how *FAM46A* is regulated in different cell states, as well as understanding if the hot spot region has a functional role in adipocytes.

2.3. Results

2.3.1. Optimisation of CaptureC digestion efficiency in SGBS cells

In order to observe the chromatin conformation of pre-adipogenic (day 0) and mature (day 12) SGBS cells, the next generation CaptureC method (Davies et al. 2016) was used (Figure 2.3.1). Briefly, cells are fixed, collected for cell lysis, and nuclei are extracted. The isolated nuclei undergo a digestion step with a 4-cutter restriction enzyme and a close proximity ligation step to create a linear DNA fragment library of interacting genomic loci, or a 3C library. This library is prepared for Illumina sequencing by addition of compatible paired-end adapters. To enrich for, or “capture”, regions of interest, biotin labelled oligonucleotides designed for specific target regions are used and pulled down using streptavidin beads. This enrichment is performed twice, aimed to decrease the amount of background signal. These captured fragments are subsequently sequenced and analysed.

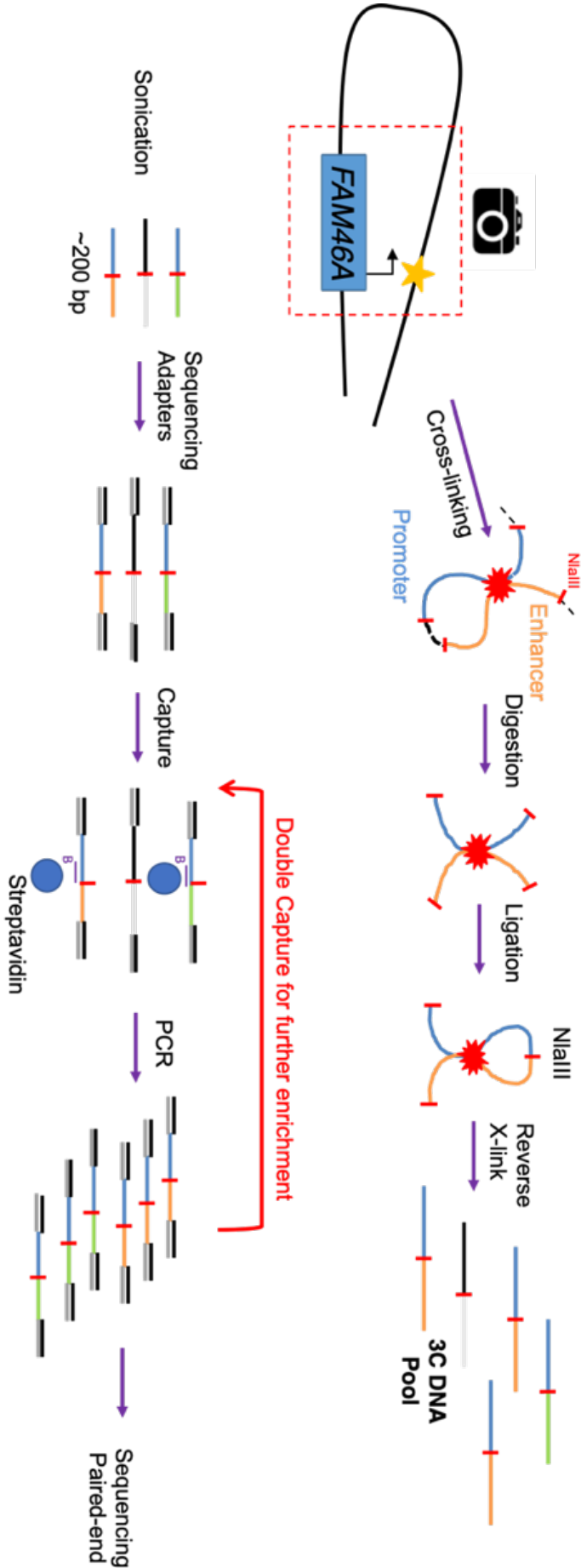


Figure 2.3.1: Overview of CaptureC method. Adapted from Davies et al. (2016) Nat. Methods.

The digestion step is an essential part of 3C methods, and thus its optimisation is vital, especially when utilising a new cell line. This is the first known time that SGBS cells have been used for CaptureC. For the first pilot experiment (Pilot 1) we followed the protocol as recommended by the authors of the method paper, using DpnII as the 4-cutter enzyme. To ensure proper digestion of the chromatin, the average digestion efficiency needed in mammalian cells has been determined to be at least at 70% (Naumova et al. 2012). Digestion efficiency is tested by collecting undigested (Control 1) and digested (Control 2) samples. These controls are quantified by quantitative polymerase chain reaction (qPCR) using one primer set that targets a restriction site junction and another at a control genomic locus, where there is no restriction site. The more the chromatin is digested, the less signal should come from the control 2 sample, thus having a higher digestion efficiency. The pilot 1 experiment yielded a digestion efficiency of 35.7% (Table 2.3.1), which is significantly lower than the expected and recommended efficiency of 70%.

To address and improve the low digestion efficiency, further optimisations were performed in scaled down test conditions. The digestion efficiencies were tested with two parameters: 1) the restriction enzyme used (DpnII vs NlaIII) and 2) the preparation of the chromatin; more specifically the method of cell lysis used to harvest chromatin. Davies et al. (2016) utilised mouse embryonic stem cells (ESCs), which were lysed by resuspension in lysis buffer in a micro-centrifuge tube. Considering the adherent and “sticky” nature of the SGBS cells when harvested, we tested two different lysate preparation methods: 1) by performing the lysis step as described in the paper, (in-tube), and 2) by adding the lysis buffer directly to the culture dish (in-plate). As shown in Table 2.3.1, it was determined that NlaIII displayed a better digestion efficiency compared to DpnII, where NlaIII had an average in-plate digestion efficiency of 82%. It was also observed that the method of cell lysis contributes greatly to the efficiency of NlaIII, where lysis performed in-tube yielded a digestion efficiency of

Chapter 2

0%. DpnII seemed unaffected by the method of lysis, where the in-plate digestion efficiency was 30% (vs 35.7% in-tube from the pilot 1 experiment).

Table 2.3.1: Digestion efficiencies from CaptureC experiments and optimisation tests.

Parameters	Pilot 1	Optimisation tests			Final	
					D0	D12
Restriction enzyme	DpnII	DpnII	NlaIII	NlaIII	NlaIII	NlaIII
Lysis method	In-tube	In-plate	In-tube	In-plate	In-plate	In-plate
Digestion efficiency	35.7%	30.0%	0.0%	82.0%	92.6%	91.5%

From these observations, it was determined the optimum conditions for proper digestion of chromatin from SGBS cells should be performed with NlaIII as the restriction enzyme, and by preparing the cell lysates using in-plate lysis, where the lysis buffer is added directly to the cell culture plate.

2.3.2. CaptureC viewpoint selection and design optimisation

Once an optimally digested 3C library was made, the next step for the CaptureC method was to design the biotinylated oligonucleotides required for the capture step (Figure 2.3.1). The designed oligonucleotide is highly dependent on the sequence of the target region for capture. For high specificity to the target region and subsequent data analysis, the oligonucleotide must to overlap the restriction site of choice, thus the design is limited to the genomic sequence around the restriction site. The method paper offers a probe design software, CapSequm, in order to optimise the design of the oligonucleotide. Our regions of interest were the *FAM46A* promoter (Figure 2.3.2; green box) and the hot spot QTL region. As this region spans almost 750 kbp, two capture viewpoints were chosen based on the most significant leptin pQTLs and *FAM46A* eQTLs (Figure 2.3.2; blue and red boxes).

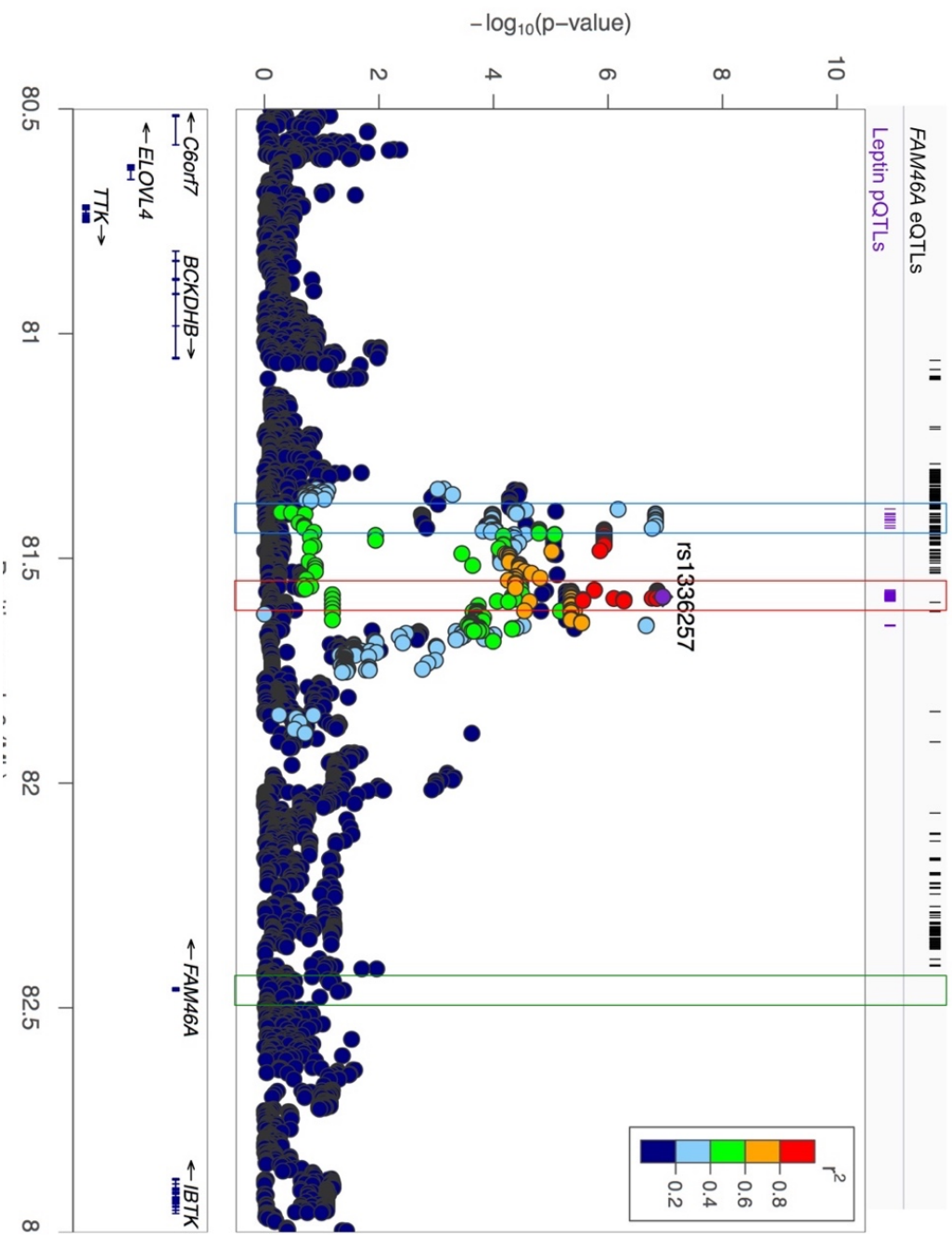


Figure 2.3.2: Association plot for pQTLs associated to changes in plasma leptin during low-calorie diet (LCD) intervention.

Each dot is an associated SNP, where the colour scale represents the linkage disequilibrium with the most significant SNP (rs1336257). The p-value of each SNP is scaled on the y-axis. Significant SNPs from pQTL and eQTL analysis are displayed on the top tracks. Coloured boxes represent the final viewpoint regions used for CaptureC.

Figure used and adapted from Carayol et al. (2017) Nat Comms

Chapter 2

To design the oligonucleotide probes, viewpoint coordinates are entered into CapSequm, as well as the restriction enzyme of choice. Candidate oligonucleotides undergo a BLAT search and Repeatmasker analysis to avoid duplicated regions and areas of repetitiveness. The Multi-Image Genome (MIG) viewer (McGowan et al. 2013) is an additional software used to customise the probe design. Using the standard and most stringent conditions in MIG (Density ≤ 30 and repeat length ≤ 30), four probe sets were designed for the second pilot experiment (Pilot 2) with *Nla*III digestion. Capture efficiency was determined by the final count of sequencing fragments observed in each region of interest (or capture fragments). As seen in Table 2.3.2, the probe sets for the *FAM46A* promoter captured the most efficiently with a final count of 6,447 counts (green box), compared to the three probe pairs designed QTL region, where one pair did not capture any fragments (0 counts).

A new set of probes was designed for the QTL region (Pilot 3) with an additional filtering parameter of the restriction fragment length being between 400 and 1000 bp. With this, two probe pairs were found to be suitable for capture (blue and red boxes in Table 2.3.2). The probe pair 815_574 (red box) capture more efficiently than 814_547 (blue box) with read counts of 6,072 and 1185 respectively, however, both were subsequently used for the final CaptureC experiment.

Table 2.3.2: Capture efficiency from CaptureC experiments and optimisation tests.

Probe pairs highlighted in colour (green, blue and red) correspond to the final probe pairs used. (see Figure 2.3.2 for approximate location of each probe pair).

Pilot 1				
Probe pair	FAM46A_DpnII	rs9344031_DpnII	rs115586175	rs16892315
Fragment coordinates (hg19)	chr6: 82461559-82462744	chr6: 81400407-81401283	chr6: 81432606-81433967	chr6: 81582382-81585662
Fragment length (bp)	1185	876	1361	3280
Reporter fragments (final count)	328	10	2652	3
Pilot 2				
Probe pair	FAM46A_NIaIII	rs9344031_NIaIII	rs985736	rs12175546
Fragment coordinates (hg19)	chr6: 82461856-82462915	chr6: 81400872-81401073	chr6: 81428751-81428895	chr6: 81583800-81584111
Fragment length (bp)	1059	201	144	311
Reporter fragments (final count)	6447	17	0	1
Pilot 3				
Probe pair	814_707	814_546	815_416	815_574
Fragment coordinates (hg19)	chr6: 81405587-81406294	chr6: 81399707-81400253	chr6: 81592850-81593266	chr6: 81588964-81589538
Fragment length (bp)	707	546	416	574
Reporter fragments (final count)	109	1185	0	6072

2.3.3. CaptureC validation

Once the method was optimised for SGBS cells, four and three biological replicate experiments were performed for day 0 (D0) and day 12 (D12) SGBS cells respectively. Figure 2.3.3 shows how the CC3Analyzer filters the sequencing reads used for the final CaptureC result. Briefly, we consider three categories of reads: capture, exclusion and reporter. Exclusion reads were removed from the analysis due to their proximity to the capture region, which was considered a bias. Only reads containing both a capture fragment and a reporter fragment were considered an interaction and included in the final track file.

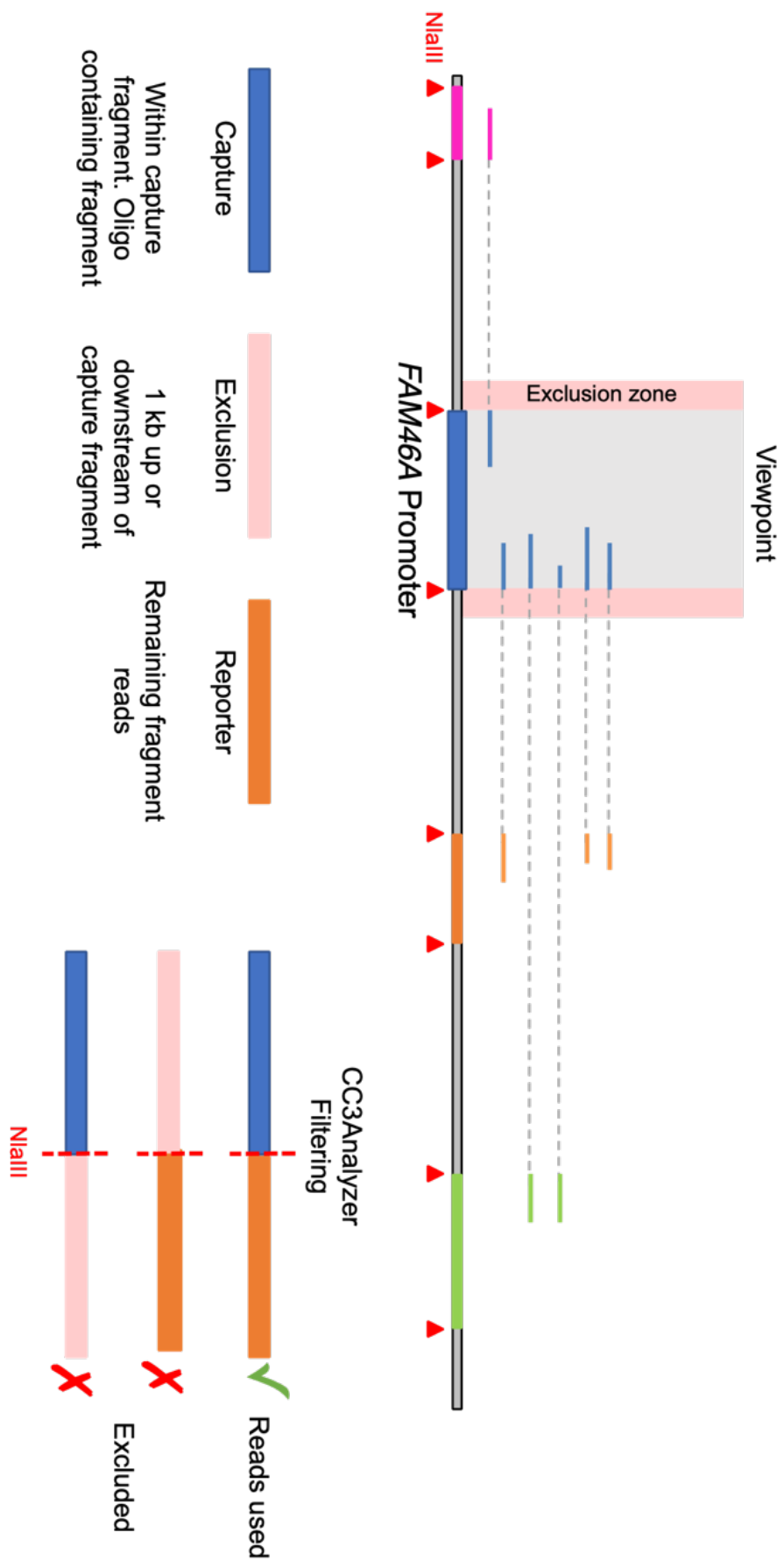


Figure 2.3.3: Overview of CC3Analyzer filtering. Adapted from Davies et al. (2016) Nat. Methods.

To test the analysis pipeline and create a reference set of read statistics, raw data from the original publication was used and reproduced. It should be noted that the publication used ESC from mice and targeted the Hba gene, which has known and validated enhancers (Tolhuis et al. 2002). Table 2.3.3 shows the final read statistics in comparison to the statistics of the original method paper. When comparing to the publication data, we observe the number of capture and reporter reads are significantly less in our SGBS samples, with the smallest number of reads being found in the D12 samples. The final percentage of reads being used in our analysis is 7.3 – 9.3% vs 15.6% being used in the Davies paper. Considering the different cell type (ESC vs SGBS) and target regions, the CaptureC dataset produced was considered acceptable and thus used for further validation and analysis.

Table 2.3.3: CaptureC read statistics from the original publication (Davies et al. (2016)) and the average values from D0 and D12 SGBS samples. Reported from the CC3Analyzer script.

CaptureC - Statistics (reads)	Davies et al.	D0 Average	D12 Average
Restriction enzyme	DpnII	NIaIII	NIaIII
Unmapped fragments SAM	2,056,490	1,157,759	924,980
Mapped fragments SAM	20,559,988	9,052,902	7,079,767
Proximity exclusion	778,119	310,850	299,039
Reporter fragments	13,217,953	7,089,083	5,493,126
Reads entering analysis	14,114,820	6,478,123	5,196,534
Capture containing reads	6,510,495	1,607,742	1,244,141
Capture and reporter	2,205,992	619,725	394,123
No capture or reporter (excluded)	11,908,828	5,858,397	4,802,411
Reads used in analysis	2,205,992	619,725	394,123
Percentage (entering/used)	15.6%	9.3%	7.3%

The tracks on Figure 2.3.4 show the final interaction map between the *FAM46A* capture viewpoint and its surrounding region on chromosome 6 for three replicates in SGBS cells at D0 (lime green tracks) and D12 (dark green tracks). To validate our experiment, Hi-C data from human muscle

Chapter 2

(Schmitt et al. 2016) was overlaid with the CaptureC data. As seen in Figure 2.3.4, when comparing the up- and downstream regions at equal linear distance from the FAM46A capture viewpoint, we found stronger interactions between the FAM46A capture viewpoint and its TAD of the Hi-C map, as compared to outside its TAD. This supports the quality of our interaction data capture. We can also observe the aforementioned lower amount of reads in the D12 sample as compared to D0. This could in part be due to the fact that in the mature state, more lipids are present in the SGBS cells and partly prevent efficient capturing of the interactions. Despite the reduction in reads entering the analysis, the CaptureC signal obtained from our SGBS samples overlap with known interaction regions, and thus confirms a successful capture experiment performed in SGBS pre- and mature adipocytes.

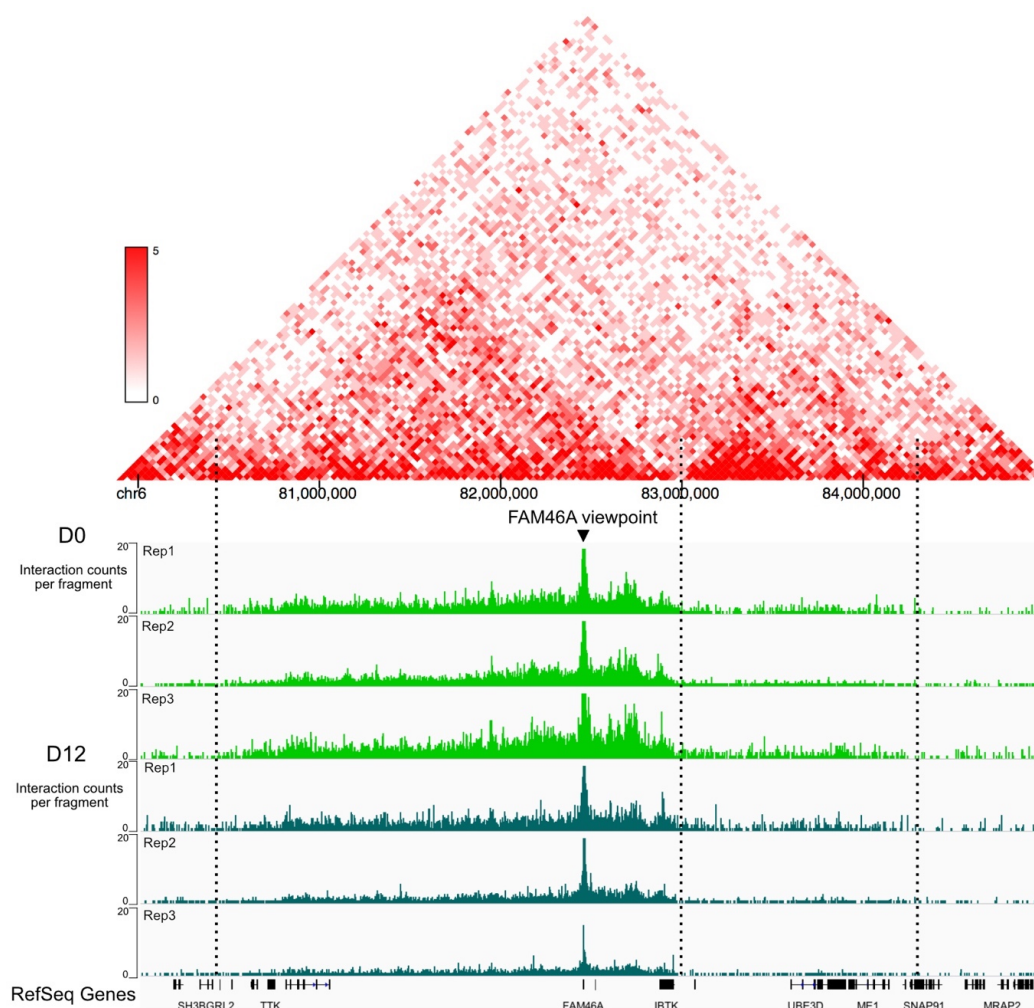


Figure 2.3.4: FAM46A viewpoint CaptureC interaction profile from three replicates of Day 0 and Day 12 SGBS samples.

HiC plot from muscle (Schmitt et al. 2016) was acquired using the 3DGenome Browser (Y. Wang et al. 2018) for validation of the interaction profile of the region. Dashed lines indicate the borders of annotated TADs.

2.3.4. Interaction peak calling using peakC

CaptureC is a powerful method to identify the interaction profiles of a region of interest with the intention to discover a distant regulatory locus. However, consistent peak calling to find interesting candidates with chromosome confirmation data has still not been properly established. Visual inspection of the interaction peak profiles can offer some indication, although this is dependent on highly reproducible and strong interactions to be present. Recently, a peak calling package specifically for 4C (a modified version of 3C) and CaptureC data, named peakC, has been developed and offers robust peak calling with statistical analysis (Geeven

Chapter 2

et al. 2018). The peakC software was used to call peaks in the replicate CaptureC datasets for D0 and D12 for all three viewpoints: *FAM46A* promoter (Figure 2.3.5, green box), QTL region 1 (Figure 2.3.5, blue box) and QTL region 2 (Figure 2.3.5, red box). By averaging the replicate datasets, the peakC software was able to call significant peaks for all the D0 viewpoints (Figure 2.3.5A; significant peaks highlighted in red). However, only the *FAM46A* viewpoint had significant peaks in the D12 dataset (Figure 2.3.5B). Due to the QTL viewpoints having a small number of peaks called in the D0 samples (25 peaks for both viewpoints vs 190 for *FAM46A*), and none in the D12 samples, they were excluded from further analysis. When observing differences in the viewpoints between D0 and D12, significant peak signals are being lost at D12, suggesting a shift in the interaction landscape during differentiation.

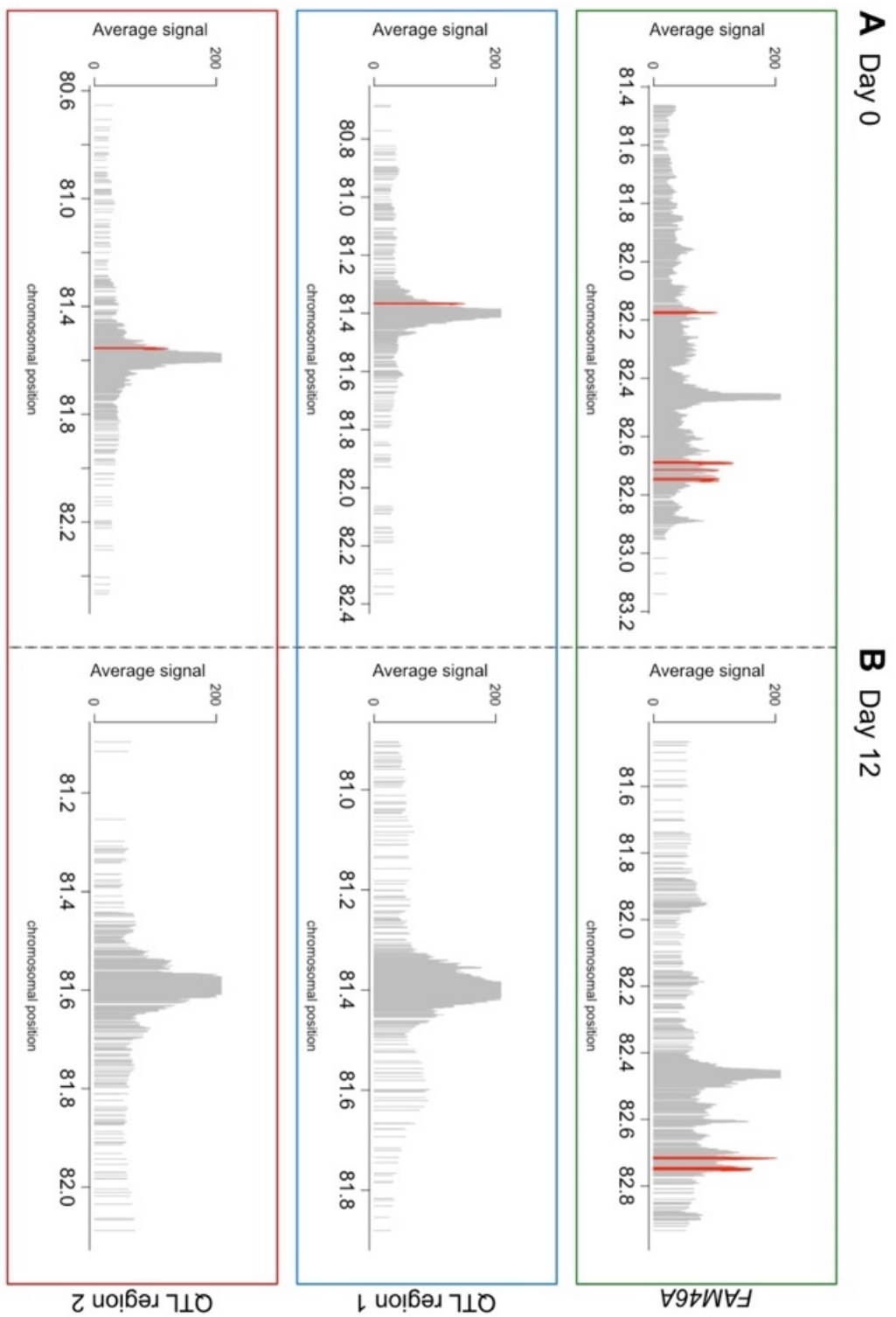


Figure 2.3.5: Peak calling plots of CaptureC data from A) D0 and B) D12 SGBS samples using peakC

Each plot is a different viewpoint (FAM46A, QTL region 1 and QTL region 2) produced using the peakC software (Geeven et al. 2018). Grey peaks are average peaks from replicate samples and peaks considered significant are highlighted in red.

Chapter 2

2.3.5. Analysis of *FAM46A* viewpoint for regulatory elements

To understand the relevance of our significant interactions called by peakC, a comparative analysis using other datasets (in-house and publicly available data) was applied. Regions of interest were determined by overlapping signal signatures across these datasets. Specifically, regions of open chromatin were analysed using data from ATAC-seq (Assay for Transposase-Accessible Chromatin using sequencing) in SGBS cells (data generated by Nele Gheldof, unpublished). Areas of open chromatin are known to indicate regions for regulatory elements, such as TFBSs or enhancers (Tsompana & Buck 2014). Other indicators of transcriptional activity are histone modifications, such as H3K27Ac, which is known for identifying active enhancers. Figure 2.3.6 shows the overlapping data over a 700 kbp region around the *FAM46A* viewpoint, covering the called peaks from the peakC analysis. covering the called peaks from the peakC analysis.

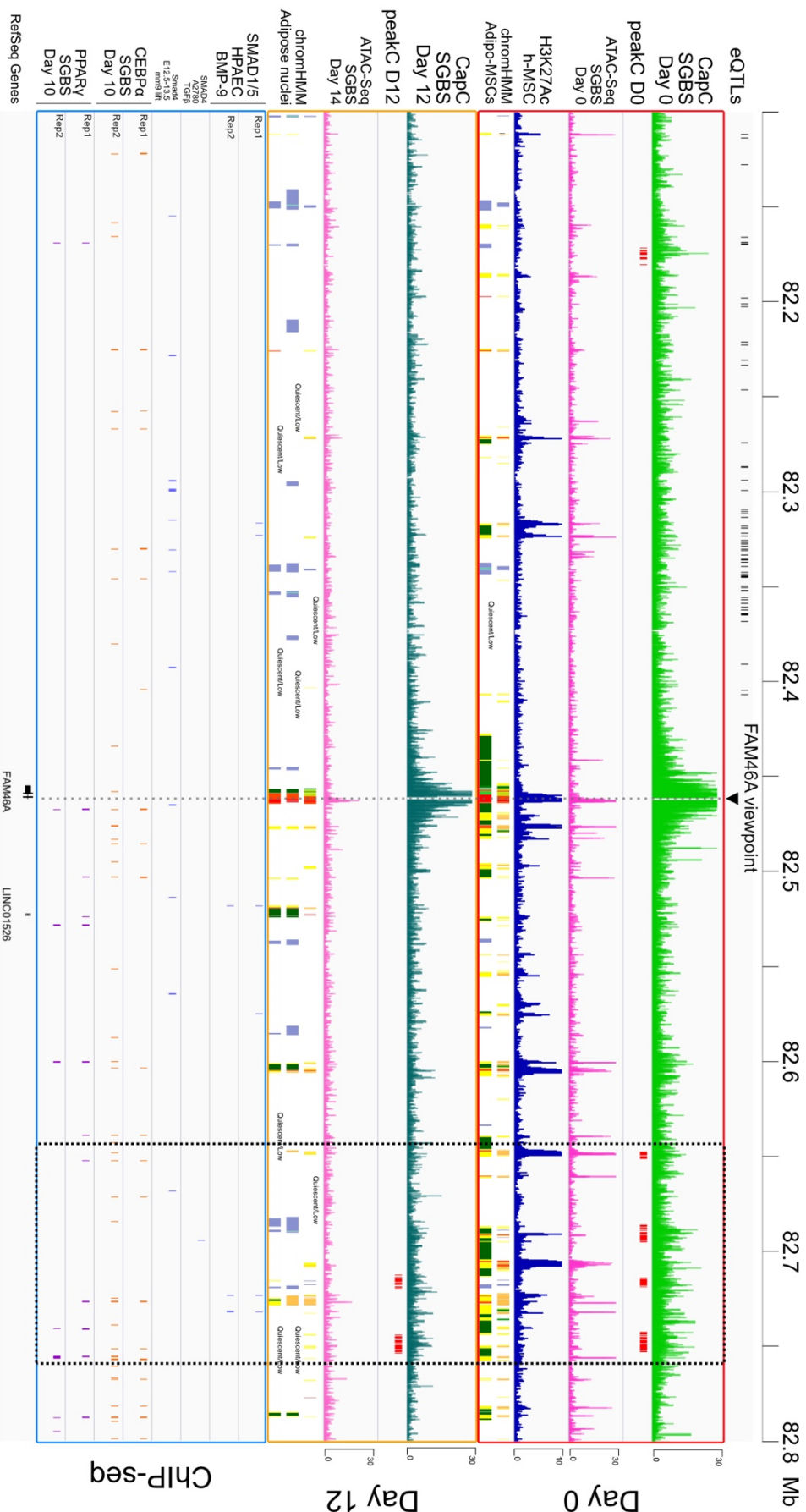


Figure 2.3.6: Comparative analysis of 700 kbp locus around FAM46A viewpoint (arrow and grey dashed line). Datasets include CaptureC data (this study, green and teal tracks) with peakC called peaks (red tracks), ATAC-seq performed in differentiating SGBS cells (Nele Gheldof, unpublished, pink tracks), H3K27Ac ChIP-seq performed in human mesenchymal stem cells (h-MSCs) from Rauch et al. (2019) Nat. Gen. (dark blue track), chromHMM data from the WashU Epigenome Browser v52.0.0 (Li et al., (2019) Nucleic Acids Research) in adipose derived mesenchymal stem cells (Adipo-MSCs) and adipose nuclei (red = active TSS, orange = flanking active TSS, green = transcription, yellow = enhancer, blue = heterochromatin) and ChIP-seq data (blue box) from SMAD1/5 in human pulmonary artery endothelial cells (HPAECs) from Morikawa et al. (2019) Sci Signal, SMAD4 in TGFβ1 stimulated A2780 cells (ovarian cancer) from Kennedy et al. (2011) PLOS ONE, mm9 liftOver Smad4 in E12.5-13.5 mice forelimbs from Yan et al. (2018) J Biol Chem and CEBPα and PPARγ in differentiated SGBS cells from Schmidt et al. (2011) BMC Gen.

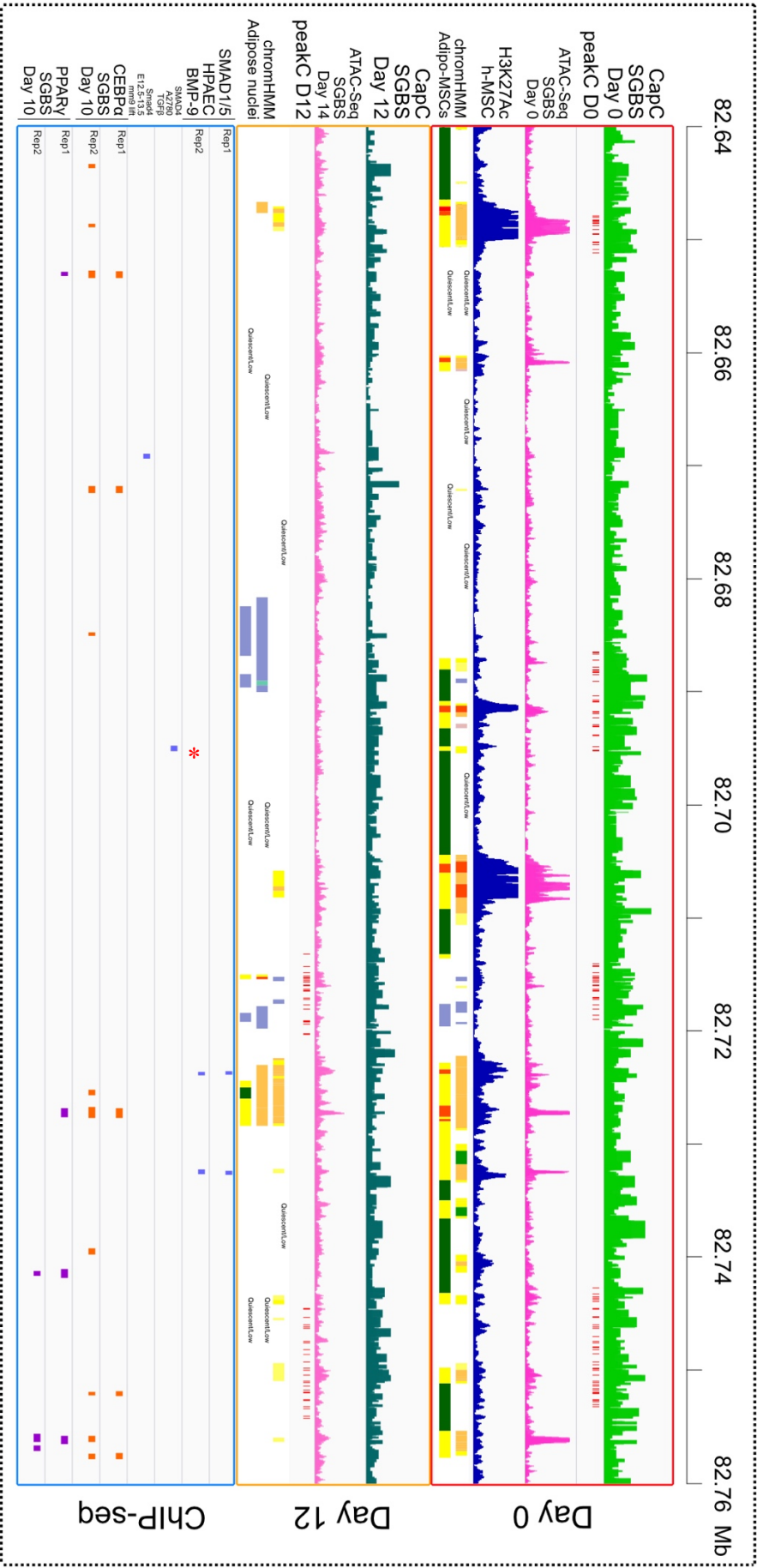


Figure 2.3.7: Zoom in of comparative analysis (125 kbp) (Figure 2.3.6 dashed box)

Datasets include CaptureC data (this study, green and teal tracks) with peakC called peaks (red tracks), ATAC-seq performed in differentiating SGBS cells (Nele Gheldof, unpublished, pink tracks), H3K27Ac ChIP-seq performed in human mesenchymal stem cells (h-MSCs) from Rauch et al. (2019) Nat. Gen. (dark blue track), chromHMM data from the WashU Epigenome Browser v52.0.0 (Li et al., (2019) Nucleic Acids Research) in adipose derived mesenchymal stem cells (Adipo-MSCs) and adipose nuclei (red = active TSS, orange = flanking active TSS, green = transcription, yellow = enhancer, blue = heterochromatin) and ChIP-seq data (blue box) from SMAD1/5 in human pulmonary artery endothelial cells (HPAECs) from Morikawa et al. (2019) Sci Signal, SMAD4 in TGFβ1 stimulated A2780 cells (ovarian cancer) from Kennedy et al. (2011) PLoS ONE, mm9 liftOver Smad4 in E12.5-13.5 mice forelimbs from Yan et al. (2018) J Biol Chem and CEBPα and PPARγ in differentiated SGBS cells from Schmidt et al. (2011) BMC Gen.

When comparing solely the ATAC-seq data between D0 and D12, we observe a more open chromatin signature at D0 around the *FAM46A* promoter, with more peaks appearing (Figure 2.3.6, pink tracks). At D12 ATAC-seq peaks observed at D0 are lost or the signal is reduced, indicating less open chromatin in mature SGBS adipocytes. The H3K27Ac Chromatin immunoprecipitation sequencing (ChIP-seq) data (Figure 2.3.6, dark blue track) overlaps well with the D0 ATAC-seq data, suggesting a reduction in enhancer signatures. It should be noted that this data was generated from human MSCs (h-MSCs), which were derived from bone marrow. Bone marrow-MSCs (BM-MSCs) are the pluripotent cell stage previous to pre-adipocytes, which then are able to terminally differentiate into adipocytes.

As peakC called most peaks upstream of the viewpoint, we scanned this region for any interesting overlays (Figure 2.3.7 and dashed box Figure 2.3.6). Using chromHMM tracks from adipose tissue derived MSCs (adipo-MSCs) and adipose nuclei to predict regulatory elements, we observed that peakC clusters from D0 overlap on multiple predicted active sites. There was also a general reduction in predicted regulatory elements when comparing the adipo-MSCs and the adipose nuclei (Figure 2.3.6, chromHMM tracks). The D0 peakC data has a tendency to overlap on TSSs (red) and enhancers (yellow), particularly in clusters that are not found in the D12 sample. The green annotations are indicative of active transcription, and are more enriched in adipo-MSCs, close to the peakC clusters. The peakC clusters found in the D12 sample mainly overlap on predicted heterochromatin regions and enhancers. Heterochromatin is a form of DNA that is tightly packed and associates with repressed gene expression. There are more heterochromatin annotations in the adipose nuclei, complementing the ATAC-seq data from mature adipocytes that have less accessible chromatin in this region.

Chapter 2

FAM46A has previously been found to interact with SMAD proteins (Colland et al. 2004; Watanabe et al. 2018). Thus, we compared publicly available ChIP-seq datasets baiting SMAD1, 2 and 4 (Figure 2.3.7, blue “ChIP-seq” box) were compared to the other datasets to determine if any SMAD proteins are involved in the regulation of *FAM46A* expression by binding to the promoter or any called interaction sites (Figure 2.3.7, red “peakC” tracks). Interestingly, SMAD4 data from TGFβ1-stimulated ovarian cancer cells have a called peak overlap with an interaction region close to a predicted TSS region (Figure 2.3.7, red asterisk). SMAD1/5 and Smad4 (LiftOver from mm19) called peaks do not overlap with any sites of interest. Furthermore, PPAR γ and CEBP α ChIP-seq data from mature SGBS cells was studied to check if these master regulators of adipogenesis may also be involved with the regulation of *FAM46A* expression (Figure 2.3.7, orange and purple tracks). However, most called ChIP-seq peaks from these two datasets did not overlap with the peakC data, except for one CEBP α peak set (Figure 2.3.7, blue asterisk), which only overlaps with the peakC data in the mature adipocyte datasets.

By compiling all the different genomic datasets, we observed a general trend of mature adipocytes (D12) having signatures of reduced transcriptional activity in the region around the *FAM46A* promoter when compared to pre-adipogenic cells (D0).

2.3.6. Motif enrichment and pathway analysis of interaction peaks

As the selected ChIP-seq datasets revealed very few potential TF binding to the *FAM46A* interactome region, a motif enrichment analysis was performed to explore other potential transcriptional regulators that may bind in the region. Due to the D0 sample shows a more active profile compared to the D12, the motif enrichment was performed using the called peaks from D0 only. As the recommended peak size for motif scanning is 200 bp, the D0 peak file (containing 190 peaks) was filtered for any overlapping peak regions, resulting in 82 peaks as an input for the

analysis. To perform the known motif enrichment analysis, the findMotifsGenome.pl tool was used from HOMER (Heinz et al. 2010). GENRE (Mariani et al. 2017) was also used in the analysis pipeline in order to create an appropriate background file for the differential motif discovery, which subsequently contained 78 sequences. After running the findMotifsGenome.pl, 76 motifs were identified with a p-value <0.01 (out of a total of 364 motifs identified). The top 20 output results from the analysis can be found in Table 2.3.4 ranked by p-value.

Table 2.3.4: Top 20 motif enrichment results from HOMER ranked by p-value

Known motif TF	Consensus	P-value	Log P-value	q-value (Benjamini)	# of Target Sequences with Motif	% of Target Sequences with Motif	# of Background Sequences with Motif	% of Background Sequences with Motif
Pit1	ATGCATAATTCA	1.00E-04	-1.15E+01	0.0037	8	9.76%	1.8	2.30%
Gfi1b	MAATCACTGC	1.00E-04	-9.36E+00	0.0079	7	8.54%	1.9	2.36%
SPI1	YCWGGAATGY	1.00E-04	-1.13E+01	0.0037	13	15.85%	3.8	4.81%
NR2F2	AGRGGTCA	1.00E-04	-9.68E+00	0.0076	12	14.63%	4	5.01%
Isl1	CTAATKGV	1.00E-04	-9.23E+00	0.0079	33	40.24%	18	22.62%
Gata1	SAGATAAGRV	1.00E-03	-7.36E+00	0.0288	6	7.32%	1.1	1.44%
Erra	CAAAGGTCAG	1.00E-03	-8.99E+00	0.0079	17	20.73%	6.7	8.46%
Lhx3	ADBTAATTAR	1.00E-03	-7.69E+00	0.0238	30	36.59%	16.5	20.79%
Smad2	AGGTGHCAGACA	1.00E-02	-5.53E+00	0.1004	5	6.10%	0	0.00%
HNF4a	CARRGKBAAAAGTYCA	1.00E-02	-5.53E+00	0.1004	5	6.10%	1.1	1.39%
Esrrb	KTGACCTTGA	1.00E-02	-5.53E+00	0.1004	5	6.10%	1.6	1.96%
Gata2	BBCTTATCTS	1.00E-02	-5.33E+00	0.1004	7	8.54%	2.8	3.47%
Oct6	WATGCAAATGAG	1.00E-02	-5.33E+00	0.1004	7	8.54%	2.8	3.53%
Sox17	CCATTGTTYB	1.00E-02	-5.54E+00	0.1004	9	10.98%	3.1	3.85%
DMC1	ADGGYAGYAGCATCT	1.00E-02	-5.54E+00	0.1004	9	10.98%	3.7	4.68%
TEAD2	CCWGGAATGY	1.00E-02	-5.54E+00	0.1004	9	10.98%	3.9	4.91%
Prop1	NTAATBNAATTA	1.00E-02	-6.72E+00	0.0441	15	18.29%	6.8	8.53%
HOXD13	NCYAATAAAA	1.00E-02	-6.72E+00	0.0441	15	18.29%	6.9	8.70%
Mef2b	GCTATTTTTGGM	1.00E-02	-6.18E+00	0.0687	16	19.51%	7.2	9.12%
GATA3	AGATAASR	1.00E-02	-5.63E+00	0.1004	21	25.61%	12	15.08%

Chapter 2

As a primary interpretation of the HOMER motif enrichment, a pathway analysis was performed on the top 69 significant TFs motifs (p-value <0.01 and % of background <25%) using Metascape (Y. Zhou et al. 2019). Table 2.3.5 and Figure 2.3.8 below show the results of the gene set enrichment. The top pathways include chordate embryonic development and cell fate commitment. Interestingly, the SMAD 2/3 pathway is included, which was also one of the top scoring enriched motifs found in the HOMER analysis (Table 2.3.4, grey highlight). A multitude of other differentiation pathways are significantly enriched, such as stem cell, myeloid and leukocyte differentiation (Table 2.3.5). The observed enrichment of pathways related to cell fate determination and differentiation (such as stem cell, myeloid and leukocyte differentiation) suggests a potential regulation of *FAM46A* during differentiation events in general.

Table 2.3.5: Top 20 enriched pathways from Metascape

Count = number of genes in input gene list found in pathway

% = percentage of Count in pathway (count/total number of genes in pathway).

GO	Category	Description	Count	%	Log10(P)	Log10(q)
GO:0043009	GO Biological Processes	chordate embryonic development	28	40.58	-25.61	-21.29
GO:0045165	GO Biological Processes	cell fate commitment	21	30.43	-24.04	-20.2
GO:0048598	GO Biological Processes	embryonic morphogenesis	20	28.99	-15.73	-12.56
M2	Canonical Pathways	PID SMAD2 3NUCLEAR PATHWAY	11	15.94	-15.31	-12.17
GO:0003007	GO Biological Processes	heart morphogenesis	15	21.74	-15.2	-12.08
GO:0001568	GO Biological Processes	blood vessel development	20	28.99	-13.61	-10.61
GO:0007548	GO Biological Processes	sex differentiation	14	20.29	-13.55	-10.57
GO:0032870	GO Biological Processes	cellular response to hormone stimulus	19	27.54	-13.18	-10.23
hsa05202	KEGG Pathway	Transcriptional misregulation in cancer	12	17.39	-12.95	-10.01
GO:0048863	GO Biological Processes	stem cell differentiation	13	18.84	-12.45	-9.61
GO:0031016	GO Biological Processes	pancreas development	9	13.04	-11.97	-9.19
GO:0021515	GO Biological Processes	cell differentiation in spinal cord	8	11.59	-11.46	-8.72
GO:0030099	GO Biological Processes	myeloid cell differentiation	14	20.29	-10.93	-8.27
GO:0002521	GO Biological Processes	leukocyte differentiation	15	21.74	-10.83	-8.19
GO:0001709	GO Biological Processes	cell fate determination	7	10.14	-10.57	-7.97
GO:0050673	GO Biological Processes	epithelial cell proliferation	14	20.29	-10.51	-7.91
GO:0007423	GO Biological Processes	sensory organ development	15	21.74	-10.47	-7.88
GO:0048483	GO Biological Processes	autonomic nervous system development	7	10.14	-10.42	-7.85
GO:0071363	GO Biological Processes	cellular response to growth factor stimulus	16	23.19	-10.02	-7.51
GO:0001763	GO Biological Processes	morphogenesis of a branching structure	10	14.49	-9.63	-7.15

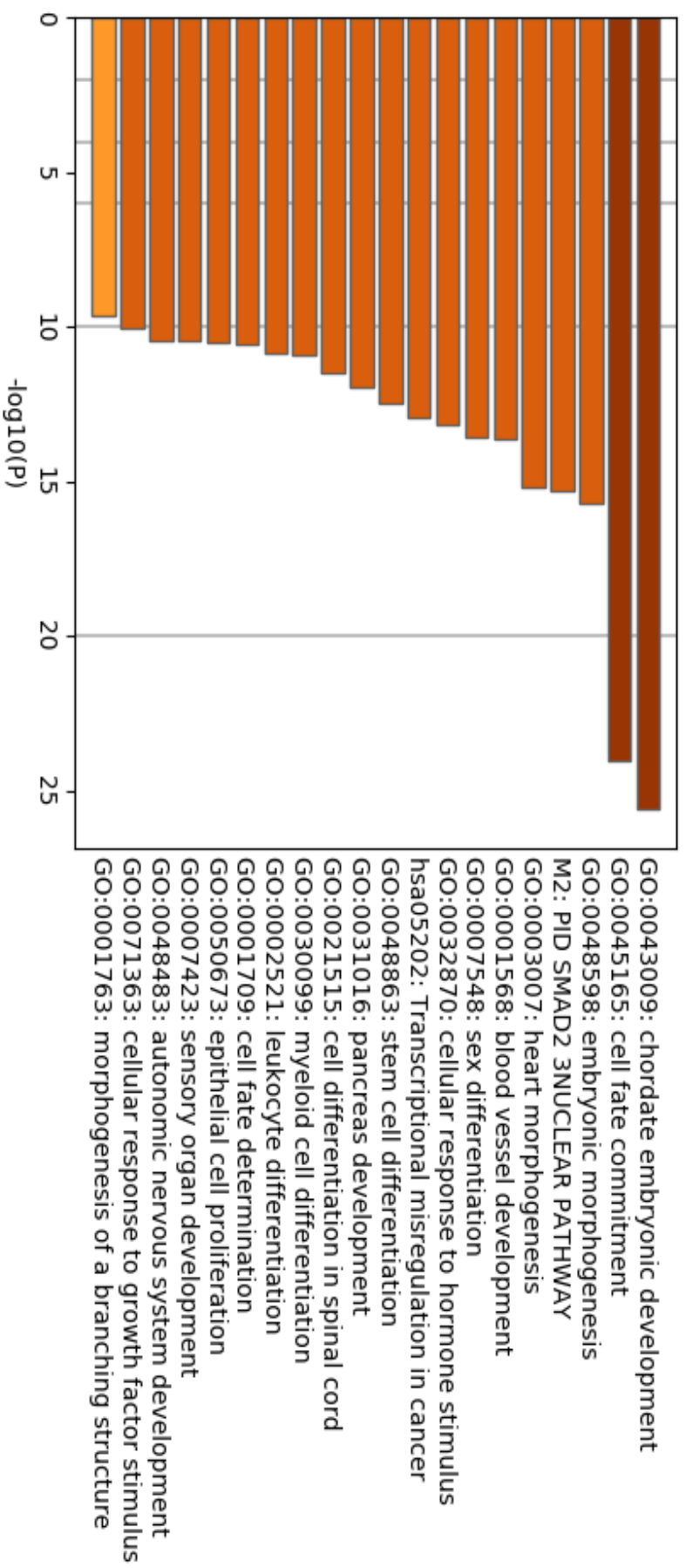


Figure 2.3.8: List of top 20 enriched pathways by p-value.

Figure extracted from the Metascape analysis report. The darker the colour, the more significantly the pathway is enriched. More details on each pathway and the analysis are found in Table 2.3.5.

To understand how these TFs and enriched pathways are involved in SGBS cells, microarray data from D0, D8 and D14 SGBS cells (data generated by Nele Gheldof, unpublished) was analysed to determine which TF are expressed in our cell model. Out of the 69 significant genes discovered by the HOMER motif enrichment analysis, 23 genes were expressed at a normalised expression (NE) above 6.60 at D0. Table 2.3.6 ranked the 23 genes by fold-change (FC) when comparing D14 vs D0. The genes highlighted in the table are also found in the top 20 enriched motifs from the HOMER analysis to emphasise their potential significance. One of the top scoring up regulated gene during SGBS adipogenesis was ESRRA (estrogen-related receptor A), which has been suggested to be a metabolic regulator (Ijichi et al. 2007). The most down regulated, TEAD2, is primarily involved in Hippo-signalling. However, studies in murine adipocytes (3T3-L1) have found that knock down of four TEAD proteins stimulates adipogenesis, and TEAD4 represses adipogenesis in a TEAD4-VGLL4-CtBP2 repressor complex that targets promoters of regulators of adipogenesis, such as PPAR γ (W. Zhang et al. 2018).

The motif enrichment and downstream analyses have shown that peaks interacting with the *FAM46A* promoter enrich for motifs targeted by TFs involved in cell fate and SMAD pathways. In SGBS cells these enriched TFs may be involved in adipogenic regulation and metabolism.

Chapter 2

Table 2.3.6: Expression of 23 genes from HOMER motif analysis in SGBS pre-adipocytes (D0), maturing adipocytes (D8) and mature adipocytes (D14) ranked by FC. Highly enriched genes (top 20 of HOMER analysis) are highlighted in grey.

Gene	ID	D0 Average NE	D8 Average NE	D14 Average NE	FC (D14 vs D0)	Log2FC
RXRA	ILMN_1687315	9.46	11.09	10.92	1.15	0.207
ESRRA	ILMN_1774272	7.21	8.26	8.44	1.17	0.226
AR	ILMN_1767351	7.02	7.47	7.65	1.09	0.125
USF2	ILMN_1756696	7.37	7.62	7.85	1.06	0.090
CEBPB	ILMN_1693014	11.14	11.13	11.4	1.02	0.034
USF1	ILMN_2353240	6.75	6.97	6.93	1.03	0.038
MEF2C	ILMN_1742544	6.78	7.03	6.92	1.02	0.030
NR2F2	ILMN_1745785	6.85	6.87	6.92	1.01	0.015
NPAS2	ILMN_1765558	6.87	6.97	6.87	1	0.001
ATF2	ILMN_1748271	6.8	6.67	6.79	1	-0.002
PROP1	ILMN_1730492	6.62	6.62	6.59	1	-0.005
NF1	ILMN_1726387	6.72	6.68	6.69	1	-0.006
SIX1	ILMN_1812702	7.1	7.22	7.07	1	-0.006
ISL1	ILMN_1763390	6.63	6.59	6.59	0.99	-0.007
SOX4	ILMN_1815745	8.43	8.75	8.32	0.99	-0.018
SOX15	ILMN_1784217	6.71	6.55	6.52	0.97	-0.040
ATF1	ILMN_1801923	6.92	6.76	6.73	0.97	-0.042
ZEB1	ILMN_1762231	6.71	6.55	6.49	0.97	-0.050
ARNT	ILMN_1762582	6.96	6.82	6.7	0.96	-0.055
TEAD2	ILMN_1682781	9.19	8.81	8.88	0.97	-0.050
TEAD4	ILMN_1705301	7.32	6.88	6.67	0.91	-0.133
Jun	ILMN_1806023	9.17	8.58	8.38	0.91	-0.130
TCF4	ILMN_1814194	8.28	7.37	7.32	0.88	-0.177

2.4. Discussion

Elucidating a functional role for variants identified from GWAS or QTL analyses is rare (as challenging) and requires extensive follow-up studies. In this study, a hot spot region on chromosome 6 was explored for potential regulatory activity, specifically in the regulation of downstream gene FAM46A, in human pre-adipogenic and mature adipocytes. A 3C experiment using the CaptureC method was optimised and performed in SGBS cells to understand how FAM46A is regulated, and if the hot spot region contributed to this mechanism.

3C methods are dependent on high cell numbers and high digestion efficiencies for optimal results and interaction maps. Optimisation for the cell type of interest is vital. From our study, we observed that digestion efficiencies improved in SGBS cells when lysed as a monolayer in the cell culture dish (vs collected and lysed in suspension), and when using *NlaIII* as the digestion restriction enzyme (vs *DpnII*). SGBS cells are adherent pre-adipogenic cells with fibroblast morphology (Wabitsch et al. 2001). When handling these cells, it was observed that they have a “sticky” nature when scraped from the plate, making homogenisation of the cell lysate difficult and a long process. This clumping of cells may have contributed to the reduction in resolution of the pilot CaptureC data by including intact cells during centrifugation and digestion steps, reducing the total amount of nuclei digested. Lysis as a monolayer vastly improved the digestion efficiency, perhaps by allowing more cells to be lysed and thus making the chromatin more accessible for digestion.

Further optimisations were required during the design of the biotinylated oligonucleotides that target the regions of interest; a unique aspect of the CaptureC method. Unfortunately, these oligos are dependent on the region of interest, as well as the restriction enzyme used since the oligonucleotide must contain the restriction site sequence. The restrictive nature of the DNA sequence makes this step difficult to optimise, even

Chapter 2

when using the recommended oligonucleotide web portal, Capsequm (Davies et al. 2016). From our results, probes designed for the *FAM46A* promoter (both under DpnII and NlaIII digestion) captured efficiently. The probes designed for the hot spot region required further optimisation. Since the region was almost 750 kbp, and 4-cutter fragments average around 250 bp, multiple probes could be, and were, designed. We found restriction fragment size to be an important factor, where the best captures occurred when regions were larger than 400 bp. Since probe pairs are 120 bp in length, perhaps overlapping regions on the probes disrupted the capture, thus capturing better when probes have no overlaps. However, certain probes passing all filtering steps, including fragment size, did not capture efficiently (Table 2.3.3, Pilot 3, probe pairs 814_707 and 815_416). It is difficult to elucidate why this is. Perhaps the DNA sequence in this region does not favour capture, or the chromatin is harder to access in these areas. ATAC-seq data from SGBS cells shows that the hotspot region is not highly accessible with low signal signatures appearing in this region (data not shown).

Using the aforementioned optimisations, the final CaptureC results showed comparable read statistics to the data produced in the original publication. However, the capture containing reads from our study was reduced compared to the *Hba* example, which may be a reflection of the capture efficiency of the probes or the strength of the interaction landscape around the *Hba* example. Interestingly, in our data the capture containing reads did not differ much from D0 to D12, despite the reduction in total mapped reads (two million less in D12 vs D0) (Table 2.3.3). In general, the D12 samples had less reads entering and exiting the analysis pipeline compared to D0. This may be due to the chromatin preparation and the differences in the samples. Since D12 samples are mature adipocytes containing lipid droplets, these droplets may have interrupted the various steps of the library preparation (Figure 2.3.1). For example, during centrifugation lipid containing cells would float to the surface,

making harvesting difficult. This lipid interruption in general would lead to less concentrated 3C libraries, and thus reduced reads in the final sequencing step.

Calling relevant and significant interactions in 3C data has not been standardised, with multiple methods currently available (McCord et al. 2020). Probing any region on the genome will create an interaction profile due to co-migration of near-by fragments during the cross-linking step, creating a random contact frequency map of decreasing signal from the viewpoint. Therefore, replicate experiments are crucial for identifying “true” interactions. The peakC (Geeven et al. 2018) software addresses these limitations by modelling the background distribution of random signal, taking into account the profile of the background rather than the total read counts. Since this method uses the distance from the viewpoint as a variable in the model, long distance peak calling is a limitation of this software. One of the furthest range *cis*-regulatory elements is a *Shh* enhancer, which was found >100 kbp away from the *Sox2* promoter (H. Y. Zhou et al. 2014). This is an exceptional case with a strong interaction signal. Less frequent interactions are more difficult to identify, as these signals are too weak to be filtered from the background signal, especially from a distance. When analysing our CaptureC data with peakC, we found that most of the interaction signals were occurring downstream of the *FAM46A* promoter instead of at the upstream hot spot region, as hypothesised (Figure 2.3.5). When targeting the hot spot region, there were significant interaction fragments close to the viewpoint in the D0 sample for both viewpoints, and none in the D12, and none close to the *FAM46A* promoter. As seen from the ATAC-seq and public H3K27Ac data, we observe that this region is one of low accessibility and activity, which may explain the lack of significant interactions in this region and to the *FAM46A* promoter (data not shown). However, this may also be due to a limitation of the peak calling, where the hot spot region is too distal for the model to detect the hypothesised interaction. This interaction may also be

Chapter 2

weak, making it more difficult to detect above the background signal, which is used as variable in the peakC peak calling software. The authors of the peakC software also note that the quality of the 3C template is an important factor for efficient peak calling, which may also explain why the D12 sample had less significant interactions detected vs the D0. Another explanation for the lack of significant interactions could be due to the large size of the hot spot region. CaptureC is dependent on the restriction enzyme of choice, which in turn limits the probing regions to specific DNA fragments flanked by restriction sites. The average spacing of restriction sites of a 4-cutter enzyme is 250 bp. Considering the hot spot region was almost 750 kbp long, choosing where to probe the region for capture can be challenging, especially with the probe design issues that had to be optimised. Perhaps we missed the “real” interaction region and were not probing at the correct interaction point, thus explaining the lack of interactions found. It is also important to consider that this hot spot region was found with QTL analyses and, although powerful, does not necessarily pinpoint the functional variants causing the observed phenotype. As the region upstream of the *FAM46A* promoter revealed significant interaction peaks when probing for the promoter, perhaps these loci are in linkage disequilibrium with the hot spot region, which would have been an interesting follow-up analysis.

Considering this upstream region from the *FAM46A* promoter contained the most significant interactions, the subsequent analysis was focused here. Comparative analysis of the significant peaks with publicly available genomic datasets and SGBS ATAC-seq data revealed that pre-adipogenic cells have a more accessible chromatin profile, and more predicted regulatory elements indicative of gene activity and transcription compared to differentiated cells (Figure 2.3.6 and 2.3.7). This could once more be an example of mature adipocytes containing lipid droplets producing lower quality data. However, considering that *FAM46A* expression decreases during adipocyte differentiation (Chapter 3, Figure 3.3.1), the reduction in

interactions and open chromatin does suggest regulation of differential *FAM46A* expression during differentiation.

The comparative analysis with ChIP-seq data sets found an overlapping SMAD4 binding site with our peakC data, close to a region of open chromatin and H3K27Ac activity (Figure 2.3.7, red asterisk). This experiment was performed in human ovarian cancer cells stimulated with TGFβ1, where unstimulated cells showed no binding at this site (data not shown). SMAD4 is part of the TGFβ and BMP signalling pathways, which regulates a whole array of cellular processes, from apoptosis to fibrosis (Massagué & Wotton 2000). These two pathways have opposing roles in adipogenesis. BMP proteins have been used to induce adipogenesis, while TGFβ has been found to have an inhibitory role via SMAD3 signalling (Li & Wu 2020). SMAD4 recruits the phosphorylated SMAD complexes for translocation into the nucleus to regulate target genes (Li & Wu 2020). As SMAD4 is included in both pathways, it is unclear how SMAD4 is involved in adipogenesis or *FAM46A* regulation. Previous studies found Fam46a involved in the BMP pathway (Watanabe et al. 2018), interacting physically with Smad4. However, FAM46A was also identified as an interaction partner of SARA, a recruitment protein for TGFβ receptor II (Colland et al. 2004), suggesting a dual role for FAM46A in both pathways, similar to SMAD4. To understand if the *FAM46A* promoter is a target gene of SMAD TF complexes, it would be interesting to perform a follow-up SMADs ChIP-seq study in differentiating SGBS cells under different ligand stimulations of the BMP and TGFβ pathways. The interaction landscape of the promoter could also once more be observed under these various stimulations, to discover if various interactions are ligand dependent.

HOMER was used as a motif enrichment tool with the significant peaks from the CaptureC data to understand what TFs could potentially bind the loci that interact with the promoter and in turn regulate *FAM46A*

Chapter 2

expression. Using the gene names of the TFs that recognise these motifs, a pathway analysis from Metascope found these TFs were involved in multiple development and differentiation pathways (Table 2.3.5 and Figure 2.3.8). Interestingly, the SMAD2/3 pathway was present as an enriched pathway, which is involved in the aforementioned TGF β -signalling pathway. When exploring the enriched TFs that were being differentially expressed in SGBS cells, we found *RXRA* and *ESRRA* were the top down-regulated, while *TEAD2* and *TEAD4* were down-regulated during adipogenesis (Table 2.3.6). RXR is the heterodimeric partner for PPAR γ to bind DNA to regulate downstream genes required for adipogenesis (Nakachi et al. 2008). *ESRRA* has been found to regulate genes involved in adipocyte differentiation, including PPAR γ (Ijichi et al. 2007). It is difficult to determine without more follow-up studies if these down-regulated TFs target the *FAM46A* promoter to repress expression, either by direct binding or to a repressor element in the interaction regions. The TEA domain family (TEAD) are known to be involved in the Hippo signalling pathway, and was recently shown to have a repressive role in adipogenesis (W. Zhang et al. 2018). As this family of genes follow the same expression pattern of *FAM46A* during adipogenesis, it can be suggested that these TFs up-regulate the expression of *FAM46A* in order to repress adipogenesis and keep cells in a pre-adipogenic or stem-like state. If this repression cascade of TFs is induced by TGF β -signalling, where SMAD4 is an additional regulator, cannot be said without the aforementioned proposed follow up studies.

The interesting observations that a SMAD4 binding site is found during TGF β stimulation, which is an anti-adipogenic treatment, the enriched SMAD2/3 motifs in the called interaction peaks and the active signature of the *FAM46A* promoter on D0 vs D12 cells leads to the possible role of the TGF β pathway as a regulator of *FAM46A* expression. *FAM46A* is highly expressed in non-differentiated SGBS cells and this expression

reduces as the cell differentiates (Chapter 3, Figure 3.3.1), indicating that the *FAM46A* protein has a potential role in precursor cells. Proposed follow-up experiments to further understand this possible regulatory role of the TGF β signalling pathway would include interruption of the expression of SMAD proteins (2/3 and 4) by RNA interference to observe changes in *FAM46A* expression and the interaction profile of the promoter, using CaptureC, with and without TGF β stimulation. As we also observed enriched motifs of other anti-adipogenic TFs, such as TEADs, it would additionally be interesting to knockdown these proteins to understand if these TFs are also involved in the regulation of the *FAM46A* promoter during adipocyte differentiation.

In conclusion, the optimal conditions for CaptureC were established in SGBS cells. When performing CaptureC around the *FAM46A* promoter we observe a significant interaction profile downstream of the promoter. These interaction peaks enrich for TF motifs in which their respective TFs are involved in pathways such as cellular differentiation and cell fate determination, including SMAD pathways, which have a dual role in adipogenesis. TFs with enriched motifs at *FAM46A* promoter interaction peaks, such as RXRA and ESRRA could potentially be repressing *FAM46A* expression during adipogenesis, while TEAD TFs may up-regulate *FAM46A* expression in pre-adipocytes (D0 cells). Taken together, these results suggest *FAM46A* expression is required to maintain the pre-adipogenic state and this may be dependent on TGF β -BMP pathways.

Chapter 2

2.5. Materials and methods

2.5.1. Cell culture

SGBS pre-adipocyte cells were used as a model for human adipocyte differentiation and were cultured as previously described (Wabitsch et al. 2001). Briefly, pre-adipocyte cultures were maintained and expanded in Complete Medium consisting of DMEM/F12 medium (Gibco) supplemented with 10% FBS (Gibco), 33 mM biotin (Sigma-Aldrich), 17 mM pantothenate (Sigma-Aldrich), and 100 U/mL penicillin-streptomycin (Gibco). Differentiation into adipocytes was induced using 0F basal medium (Complete Medium without FBS) supplemented with 10 mg/mL transferrin, 1 mM recombinant human insulin, 1 mM cortisol, 2 nM tri-iodothyronine (T3), 250 nM dexamethasone, 2.5 mM IBMX, and 10 mM rosiglitazone (Cayman Chemical). All compounds were from Sigma-Aldrich unless stated otherwise. After four days of induction, a maintenance medium was used consisting of 0F basal medium supplemented with 10 mg/mL transferrin, 1 mM human insulin, 1 mM cortisol, and 2 nM T3. Mature adipocytes were assayed at day 12 of differentiation for CaptureC.

2.5.2. CaptureC library preparation: 3C libraries

CaptureC libraries were prepared as previously described (Davies et al. 2016). To isolate nuclei from SGBS pre-adipocytes (D0) and mature adipocytes (D12), 20 million SGBS cells were grown in 15 cm cell culture dishes. Cells were washed with 0F medium and fixed with 10 mL of 2% paraformaldehyde in 0F basal culture media (80-90% confluency on day 0) for 10 min on a rocking platform. The fixation was quenched with the addition of glycine (final concentration 0.5 M) for 5 min. The fixation solution was removed, and the plates were washed twice with ice cold PBS. To lyse the cells, 3 mL of lysis buffer (10 mM Tris-HCl pH 8, 10 mM NaCl, 0.2% Nonidet P40 and 1X Protease Inhibitor Cocktail (ThermoFisher)) was added to each plate and placed on a rocking

platform at 4°C for 30 min. Cells were subsequently scraped on ice and pooled together in a 50 mL Falcon tube. Once harvested, cells were centrifuged at 2500g at 4°C for 10 min to remove cell debris. The nuclei pellet was snap frozen in liquid nitrogen and stored at -80°C until ready for digestion.

Pelleted nuclei were thawed at room temperature (RT) and washed once by resuspension in PBS (at RT). The suspension was transferred to a 1.5 mL Eppendorf tube and centrifuged for 5 min at 14,000g at RT. The washed pellet was re-suspended in 1X CutSmart buffer (NEB) and transferred to a 1 mL Dounce homogeniser (Hamilton) placed on ice for homogenisation (45 – 55 strokes). The homogenate was then transferred to a clean 1.5 mL Eppendorf tube. The homogeniser was rinsed with an additional 400 µL buffer and added to the tube. The homogenate was centrifuged once more for 5 min at 14,000g at RT. 650 µL 1X CutSmart buffer was added and the pellet was re-suspended. Homogenisation was repeat if pellet did not re-suspend easily. Three parallel digestion reactions were set up (200 µL each), including a small undigested control sample (50 µL), with the addition of 60/15 µL 10X CutSmart buffer, 10/2.5 µL 20% SDS (Bio-Rad), 434/116 µL nuclease-free water, 66/16 µL Triton X-100 (Sigma-Aldrich) and 500 U of NlaIII restriction enzyme (NEB). The digestion was performed on a shaking thermomixer at 37°C overnight, where an additional 500 U of enzyme was added after 5 – 6 hours of initial digestion. After overnight incubation, another 500 U of NlaIII was added and incubated for another 3 – 5 hours. A digestion control was collected by pooling 100 µL of lysate from each digestion reaction. Both digestion controls were treated and extracted in a scaled-down version of the protocol below. The restriction enzyme was heat inactivated at 65°C for 20 min, and the digests were cooled on ice. The ligation reactions were prepared (500 µL nuclease-free water, 134 µL of 10X ligation buffer and 8 µL 30 U/ µL T4 ligase (ThermoFisher)) and incubated overnight at 16°C on the Thermomixer. The digestion control

Chapter 2

samples had 3 mL Proteinase K (Ambion, 20 mg/mL) added and were incubated overnight at 65°C overnight. To de-crosslink the samples, 5 mL of Proteinase K was added to each reaction, which were incubated overnight at 65°C. Each reaction was treated with RNase-A (ThermoFisher) at 37°C for 30 min. Reactions were pooled together in a 15 mL Falcon tube and DNA was purified utilising a phenol-chloroform and chloroform extraction. To precipitate the DNA, the extracted samples were added to a 50 mL Falcon tube with 7 mL nuclease-free water, 1 mL 3M sodium acetate, and 35 mL of 100% ethanol and placed in -80°C overnight. The precipitate DNA was pelleted by centrifugation at 2,200g at 4°C for 45 min. The pellet was then washed with 70% ethanol and dried at RT. The 3C library was reconstituted in nuclease-free water and stored at -20°C.

2.5.3. CaptureC library preparation: sequence adaptors and capture

To prepare the 3C libraries for capture and sequencing, 3 mg of purified library was sonicated into 200 bp fragments using the Covaris E220 (duty cycle 10%; intensity 5; cycles per burst 200; 6 cycles of 60 seconds). Quality controls were performed using the Agilent 2200 TapeStation (High Sensitivity). 1 mg of sheared 3C library was used for end repair and adaptor ligation using NEBNext reagents (NEB; E7370 and E7335). Indexes were added using NEB index primers with 8 PCR cycles using the Herculase II PCR kit (Agilent). Size selection and purification steps between each reaction was done with Clean NGS beads (Clean NA).

1 – 2 mg of indexed library was used for the capture step. Biotinylated DNA oligonucleotides were diluted to 2.9 mM and mixed at a ratio of 1:1. The Nimblegen SeqCap EZ kits were used (Roche; 05634261001, 07145594001 and 06777287001) for the hybridisation steps. Libraries were mixed at a 1:1 ratio by mass with 5 mg COT DNA, HE universal oligo and HE index oligo (both 1,000 pmol) and vacuum centrifuged at 45°C until dry. The DNA was re-suspended in hybridisation buffer and

component A and denatured at 95°C for 10 min. The libraries were added to 13 pmol of preheated probe mix and incubated at 47°C for up to 72 hours. Captured fragments were recovered with 100 mL of streptavidin beads (ThermoFisher), which were pre-washed twice with SeqCap EZ bead wash buffer (using a magnet scaffold to capture the beads), at 47°C for 45 min whilst shaking. The beads were then washed with SeqCap EZ buffers I, stringent wash, II and III. After the wash steps, PCR-grade water was added to the beads. The captured material was PCR amplified using the beads as a template (7 – 9 cycles). A second round of hybridisation (up to 24 hours) was performed with 1 – 2 mg of captured library. Library quality and quantity was assessed with the Agilent 2200 TapeStation and Qubit dsDNA kit (ThermoFisher).

2.5.4. Sequencing parameters

The pool of libraries was loaded on a MiSeq flowcell (Illumina; 300 np, V2, MiSeq4) at 8 pM with a 5% PhiX spike. 3 to 5 million 151 bp pair end reads was produced per sample.

2.5.5. Data analysis

Raw sequencing reads were quality checked using fastQC (Andrews 2010). To remove sequencing adapters, trim_galore (Krueger 2012) was used with the paired parameter including both reads of each library. Once trimmed and the removal checked again with fastQC, paired-end reads were combined using FLASH (Magoč & Salzberg 2011) with interleaved parameters. To ensure full inclusivity of reads, paired and unpaired reads were concatenated in the same output file. To create an *in silico* NlaIII digestion library of the reads, a modified version of the DpnII2E.pl script (Davies et al. 2016) was used, named NlaIII2E.pl. This script scans each paired read to find the appropriate restriction site sequence and separates the reads at the site, while still recognizing the partner fragment. Once the reads had been separated by the restriction site, the reads were mapped to the hg19 genome using bowtie2 (Langmead & Salzberg 2012) with the

Chapter 2

follow parameters: -p 1 -m 2 --best --strata --sam --chunkmb 256. With the subsequent mapped SAM file, the data was processed with the CCAalyzer3.pl script (Davies et al. 2016) to identify the interactome map of the bait of interest. Wig files were viewed using IGV and the UCSC genome browser (Kent et al. 2002).

Follow-up analyses were performed using the R script peakC (Geeven et al. 2018) to identify significant peaks in a 1000e3 MB window by combining biological replicates. All default settings were used. The output was a list of significant regions, which was used for HOMER analysis (Heinz et al. 2010) for identification of potential transcriptional factor motifs/binding sites at our peaks with default settings, except for the background (-bg). A background file was created using GENRE (Mariani et al. 2017) using called interaction peaks with a 200 bp border as the input file. HOMER's output list of significant TFs motifs (p-value <0.01 and background percentage <25%) was manually screened for the gene names of TFs that recognise the motif. This gene list was used with Metascape (Y. Zhou et al. 2019) for a pathway analysis.

Chapter 3:

Functional characterisation of FAM46A in human adipocytes and murine MSCs

3. Functional characterisation of FAM46A in human adipocytes and murine MSCs

3.1. Abstract

FAM46A is a NTase protein that is ubiquitously expressed across tissues, and currently its function is unknown. Recently, a potential role for FAM46A was suggested in adipose tissue, where FAM46A is differentially expressed during adipogenesis and weight-loss (Carayol et al. 2017). In order to follow up on this study, we attempted to characterise FAM46A in human SGBS cells and murine MSCs during cell differentiation.

As previously observed, FAM46A expression decreased during adipogenesis in SGBS cell and primary APCs samples. Fractionation of adipose tissue and subsequent qPCR analysis showed FAM46A expression was enriched in the SVF of adipose tissue. Complementary data from Tabula muris (Tabula Muris Consortium et al. 2018) also showed an enrichment of Fam46a expression in the MSCs found in fat tissue. RNA interference of FAM46A in vitro lead to a decrease in adipogenesis markers and a decrease in lipid accumulation in differentiating SGBS adipocytes. In murine C3H10T1/2 MSCs, over-expression of Fam46a caused the opposite effect, where adipogenesis markers increased in expression. Primary BM-MSCs had a high expression of Fam46a with a dynamic expression profile during adipo- and osteogenesis. Treatment with small interfering RNA (siRNA) targeting Fam46a in BM-MSCs led to an increase in the expression of MSC marker, Vim.

In conclusion, the in vitro characterisation of FAM46A showed a potential regulatory role during differentiation, however, this may be

biased by the cell systems utilised and requires more in depth study.

3.2. Introduction

3.2.1. Cellular models for adipogenesis

Adipogenesis is a complex process that consist of progenitor cells, such as MSCs or APCs, maturing into a unilocular lipid bearing cell (Rosen & MacDougald 2006). Due to the current obesity epidemic, there has been emphasis on studying the process of adipogenesis, as this can lead to possible therapeutics targeting the adipose tissue for management of this disease. Expansion of the adipose tissue occurs either by hypertrophy, where the cells expand as lipid accumulates, or hyperplasia, where new adipocytes arise from differentiation of APCs within the tissue (Haczeyni et al. 2018). To study adipogenesis, biopsies of adipose tissue can be treated with collagenase to extract the APCs. in vitro induction of differentiation can be performed with a compound cocktail of insulin, a glucocorticoid (i.e. dexamethasone), an intracellular cAMP stimulator (i.e. methylisobutylxanthine) and serum (Scott et al. 2011). PPAR γ agonists, such as rosiglitazone, are also commonly used to induce differentiation (Bartelt & Heeren 2014). However, primary cell culture of APCs is not always accessible and requires preparation of collected adipose tissue. This process is lengthy and is dependent on the size the adipose tissue sample for sufficient APCs to work with. Fortunately, there are a number of precursor cell lines with the potential to differentiate into adipocytes that have been established as appropriate models to study adipogenesis.

Currently, the most common models for studying adipogenesis and adipocyte biology are the following. 3T3-L1 cells are immortalised murine pre-adipocytes that were discovered in the 1970s due to their spontaneous capability to differentiate and accumulate lipids (Green & Meuth 1974). More than 8,000 articles involving 3T3-L1 cells have been published to date, making it the most popular cell

model for adipogenesis (Sadie-Van Gijzen 2019). As seen in Chapter 2, our chosen cell model for adipogenesis were Simpson-Golabi-Behmel syndrome or SGBS cells, which is another popular cell line. This model for human adipocytes derives from a male infant patient with SGBS, a rare genetic disorder characterised by overgrowth and tissue dysplasia (Neri et al. 2013). Pre-adipocytes were prepared from a subcutaneous adipose tissue sample from this patient, and these cells have a high generation capacity without immortalisation (Wabitsch et al. 2001).

As adipocytes are derived from MSCs, cell lines classed as MSCs are also commonly used models for adipogenesis. C3H10T1/2 cells are murine pluripotent MSCs, also discovered in the 1970s that are able to differentiate into adipocytes, as well as osteoblasts and chondrocytes (L. Zhao et al. 2009). Like 3T3-L1 cells, they are an immortalised line, and have been established as an alternative to primary BM-MSCs.

3.2.2. Research question

As *FAM46A* was found to be differentially expressed in subcutaneous adipose tissue during a weight-loss event (Carayol et al. 2017), we were interested to understand the potential role of *FAM46A* in adipose tissue biology. Other than the Carayol et al. study, there is no characterisation of *FAM46A* in an *in vitro* cell model for adipogenesis. Potential functional roles of *FAM46A* include interaction studies, where a yeast-two hybrid screen for Smad-signalling proteins identified Fam46a as a potential interaction partner for SARA (Colland et al. 2004) and immunoprecipitation assays in developing *Xenopus* embryos found that Fam46a interacted Smad1 and Smad4 (Watanabe et al. 2018), in the TGF β -signalling pathway. Both of these pathways have regulatory functions in adipogenesis. Therefore, we attempted to

Chapter 3

characterise the function of FAM46A during adipogenic differentiation, primarily with human adipocyte cell model SGBS, and secondly with murine MSCs, to further understand the role of this relatively unknown protein.

3.3. Results

3.3.1. FAM46A expression in adipocytes

To understand the pattern of expression of *FAM46A* in SGBS cells during adipogenesis, two different methods of gene expression detection were analysed: in-house microarray data (data produced by Nele Gheldof, unpublished) and qPCR data collected during adipogenesis. Both assays reveal a trend of *FAM46A* expression decreasing as cells differentiate into mature adipocytes in SGBS cells (Figure 3.3.1). When observing microarray data from primary human adipocytes samples, the same reduction in expression is observed, showing this differential expression also occurs in other adipocyte cell models and is not unique to SGBS cells.

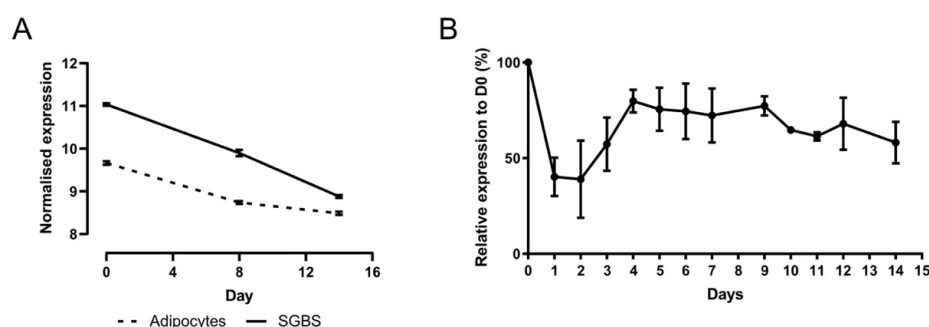


Figure 3.3.1: Gene expression pattern of *FAM46A* in SGBS cells and human adipocytes.

A) *FAM46A* microarray normalised expression data in human primary adipocytes and SGBS cells. Cells were collected at day 0 (before induction of adipogenesis), day 8 and day 14 of adipogenesis.

B) qPCR adipocyte differentiation time course of *FAM46A* expression in SGBS cells from day 0 to day 14. Expression ($2^{-\Delta Ct}$) is plotted relative to day 0.

Datapoints are plotted as mean \pm SEM.

3.3.2. Disruption of *FAM46A* expression during adipogenesis

Since *FAM46A* shows a differential expression pattern during adipogenesis, we were interested in understanding if *FAM46A* had a regulatory role in this process. To study this, siRNAs targeting *FAM46A* mRNA were used to reduce *FAM46A* expression and observe the effect on the expression of key regulators of

Chapter 3

adipogenesis; *PPAR γ* and *CEBPA*. In addition, over-expression plasmids were used to see if an excess of *FAM46A* expression would also have an effect on adipogenesis. SGBS cells were transfected with the constructs two days prior to induction of adipogenesis (day -2), and samples were collected for qPCR analysis at day 0 and day 6 of adipogenesis.

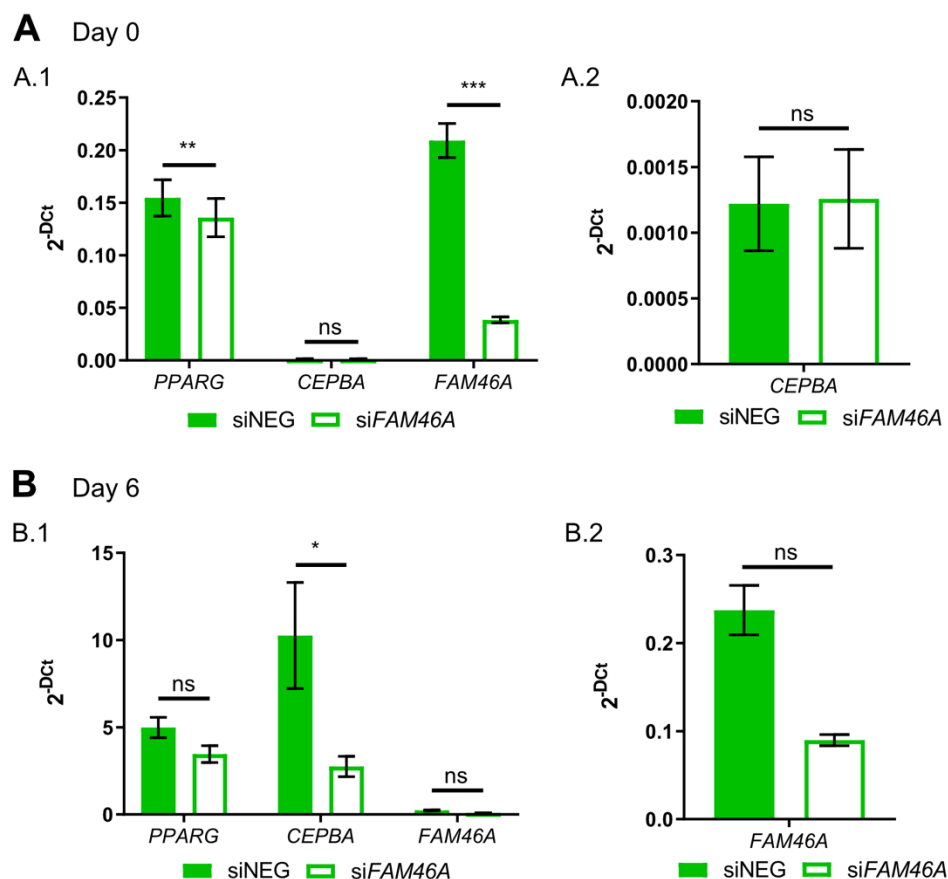


Figure 3.3.2: qPCR plots of *PPAR γ* , *CEBPA* and *FAM46A* in SGBS cells treated with siRNA targeting *FAM46A* expression.

A) Day 0 SGBS samples two days after transfection with siRNA targeting *FAM46A* (siFAM46A) and a negative control siRNA (siNEG).

B) Day 6 differentiated SGBS samples transfected with siRNA targeting *FAM46A* (siFAM46A) and a negative control siRNA (siNEG).

Expression is represented in $2^{-\Delta Ct}$ relative to housekeeping gene *TBP*. Bars are plotted as mean \pm SEM. Statistical tests are unpaired student t-tests with Welch correction. *p-value $p < 0.05$, **p-value $p < 0.01$, ***p-value $p < 0.001$, ****p-value $p < 0.0001$.

The siRNA constructs (ThermoFisher) used to target *FAM46A* expression were able to significantly reduce expression on day 0 (Figure 3.3.2A.1), however, this effect begins to diminish by day 6

of differentiation (Figure 3.3.2B.2), where it is no longer significantly different from the control siRNA (siNEG). The expression of *PPAR γ* on day 0 is affected by the siRNA treatment, where siFAM46A samples present a reduced *PPAR γ* expression (Figure 3.3.2A.1). *CEBPA* expression is too low at day 0 to observe any effects (Figure 3.3.2A.2). On day 6, however, *PPARG* and *CEBPA* expression increases as adipocytes mature, and we are able to capture a significant reduction in *CEBPA* expression in the day 6 samples (Figure 3.3.2B.1). *PPARG* expression also shows a similar trend, however, it is no longer considered significant.

A second transfection was performed using a CMV-controlled over-expression plasmid for FAM46A and an empty CMV plasmid as the negative control (OriGene) to observe if up-regulation of FAM46A would also affect the expression of the adipogenesis regulators. Overall, the over-expression had no significant effect (Figure 3.3.3). The plasmid was inconsistent with the level of over-expression between replicate samples, as seen from the wide error margin when probing for FAM46A expression (Figure 3.3.3A). There was no significant difference in *PPARG* and *CEBPA* expression between the different plasmid treatments, nor on the different days.

Chapter 3

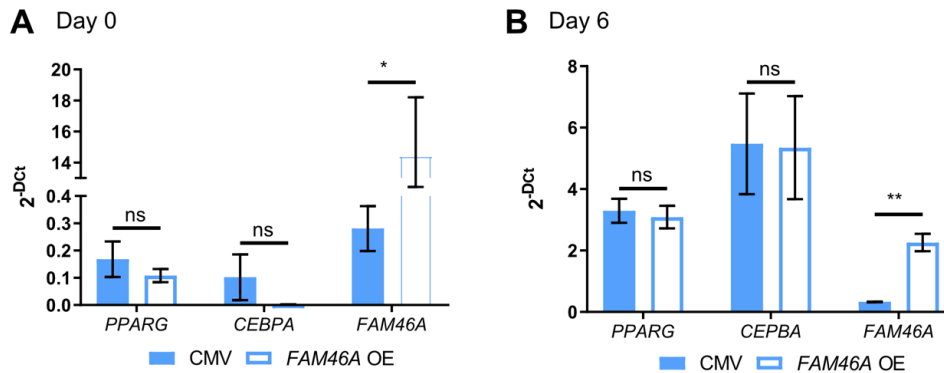


Figure 3.3.3: qPCR plots of *PPAR*_γ, *CEBPA* and *FAM46A* in SGBS cells treated with CMV-controlled expression plasmids for *FAM46A*.

A) Day 0 SGBS samples two days after transfection with *FAM46A* over-expression plasmid (*FAM46A* OE) and an empty CMV plasmid as a negative control (CMV).

B) Day 6 differentiated SGBS samples transfected with *FAM46A* over-expression plasmid (*FAM46A* OE) and an empty CMV plasmid as a negative control (CMV).

Expression is represented in $2^{-\Delta C_t}$ relative to housekeeping gene *TBP*. Bars are plotted as mean \pm SEM. Statistical tests are unpaired student t-tests with Welch correction. *p-value $p < 0.05$, **p-value $p < 0.01$, ***p-value $p < 0.001$, ****p-value $p < 0.0001$.

To further assess the effect of *FAM46A* disruption on adipogenic functionality, the lipid accumulation of differentiated SGBS was quantified using a fluorescent microscopy-based method developed by Dr. Magda Zachara in the Deplancke lab (see section 3.5 for detailed method). Cells were assayed on day 6 of differentiation. The lipid accumulation of each treatment is calculated by quantifying the fluorescent lipid signal and normalising it to the fluorescent nuclei signal (AdipoScore). As shown in Figure 3.3.4, the si*FAM46A* treatment had a significant effect on lipid accumulation, where it is reduced in comparison to the siNEG treatment. Similar to the qPCR experiment, the over-expression of *FAM46A* had no effect when compared to the empty CMV plasmid. However, both of these constructs had a low AdipoScore, even with similar cell numbers to the siRNA treatment (Figure 3.3.4A.2), indicating the constructs may be perturbing the adipogenic potential of the SGBS cells.

In summary, experiments in SGBS cells showed that a reduction of *FAM46A* in pre-adipocytes reduces the expression of master

regulators *PPAR γ* and *CEBPA* during adipogenesis. This effect was verified by observing a reduction in lipid accumulation when *FAM46A* was knocked down, suggesting FAM46A may have a role in adipogenesis.

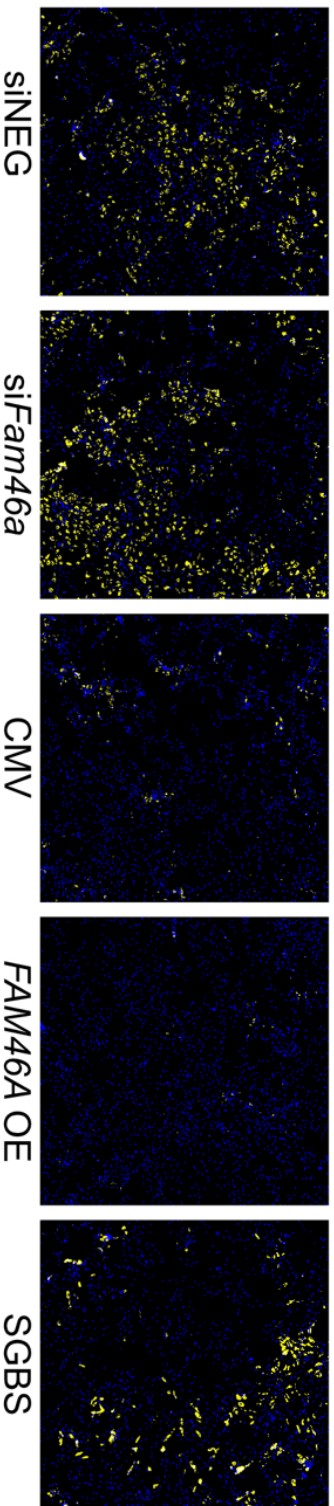
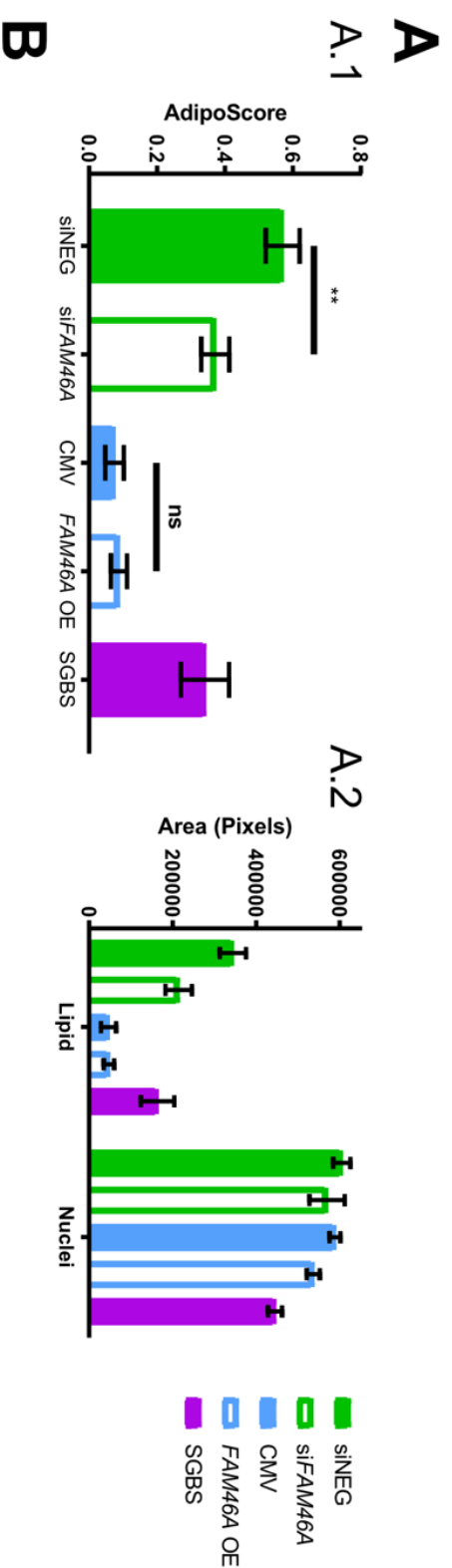


Figure 3.3.4: Lipid quantification of day 6 SGBS adipocytes under different FAM46A disruptions

A.1) AdipoScore of SGBS cells treated with siRNAs (control siNEG and siFAM46A) and CMV-controlled over-expression plasmids (control empty CMV (CMV) and CMV-FAM46A over-expression (FAM46A OE)). Control untreated SGBS cells were also assayed.

A.2) Absolute area values (in pixels) of fluorescent signal from lipids (BODIPY) and nuclei (Hoechst).

B) Microscopy images of example wells from each treatment (siNEG, siFAM46A, CMV, FAM46A OE and untreated SGBS). Yellow fluorescent signal is the lipid stain (BODIPY) and blue fluorescent signal is the DNA stain (Hoechst).

Bars are plotted as mean \pm SEM. Statistical tests are unpaired student t-tests with Welch correction. *p-value $p < 0.05$, **p-value $p < 0.01$, ***p-value $p < 0.001$, ****p-value $p < 0.0001$.

3.3.3. Expression in human adipocyte samples and mouse depots

Since *FAM46A* is differentially expressed during adipogenesis, where it is most expressed in the pre-adipocyte and early stages, we were interested in studying *FAM46A* expression patterns in the different cell populations in WAT. WAT contains several different cell types, where the majority of the tissue is the lipid-bearing white adipocytes. The remaining cells can be found in the SVF, which contains several immune cells types, committed pre-adipocytes and precursor cells (Ehrlund et al. 2016). In collaboration with Dr. Nathalie Viguerie at the I2MC (Inserm Toulouse), subcutaneous adipose tissue biopsies were taken from human patients for qPCR analysis. The whole adipose tissue (AT) was separated into two separate fractions; isolated adipocytes (IA) and the SVF. When analysing *FAM46A* expression in each fraction, the SVF had the highest expression of *FAM46A* when compared to the IA (Figure 3.3.5). This complements the SGBS data, where pre-adipocytes have a higher expression compared to mature adipocytes.

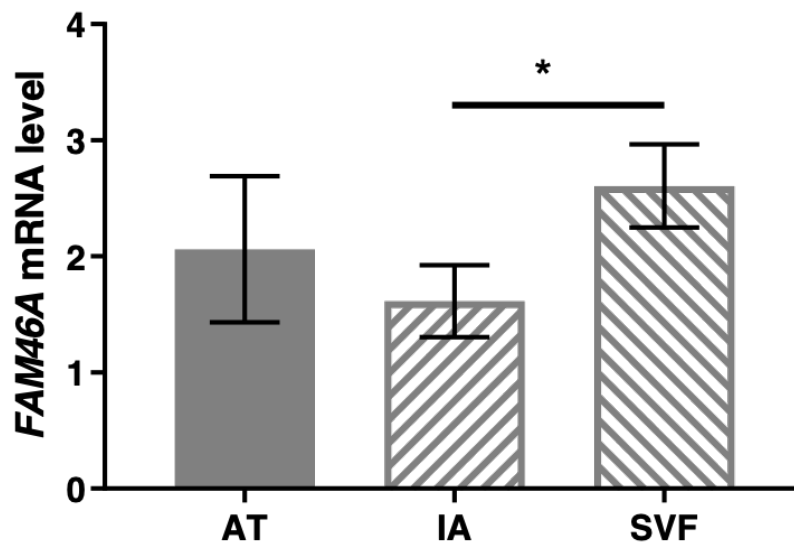


Figure 3.3.5: FAM46A expression in human subWAT samples

Adipose tissue (AT) was fractionated into isolated adipocytes (IA) and the stromal vascular fraction (SVF) and FAM46A expression was measured by qPCR. In collaboration with Dr. Nathalie Viguerie at the I2MC (Inserm Toulouse)

Bars are plotted as mean ± SEM. Statistical tests are unpaired student t-tests with Welch correction. *p-value $p < 0.05$, **p-value $p < 0.01$, ***p-value $p < 0.001$, ****p-value $p < 0.0001$.

To verify this expression pattern in other systems, we used single-cell RNA sequencing (scRNA-seq) data from the public resource *Tabula Muris* (Mouse Atlas) (Tabula Muris Consortium et al. 2018). Using the FACS-based transcript analysis in mouse fat tissue, we observed that *Fam46a* is highly expressed in the MSC population, and the myeloid population (Figure 3.3.6). Myeloid cells are progenitor cells originating from hematopoietic stem cells that are able to differentiate into several immune cells, such as macrophages. There is controversial evidence that that adipocytes can also originate from the myeloid lineage (Sanchez-Gurmaches & Guertin 2014). From this, *Fam46a* appears to be highly expressed in progenitor type cells, including pre-adipocytes, MSCs and myeloid cells.

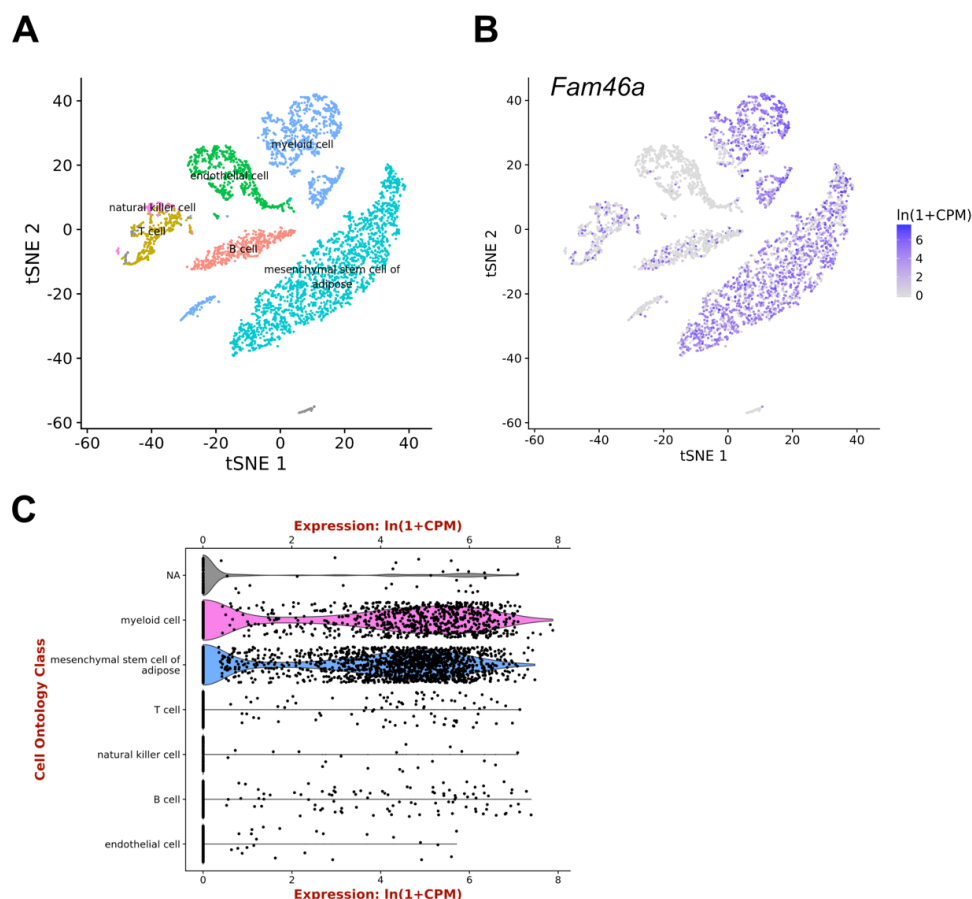


Figure 3.3.6: *Fam46a* expression in murine fat tissue (Tabula Muris)

A) tSNE analysis of different populations found in fat tissue labelled by cell ontology class

B) tSNE analysis of different populations found in fat tissue labelled with *Fam46a* expression patterns.

C) *Fam46a* expression levels of each cell separated by cell ontology class.

All figures were taken from Tabula Muris Consortium et al. (2018)

3.3.4. *Fam46a* expression in murine MSCs

As the data from *Tabula muris* showed an enrichment of *Fam46a* expression in progenitor cells, we were interested in studying the role of *Fam46a* in an MSC cell line, as SGBS cells have already committed to the adipocyte lineage. Murine C3H10T1/2 cells have been shown to be pluripotent stem cells, able to differentiate into adipocytes, chondrocytes and osteoblasts (L. Zhao et al. 2009). We also tested primary BM-MSCs from C57BL/6J mice (special thanks to Frédérica Schyrr (EPFL) for the protocol) as an additional model. To validate these cell models, we differentiated cells into adipocytes and osteoblasts while monitoring the expression of key

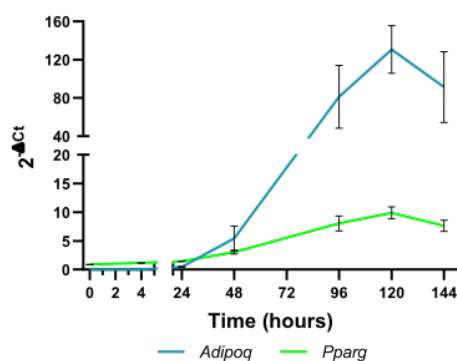
Chapter 3

regulators of each process, and *Fam46a*, with qPCR. Each qPCR experiment and analysis were performed by Patricia Leone (NIHS).

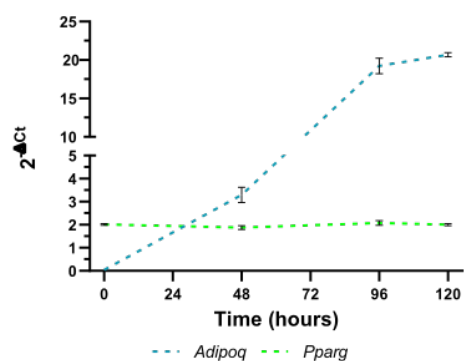
As seen in Figure 3.3.7A, adipogenic markers Adipoq and PPAR γ were used as indicators of differentiation. Adipoq codes for adiponectin, an adipokine secreted by mature adipocytes (Esmaili et al. 2020). C3H10T1/2 cells show an increase in both markers as cells differentiate, while BM-MSCs only show an increase in Adipoq. ORO staining of C3H10T1/2 cells confirm that the cells are accumulating lipids (Figure 3.3.8A). As for *Fam46a* expression, C3H10T1/2 cells display a relative low expression (Figure 3.3.7A.3/B.3). Surprisingly, *Fam46a* expression in C3H10T1/2 cells, while low, primarily show an increase in expression which starts to decrease at day 5 of differentiation. Conversely, BM-MSCs have a much higher expression (Figure 3.3.7A.4/B.4) which begins to decrease at day 4 after induction. During osteogenesis, Runx2, considered the master regulator of osteogenesis (Komori 2019), increases during differentiation but uniquely in the BM-MSC samples (Figure 3.3.7B.2). Interestingly, in C3H10T1/2 cells PPAR γ is increasing in expression during osteogenesis. To confirm the osteogenic lineage, ALP activity assays and staining was performed on C3H10T1/2 cells (Figure 3.3.8B). ALP activity increases as cells differentiate, and 5-bromo-4-chloro-3-indolyl phosphate/nitro blue tetrazolium (NBT), a substrate of ALP, staining confirms this activity. Interestingly, the osteoblast-induced cells also accumulated lipids (arrows on Figure 3.3.8B), although not as prominently as the adipogenic C3H10T1/2.

A Adipogenesis

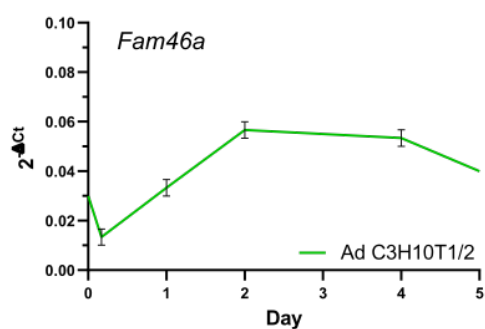
A.1



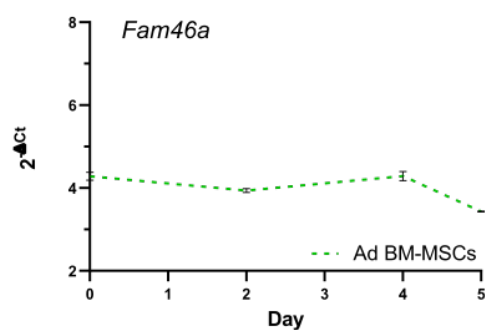
A.2



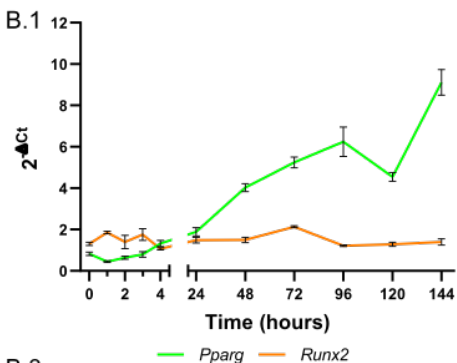
A.3



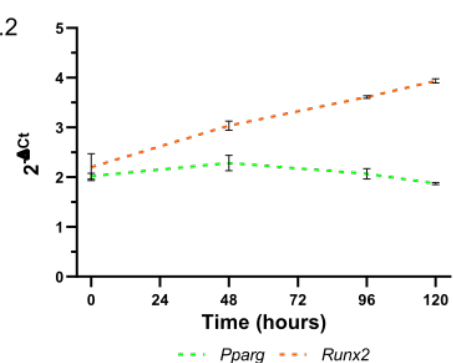
A.4

**B** Osteogenesis

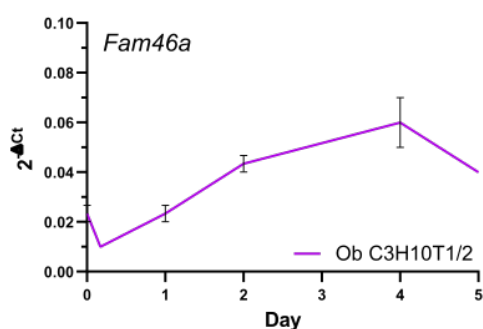
B.1



B.2



B.3



B.4

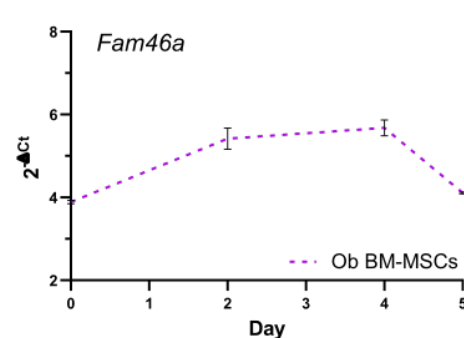


Figure 3.3.7: qPCR analysis of markers in A) adipogenic and B) osteogenic differentiation of C3H10T1/2 (solid lines) and BM-MSCs (dashed lines).

Expression is represented in $2^{-\Delta Ct}$ relative to housekeeping gene *Tbp*. Bars are plotted as mean \pm SEM. Statistical tests are unpaired student t-tests with Welch correction. *p-value $p < 0.05$, **p-value $p < 0.01$, ***p-value $p < 0.001$, ****p-value $p < 0.0001$.

Chapter 3

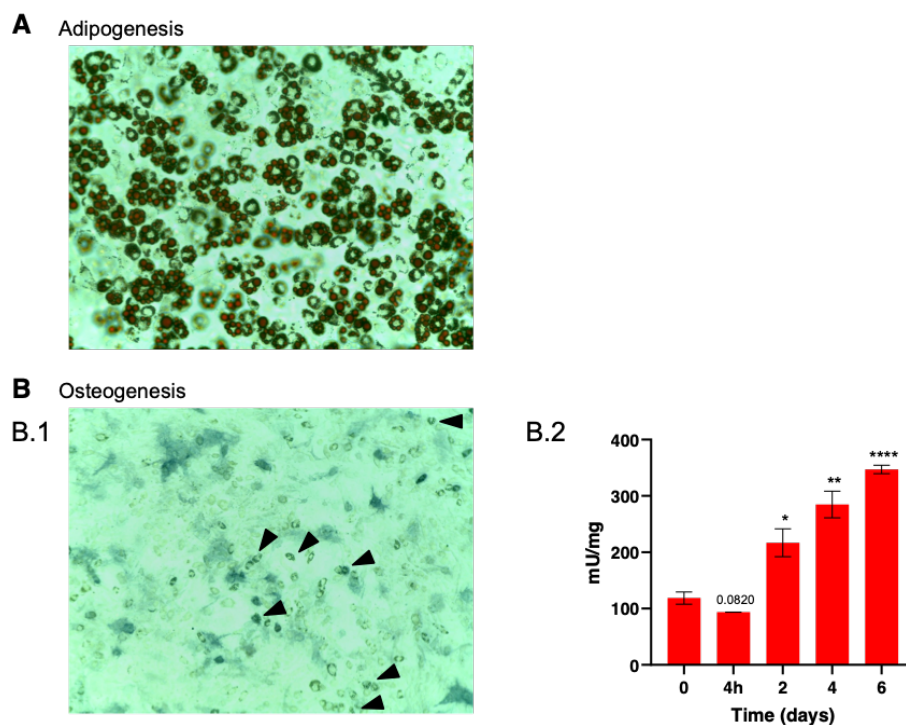


Figure 3.3.8: A) ORO staining on day 6 of adipogenic differentiation in C3H10T1/2 cells

B) ALP activity in C3H10T1/2 osteoblasts

B.1) SIGMAFAST BCIP/NBT staining for ALP activity on day 6 of osteogenesis

B.2) ALP activity monitoring during osteogenesis (day 0 to day 6). Arrows indicate lipid droplets.

Bars are plotted as mean \pm SEM. Statistical tests are unpaired student t-tests with Welch correction compared to day 0. *p-value $p < 0.05$, **p-value $p < 0.01$, ***p-value $p < 0.001$, **** p-value $p < 0.0001$.

3.3.5. Disruption of *Fam46a* expression in murine MSCs

Similar to the experiments performed with SGBS cells, we were interested in monitoring any effects of knocking down or over-expressing *Fam46a* in the C3H10T1/2 cells during adipo- and osteogenesis. Due to limited access to primary BM-MSCs, we only performed RNA interference with these cells with no induction of differentiation.

Figure 3.3.9 is a summary of the results in C3H10T1/2 cells. During adipogenesis, knock down of *Fam46a* significantly reduced the expression of *Adipoq* (adiponectin) on day 2 of differentiation (Figure 3.3.9A.1). However, the expression of *Fam46a* was no longer disrupted by the siRNA on day 2. The over-expression of *Fam46a* caused the opposite effect, where *Adipoq* expression

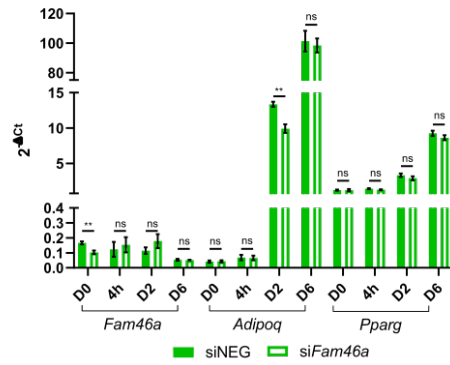
increased on day 2 and day 6 compared to the control (Figure 3.3.9A.2). We also observed an increase in *PPAR γ* expression on day 6 when *Fam46a* was over-expressed. As for the osteogenic differentiation, no markers were significantly different to the control samples in either treatment, even with significant disruptions in *Fam46a* expression (Figure 3.3.9B).

As the C3H10T1/2 data only showed minimal effects on differentiation markers, we were interested in observing the effect of *Fam46a* knockdown on MSC markers. Due to the differential expression *Fam46a* and enrichment in stromal populations, we hypothesised that *Fam46a* may be involved in maintaining stem or precursor state of MSCs (and other progenitors), perhaps by blocking the transition to epithelial cells. Epithelial-mesenchymal transition or mesenchymal-epithelial transition (EMT/MET) occurs during embryonic development, and has been studied in epithelial-derived tumour malignancies (Nieto et al. 2016). EMT and MET is a dynamic process, where cells can remain in a semi-transition state, which can be induced *in vitro*. BM-MSCs were treated with an siRNA targeting *Fam46a* and assayed two and six days after transfection (D2 and D6). No differentiation cocktails were added. As observed previously, *Fam46a* expression is much higher in BM-MSCs compared to C3H10T1/2 cells, and the siRNA was able to significantly reduce the expression, even up to six days after transfection (Figure 3.3.10). MSC marker vimentin (*Vim*), or N-cadherin, increased during *Fam46a* knock down at day 2, however, this was no longer observed by day 6. Epithelial marker *Cdh1*, or E-cadherin, expression also increased while *Fam46a* expression was reduced, however, this gene is very lowly expressed, with Ct >30 ($2^{-\Delta\text{Ct}}$ in siNEG <0.002).

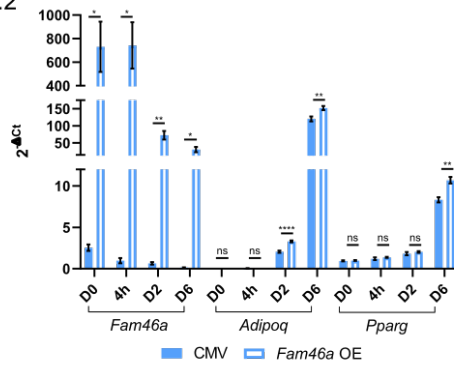
Chapter 3

A C3H10T1/2 Adipogenesis

A.1

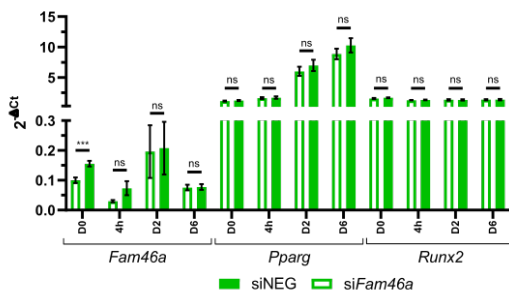


A.2



B C3H10T1/2 Osteogenesis

B.1



B.2

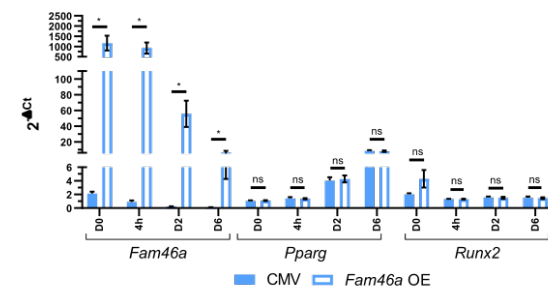


Figure 3.3.9: A) Adipogenic and B) osteogenic differentiation of C3H10T1/2 treated with siRNA and over-expression plasmids. Cells were transfected 2 days before induction (D0). Cells were collected and assayed at D0, 4h, D2, and D6.

A.1) qPCR plot of *Fam46a*, *Adipoq* and *PPAR γ* expression in cells treated with siRNA targeting *Fam46a* and negative control (siNEG)

A.2) qPCR plot of *Fam46a*, *Adipoq* and *PPAR γ* expression in cells treated with CMV-controlled *Fam46a* over-expression plasmid and empty CMV plasmid (control).

B.1) qPCR plot of *Fam46a*, *PPAR γ* and *Runx2* in cells treated with siRNA targeting *Fam46a* and negative control (siNEG)

B.2) qPCR plot of *Fam46a*, *PPAR γ* and *Runx2* in cells treated with CMV-controlled *Fam46a* over-expression plasmid and empty CMV plasmid (control).

Expression is represented in $2^{-\Delta C_t}$ relative to housekeeping gene *Tbp*. Bars are plotted as mean \pm SEM. Statistical tests are unpaired student t-tests with Welch correction. *p-value $p < 0.05$, **p-value $p < 0.01$, ***p-value $p < 0.001$, ****p-value $p < 0.0001$.

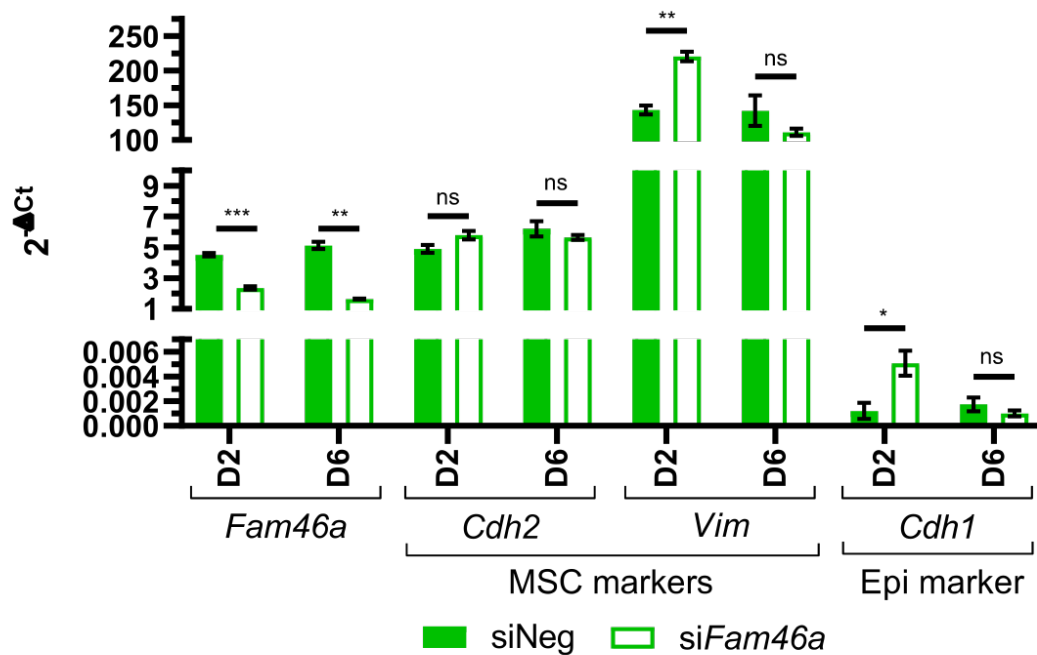


Figure 3.3.10: qPCR plot of BM-MSCs treated with siRNA targeting *Fam46a* and negative control siRNA (siNEG). Cells were transfected (D0) and samples were collected on two and six days after (D2 and D6) for analysis.

Expression is represented in $2^{-\Delta C_t}$ relative to housekeeping gene *Tbp*. Bars are plotted as mean \pm SEM. Statistical tests are unpaired student t-tests with Welch correction. *p-value $p < 0.05$, **p-value $p < 0.01$, ***p-value $p < 0.001$, ****p-value $p < 0.0001$.

3.4. Discussion

FAM46A is an RNA binding protein with a NTase domain, predicted to modify the 3' ends of mRNA transcripts. However, few studies have elucidated the role of this protein, and only one previous publication found a possible regulatory role for FAM46A in adipocytes (Carayol et al. 2017). In this study, we explored the expression of FAM46A in different *in vitro* cell systems and attempted to understand its function by gene expression interference. In SGBS cells and primary cell cultures we observed a decrease in *FAM46A* expression during adipogenesis. When using SGBS as our primary cell model, decreasing the expression of *FAM46A* by RNA interference lead to a decrease in expression of *PPAR γ* in the pre-adipocyte state (D0) (Figure 3.3.2A). As cells matured to day 6, this trend continued, as well as decreasing the expression of *CEBPA* (Figure 3.3.2B). This reduction in adipogenic signature continued to be observed when quantifying lipid accumulation, where si*FAM46A* treated cells accumulated less lipids than the negative control (siNEG) (Figure 3.3.4). Over-expression of *FAM46A* affected the expression of the adipogenic master regulators and lipid accumulation by an overall reduction in gene expression and lipids quantified (Figure 3.3.3 and 3.3.4). However, this reduction appeared in both the *FAM46A* over-expression samples and the empty CMV control. This indicates the plasmid backbone may be causing a certain cell stress that is beyond the over-expression of *FAM46A*. Previous studies with CMV-controlled over-expression constructs also observed perturbations in the cell state, even with the empty control plasmid (Stepanenko & Heng 2017).

Using fractionated human subcutaneous biopsy samples (in collaboration with Inserm in Toulouse), *FAM46A* expression was

enriched in the SVF of the adipose tissue (Figure 3.3.5). This enrichment was also observed in scRNA-seq data in mouse from *Tabula muris*, in fat tissue (Figure 3.3.6). To follow-up on the population enrichment results, C3H10T1/2 cells and primary BM-MSCs were utilised. *Fam46a* expression differed between the different cell lines, where BM-MSCs showed a high expression compared to C3H10s, where *Fam46a* was lowly expressed (Figure 3.3.7). When observing the expression profile during osteogenesis, *Fam46a* expression increased in both the C3H10T1/2 cells and BM-MSCs (Figure 3.3.7). During adipogenesis, *Fam46a* expression increased in the C3H10T1/2, but decreased in the BM-MSCs. C3H10T1/2 showed an expected increase *PPAR γ* and *Adipoq* in during adipogenesis (Figure 3.3.7A), however, master regulator *Runx2* during osteogenic differentiation did not increase and was stably expressed (Figure 3.3.7B). *PPAR γ* was increasing, and ALP assay also showed an increase in ALP activity during osteogenesis (Figure 3.3.8B.2). Staining of C3H10T1/2 revealed expected phenotypes, where C3H10T1/2 adipocytes contained lipids stained with ORO and osteoblasts had positive staining for ALP activity (Figure 3.3.8). However, when observing the “osteogenic cells” lipid droplets were visible. A possible explanation could be that *PPAR γ* was being expressed during osteogenesis (perhaps from dexamethasone being included in the osteogenic cocktail), thus promoting an adipogenic state. In normal osteogenic conditions, *PPAR γ* would be repressed by *Runx2* (Ge et al. 2016), which in our C3H10T1/2 cell culture, was expressed, but did not increase during osteogenesis. Gene expression perturbation of *Fam46a* in C3H10T1/2 cells showed a decrease of *Adipoq* expression when treated with si*Fam46a* during adipogenesis (Figure 3.3.9). Over-expression of *Fam46a* showed the inverse effect, where *Adipoq*

Chapter 3

expression significantly increased, as well as *PPAR γ* , which complements the SGBS data. Osteogenesis showed no significant effects during either siRNA treatment or over-expression. However, considering the lowered *Runx2* expression and high *PPAR γ* expression, these cells may not be in an osteogenic state.

BM-MSCs had high expression of all markers during both differentiation states. Unlike the C3H10T1/2 cells, *Runx2* was increasing during osteogenic stimulation (Figure 3.3.10). However, *PPAR γ* remained stable during both adipogenesis and osteogenesis. Due to limited numbers of BM-MSCs, we were unable to stain these cells for lipids or ALP activity. BM-MSCs were transfected with the different constructs and the effect of RNA interference was studied on genes involved in EMT. MSC marker *Vim* and epithelial marker *Cdh1* had significant increases in expression two days after transfection. However, *Cdh1* was extremely lowly expressed (Ct values >30) and cannot be reliably evaluated.

Overall, this *in vitro* study indicates that FAM46A may have a role in adipogenesis regulation, by being a pro-adipogenic factor (decrease/increase in *FAM46A* expression leads to decrease/increase in adipogenic markers). However, this is contrary to the results showing that *FAM46A* expression decreases during adipogenesis, as well as the enrichment of *FAM46A* expression in precursor cell populations. The *in vitro* systems we used, especially the C3H10T1/2, have their limitations as cell lines and this study shows that proper validation of the system should be performed before any proper conclusions can be made.

3.5. Materials and methods

3.5.1. SGBS cell culture

SGBS pre-adipocyte cells were used as a model for human adipocyte differentiation and were cultured as previously described (Wabitsch et al. 2001). Briefly, pre-adipocyte cultures were maintained and expanded in Complete Medium consisting of DMEM/F12 medium (Gibco) supplemented with 10% FBS (Gibco), 33 mM biotin (Sigma-Aldrich), 17 mM pantothenate (Sigma-Aldrich), and 100 U/mL penicillin-streptomycin (Gibco). Differentiation into adipocytes was induced using 0F basal medium (Complete Medium without FBS) supplemented with 10 mg/mL transferrin, 1 mM recombinant human insulin, 1 mM cortisol, 2 nM tri-iodothyronine (T3), 250 nM dexamethasone, 2.5 mM IBMX, and 10 mM rosiglitazone (Cayman Chemical). All compounds were from Sigma-Aldrich unless stated otherwise. After four days of induction, a maintenance medium was used consisting of 0F basal medium supplemented with 10 mg/mL transferrin, 1 mM human insulin, 1 mM cortisol, and 2 nM T3. Mature adipocytes were assayed from day 6-14, depending on the assay.

3.5.2. SGBS qPCR

Total RNA from SGBS cells was extracted using the RNeasy Plus Mini kit (Qiagen). 500 ng of total RNA was reverse transcribed using the High Capacity cDNA Reverse Transcription kit (Applied Biosystems). To perform the qPCR the SYBR Green kit LightCycler 1536 DNA Green Master kit was used with the LightCycler Instrument (both from Roche).

3.5.3. Transfection of siRNA and CMV-plasmids

siRNA targeting human *FAM46A* (*siFAM46A*) and control siRNA (*siNEG*) were ordered from ThermoFisher Scientific (predesigned

Chapter 3

Silencer Select siRNAs). Over-expression plasmids pCMV6 empty vector (CMV control) or *FAM46A* full human cDNA ORF Clone were ordered from OriGene. To introduce the constructs into SGBS cells, the Neon Transfection System (Invitrogen) was used with the following parameters: 1100 V, 20 ms, 1 pulse. 25 nM of siRNA and 500 ng of CMV-plasmid was introduced into 300,000 – 500,000 cells per sample. Cells were assayed after 48 hours.

3.5.4. Lipid quantification

Based on protocol from Dr. Magda Zachara. After transfection, SGBS cells were plated in a 96-well dark microscopy plate and differentiation was induced. On day 6 cells were fixed with 4% formaldehyde and washed with PBS. Nuclei were stained with Hoechst and lipids were stained with BODIPY (boron-dipyrromethene, Invitrogen). Images of each well were taken with the Operetta microscope (Perkin Elmer), an automated imaging system. Each well had 25 images taken in order to cover as much surface area of the well as possible. Image analysis comprised of an ImageJ algorithm developed by Dr. Magda Zachara in collaboration with the EPFL BIOP imaging facility to quantify the nuclei and lipid fluorescent signal in pixels.

3.5.5. C3H10T1/2 cell culture

C3H10T1/2 (ATCC) cell cultures were maintained in DMEM media supplemented with 10% FBS, 1% GlutaMAX™ and 100 U/mL penicillin-streptomycin (all from Gibco). Adipogenic differentiation was induced with maintenance media supplemented with 10 mM rosiglitazone (Cayman Chemical) and 200 nM insulin (Sigma). Osteogenic differentiation was induced with α MEM media (Gibco) supplemented with 10% FBS, 1% GlutaMAX™, 100 U/mL penicillin-streptomycin (all from Gibco), 100 nM dexamethasone,

50 mg/mL ascorbic acid and 10 mM b-GP (Sigma). Differentiation medias were replaced every 2 days and assayed up to day 6.

3.5.6. BM-MSCs cell culture

Primary BM-MSCs from C57BL/6J were collected with the kind help of Frédérica Schyrr who provided the following protocol. Tibiae and femur were collected and washed. Bone tips were removed and the remaining bone was stored in cold PBS. Bone marrow was flushed from the bone using a syringe and α MEM media (Gibco) supplemented with 10% FBS. Flushed bones were cut into small pieces using a scalpel and transferred with media into a 15 mL Falcon tube. Bone chips were digested with 1% collagenase I for 1 hour at 37°C. Detached cells and bone chips were washed and re-suspended in PBS supplemented with 2% FBS (Gibco). Using a mesh filter, BM-MSCs were isolated from bone chips and plated. Cells were maintained in α MEM media supplemented with 10% FBS and 100 U/mL penicillin-streptomycin (all from Gibco). Induction of differentiation (adipo- and osteogenesis) was the same cocktails as C3H10T1/2 cells.

3.5.7. Murine MSCs qPCR

Total RNA from cells was extracted using the RNeasy Tissue Kit (Qiagen). 300-500 ng of total RNA was reverse transcribed using the qScript cDNA SuperMix (Quanta Biosciences). To perform the qPCR the SYBR Green kit LightCycler 1536 DNA Green Master kit was used with the LightCycler Instrument (both from Roche).

3.5.8. Lipofectamine transfection of siRNA and CMV-plasmids

C3H10T1/2 and primary BM-MSCs were transfected with different constructs using Lipofectamine 3000 (Invitrogen) according to the user manual. siRNA targeting murine Fam46a and control siRNA

Chapter 3

(siNEG) were ordered from ThermoFisher Scientific (predesigned Silencer Select siRNAs). Over-expression plasmids pCMV6 empty vector (CMV control) or Fam46a full mouse cDNA ORF clone were ordered from OriGene. 50 nM of siRNA and 500 ng of CMV-plasmid were introduced to cells plated in a 12-well format. After 48 hours, cells were assayed before and after induction of differentiation (up to day 6).

3.5.9. ORO and ALP activity staining

C3H10T1/2 cells were stained for either lipid accumulation using Oil Red O (Sigma) or ALP activity using SIGMAFAST™ BCIP®/NBT (Sigma). For lipid accumulation, mature adipocytes (day 6) were washed in PBS and fixed in 10% formalin (Sigma) for 30 min at RT. Formalin was removed and cells were washed twice in PBS. Cells were stained in ORO prepared in 60% isopropanol for 5 min. Stain was removed and cells were washed twice in PBS and imaged in PBS. ALP activity staining was performed according to the manufacturers instructions (Sigma).

Chapter 4:

Functional characterisation of Fam46a *in vivo* under metabolic challenge

4. Functional characterisation of Fam46a *in vivo* under metabolic challenge

4.1. Abstract

Mice are one of the most commonly used animal models to study metabolism. With adipose tissue being considered an endocrine organ, it is important to understand how an entire mammalian system is affected by a specific gene KO. Using a full body Fam46a KO mouse model, we studied the potential metabolic contribution of Fam46a *in vivo*.

Wildtype (WT) and KO mice were fed a standard or HFD for 13 weeks. We observed in our study that the KO mice presented a smaller body size and weight, with increased plasma ALP, as previously described. During the HFD intervention, the KO mice gained less weight as compared to the WT. The KO mice also presented a phenotype where glucose clearance, energy expenditure, food intake, respiratory exchange rate (RER) and liver TG accumulation showed signs of HFD consumption as compared to the standard chow diet. The WT mice at end study were obese, while the KO mice displayed a healthier metabolic profile. Histological analysis of WAT in the KO exhibited significant reduction in adipocyte size in both subWAT and visWAT depots. Transcriptomic analysis of the subWAT and visWAT during the HFD revealed discrepancies in expression profiles between these two depots. In the KO, we observed an up regulation of genes expressed in ribosome related pathways in the subWAT, whereas the visWAT showed an up regulation of energy expenditure pathways, such as fatty acid metabolism, and lipid metabolic processes. Both tissues had down regulation of cell components processes, like ECM. An enrichment of TGF β and BMP pathways was also observed in both depots.

In conclusion the Fam46a KO mice appeared to be protected from diet induced obesity (DIO) with improved metabolic profiles. Gene expression analysis showed an enrichment of interesting pathways related to metabolism, indicating the KO depots are more metabolically active. However, considering the baseline health the KO mouse and having compared it to an obese WT mouse, certain biases may have occurred. Despite this, similarities in mouse models of Smad3 and 4 indicate a potential dual role regulatory role for Fam46a in the TGF β and BMP pathways.

Chapter 4

4.2. Introduction

4.2.1. Mouse models for metabolic studies

Animal models provide a powerful tool to study biological questions, particularly ones concerning specific pathologies. Mice have been indispensable models to understand the underlying mechanisms of complex metabolic diseases, such as obesity and related metabolic syndromes (Lutz 2020). Several transgenic mouse models are available to study full body metabolism in disease states. Classic transgenic mouse models involve the removal, or KO, of a gene of interest. Well-known metabolic mouse models are frequently used, such as *ob/ob* mice (disruption of the leptin gene) for obesity (Coleman 1978) and *db/db* mice who have insulin resistance (Chen et al. 1996). Without any genetic modifications, certain mouse strains are also susceptible to DIO, such as C57BL/6 (Nishikawa et al. 2007). Under HFD, these mice develop obesity, which includes symptoms such as hyperinsulinemia and hyperglycaemia. In order to understand mechanisms that combat metabolic syndromes, transgenic KO mice have been challenged with an HFD to observe any resistance to obesity and the associated co-morbidities. This may lead to further understanding of the gene of interest, and how it contributes to the development of metabolic syndrome (Fuchs et al. 2018).

4.2.2. Skeletal dysplasia model is a *Fam46a* loss of function mutation

Our gene of interest, *Fam46a*, has limited knowledge concerning its function in vivo. The most extensive study to date was a mouse model involving a nonsense mutation in the *Fam46a* gene (Diener et al. 2016). N-ethyl-N-nitrosourea (ENU) is a mutagen that is used in mice to introduce random mutations throughout the genome (Stottmann & Beier 2014). The authors analysed a series of ENU

treated mice and observed one strain that had highly elevated plasma ALP. Exome sequencing of this strain revealed a mutation that caused premature termination of translation of the *Fam46a* gene. ALPs are glycoproteins that play a role in healthy bone mineralization (Sharma et al. 2014). Elevated plasma ALP in human patients have been used as marker to diagnose bone and liver diseases (Epstein et al. 1986), and may be an indication of active bone formation due to increased osteoblast activity (Saraç & Saygılı 2014). Heterozygous (*Fam46a*^{E157*/+}) mice from this study also presented elevated plasma ALP, but were otherwise healthy, while the homozygous (*Fam46a*^{E157*/-}) displayed severe bone phenotypes, including a shorter stature and deformities. *Fam46a*^{E157*/-} mice had two-fold higher plasma ALP levels compared to heterozygotes, however, both presented normal levels of other plasma bone markers (total inorganic calcium (Ca) and phosphate (P)). Other bone abnormalities of the homozygous mutant included delayed ossification, defects in intra-membrous and endochondral ossification, and thinning of the cortical bone in long bones. Additionally, the trabeculae (porous collagenous tissue found at the ends of long bones) was missing in *Fam46a*^{E157*/-} mice. From all these abnormalities, the authors concluded this *Fam46a* mutant mouse could be considered a model for skeletal dysplasia. In humans skeletal dysplasias, or osteochondrodysplasias, are generalised disorders of cartilage and bone, which includes abnormalities in the development and size of the skeleton, commonly leading to a short stature (Krakow & Rimoin 2010).

4.2.3. Research question

Although there is a published phenotyping of a full body *Fam46a* KO mouse, the authors focused more on the skeletal phenotype and disease model with no observation of the metabolic profile of

Chapter 4

this model, especially under metabolic stress. As *FAM46A*'s involvement in metabolism was proposed due to its differential expression in adipose tissue during a weight-loss event (Carayol et al. 2017), we were interested to further understand *Fam46a*'s role in whole body metabolism, focusing on characterising adipose tissue physiology under a HFD intervention. Therefore, we phenotyped a CRISPR-Cas9 full body *Fam46a* KO mouse, in collaboration with the IMG in Prague, under two different diet interventions to further characterise and understand *Fam46a*'s role *in vivo*, focusing on metabolism, more specifically in WAT function.

4.3. Results

4.3.1. Study design and modifications

In order to study the role of *Fam46a* in metabolism *in vivo*, we collaborated with Prof. Radislav Sedláček, Dr. Jan Procházka, and Goretti Aranaz Novaliches at the Institute of Molecular Genetics of the Czech Academy of Sciences (IMG Prague), who very kindly housed and performed the following studies with a full body *Fam46a* KO C57BL/6 mouse model. Two separate *ad libitum* diet interventions were applied; a standard chow diet (Figure 4.3.1A) and a HFD, where 60% of energy came from fat (Figure 4.3.1B). Due to the small and lean nature of the *Fam46a* KO mice, the mice were aged for 24 weeks on the standard diet before sacrifice to ensure sufficient adipose tissue for collection. However, the stress of various tests and the fragile bone phenotype of the KO led to a number of mice dying before the end of the study (Figure 4.3.1A). From this observation, we modified the HFD intervention to end at 21 weeks of age, meaning 13 weeks on the HFD intervention. With this modification, only one KO mouse died before the end of the study on the HFD (Figure 4.3.1B).

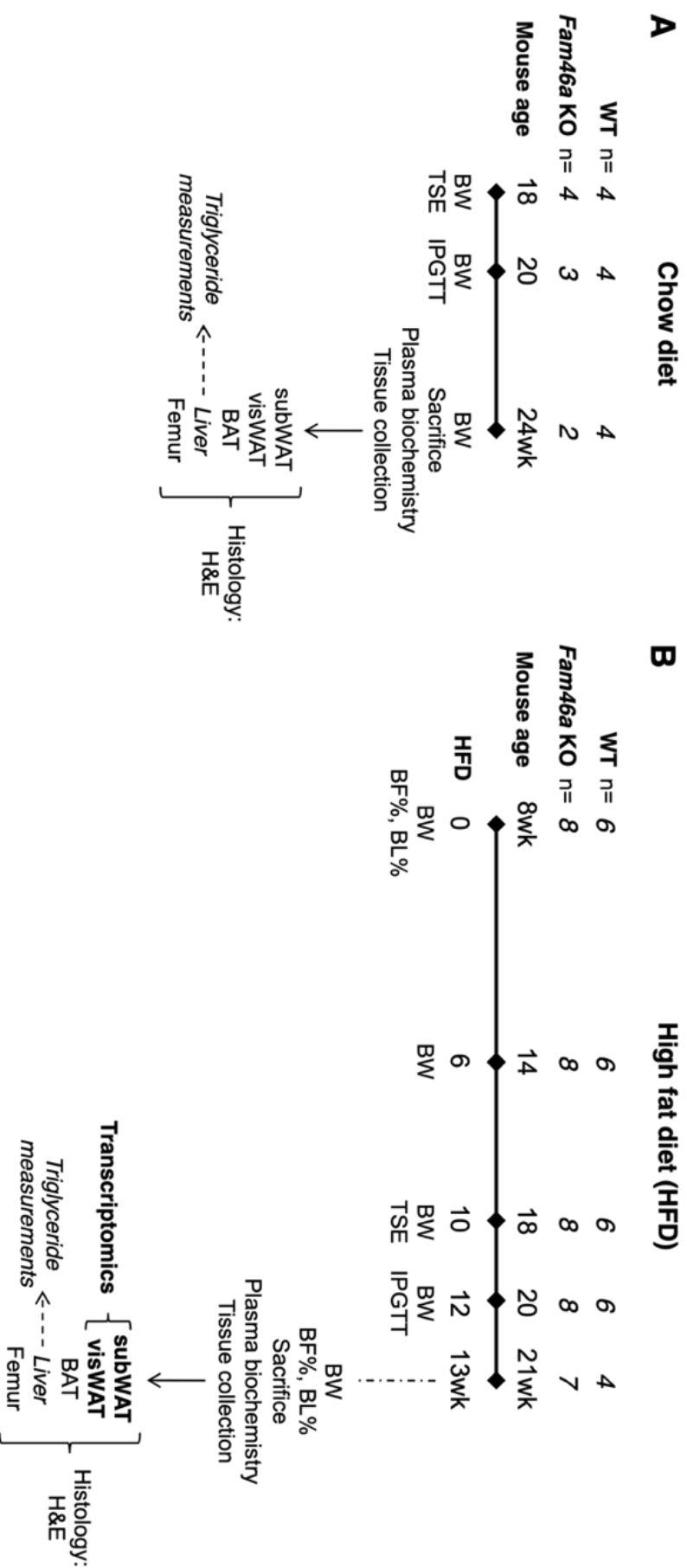


Figure 4.3.1: Overview of mouse studies on A) standard chow and B) high-fat diet (HFD).

A) Body weight (BW) was measured at 18, 20 and 24 weeks of age. Using a Phenomaster TSE, calorimetry, movement and food intake was measured at 18 weeks of age over a period of 48 hours. Intraperitoneal glucose tolerance tests (IPGTT) were performed at week 20 of age. Mice were sacrificed at 24 weeks of age in order to collect subWAT from the inguinal depot, visWAT from the gonadal/epididymal depot, brown adipose tissue (BAT), liver and femur. All tissues were prepared for haematoxylin and eosin (H&E) staining. Liver samples were measured for TG levels, and blood was collected for plasma biochemistry (liver, lipid metabolism and bone panels). All mice were male.

B) Body weight (BW) was measured at week 0, 6, 10, 12 and 13 on the HFD, where body fat mass percentage (BF%) and body lean mass percentage (BL%) was measured at the beginning and end of the study (week 0 and 13 of the HFD). Tests and BW measurements were performed at the same age as the standard diet mice (TSE week 10 of HFD, IPGTT, week 12 of HFD). At week 13 of the HFD, mice were sacrificed, and the tissues collected (same as standard diet). subWAT and visWAT samples were prepared for transcriptomic analysis (QuantSeq).

4.3.2. Body weight and body composition of WT and *Fam46a* KO mice

Body weight (BW) was measured at several time points during both interventions, weighing the mice at 18, 20 and 24 weeks of age on the standard diet and 8, 14, 18, 20 and 21 weeks of age on the HFD. Similar to what was reported previously (Diener et al. 2016), the *Fam46a* KO mice were visually smaller than the WT (data not shown). This was reflected in their BW measurements. From the start to the end of the interventions, KO mice weighed significantly less than the WT control in both diet cohorts (Figure 4.3.2A/B). The *Fam46a* KO mice were gaining weight during the HFD, however, still gaining significantly less weight compared to the WT (Figure 4.3.2C).

During the HFD, the body composition of the mice was measured using microCT scanning to distinguish the fat and lean mass measurements of the mice. Figure 4.3.3 shows the acquired measurements of fat and lean mass in the mice before and after the HFD intervention (absolute values and relative to the total measured body mass). Before the intervention at 8 weeks of age, there is no significant difference in absolute or relative fat mass between the WT and *Fam46a* KO (Figure 4.3.3A.1/A.2). The absolute lean mass is significantly reduced in the KO (Figure 4.3.3B.1). However, this effect is no longer significant when observed relative to the total body mass (Figure 4.3.3B.2). After 13 weeks of HFD, the fat mass of the WT mice, who have developed an obese phenotype, has vastly increased compared to the *Fam46a* KO, which is also significant when correcting for total body mass (Figure 4.3.3A.1/A.2). The lean mass of all mice increased during the intervention, where the WT mice have significantly more absolute lean mass (Figure 4.3.3B.1). Interestingly, when correcting for total body mass the *Fam46a* KO has significantly more lean mass.

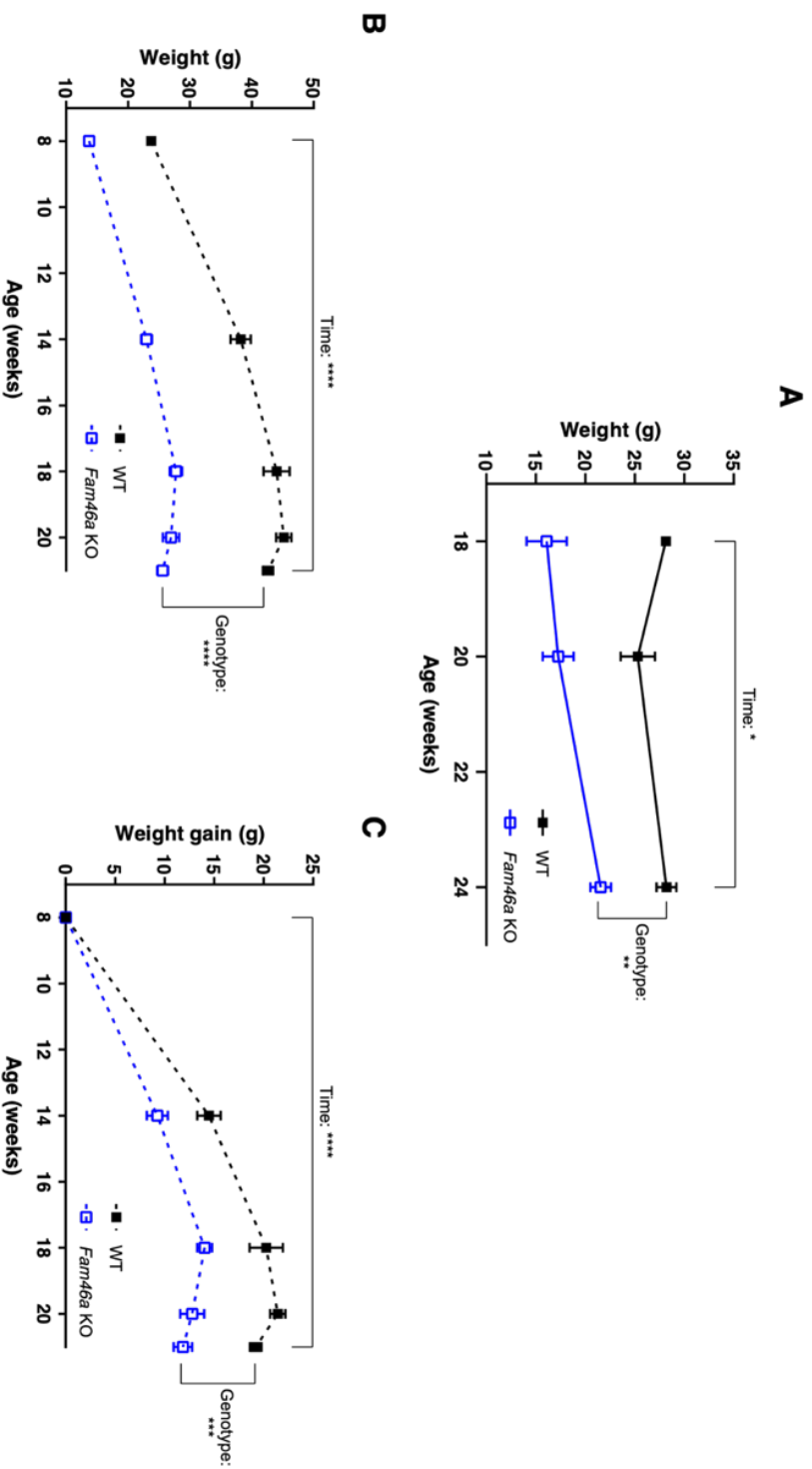


Figure 4.3.2: Body weight and body weight gain in WT and Fam46a KO mice under A) standard diet and B)C) high-fat diet (HFD).

A) Body weight monitoring on a standard diet. Weight was measured when mice were 18, 20 and 24 weeks old.

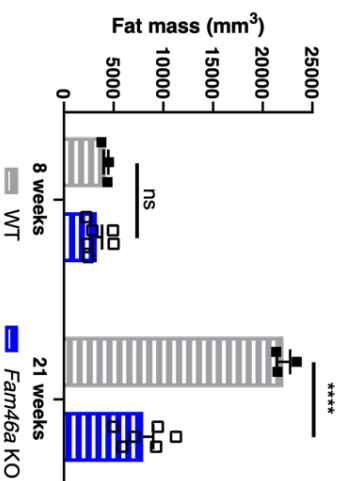
B) Body weight monitoring during 13 weeks on the HFD intervention. Weight was measured when mice were 8, 14, 18, 20 and 21 weeks old.

C) Body weight gain of mice during the 13-week HFD intervention from week 8 of age to 21 weeks.

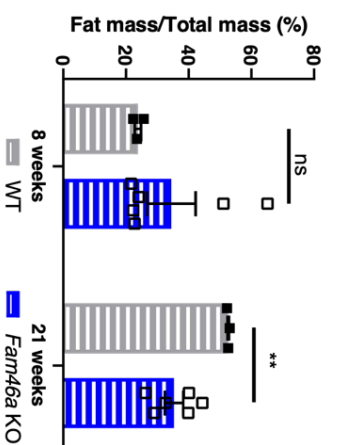
Data points are mean±SEM. Two-way ANOVAs using a mixed effect model with Bonferroni correction was used as a statistical test to compare the effect of Time and Genotype. *p-value $p < 0.05$, **p-value $p < 0.01$, ***p-value $p < 0.001$, ****p-value $p < 0.0001$.

A Fat mass

A.1

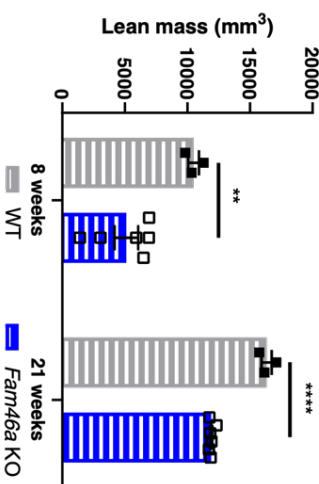


A.2



B Lean mass

B.1



B.2

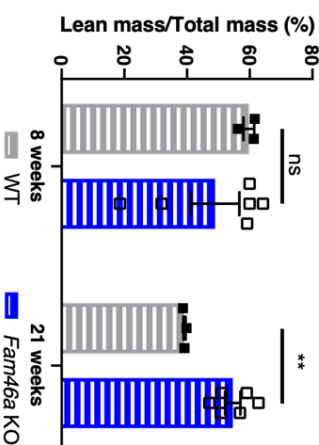


Figure 4.3.3: Body fat mass (A) and body lean mass (B) measurements in WT and Fam46a KO mice under high-fat diet (HFD).

A.1) Absolute microCT measurements of fat mass before (8 weeks old) and at the end (21 weeks old) of the HFD intervention.

A.2) Fat mass measurements over total mass measurements before (8 weeks old) and at the end (21 weeks old) of the HFD intervention.

B.1) Absolute microCT measurements of lean mass before (8 weeks old) and at the end (21 weeks old) of the HFD intervention.

B.2) Lean mass measurements over total mass measurements before (8 weeks old) and at the end (21 weeks old) of the HFD intervention.

Data points are mean±SEM. Statistical tests are unpaired student t-tests. *p-value p<0.05, **p-value p<0.01, ***p-value p<0.001, ****p-value p<0.0001.

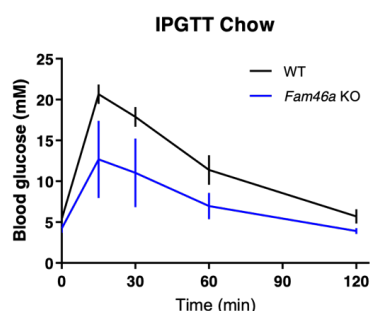
Chapter 4

4.3.3. Metabolic phenotyping: IPGTT and TSE

In order to observe patterns of glucose clearance in WT and *Fam46a* KO mice, an intra-peritoneal glucose tolerance test (IPGTT) was performed at 20 weeks of age (12 weeks on the HFD intervention). The glucose dose administered was calculated dependent on the weight of the mouse. As presented in Figure 4.3.4B, there is a significant difference in response between WT and KO mouse, where there is faster glucose clearing in the KO mouse compared to the WT on the HFD. This trend was also observed in the chow diet, however, this result was not significant, perhaps due to the low number of mice available for the test (4 WT vs 3 KO) as compared to the HFD (6 WT vs 8 KO). All mice (WT and KO) on the HFD had slower glucose clearance compared to the chow diet (Figure 3.2.4B), showing the expected effect of the HFD on glucose response (Ghatta & Ramarao 2004).

A Intraperitoneal glucose tolerance test

A.1



A.2

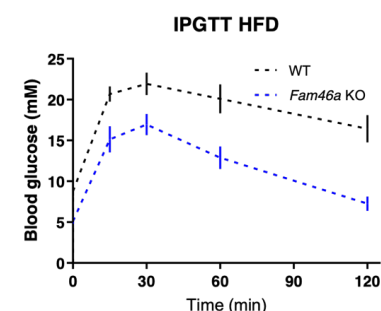
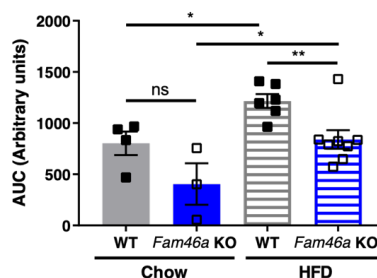
**B**

Figure 4.3.4: Intra-peritoneal glucose tolerance tests (IPGTT) in WT and Fam46a KO mice under A.1) standard diet and A.2) high-fat diet (HFD).

A.1) IPGTT curve (plasma blood glucose level measured over time) in 20 week old mice under standard diet (Chow).

A.2) IPGTT curve (plasma blood glucose level measured over time) in 20 week old mice under HFD.

B) Area under the curve (AUC) measurements for IPGTT data in both diet cohorts (Chow and HFD).

Data points are mean \pm SEM. Statistical tests are unpaired student t-tests. *p-value $p < 0.05$, **p-value $p < 0.01$, ***p-value $p < 0.001$, ****p-value $p < 0.0001$.

To further characterise the metabolic profiles of the mice, each mouse was placed in a Phenomaster TSE system for 48 hours at 18 weeks of age (10 weeks of HFD). The first 24 hours was used as an acclimation period to the cage, and the data points collected in the next 24 hours were used for the final analysis. Using indirect gas calorimetry, the mice's oxygen (O_2) consumption and carbon dioxide (CO_2) production were measured in order to calculate the RER and heat production/energy expenditure. Total food intake and average movement was also recorded.

The RER is a ratio of the volume of CO_2 (V_{CO_2}) produced and the volume of O_2 (V_{O_2}) consumed. This ratio is utilised to assess the energy fuel source, where values closer to 1 are indicative of a carbohydrate fuel source, and 0.7 is considered a predominant fat

Chapter 4

fuel source (Farinatti et al. 2016). On the standard diet, WT and KO mice have significantly different RER during the dark period, where WT mice have a RER close to 1 (Figure 4.3.5A, solid bars). When switched to a HFD, both mouse genotypes have a reduction in RER, an expected effect of the HFD (Marvyn et al. 2016), and there is no significant difference between WT and KO (Figure 4.3.5A, striped bars). The KO mice's food intake was significantly less than the WT when fed the standard diet, however, this reduction was restored with the HFD intervention (Figure 4.3.5B). Interestingly, heat production/energy expenditure was only significantly different during the HFD intervention, where *Fam46a* KO mice have an increase in energy expenditure as compared to the WT (Figure 4.3.5C). There was no significant difference in average movement between genotypes on either diet (Figure 4.3.5D).

In summary, both genotypes had expected effects of the HFD with slower glucose clearance and reduced RER. However, the KO had faster glucose clearance compared to the WT as well as reduced RER on the standard diet. Unexpectedly, an increase of heat production was measured in the KO mice during the HFD, indicating the loss of *Fam46a* may be causing a disruption in energy expenditure levels as well as protecting the mouse from the metabolic effects of the HFD intervention.

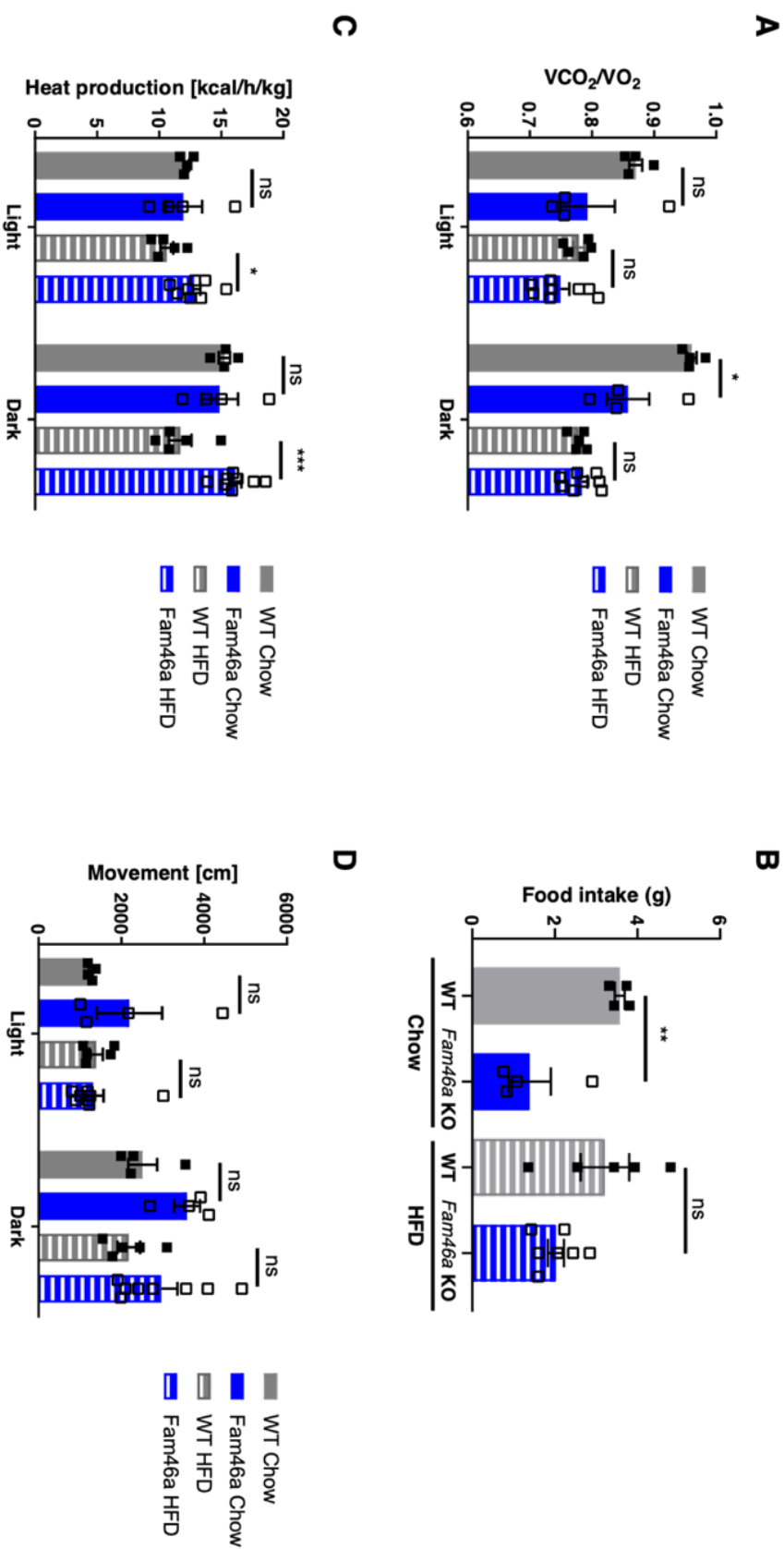


Figure 4.3.5: Phenomaster TSE measurements over a 24 hour period (after 24 hour acclimation) in WT and Fam46a KO mice under standard diet and high-fat diet (HFD).

A) Respiratory exchange rate (RER) during light and dark period (12 hours each).

B) Food intake per day under standard diet and HFD.

C) Heat production, or energy expenditure, during light and dark period (12 hours each).

D) Average movement (total distance travelled every 15 min) during light and dark period (12 hours each).

Data points are mean±SEM. Statistical tests are unpaired student t-tests. *p-value p<0.05, **p-value p<0.01, ***p-value p<0.001, ****p-value p<0.0001.

Chapter 4

4.3.4. Plasma biochemistry and histology of BAT and liver

Plasma biochemistry was performed at end study for both diets to assess the effect of the genotype and diet on different plasma marker panels. It should be noted that at end study the KO chow cohort only had two mice available for plasma biochemistry (Figure 4.3.1), therefore statistical testing was not performed. As predicted from previous observations (Diener et al. 2016), *Fam46a* KO mice had increased plasma ALP in both diets (Table 4.3.1A/B). The standard diet (Table 4.3.1A) had reduced levels of total protein and leptin in the KO. Lower plasma leptin levels are indicative of lower fat mass (Rosenbaum et al. 1997), which is exemplified in our KO mouse when combining the reduced fat mass and low leptin measurements (Figure 4.3.3A.1). Interestingly, the KO mice also presented higher levels of alanine aminotransferase (ALT) and aspartate transaminase (AST), indicative of liver damage. However, the standard error (SE) of these measurements are high due to the small number of mice available so these results are difficult to interpret. Analysis of the KO HFD cohort when compared to the WT showed a significant decrease in ALT, HDL, cholesterol and leptin, which are markers of liver damage and obesity (Table 4.3.1B). There was also a significant increase in phosphate in the KO cohort, at an average level of 1.92 mmol/dL, considered above the normal range of 1.2 – 1.8 mmol/dL (Otto et al. 2016). Phosphate is an essential mineral for bone health and homeostatic control of its levels are intricately maintained. Increased plasma phosphate can be indicative of increased bone resorption, where the bone tissue is broken down and minerals are released into the plasma (Goretti Penido & Alon 2012). Considering the skeletal dysplasia phenotype of the *Fam46a* KO mouse, this dysfunction in bone remodelling was expected.

As well as collecting plasma for analysis, BAT and liver were also collected and prepared for histological examination to assess any morphological differences. When comparing diets WT BAT and liver samples have the expected visible lipid droplets under a HFD intervention, where the BAT shows expanded lipid inclusions typically found in obese mice, while the chow diet maintains a healthy BAT morphology with smaller, uniform lipid droplets (Figure 4.3.6A). The *Fam46a* KO mice BAT and liver tissues remain similar between diets, with comparable morphology to the WT samples under the standard diet (Figure 4.3.6A/B). Images of WT HFD liver, however, show lipid accumulation within the tissue (Figure 4.3.6B.1). When measuring TG levels in the liver samples, we observe an increase in TG content under HFD as expected for both genotypes (Figure 4.3.6B.2). If we compare the genotypes, there is a trend towards WT mice having higher levels of liver TG under both diets, however, this is not considered statistically significant. Interestingly, the *Fam46a* KO mice have an increase in the TG levels in liver from the HFD, which is also observed in the clinical plasma biochemistry (not statistically significant) (Table 4.3.1B), however, it is not seen in the histological slide like the WT samples (Figure 4.3.6B.1).

By combining the plasma biochemistry and histology data, the *Fam46a* KO mice on both diets have similar profiles compared to WT mice on the standard diet. The KO mice under standard diet conditions may be experiencing liver damage, as well as a reduction in fat mass, as previously observed by body composition measurements. Conversely, KO mice that have been fed a HFD show reduced markers of liver damage and maintain a low leptin level. Markers of bone dysfunction, namely ALP and phosphate, are elevated in the KO, which may explain the severe bone phenotype described in these mice. These observations trend

Chapter 4

towards the KO model being metabolically healthier mouse under HFD when compared to the WT.

Table 4.3.1: Plasma biochemistry in WT and Fam46a KO mice under A) standard chow diet and B) HFD

SE = standard error. Statistical tests are unpaired student t-tests. *p-value $p < 0.05$, **p-value $p < 0.01$, ***p-value $p < 0.001$

A. Plasma biochemistry - Chow - 24 weeks of age				
Plasma parameters	WT	SE	Fam46a KO (n=2)	SE
Albumin (g/l)	30.35	1.03	27.99	2.47
ALP (U/L)	69.50	2.72	230.00	48.00
ALT (U/l)	25.00	2.35	39.00	18.00
AST (U/l)	44.25	2.25	99.50	65.50
Bili-T (umol/l)	2.76	0.96	3.10	0.97
Total Protein (g/L)	55.41	0.41	48.46	2.42
Urea (mmol/l)	8.18	1.00	7.05	0.26
Total cholesterol (mmol/l)	2.29	0.08	1.80	0.45
HDL (mmol/l)	1.80	0.07	1.02	0.59
LDL (mmol/l)	0.46	0.03	0.77	0.17
Lipase (U/l)	48.21	3.88	42.26	3.19
Triglycerides (mmol/l)	0.57	0.06	0.50	0.02
Calcium (mmol/l)	2.31	0.04	2.24	0.01
Phosphate (mmol/dl)	1.75	0.19	1.83	0.14
Leptin (pg/μl)	13.42	2.21	2.40	2.40

B. Plasma biochemistry - HFD - 21 weeks of age

Plasma parameters	WT (n=4)	SE	Fam46a KO (n=7)	SE	P-value	Summary
Albumin (g/l)	30.94	1.12	29.09	1.27	0.3571	ns
ALP (U/L)	56.00	4.38	164.71	14.09	0.0003	***
ALT (U/l)	85.50	21.39	18.86	2.06	0.0022	**
AST (U/l)	96.25	17.42	68.29	4.34	0.0753	ns
Bili-T (umol/l)	1.83	0.10	3.13	0.43	0.0536	ns
Total Protein (g/L)	56.92	1.00	56.01	2.35	0.7875	ns
Urea (mmol/l)	8.96	0.37	9.23	0.74	0.8036	ns
Total cholesterol (mmol/l)	5.12	0.33	3.75	0.51	0.0924	ns
HDL (mmol/l)	3.94	0.25	2.52	0.30	0.0107	*
LDL (mmol/l)	1.24	0.11	1.30	0.25	0.8570	ns
Lipase (U/l)	43.61	0.92	43.07	7.89	0.9605	ns
Triglycerides (mmol/l)	0.67	0.08	0.77	0.10	0.5251	ns
Calcium (mmol/l)	2.29	0.03	2.29	0.05	0.9433	ns
Phosphate (mmol/dl)	1.63	0.07	1.92	0.08	0.0399	*
Leptin (pg/μl)	31.75	0.69	13.22	3.38	0.0024	*

Chapter 4

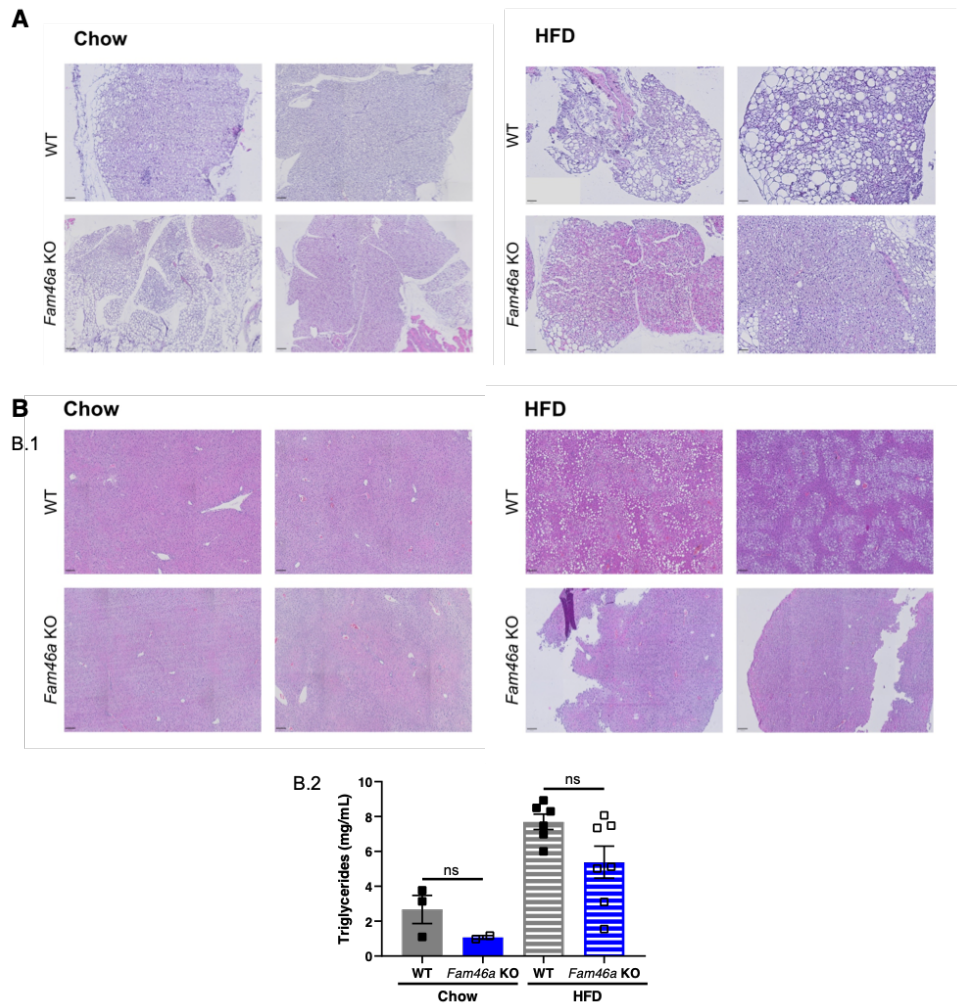


Figure 4.3.6: Haematoxylin and eosin (H&E) staining of fixed tissue after standard (chow) and HFD feeding in WT and Fam46a KO mice.

A) Brown adipose tissue (BAT) H&E stained section. Scale bar represents 100 micrometres (μm).

B.1) Liver H&E stained section. Scale bar represents 100 micrometres (μm).

B.2) Triglyceride measurements in liver tissue. Data points are mean \pm SEM. Statistical tests are unpaired student t-tests. *p-value $p < 0.05$, **p-value $p < 0.01$, ***p-value $p < 0.001$, ****p-value $p < 0.0001$.

4.3.5. subWAT and visWAT: histology and lipid quantification

As our main interest was examining the WAT of these transgenic mice, histological analysis was also performed on two different WAT depots representing subWAT and visWAT; inguinal and epididymal respectively. Images acquired of H&E stained slides were analysed using the ImageJ plugin Adiposoft (Galarraga et al. 2012), which detects individual adipocytes and quantifies their areas. We plotted the distribution and density of adipocyte size in

the different depots to compare the effect of the genotype and diet (Figure 4.3.7).

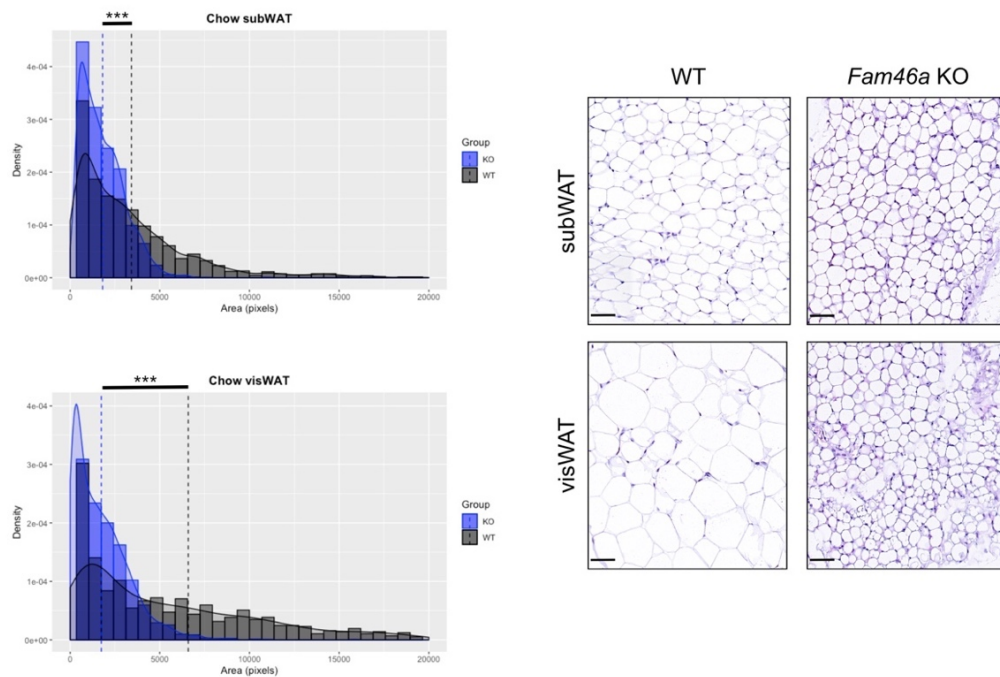
In mice on the standard chow diet, we observed that *Fam46a* KO adipocytes in both depots contain on average smaller adipocytes when compared to the WT (Figure 4.3.7A, dotted line representing mean and Table 4.3.2). Interestingly, WT visWAT adipocytes are on average larger than subWAT adipocytes, however, this difference is not as prominent in the *Fam46a* KO depots, where the mean area of both depots is approximately 2,000 pixels (Table 4.3.2). Under the HFD intervention, we observe an increase in adipocyte size for both depots and genotypes when compared to the chow, due to the increase in fat content of the diet (Figure 4.3.7B). A similar trend of smaller adipocytes in the *Fam46a* KO mice is once more observed despite the HFD. The differences in subWAT and visWAT size is now less pronounced in the WT, potentially due to the adipocytes reaching a certain limit of hypertrophy and lipid storage capacity. The *Fam46a* KO adipocytes have a more pronounced difference in size between depots, similar to what was seen in the standard diet WT samples (Figure 4.3.7B and Table 4.3.2). Overall, the *Fam46a* KO mice have smaller adipocytes when compared to the WT under both diet conditions, indicating a resistance to increased lipid accumulation in the KO adipocytes.

Table 4.3.2: Mean values of adipocyte areas (in pixels) from adipocyte size quantification in H&E stained histology slides of subWAT and visWAT under standard chow diet and HFD.

	Chow		HFD	
	subWAT	visWAT	subWAT	visWAT
WT	3,417.47	6,591.19	7,221.42	9,659.81
<i>Fam46a</i> KO	1,805.07	2,245.55	2,335.13	5,546.82

Chapter 4

A Standard diet



B High fat diet (HFD)

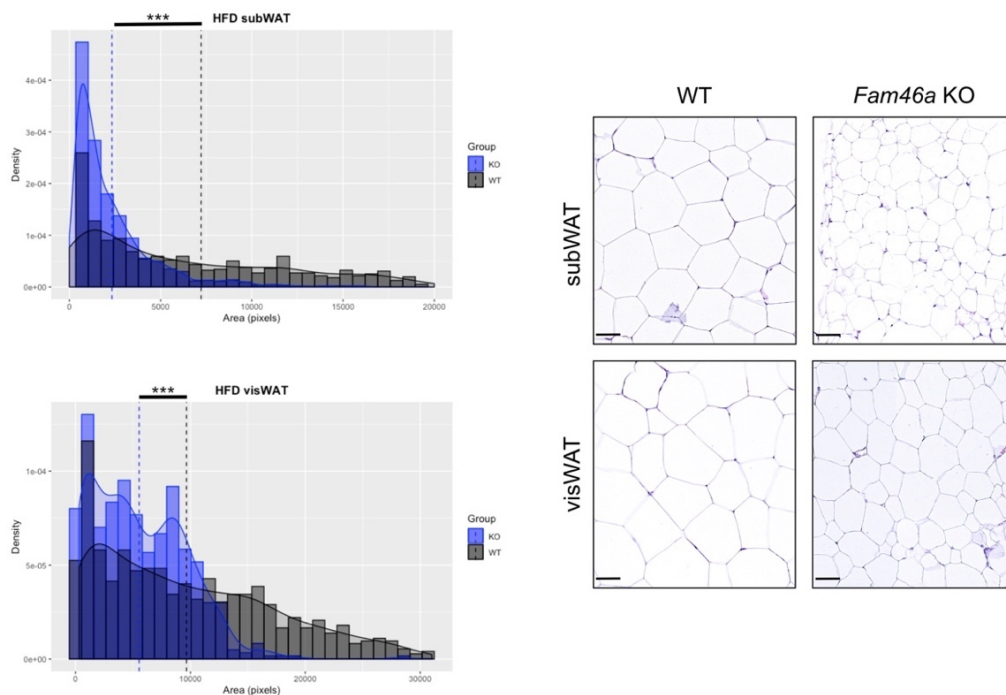


Figure 4.3.7: Adipocyte size quantification using H&E stained subWAT and visWAT sections after A) standard (chow) and B) HFD feeding in WT and *Fam46a* KO mice.

A) Distribution of adipocyte size on standard chow diet with representative subWAT and visWAT images. Scale bar represents 50 micrometres (μm).

B) Distribution of adipocyte size on HFD with representative subWAT and visWAT images. Scale bar represents 50 micrometres (μm).

Density plots represent data with mean (dashed line) and histogram of distribution of areas. Statistical tests are unpaired student t-tests. *p-value $p < 0.05$, **p-value $p < 0.01$, ***p-value $p < 0.001$

4.3.6. Transcriptomics: Differential gene expression between WT and KO

In order to further understand and characterise the effect of knocking out *Fam46a* on the expression profiles in adipose tissue, an RNA-sequencing (QuantSeq) experiment was conducted (Corley et al. 2019). As we previously observed differences in the subcutaneous vs visceral fat depots (section 4.3.5), where a loss of *Fam46a* lead to smaller adipocyte size under both diet interventions, we compared the transcriptomic profiles of samples from these two depots under HFD only. In this experiment subWAT and visWAT QuantSeq libraries were prepared, which included both male and female WAT samples at 21 weeks of age (13 weeks of HFD), for a total 24 WT samples vs 16 *Fam46a* KO samples. Using the expression data, a differential gene expression analysis at a 5% false discovery rate (FDR) threshold was performed by Dr. Alix Zollinger comparing *Fam46a* KO to WT, where the genotype, gender, tissue depot and mouse batch (sacrifice date) was corrected for in the model.

The top 200 differentially expressed genes are presented in Figure 4.3.8, where hierarchal clustering shows separation of the different genotypes and genders. The visWAT samples hierarchal clustering was similar to the subWAT depot, with the exception of two *Fam46a* KO male samples having similar expression patterns to the WT cluster (Figure 4.3.8, green group). During an initial principle component analysis (PCA), separation of the WT samples to the KO samples was observed (data not shown), therefore these sample outliers were kept for the subsequent analysis.

The subWAT analysis revealed 432 differentially expressed genes, where 176 genes and 256 genes were over-expressed and under

Chapter 4

expressed, respectively. The visWAT depot had less differentially expressed genes, with 272 genes passing the 5% FDR threshold. 132 genes were over-expressed, while 140 were under expressed.

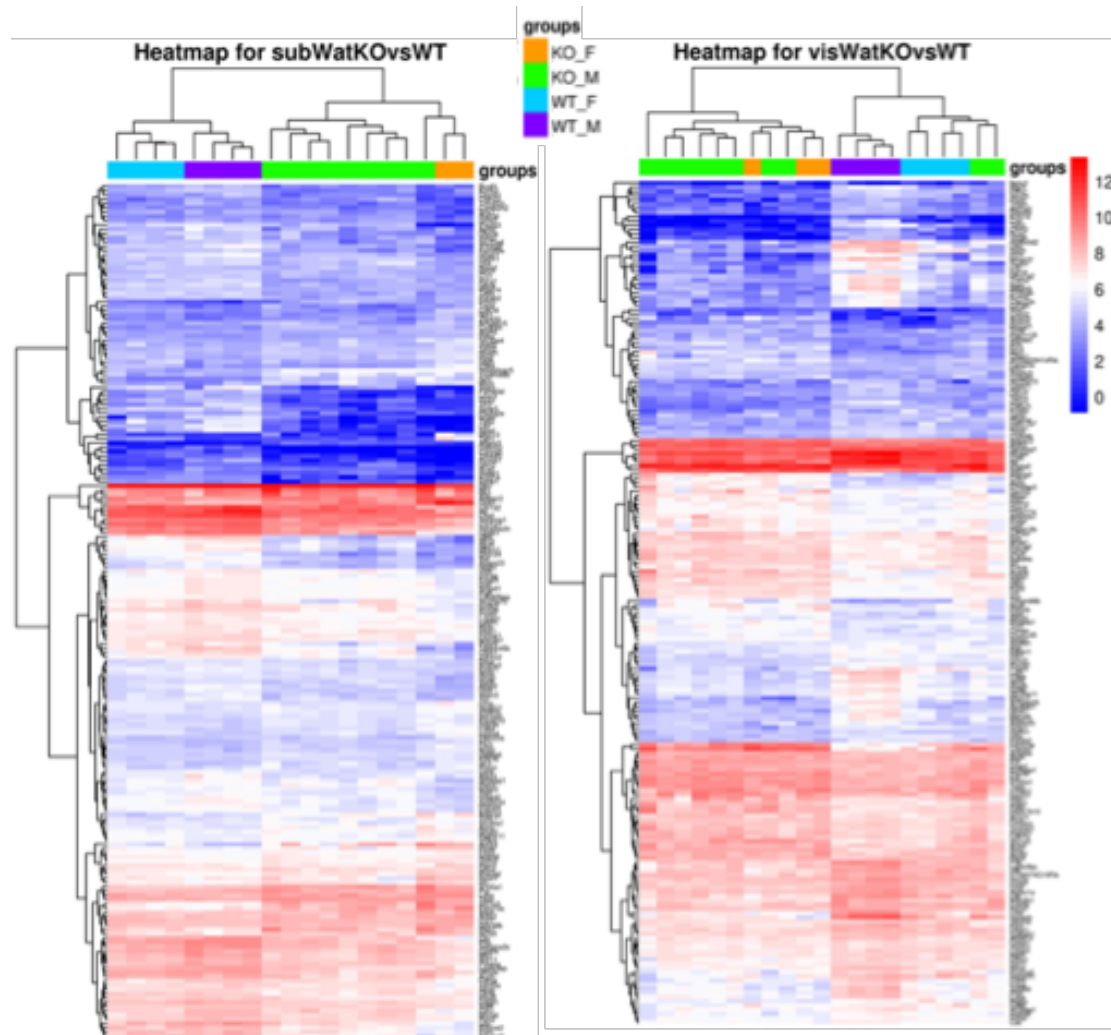


Figure 4.3.8: Heatmap of the top 200 differentially expressed genes comparing *Fam46a* KO mice to WT in subWAT and visWAT depots under HFD intervention.

Sample groups cluster by genotype (WT or KO) and gender (male (M) or female (F)).

To begin filtering the differentially expressed genes for any potentially interesting gene candidates, the top 20 most differentially expressed genes (ranked by FDR) were analysed in each depot. Tables 4.3.3A and 4.3.3B show the most significantly up regulated genes in subWAT and visWAT depots respectively. Genes highlighted in grey are the top genes found in both depots. Out of 20 top up regulated genes, seven are shared between the

two depots: *Sctr*, *Rbp7*, *Adipor2*, *Mycl*, *Fam13a*, *Acot2* and *Por*. Interestingly, *Rbp7* is a gene target of adipogenesis master regulator PPAR γ (Zizola et al. 2008) and is involved in the activation of thermogenesis in BAT via binding of its substrates, such as retinol (Bonet et al. 2015). *Adipor2* is one of the receptors for adiponectin, an adipokine involved in fatty acid oxidation and glucose uptake (Yamauchi et al. 2003), and *Fam13a* is a novel candidate gene as a regulator of adipocyte differentiation (Fathzadeh et al. 2020), indicating the potential involvement of *Fam46a* in metabolic related functions.

Chapter 4

Table 4.3.3: Top 20 up regulated differentially expressed genes in A) subWAT and B) visWAT.

logFC = log fold-change; logCPM = log counts per million reads mapped; FDR = false discovery rate `P-values were corrected for multiple testing (FDR) using Benjamini and Hochberg correction methods

A. Top-20 up regulated genes subWAT KO vs WT

Gene	logFC	logCPM	PValue	FDR
<i>Sctr</i>	4.81	4.60	5.64E-08	5.51E-05
<i>S100a1</i>	2.44	8.91	4.76E-08	5.51E-05
<i>Rbp7</i>	3.66	8.33	1.15E-07	8.79E-05
<i>Nle1</i>	2.90	3.38	5.71E-07	2.77E-04
<i>Clk2</i>	1.83	5.28	1.03E-06	3.88E-04
<i>Nop58</i>	1.56	6.98	1.21E-06	4.13E-04
<i>Slc1a3</i>	3.66	8.36	1.35E-06	4.44E-04
<i>Adipor2</i>	2.28	9.97	1.62E-06	5.16E-04
<i>Mycl</i>	4.55	4.94	1.83E-06	5.64E-04
<i>Ly6a</i>	1.90	8.85	2.95E-06	7.92E-04
<i>Fam13a</i>	3.89	8.76	4.62E-06	1.08E-03
<i>Acot2</i>	1.82	8.33	5.83E-06	1.26E-03
<i>Cdk4</i>	1.68	5.70	9.13E-06	1.79E-03
<i>Podxl</i>	2.92	6.89	1.13E-05	2.07E-03
<i>Por</i>	1.91	8.77	1.13E-05	2.07E-03
<i>Sertad1</i>	1.65	5.54	1.40E-05	2.34E-03
<i>Gatb</i>	2.16	4.47	1.75E-05	2.66E-03
<i>Josd2</i>	2.16	5.06	2.39E-05	3.45E-03
<i>Cct7</i>	1.08	8.49	2.48E-05	3.45E-03
<i>Muc13</i>	7.05	2.34	2.57E-05	3.46E-03

B. Top-20 up regulated genes visWAT KO vs WT

Gene	logFC	logCPM	PValue	FDR
<i>Fam13a</i>	4.52	8.76	2.48E-08	8.43E-05
<i>Mycl</i>	5.21	4.94	5.47E-08	9.39E-05
<i>Asns</i>	2.88	6.44	7.64E-08	1.11E-04
<i>Sik2</i>	2.31	7.10	1.44E-07	1.63E-04
<i>Grb10</i>	1.78	6.54	3.25E-07	2.76E-04
<i>Acot2</i>	1.87	8.33	3.98E-07	3.12E-04
<i>Fbxo21</i>	1.84	8.08	9.27E-07	5.25E-04
<i>Adipor2</i>	2.03	9.97	1.06E-06	5.69E-04
<i>Rbp7</i>	2.82	8.33	1.13E-06	5.69E-04
<i>Hoxa7</i>	1.54	6.62	1.45E-06	6.21E-04
<i>Tmem88b</i>	3.30	5.46	1.46E-06	6.21E-04
<i>Crls1</i>	2.88	6.04	2.13E-06	8.05E-04
<i>Cyp2e1</i>	4.98	9.59	2.02E-06	8.05E-04
<i>Ebf2</i>	1.98	6.54	2.27E-06	8.27E-04
<i>Klhl2</i>	1.59	6.88	2.42E-06	8.52E-04
<i>Adgrg2</i>	4.18	6.75	2.60E-06	8.84E-04
<i>Sctr</i>	3.39	4.60	2.97E-06	9.77E-04
<i>Sqor</i>	1.82	5.84	3.15E-06	1.00E-03
<i>Gsta3</i>	2.12	7.26	3.25E-06	1.00E-03
<i>Por</i>	1.80	8.77	4.33E-06	1.17E-03

As for the top down regulated genes (Table 4.3.4), six out of the 20 genes were shared with in both depots, *Fam46a/Tent5a* (Table 4.3.4, light green highlight) being one of the genes as well as *Pltp*, *Sycp3*, *Lbp*, *Ubd*, and *Pcolce2*. Both *Pltp* and *Lbp* are lipid transfer proteins belonging to the same family (Kirschning et al. 1997). High levels of plasma PLTP are found in obese, insulin resistant and type II diabetes patients (Dullaart et al. 2012) and increased LBP in adipose tissue is connected to impaired adipocyte differentiation and inflammation signatures typically found in obese tissue (Moreno-Navarrete et al. 2013). *Ubd* is known as a regulator of inflammatory response and fat metabolism, where a KO mouse model showed reduced fat mass and extended life span (Canaan et al. 2014). Short hairpin RNA (shRNA) experiments targeting the *Ubd* gene showed suppression of adipogenesis, indicating a role in adipogenic differentiation (C. Zhao et al. 2018). Once more the top differentially expressed genes, in this case that are down-regulated, have seemingly relevant functions in pathways related to metabolism and adipogenic function, indicating that *Fam46a* may also play a key role in these biological processes.

Chapter 4

Table 4.3.4: Top 20 down regulated differentially expressed genes in A) subWAT and B) visWAT.

*logFC = log fold-change; logCPM = log counts per million reads mapped; FDR = false discovery rate
P-values were corrected for multiple testing (FDR) using Benjamini and Hochberg correction methods*

A. Top-20 down regulated genes subWAT KO vs WT

Gene	logFC	logCPM	PValue	FDR
<i>Gpc4</i>	-3.25	4.76	2.74E-09	1.47E-05
<i>Pltp</i>	-3.06	8.67	2.88E-09	1.47E-05
<i>Sycp3</i>	-7.24	4.63	8.28E-09	2.31E-05
<i>Adam23</i>	-3.49	5.36	9.06E-09	2.31E-05
<i>Lbp</i>	-3.75	8.26	1.58E-08	3.22E-05
<i>Ubd</i>	-7.14	6.61	2.22E-08	3.78E-05
<i>Gnao1</i>	-3.97	5.31	3.50E-08	5.10E-05
<i>Pam</i>	-3.20	7.74	5.94E-08	5.51E-05
<i>Ostm1</i>	-2.41	4.95	5.07E-08	5.51E-05
<i>Dcbld2</i>	-2.12	6.21	1.21E-07	8.79E-05
<i>Tent5a</i>	-1.78	7.42	1.09E-07	8.79E-05
<i>Serping1</i>	-2.39	10.30	1.84E-07	1.25E-04
<i>Gm2a</i>	-2.29	8.47	2.35E-07	1.50E-04
<i>Arap2</i>	-2.62	5.21	2.76E-07	1.65E-04
<i>Olfm1</i>	-3.57	6.99	3.36E-07	1.90E-04
<i>Insig1</i>	-4.73	4.32	4.04E-07	2.17E-04
<i>Cyb5b</i>	-2.78	8.39	5.07E-07	2.58E-04
<i>Pcolce2</i>	-3.47	7.29	6.75E-07	3.13E-04
<i>Nkd1</i>	-4.06	6.04	7.69E-07	3.27E-04
<i>Serpina3</i> <i>n</i>	-2.55	9.47	7.68E-07	3.27E-04

B. Top-20 down regulated genes visWAT KO vs WT

Gene	logFC	logCPM	PValue	FDR
<i>Ubd</i>	-6.70	6.61	1.97E-09	2.01E-05
<i>Sycp3</i>	-5.45	4.63	1.16E-08	5.89E-05
<i>Ctsd</i>	-2.18	11.11	3.45E-08	8.79E-05
<i>Pltp</i>	-2.30	8.67	5.53E-08	9.39E-05
<i>Fbln7</i>	-4.66	3.83	1.08E-07	1.38E-04
<i>Tph2</i>	-10.19	4.06	2.72E-07	2.52E-04
<i>Lbp</i>	-2.76	8.26	2.56E-07	2.52E-04
<i>Saa3</i>	-6.96	5.76	4.42E-07	3.22E-04
<i>Mfge8</i>	-2.06	11.94	5.24E-07	3.56E-04
<i>Ctss</i>	-3.47	7.41	8.04E-07	5.13E-04
<i>Gusb</i>	-2.01	6.08	9.24E-07	5.25E-04
<i>Kcnj14</i>	-9.18	1.96	1.17E-06	5.69E-04
<i>Dot1l</i>	-1.21	7.00	1.39E-06	6.21E-04
<i>Lrpap1</i>	-1.20	7.04	2.12E-06	8.05E-04
<i>Pcolce2</i>	-2.66	7.29	3.77E-06	1.10E-03
<i>Hexa</i>	-1.77	6.97	3.69E-06	1.10E-03
<i>Pld3</i>	-2.23	6.54	4.13E-06	1.17E-03
<i>Tent5a</i>	-1.24	7.42	4.35E-06	1.17E-03
<i>Arhgap25</i>	-2.80	5.14	6.27E-06	1.60E-03
<i>Gba</i>	-1.55	5.24	8.94E-06	2.17E-03

4.3.7. Pathway analysis: exploratory and specific pathways

Although the differential gene expression analysis revealed interesting and highly significant gene candidates, it is important to take the entire gene expression profile into consideration. Therefore, a pathway analysis was performed on the differentially expressed genes (FDR <5%), to identify which molecular pathways are up or down-regulated when *Fam46a* expression is lost. While on the individual differential gene expression level we highlighted similarities between the two depots, the exploratory GO pathway analysis (Mi et al. 2019) revealed a discrepancy in pathways between the subWAT and the visWAT.

Although there were more significantly differentially expressed genes in the subWAT depot than the visWAT (432 vs 272 total genes), the subWAT pathway analysis had less pathways overall. Only two pathways were found to be enriched when analysing the 176 down-regulated genes (Table 4.3.5A); ribosome biogenesis and ribonucleoprotein complex biogenesis. Considering the function of the Fam46 family of proteins include nucleotide binding and RNA 3'-end modifications (Warkocki et al. 2018), the up-regulation of these pathways could be a compensation mechanism for the lack of *Fam46a*. When analysing the 256 down-regulated genes, the top 20 GO pathways involve mostly cell component (CC) subontologies, such as the extracellular space, endoplasmic reticulum and cell membrane (Table 4.3.5B). More relevant pathways to metabolic function included a down-regulation of 39/1318 genes in the lipid metabolic process, and 16/324 genes in the lipid catabolic process (Table 4.3.5B).

Chapter 4

Table 4.3.5: Top (up to 20) up and down enriched GO pathways using differentially expressed genes (KO vs WT) in subWAT samples. Grey highlights indicate same pathway found in the visWAT pathway analysis (Table 4.3.5). Green highlight indicates same pathway but with different directionality (up vs down) found in the visWAT pathway analysis.

Ontology (Ont) classification: BP = biological process; CC = cellular component; MF = molecular function

Up/Down = number of differentially expressed genes found in samples involved in pathway

A. Up regulated GO pathways for contrast subWAT KO vs WT

	Term	Ont	Up	p_adj_up
GO:0042254	ribosome biogenesis	BP	16	0.00450
GO:0022613	ribonucleoprotein complex biogenesis	BP	18	0.03616

B. Down regulated GO pathways top-20 for contrast subWAT KO vs WT

	Term	Ont	Down	p_adj_down
GO:0005576	extracellular region	CC	62	1.92E-06
GO:0044421	extracellular region part	CC	53	1.92E-06
GO:0005615	extracellular space	CC	47	2.44E-06
GO:0005783	endoplasmic reticulum	CC	57	4.24E-05
GO:0012505	endomembrane system	CC	95	2.75E-04
GO:0016020	membrane	CC	152	2.75E-04
GO:0031224	intrinsic component of membrane	CC	100	3.87E-04
GO:0016021	integral component of membrane	CC	96	7.67E-04
GO:0099240	intrinsic component of synaptic membrane	CC	14	7.67E-04
GO:0099699	integral component of synaptic membrane	CC	13	1.49E-03
GO:0031226	intrinsic component of plasma membrane	CC	36	6.35E-03
GO:0006629	lipid metabolic process	BP	39	1.22E-02
GO:0031012	extracellular matrix	CC	20	1.31E-02
GO:0099055	integral component of postsynaptic membrane	CC	10	1.45E-02
GO:0016042	lipid catabolic process	BP	16	1.48E-02
GO:0005887	integral component of plasma membrane	CC	33	1.48E-02
GO:0062023	collagen-containing extracellular matrix	CC	17	1.48E-02
GO:0044425	membrane part	CC	113	1.98E-02
GO:0098936	intrinsic component of postsynaptic membrane	CC	10	1.98E-02
GO:0097060	synaptic membrane	CC	19	2.02E-02

For the visWAT pathway analysis we observed 11 up regulated GO pathways, where the lipid metabolic process pathway was the most significant (Table 4.3.6A, light green highlight). Interestingly, this pathway was also enriched in the subWAT, however, in the down regulated pathways (Table 4.3.5B, light green highlight), underlining a discrepancy between the depots. Other up regulated pathways in the visWAT included relevant metabolic biological processes, such as multiple lipid metabolic processes, fatty/carboxylic acid metabolic processes and mitochondrion (Table 4.3.6A). As for the down regulated pathway analysis, extracellular space and membrane cell component subontologies were enriched, similar to the subWAT depot (Table 4.3.6B, grey highlights). The most significant down regulated molecular function pathways involve hydrolase activity. Hydrolase enzymes are required for the catabolism of TGs into fatty acids (FAs) and glycerol, also known as lipolysis (Lass et al. 2011). Several biological process pathways involved in inflammation were also enriched, such as defence response, inflammatory response and regulation of interleukin-6 production (Table 4.3.6B). These pathways are important in the pathology of obesity, as this disease is characterised by an increase in inflammation in adipose tissue (Ellulu et al. 2017).

Table 4.3.6: Top (up to 20) A) up and B) down enriched GO pathways using differentially expressed genes (KO vs WT) in visWAT samples. Grey highlights indicate same pathway found in the subWAT pathway analysis (Table 4.3.4). Green highlight indicates same pathway but with different directionality (up vs down) found in the subWAT pathway analysis.

Ontology (Ont) classification: BP = biological process; CC = cellular component; MF = molecular function

Up/Down = number of differentially expressed genes found in samples involved in pathway

A. Up regulated GO pathways for contrast visWAT KO vs WT

	Term	Ont	Up	p_adj_up
GO:0006629	lipid metabolic process	BP	31	3.30E-05
GO:0044255	cellular lipid metabolic process	BP	27	3.30E-05
GO:0032787	monocarboxylic acid metabolic process	BP	20	8.30E-05

Chapter 4

GO:0006631	fatty acid metabolic process	BP	16	1.40E-04
GO:0019752	carboxylic acid metabolic process	BP	22	5.70E-03
GO:0008610	lipid biosynthetic process	BP	17	5.70E-03
GO:0043436	oxoacid metabolic process	BP	22	9.70E-03
GO:0006082	organic acid metabolic process	BP	22	1.20E-02
GO:0005739	mitochondrion	CC	34	1.30E-02
GO:0019216	regulation of lipid metabolic process	BP	13	1.40E-02
GO:0044281	small molecule metabolic process	BP	31	2.50E-02

B. Down regulated GO pathways top-20 for contrast visWAT KO vs WT

	Term	Ont	Down	p_adj_down
GO:0004553	hydrolase activity, hydrolyzing O-glycosyl compounds	MF	11	2.60E-06
GO:0016798	hydrolase activity, acting on glycosyl bonds	MF	11	2.10E-05
GO:0005615	extracellular space	CC	30	8.00E-05
GO:0005773	vacuole	CC	20	8.00E-05
GO:0005576	extracellular region	CC	38	1.00E-04
GO:0044421	extracellular region part	CC	32	2.40E-04
GO:1901136	carbohydrate derivative catabolic process	BP	10	2.80E-04
GO:0000323	lytic vacuole	CC	17	3.00E-04
GO:0044425	membrane part	CC	74	3.60E-04
GO:0031224	intrinsic component of membrane	CC	62	4.40E-04
GO:0006952	defense response	BP	30	4.40E-04
GO:0016021	integral component of membrane	CC	60	5.70E-04
GO:0005764	lysosome	CC	16	9.80E-04
GO:0016020	membrane	CC	89	1.00E-03
GO:0007611	learning or memory	BP	12	1.70E-03
GO:0006954	inflammatory response	BP	19	2.30E-03
GO:0002376	immune system process	BP	38	5.20E-03
GO:0050890	cognition	BP	12	5.20E-03
GO:0032675	regulation of interleukin-6 production	BP	9	5.20E-03
GO:0006517	protein deglycosylation	BP	4	6.80E-03

A KEGG pathway analysis was also performed, which only found enriched pathways in the visWAT depot (Table 4.3.7). Similar to the GO pathway analysis, an up-regulation of metabolic pathways is

observed, as well as valine, leucine and isoleucine degradation (Table 4.3.7A). These amino acids are known as branched chain amino acids (BCAA), whose homeostasis is dependent on adipose tissue (Herman et al. 2010). Interestingly, plasma levels of BCAA have been used as markers of obesity, insulin resistance and type II diabetes (Newgard et al. 2009). The down-regulated KEGG pathways included lysosome and phagosome pathways (Table 4.3.7B), which was also enriched in the GO analysis (Table 4.3.6B). Pathways not found in the previous GO analysis were glycan degradation pathways and complement and coagulation cascades. Proteoglycans are made up of glycosaminoglycan chains attached to a core protein and is an integral part of the ECM. They have been found to regulate a plethora of processes, such as differentiation and inflammation (Pessentheiner et al. 2020). The complement cascade is the main cascade of the innate immunity response and has been studied for its interaction with the coagulation system and their roles in the inflammation response (Oikonomopoulou et al. 2012).

Table 4.3.7: Top A) up and B) down enriched KEGG pathways using differentially expressed genes (KO vs WT) in visWAT samples

N = number of genes in pathway; Up/Down = number of differentially expressed genes found in samples involved in pathway

**A. Up regulated KEGG pathways for contrast visWAT
KO vs WT**

	Pathway	N	Up	p_adj_up
path:mmu01100	Metabolic pathways	919	29	0.0016
path:mmu00280	Valine, leucine and isoleucine degradation	39	6	0.0016

**B. Down regulated KEGG pathways for contrast visWAT KO vs
WT**

	Pathway	N	Down	p_adj_down
path:mmu04142	Lysosome	88	13	5.80E-08

Chapter 4

path:mmu00511	Other glycan degradation	16	6	7.20E-06
path:mmu00531	Glycosaminoglycan degradation	12	4	1.70E-03
path:mmu04145	Phagosome	95	7	2.70E-02
path:mmu04610	Complement and coagulation cascades	48	5	3.20E-02

The final pathway analysis was a selected GO pathway analysis, incorporating relevant pathways from literature, such as TGF β and BMP signalling, as well as our own observations, such as stem cell differentiation and leptin signalling (see Chapter 2 and 3). The same 19 pathways were analysed in both the subWAT depot (Table 4.3.8) and the visWAT depot (Table 4.3.9). Both depots had the same six pathways as the most significant when ranking by FDR.Mixed, meaning the directionality of the pathway is not considered (Tables 4.3.8 and 4.3.9, highlighted sections). Pathways of note were the TGF β receptor and BMP signalling pathways considering Fam46a's previous involvement in these pathways (Colland et al. 2004; Watanabe et al. 2018). If direction is considered, the TGF β receptor signalling pathway goes down, while the BMP signalling pathway goes up. However, this does not always pass the FDR threshold of 5% to be considered significant when direction is taken into account. The leptin signalling pathway is not significantly enriched in either depot, while stem cell differentiation is up regulated both, however, it is only considered significant in the visWAT samples. The only discrepancy found between the two depots involves the negative regulation of cell proliferation where, if directionality is considered, the subWAT enriches for a down regulation of this pathway while in the visWAT it is up regulated. There are 242 genes involved in this pathway, thus directionality with statistical significance of genes in one direction may be difficult to determine.

Table 4.3.8: Gene set enrichment analysis of selected GO pathways/terms in subWAT

*N*Gene = Number of genes in pathway; *P*Value/*F*DR = significance when direction is considered (up or down); *P*Value/*F*DR.Mixed = significance when direction is not considered. *F*DR.Mixed = significance when direction is not considered.

Gene set enrichment test of selected GO terms: subWAT

GO ID	NGene	Direction	PValue	FDR	PValue.Mixed	FDR.Mixed	Pathway Name
GO:0032870	28	Up	0.031	0.29	7.40E-09	1.40E-07	cellular response to hormone stimulus
GO:0001568	35	Down	0.077	0.37	4.20E-07	4.00E-06	blood vessel development
GO:0007179	46	Down	0.01	0.19	1.00E-06	4.90E-06	TGFb receptor signaling pathway
GO:0008285	242	Down	0.195	0.46	1.00E-06	4.90E-06	negative regulation of cell proliferation
GO:0030509	41	Up	0.443	0.70	1.50E-03	5.80E-03	BMP signaling pathway
GO:0010035	14	Up	0.238	0.50	5.60E-03	1.80E-02	response to inorganic substance
GO:0048863	22	Up	0.058	0.37	5.30E-02	1.40E-01	stem cell differentiation
GO:0007389	20	Up	0.425	0.70	7.30E-02	1.70E-01	pattern specification process
GO:0006352	17	Up	0.173	0.46	1.20E-01	2.60E-01	DNA-templated transcription, initiation
GO:0045165	27	Down	0.973	0.99	4.00E-01	7.70E-01	cell fate commitment
GO:0048608	3	Down	0.133	0.46	1.00E+00	1.00E+00	reproductive structure development
GO:0033210	10	Down	0.174	0.46	7.80E-01	1.00E+00	leptin-mediated signaling pathway
GO:0045995	12	Up	0.347	0.66	8.80E-01	1.00E+00	regulation of embryonic development
GO:0061614	4	Down	0.521	0.76	1.00E+00	1.00E+00	pri-miRNA transcription by RNA polymerase II
GO:0061061	6	Up	0.578	0.79	9.90E-01	1.00E+00	muscle structure development
GO:0001708	15	Up	0.621	0.79	8.40E-01	1.00E+00	cell fate specification
GO:0007423	1	Down	0.775	0.92	7.70E-01	1.00E+00	sensory organ development
GO:0048598	14	Up	0.879	0.98	9.40E-01	1.00E+00	embryonic morphogenesis
GO:0034616	6	Down	0.987	0.99	1.00E+00	1.00E+00	response to laminar fluid shear stress

Chapter 4

Table 4.3.9: Gene set enrichment analysis of selected GO pathways/terms in visWAT

NGene = Number of genes in pathway; *PValue/FDR* = significance when direction is considered (up or down); *PValue/FDR.Mixed* = significance when direction is not considered.

Gene set enrichment test of selected GO terms: visWAT

GO ID	NGenes	Direction	PValue	FDR	PValue.Mixed	FDR.Mixed	Pathway Name
GO:0032870	28	Up	0.02	0.18	2.30E-08	4.50E-07	cellular response to hormone stimulus
GO:0008285	242	Up	0.26	0.45	6.80E-05	6.50E-04	negative regulation of cell proliferation
GO:0048863	22	Up	0.04	0.19	1.30E-04	8.50E-04	stem cell differentiation
GO:0030509	41	Up	0.31	0.45	1.60E-03	6.30E-03	BMP signaling pathway
GO:0007179	46	Down	0.27	0.45	1.50E-03	6.30E-03	TGFβ receptor signaling pathway
GO:0001568	35	Down	0.61	0.74	2.40E-02	7.70E-02	blood vessel development
GO:0010035	14	Up	0.01	0.18	6.10E-02	1.70E-01	response to inorganic substance
GO:0045995	12	Up	0.04	0.19	1.30E-01	3.00E-01	regulation of embryonic development
GO:0007389	20	Up	0.11	0.35	2.00E-01	3.70E-01	pattern specification process
GO:0007423	1	Up	0.21	0.45	2.10E-01	3.70E-01	sensory organ development
GO:0045165	27	Up	0.37	0.50	1.90E-01	3.70E-01	cell fate commitment
GO:0033210	10	Up	0.27	0.45	6.70E-01	9.90E-01	leptin-mediated signaling pathway
GO:0006352	17	Up	0.70	0.78	6.30E-01	9.90E-01	DNA-templated transcription, initiation
GO:0001708	15	Up	0.08	0.29	8.20E-01	1.00E+00	cell fate specification
GO:0034616	6	Up	0.19	0.45	9.50E-01	1.00E+00	response to laminar fluid shear stress
GO:0061614	4	Up	0.29	0.45	7.70E-01	1.00E+00	pri-miRNA transcription by RNA polymerase II
GO:0048598	14	Up	0.63	0.74	8.60E-01	1.00E+00	embryonic morphogenesis
GO:0061061	6	Up	0.90	0.93	1.00E+00	1.00E+00	muscle structure development
GO:0048608	3	Down	0.93	0.93	1.00E+00	1.00E+00	reproductive structure development

To summarise, the individual differential gene expression analysis and the pathway analysis show that both the WAT depots have an up regulation of genes involved in relevant metabolic processes such as thermogenesis, catabolism of BCAAs, and regulation of adipocyte differentiation. Down regulated genes are implicated pathways commonly dysregulated in obesity such as HDL maturation, inflammation and ECM pathways. The subWAT exclusively had an up regulation of pathways involved in ribosome biogenesis. The lipid metabolic process pathway was down regulated in subWAT, which surprisingly was up regulated in visWAT. The effect of the *Fam46a* KO is seemingly affecting the metabolic state of these adipose tissue depots, targeting particular pathways depending on the depot.

Chapter 4

4.4. Discussion

Although a previously published study had explored the effect of a full body Fam46a KO mouse model (Diener et al. 2016), our study focused on the metabolic phenotyping of this mouse model under standard diet and a HFD challenge. A particular focus was placed on the adipose tissue, as FAM46A is differentially expressed in obese and post weight-loss subWAT in humans (Carayol et al. 2017).

Overall, our results indicate that the loss of Fam46a under HFD protects the mouse against obesity and related metabolic dysfunction. The KO, as previously observed, was smaller in size and weight compared to the WT, yet able to gain weight during the HFD (Figure 4.3.2). Accompanying body composition data complements this difference, where the KO had lower fat and lean mass than the WT before and after the HFD (Figure 4.3.3). Metabolic profiling of the KO mice show they had significantly better glucose clearance (Figure 4.3.4), reduced RER on the standard diet and increased energy expenditure during the HFD intervention. When compared to the WT levels of biomarkers in the plasma indicated a metabolically healthier mouse under HFD with decreased levels of ALT, HDL and leptin. The histological analysis of the different tissues shows a similar trend, where the obese WT mice have expanded lipids in their BAT and lipid deposits in the liver indicative of an obese phenotype (Figure 4.3.6A/B). The HFD KO mice do not show these morphologies, however, TG measurements of the liver tissue show that these mice livers have higher levels of TGs than the chow WT and KO (Figure 4.3.6B.2) as well as markers of obesity in the plasma (Table 4.3.1), indicating the diet has an effect on the health of these mice.

Although we observed these significant differences between the WT and KO genotypes that suggest obesity resistance and a healthier metabolic outcome in the KO mice, it is important to consider the smaller stature of the KO mice, which occurs since 3 weeks of age (data not shown), as well as the overall health of these mice, as a full body loss of *Fam46a* in our mice lead to high mortality (Figure 4.3.1A) and small litters (data not shown). The average food intake was also less when fed standard chow, however, this improved with the HFD (Figure 4.3.5B). This improvement may be due to the nature of the food, as chow pellets are harder than ones used for HFD. With the fragile bone and teeth phenotype of the KO mice, the softer nature of the HFD pellets, and the tendency to crumble to the bottom of the cage, perhaps made it easier for the mice to consume food under the HFD intervention. This reduction in chow food intake may also explain the RER result, where chow diet fed KO mice tend to use fat as a fuel source when fed a chow diet (Figure 4.3.5A). RER reduction is commonly seen when mammals are in a fasted state (Compher et al. 2006; Marvyn et al. 2016) and considering this significant difference in RER is no longer observed on the HFD, the reduced food intake may be causing this effect. It is also important to consider that near end study, the WT mice are extremely obese, weighing nearly twice as much as the KO (Figure 4.3.2B). As many of these tests are normalised by total body weight, such as the IPGTT and certain TSE measurements, the final results may be misinterpreted due to these calculated ratios. For example, when comparing heat production, we observe significant differences during the HFD (Figure 4.3.5C). If we remove the weight correction, there is no significant difference between the genotypes (data not shown). Packard and Boardman (1999) have reviewed these issues associated with total body weight normalisation and have

Chapter 4

suggested to utilise graphical analysis of the data and the multiple regression (ANCOVA) instead to remove the effects of body size. It should also be noted that obesity itself has a negative effect on energy expenditure (Choi et al. 2015), therefore the increase in heat production could also be due to the KO mouse's non-obese nature.

When interpreting the transcriptomics data on the different WAT depots during the HFD, interesting differences emerged between the genotypes and the depots themselves. When Fam46a is missing, the subWAT depot showed a unique up regulation of genes involved in ribosome biogenesis (Table 4.3.5A). Considering the Fam46a protein contains a NTase domain, which have been found to modify 3' ends of RNA (Kuchta et al. 2016; Warkocki et al. 2018), this could be a possible compensation mechanism for the loss of Fam46a. Interestingly, none of the other paralogs of Fam46a were significantly down-regulated in our dataset, indicating they may not compensate for one another further solidifying the unique function of each protein in the FAM46 family (data not shown). The up regulated pathways in visWAT were all metabolically related, such as lipid and fatty acid metabolic processes, and mitochondrion (Table 4.3.6A), suggesting that KO visWAT is more metabolically active and the mice are expending more energy. This is complemented by the metabolic cage data, where heat production is significantly higher in the KO mice under HFD (Figure 4.3.5C). Additionally, Rbp7 was also a top up regulated gene in both depots. Retinol-binding protein 7 (Rbp7), or cellular retinol-binding protein type III (CrbpIII), is involved in the metabolism of retinol or vitamin A (Quadro et al. 2003). Interestingly, proper transport of retinol metabolites (retinoids) is important for adipose tissue browning and adaptive thermogenesis under cold exposure (Fenzl et al. 2020). Adipor2 was also up

regulated in the KO tissues, which also is involved in pathways of fatty acid oxidation and glucose uptake via adiponectin signalling (Yamauchi et al. 2007). This suggests that loss of Fam46a in these WAT depots may be transitioning to a more beige or brown state, which ultimately is leading to their healthier metabolic profiles and leaner bodies. Why the down-regulated pathways are different in subWAT vs visWAT may be due to the different roles of the depots. It is widely accepted that subWAT and visWAT, despite their morphological similarities, have different metabolic and functional profiles (Saely et al. 2012). In metabolic syndrome patients, the visWAT depots are more indicative of their health status than the subWAT depots. Obese mice are more likely to show hypertrophic visWAT, while subWAT is more likely to undergo hyperplasia, indicative of the different precursor capacities, where subWAT APCs are more adipogenic (Esmaili et al. 2020).

Interestingly, we also observed up regulated expression of novel anti-adipogenic regulator Fam13a in the KO (Table 4.3.3). Although it has a similar nomenclature to Fam46a, Fam13a is involved in Rho GTPase signalling and variations in its locus have been genetically associated with several metabolic conditions (Shungin et al. 2015; Loos & Kilpeläinen 2018). KO mice models of Fam13a have accelerated adipogenic differentiation potential in isolated APCs, as well as a decrease in the visWAT/subWAT ratio while still gaining weight on a HFD (Fathzadeh et al. 2020). The authors propose that Fam13a has anti-adipogenic properties that leads to an unhealthy body fat mass distribution between fat depots, where increased expression leads to insulin resistance. Perhaps Fam46a also has anti-adipogenic potential that is being compensated by the increase in Fam13a expression, which also affects the different adipogenic potentials of the different depots.

Chapter 4

As for the down-regulated genes/pathways, both depots had structural components, such as extracellular region and membrane enriched, as well as pathways involved in immune and inflammation response (Table 4.3.5). ECM is an essential component of adipose tissue as it is required to maintain structure during remodelling of the tissue under stress conditions, such as obesity (Ruiz-Ojeda et al. 2019). It is a complex scaffold of several proteins, polysaccharides and proteoglycans. As ECM remodelling and subsequent inflammation is indicative of metabolically distressed adipose tissue, the KO depots appear once more to have a more healthy and stable adipose tissue state.

Although the KO mouse has presented promising results in improving metabolic health, it is once more important to note the differences in metabolic state of the mice in our study. As mentioned with the phenotyping data, the WT mice are considered obese under the HFD while the KO mice remain lean. Choi et al. (Choi et al. 2015) characterised and performed transcriptomic analysis on visWAT from mice under an HFD intervention and observed similar pathway enrichments to the ones in our study, such as mitochondria and fatty acid oxidation pathways. They suggest these pathways are down regulated under HFD due to a higher energy expenditure in standard diet chow. They also report an up regulation in pathways involved in immune response and inflammation, which were enriched in our datasets as well. Promising candidate gene *Fam13a* was also highly down regulated in their study. As the KO mice are not obese, the metabolic phenotyping and pathways analysis may unfortunately be a result of comparing obese vs non-obese tissue instead of the effect of the genotype. However, a novel phenotype in our study was the increase in plasma phosphate during the HFD. Phosphate is a key mineral in bone health and regulation of body phosphate

homeostasis includes intestinal absorption of dietary phosphate and uptake/intake between the plasma and essential phosphate organs (bone and kidney) (Marks et al. 2010; Goretti Penido & Alon 2012). Increased levels in the plasma can indicate bone resorption, which involves osteoclast degradation for regulation of bone mass (Teitelbaum 2000). This is not an unexpected observation and is highlighting the potential role of Fam46a in the regulation of bone formation.

The selected GO pathway analysis presented two potentially interesting pathways, namely the TGF β receptor and BMP signalling pathways. These particular signalling pathways have interesting inverse roles in the regulation of adipogenesis, where TGF β tends to inhibit adipogenesis while BMP promotes adipogenesis (Tan et al. 2012), with certain exceptions. Both pathways involve specific ligands and phosphorylation of downstream Smad proteins that form a TF complex that regulate the expression of specific target genes. A downstream scaffold Smad protein, Smad4, is shared between these pathways. Interestingly, a limb specific Smad4 KO causes chondrodysplasia in mice (Yan et al. 2018), displaying similar limb phenotypes to the Fam46a KO mouse. Chondrodysplasia is a type of skeletal dysplasia with short stature and dysregulated endochondral growth (Irving et al. 2008). The authors observe reduced expression of chondrocyte differentiation markers in their KO model and suggest that Smad4 is required for proper chondrocyte maturation by regulating Runx2, a master regulator of osteogenesis (Komori 2019). A KO mouse model of Smad3, a binding partner of Smad4 upstream of the signalling pathway, revealed a strikingly similar metabolic profile to our KO mouse model (Tan et al. 2011). On a standard diet this mouse has reduced BW and length, reduced adiposity, improved glucose tolerance and the WAT showed an up

Chapter 4

regulation of genes involved in FA uptake and oxidation. It also appeared to be resistant to DIO. The authors conclude that Smad3 inhibits antiadipogenic factor CHOP-10, which in turn inhibits C/EBP β . This TF acts on the promoters of PPARs: α , δ and γ . PPAR γ is a master regulator of adipogenesis and PPAR α/δ have anti-obesogenic properties, where their expression leads to increased fat metabolism and β -oxidation (Y.-X. Wang et al. 2003; Barish et al. 2006). This suggested mode of action of Smad3 in adipocytes goes against the consensus that TGF β tends to inhibit adipogenesis. However, more studies show there is some discrepancy between these two SMAD pathways in adipocytes and may be dependent on the lineage of the precursor and specific ligands utilised for induction (Figure 4.4.1). For example, it was discovered that ligand TGF β 3 is a regulator of adipocyte number by promoting the proliferation of adipocyte progenitor cells, enabling hyperplasia (Petrus et al. 2018).

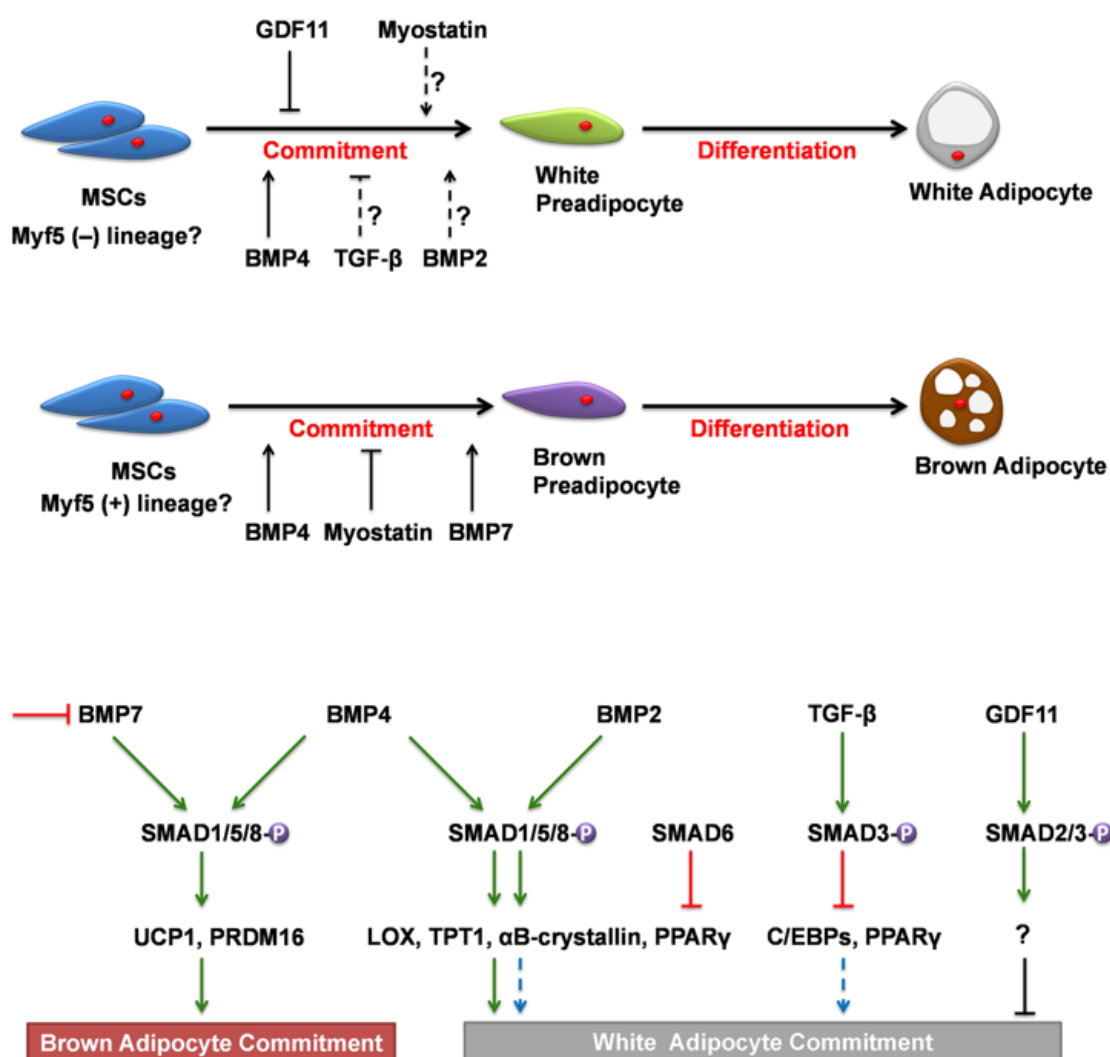


Figure 4.4.1: Summary of TGF-β/SMAD signalling action on the adipocyte commitment of MSCs.

Taken from Li & Wu (2020).

Dashed lines represent inconclusive interactions from discrepancies in different studies.

In conclusion, any differences in the metabolic profiles of the *Fam46a* KO mice could be explained by their small statures, decrease in food intake and overall non-obese phenotype. This consideration of the obese vs non-obese phenotypes should be taken into account and more follow-up studies are needed to fully understand the role of *Fam46a* in adipose tissue biology. In future studies, it would be more beneficial to do an adipose tissue specific *Fam46a* KO to remove any bias of the bone phenotype, which hopefully improves the food intake. It could also be interesting to repeat the phenotyping in the heterozygous genotype of the KO mice, as the previous study observed that the heterozygous mice had no bone abnormalities, yet a

Chapter 4

significant increase in plasma ALP compared to WT, indicating the importance of *Fam46a* in bone health. Expression of Smad proteins, such as *Smad4* or *Smad3*, in adipose tissue could also be perturbed to observe if this affects *Fam46a* expression and regulation. It would have been interesting to perform further adipose tissue phenotyping by weighing the different depots to compare the ratio of subWAT vs visWAT, as well as compare the adipogenic potential of each depots APCs to understand if *Fam46a* plays a functional role in adipogenesis and adiposity, like *Smad3*.

4.5. Materials and methods

4.5.1. Animal experiments

Animals were maintained in individually ventilated cages in a room with controlled temperature (22 ± 2 °C) and humidity under a 12 hr light/12 hr dark cycle and with standard chow diet (Standard diet 1314, Altromin) and HFD (D12492, ssniff Spezialdiäten GmbH) and drink *ad libitum*. Animals were closely followed-up by the animal caretakers and the experimenters, with regular inspection by a veterinarian, as per the standard health and animal welfare procedures of the local animal facility. No statistical method was used to predetermine sample size. All animal experiments were approved by the Animal Ethics Committee of the Czech Academy of Sciences (primary screen project number: 62/2016 and secondary screen project number: 45/2017) and were performed according to Czech guidelines for the Care and Use of Animals in Research and Teaching.

4.5.2. Generation of *Fam46a* KO mouse

Fam46a KO mouse line was created using CRISPR/Cas9 method with gRNA targeting exon 2 (5'AGCTTCGGCGACCGCTGCT 3') in C57BL/6N background. *Tent5a^{em1(IMPC)Ccpz}* mouse line was obtained with 26nt loss of function deletion in exon 2 (c.436_462del26). For genotyping, DNA was extracted from tails of 4 weeks old mice using the DirectPCR DNA Extraction System (Viagen, USA) and it was used as template for PCR with primers Forward: 5'-CTTTCCAGATGCATCAGAGATAC-3', Reverse: 5'-CCTCACACCTTCACAATCAG-3'.

4.5.3. Mouse diets: chow and HFD compositions

Mice were fed *ad libitum* pellets of standard diet, where 10% kcal was from fat. At 8 weeks of age, mice on the HFD intervention were fed high-fat food pellets composed of 60% kcal fat, 20% kcal

Chapter 4

carbohydrate and 20% kcal protein. The mice were fed this diet for 13 weeks.

4.5.4. Mouse body weight and composition measurements

Body weight measurements were performed at 18, 20 and 24 weeks of age on the standard diet. During the HFD intervention, body weight was measured at 8, 14, 18, 20 and 21 weeks of age (0, 6, 10, 12 and 13 weeks of HFD).

Whole body composition was measured at week 8 and week 21 of age during the HFD intervention. Body composition analysis by MicroCT was performed in HFD mice at the beginning of the experiment at 8 weeks old and at the end of experiment at 21 weeks old. Mice were anaesthetised by injecting 20% Zoletil – 0.03g/kg tiletamin and 0.03g/kg zolezepam (Virbac, France) intramuscularly for in vivo imaging of core body, i.e. excluding the head and tail sections in SkyScan 1176 (Bruger, Belgium) at a resolution of 35 μm per voxel (0.5 mm Al filter; voltage, 50 kV; current, 160 μA ; exposure, 200 ms; rotation, 0.7° and round scan). Reconstruction was performed in an InstaRecon 2.0.4.0 (Bruker, Belgium) with the following parameters set up to smoothing = 7, ring artifact reduction = 5, beam hardening correction = 10%, and defect pixel masking threshold = 5%. The range of intensities was set from 0.004 AU to 0.02 AU. Regions of interest for fat and lean mass analysis were selected in CT analyser 1.18.4.0 (Bruker, Belgium). Fat and lean mass were segmented from background by the Otsu method. Three-dimensional analysis was performed to describe changes in fat and lean volume in absolute and relative values.

4.5.5. Indirect calorimetry with TSE Phenomaster

Using the TSE PhenoMaster system (TSE Systems, Germany), mouse locomotion activity, feeding behaviour, oxygen consumption,

and carbon dioxide production were measured at 18 weeks of age. Mice were weighed before entering the cage system with food and water available *ad libitum*. A 24-hour measured acclimation period was done. The mice were housed for a total of 48 hours, with controlled 12h dark/12h light cycles. Mouse body weight was measured when removed from the cage.

4.5.6. Intra-peritoneal glucose tolerance test

IPGTTs were performed at week 20 of age on the standard diet and HFD. Animals were fasted overnight then weighed and fasted blood glucose levels were measured ($t = 0$). Blood was taken from the lateral tail vein with a small incision. The volume of the 20% glucose administered is determined according to the body weight (2 g of glucose/kg body mass). The glucose solution was administered inter-peritoneally by injection. Blood glucose was measured at 15, 30, 60 and 120 minutes after injection.

4.5.7. Plasma biochemistry measurements

Blood samples from 21 weeks old, HFD fed male mice were collected after overnight of fasting. On the other hand, blood samples from 24 weeks old, standard chow-fed male mice were collected after 6 hours of fasting. Mice were anaesthetised with isoflurane and were bled via the retrobulbar sinus using glass capillaries. Samples were collected in Lithium-Heparin coated tubes (KABE cat # 078028). After collection, each tube was mixed by gentle inversion and then kept on RT until centrifugation. Samples were centrifuged within 1 hour of collection at $5000 \times g$, for 10 minutes at 8°C . Once separated from the cells, plasma samples were analysed on the Beckman AU480 biochemical analyser for liver function (Albumin, ALP, ALT, AST, Bilirubin, Total protein, Urea), lipid metabolism (Cholesterol, Glycerol, Lipase, TGs) and bone profile (Ca, P, Albumin, ALP) parameters. Leptin

Chapter 4

was measured from plasma using Mouse Leptin DuoSet ELISA kit following the manufacturer's instructions (R&D Systems DY498, Minneapolis, USA).

4.5.8. Histological preparation and analysis

After sacrifice, mouse tissues were dissected and prepared for histological and transcriptomic analysis. Adipose tissue depots (inguinal (subWAT), epididymal (visWAT) and BAT) were collected and fixed in 10% formalin, rinsed in PBS and embedded in paraffin. Liver samples were snap-frozen in liquid nitrogen and cryopreserved. Femurs were fixed in 4% paraformaldehyde (PFA) for 48-72 hours and subsequently decalcified in OSTEOSOFT (Merck) for 10 days. The dehydrated tissue was embedded in paraffin. Paraffin blocks and cryopreserved tissue were sectioned and stained with haematoxylin and eosin (H&E) for image acquisition using the Olympus Slide Scanner V120. Images were analysed using QuPath and ImageJ. Adipocyte size was measured using ImageJ plugin Adiposoft, using 6-10 images per depot. Downstream analysis and graphs of quantification were made in R.

4.5.9. Library preparation for QuantSeq and sequencing

Cryopreserved adipose tissue samples (subWAT and visWAT) were prepared for RNA sequencing (RNA-seq) using the QuantSeq method (Corley et al. 2019). Total RNA was extracted from the tissue using the Agencourt RNAdvance Tissue Kit (Beckman Coulter) and quality was assessed using the Fragment Analyser and Quant It Ribogreen (Life Technologies) for concentration. 200 to 300 ng of RNA was used to prepare the QuantSeq libraries with the QuantSeq 3' mRNA-Seq Library Prep Kit HT for Illumina (Lexogen) following user instructions.

The sequencing was performed on HiSeq 2500 with Rapid V2 chemistry SR 65 cycles (Illumina), loaded at 9pM and with 3% Phix.

4.5.10.Quantseq data filtering and normalisation

The sequencing run was demultiplexed with Bcl2FastQ using an in-house analysis pipeline (UPB pipeline; Demux2696). After demultiplexing, per sample sequencing data quality was evaluated with the FastQC software. Raw counts were obtained from sequences by mapping to the mouse reference genome using STAR v.2.5.3 (Dobin A, 2012) and counting using htseq-count v.0.6.1 (Anders S, 2014). Data was filtered by selecting genes with at least 15 reads in at least 2 samples. The filtering step was performed on the CPM values which took the library size into account (threshold of 3.151 on the CPM values). Genes with no annotation were discarded. Data was normalised using the TMM method.

Final conclusions and outlook

Final conclusions and outlook

Our interest in FAM46A arose from a genetic association study that showed a hotspot region of variants associating with the expression of downstream gene *FAM46A*. As this association was found during a weight-loss intervention study and the eQTL analysis was performed in RNAseq data acquired from subWAT biopsies, we aimed to understand the potential role and regulation of FAM46A in adipocyte biology. Our study involved three different approaches: investigating the regulatory landscape of the *Fam46a* promoter (Chapter 2), *in vitro* expression interference of *Fam46a* in differentiating adipocytes (Chapter 3) and *in vivo* characterisation of a full body KO mouse model (Chapter 4).

The first part of our study showed that the *Fam46a* promoter is interacting with proximal upstream loci. These regions contained motifs for a number of TFs involved in the regulation of adipogenesis, such as TEADs and ESRRA. Pathway analysis of the TFs interacting with these motif regions enrich for pathways such as cell fate commitment and the SMAD2/3 TGF β -signalling pathway, known to repress adipogenesis (Li & Wu 2020). Comparative analysis with ChIP-seq datasets revealed SMAD4 binds to an interaction region of *Fam46a* under TGF β 1 stimulation indicating *Fam46a* may be a target gene of SMAD4 during induction of TGF β -signalling pathway. The *in vitro* study produced controversial results, where siRNA knockdown of *Fam46a* delayed lipid accumulation after induction of adipogenesis. Expression of master regulators *PPAR γ* and *CEBPA* were also reduced, indicating FAM46A may be a pro-adipogenic factor, although scRNAseq data would suggest it is required for maintaining stem state due to an enriched expression of *Fam46a* in stem or stem-like cells. In the final *in vivo* study, where a *Fam46a* KO mouse was metabolically phenotyped, we observed that the KO appeared more metabolically healthy, with reduced fat mass and improved glucose clearance. Transcriptomic analysis of two WAT depots suggested that the KO adipose tissue is more metabolically active and

enriches for pathways involved in stem cell differentiation and BMP/TGF β -signalling, mirroring what was previously observed in the first part of our study.

Overall, previous studies and the results indicate that *Fam46a* may be a novel player in the opposing BMP/TGF β -signalling pathways and its expression regulated by these pathways via SMAD4. Considering KO of *Smad4* in mice mimics the *Fam46a* phenotype and there was a potential binding site close to the *FAM46A* promoter, we propose that expression and binding of SMAD4 activates expression of *FAM46A* when the TGF β -signalling pathway is activated. However, SMAD4 is shared between the two BMP/TGF β -signalling pathways, therefore it would be interesting to further elucidate which pathway, or both, activates the expression of *FAM46A* and if this contributes to the stem state of a cell. It would also be interesting to understand how SMAD4 and *FAM46A* interact and if this complex is somehow a dual-regulator between the BMP/TGF β -signalling pathways. Although, the *Smad3* KO mouse has a similar metabolic phenotype to the *Fam46a* KO, which is part of the anti-adipogenic TGF β -signalling pathway, it is perhaps more likely *Fam46a* belongs to this pathway vs BMP.

Future studies would require further optimisations of *in vitro* and *in vivo* systems that produce cleaner and more reliable data to reach any plausible conclusions. Considering the severe phenotype of the KO mouse, understanding the function of *Fam46a* could have potential therapeutic benefits in adipose tissue diseases, such as obesity, as well as skeletal dysplasia.

References

References

- Anderson, J.W. et al., 2001. Long-term weight-loss maintenance: a meta-analysis of US studies. *The American journal of clinical nutrition*, 74(5), pp.579–584.
- Andrews, S., 2010. FastQC: A Quality Control Tool for High Throughput Sequence Data. Available at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
- Barish, G.D., Narkar, V.A. & Evans, R.M., 2006. PPAR delta: a dagger in the heart of the metabolic syndrome. *The Journal of clinical investigation*, 116(3), pp.590–597.
- Bartelt, A. & Heeren, J., 2014. Adipose tissue browning and metabolic health. *Nature reviews. Endocrinology*, 10(1), pp.24–36.
- Besio, R. et al., 2019. Bone biology: insights from osteogenesis imperfecta and related rare fragility syndromes. *The FEBS journal*, 286(15), pp.3033–3056.
- Bessesen, D.H., 2008. Update on obesity. *The Journal of clinical endocrinology and metabolism*, 93(6), pp.2027–2034.
- Billon, N. et al., 2007. The generation of adipocytes by the neural crest. *Development (Cambridge, England)*, 134(12), pp.2283–2292.
- Birsoy, K., Festuccia, W.T. & Laplante, M., 2013. A comparative perspective on lipid storage in animals. *Journal of cell science*, 126(Pt 7), pp.1541–1552.
- Bjørndal, B. et al., 2011. Different adipose depots: their role in the development of metabolic syndrome and mitochondrial response to hypolipidemic agents. *Journal of obesity*, 2011, p.490650.
- Bonet, M.L. et al., 2015. Carotenoids and their conversion products in the control of adipocyte function, adiposity and obesity. *Archives of biochemistry and biophysics*, 572, pp.112–125.
- Canaan, A. et al., 2014. Extended lifespan and reduced adiposity in mice lacking the FAT10 gene. *Proceedings of the National Academy of Sciences of the United States of America*, 111(14), pp.5313–5318.
- Cannon, B. & Nedergaard, J., 2004. Brown adipose tissue: function and physiological significance. *Physiological reviews*, 84(1), pp.277–359.
- Carayol, J. et al., 2017. Protein quantitative trait locus study in obesity during weight-loss identifies a leptin regulator. *Nature communications*, 8(1), pp.2084–14.
- Carobbio, S. et al., 2019. Brown and beige fat: From molecules to physiology and pathophysiology. *Biochimica et biophysica acta. Molecular and cell biology of lipids*, 1864(1), pp.37–50.
- Castello, A. et al., 2012. Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell*, 149(6), pp.1393–1406.
- Cawthorn, W.P., Scheller, E.L. & MacDougald, O.A., 2012. Adipose tissue stem cells meet preadipocyte commitment: going back to the future. *Journal of lipid research*, 53(2), pp.227–246.
- Chau, Y.-Y. et al., 2014. Visceral and subcutaneous fat have different origins and evidence supports a mesothelial source. *Nature Cell Biology*, 16(4), pp.367–375.

- Chen, H. et al., 1996. Evidence that the diabetes gene encodes the leptin receptor: identification of a mutation in the leptin receptor gene in db/db mice. *Cell*, 84(3), pp.491–495.
- Choe, S.S. et al., 2016. Adipose Tissue Remodeling: Its Role in Energy Metabolism and Metabolic Disorders. *Frontiers in endocrinology*, 7(Suppl 2), p.30.
- Choi, M.-S. et al., 2015. High-fat diet decreases energy expenditure and expression of genes controlling lipid metabolism, mitochondrial function and skeletal system development in the adipose tissue, along with increased expression of extracellular matrix remodelling- and inflammation-related genes. *The British journal of nutrition*, 113(6), pp.867–877.
- Christodoulides, C. et al., 2009. Adipogenesis and WNT signalling. *Trends in endocrinology and metabolism: TEM*, 20(1), pp.16–24.
- Coleman, D.L., 1978. Obese and diabetes: two mutant genes causing diabetes-obesity syndromes in mice. *Diabetologia*, 14(3), pp.141–148.
- Colland, F. et al., 2004. Functional proteomics mapping of a human signaling pathway. *Genome research*, 14(7), pp.1324–1332.
- Compher, C. et al., 2006. Best practice methods to apply to measurement of resting metabolic rate in adults: a systematic review. *Journal of the American Dietetic Association*, 106(6), pp.881–903.
- Corley, S.M. et al., 2019. QuantSeq. 3' Sequencing combined with Salmon provides a fast, reliable approach for high throughput RNA expression analysis. *Scientific reports*, 9(1), pp.18895–15.
- Davies, J.O.J. et al., 2016. Multiplexed analysis of chromosome conformation at vastly improved sensitivity. *Nature methods*, 13(1), pp.74–80.
- Dekker, J. et al., 2002. Capturing chromosome conformation. *Science (New York, N.Y.)*, 295(5558), pp.1306–1311.
- Derynck, R. & Budi, E.H., 2019. Specificity, versatility, and control of TGF- β family signaling. *Science signaling*, 12(570).
- Diener, S. et al., 2016. Exome sequencing identifies a nonsense mutation in Fam46a associated with bone abnormalities in a new mouse model for skeletal dysplasia. *Mammalian genome : official journal of the International Mammalian Genome Society*, 27(3-4), pp.111–121.
- Dina, C. et al., 2007. Variation in FTO contributes to childhood obesity and severe adult obesity. *Nature genetics*, 39(6), pp.724–726.
- Dixon, J.R. et al., 2012. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*, 485(7398), pp.376–380.
- Doyard, M. et al., 2018. FAM46A mutations are responsible for autosomal recessive osteogenesis imperfecta. *Journal of medical genetics*, 55(4), pp.278–284.
- Dullaart, R.P.F. et al., 2012. Type 2 diabetes mellitus interacts with obesity and common variations in PLTP to affect plasma phospholipid transfer protein activity. *Journal of internal medicine*, 271(5), pp.490–498.
- Ehrlund, A. et al., 2016. Transcriptional Dynamics During Human Adipogenesis and its Link to Adipose Morphology and Distribution. *Diabetes*, p.db160631.

- Ellulu, M.S. et al., 2017. Obesity and inflammation: the linking mechanism and the complications. *Archives of medical science : AMS*, 13(4), pp.851–863.
- Epstein, E. et al., 1986. The clinical use of alkaline phosphatase enzymes. *Clinics in laboratory medicine*, 6(3), pp.491–505.
- Esmaili, S., Hemmati, M. & Karamian, M., 2020. Physiological role of adiponectin in different tissues: a review. *Archives of physiology and biochemistry*, 126(1), pp.67–73.
- Etokebe, G.E. et al., 2009. Family-with-sequence-similarity-46, member A (Fam46a) gene is expressed in developing tooth buds. *Archives of oral biology*, 54(11), pp.1002–1007.
- Farinatti, P., Castinheiras Neto, A.G. & Amorim, P.R.S., 2016. Oxygen Consumption and Substrate Utilization During and After Resistance Exercises Performed with Different Muscle Mass. *International journal of exercise science*, 9(1), pp.77–88.
- Fasshauer, M. & Blüher, M., 2015. Adipokines in health and disease. *Trends in pharmacological sciences*, 36(7), pp.461–470.
- Fathzadeh, M. et al., 2020. FAM13A affects body fat distribution and adipocyte function. *Nature communications*, 11(1), pp.1465–13.
- Fenzl, A. et al., 2020. Intact vitamin A transport is critical for cold-mediated adipose tissue browning and thermogenesis. *Molecular metabolism*, 42, p.101088.
- Ferraro, Z.M., Patterson, S. & Chaput, J.-P., 2015. Unhealthy weight control practices: culprits and clinical recommendations. *Clinical medicine insights. Endocrinology and diabetes*, 8(8), pp.7–11.
- Frayling, T.M. et al., 2007. A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science (New York, N. Y.)*, 316(5826), pp.889–894.
- Fuchs, T. et al., 2018. Animal models in metabolic syndrome. *Revista do Colegio Brasileiro de Cirurgioes*, 45(5), p.e1975.
- Galarraga, M. et al., 2012. Adiposoft: automated software for the analysis of white adipose tissue cellularity in histological sections. *Journal of lipid research*, 53(12), pp.2791–2796.
- Gaspar, R.C. et al., 2021. An update on brown adipose tissue biology: a discussion of recent findings. *American journal of physiology. Endocrinology and metabolism*, 320(3), pp.E488–E495.
- Ge, C. et al., 2016. Reciprocal Control of Osteogenic and Adipogenic Differentiation by ERK/MAP Kinase Phosphorylation of Runx2 and PPAR γ Transcription Factors. *Journal of cellular physiology*, 231(3), pp.587–596.
- Geeven, G. et al., 2018. peakC: a flexible, non-parametric peak calling package for 4C and Capture-C data. *Nucleic Acids Research*, 46(15), pp.e91–e91.
- Ghatta, S. & Ramarao, P., 2004. Increased contractile responses to 5-Hydroxytryptamine and Angiotensin II in high fat diet fed rat thoracic aorta. *Lipids in health and disease*, 3(1), pp.19–6.
- Goretti Penido, M. & Alon, U.S., 2012. Phosphate homeostasis and its role in bone health. *Pediatric nephrology (Berlin, Germany)*, 27(11), pp.2039–2048.
- Grafe, I. et al., 2018. TGF- β Family Signaling in Mesenchymal Differentiation. *Cold Spring Harbor perspectives in biology*, 10(5), p.a022202.

- Green, H. & Meuth, M., 1974. An established pre-adipose cell line and its differentiation in culture. *Cell*, 3(2), pp.127–133.
- Greenspan, P., Mayer, E.P. & Fowler, S.D., 1985. Nile red: a selective fluorescent stain for intracellular lipid droplets. *The Journal of Cell Biology*, 100(3), pp.965–973.
- Grigoraş, A. et al., 2018. Adipocytes spectrum - From homeostasia to obesity and its associated pathology. *Annals of anatomy = Anatomischer Anzeiger : official organ of the Anatomische Gesellschaft*, 219, pp.102–120.
- Grob, S. & Cavalli, G., 2018. Technical Review: A Hitchhiker's Guide to Chromosome Conformation Capture. *Methods in molecular biology (Clifton, N.J.)*, 1675(Chapter 14), pp.233–246.
- Grundy, S.M., 2015. Adipose tissue and metabolic syndrome: too much, too little or neither. *European journal of clinical investigation*, 45(11), pp.1209–1217.
- Gupta, R.K. et al., 2010. Transcriptional control of preadipocyte determination by Zfp423. *Nature*, 464(7288), pp.619–623.
- Haczeyni, F., Bell-Anderson, K.S. & Farrell, G.C., 2018. Causes and mechanisms of adipocyte enlargement and adipose expansion. *Obesity reviews : an official journal of the International Association for the Study of Obesity*, 19(3), pp.406–420.
- Harvey-Berino, J., 1999. Calorie restriction is more effective for obesity treatment than dietary fat restriction. *Annals of behavioral medicine : a publication of the Society of Behavioral Medicine*, 21(1), pp.35–39.
- Haslam, D.W. & James, W.P.T., 2005. Obesity. *Lancet (London, England)*, 366(9492), pp.1197–1209.
- Hauner, H., Wabitsch, M. & Pfeiffer, E.F., 1988. Differentiation of adipocyte precursor cells from obese and nonobese adult women and from different adipose tissue sites. *Hormone and metabolic research. Supplement series*, 19, pp.35–39.
- Heinz, S. et al., 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Molecular cell*, 38(4), pp.576–589.
- Hensrud, D.D., 2001. Dietary treatment and long-term weight loss and maintenance in type 2 diabetes. *Obesity research*, 9 Suppl 4, pp.348S–353S.
- Herman, M.A. et al., 2010. Adipose tissue branched chain amino acid (BCAA) metabolism modulates circulating BCAA levels. *The Journal of biological chemistry*, 285(15), pp.11348–11356.
- Hill, C.S., 2009. Nucleocytoplasmic shuttling of Smad proteins. *Cell research*, 19(1), pp.36–46.
- Hirsch, J. & Knittle, J.L., 1970. Cellularity of obese and nonobese human adipose tissue. *Federation proceedings*, 29(4), pp.1516–1521.
- Hollenberg, C.H. & Vost, A., 1969. Regulation of DNA synthesis in fat cells and stromal elements from rat adipose tissue. *The Journal of clinical investigation*, 47(11), pp.2485–2498.
- Hwang, I. & Kim, J.B., 2019. Two Faces of White Adipose Tissue with Heterogeneous Adipogenic Progenitors. *Diabetes & metabolism journal*, 43(6), pp.752–762.

- Ijichi, N. et al., 2007. Estrogen-related receptor alpha modulates the expression of adipogenesis-related genes during adipocyte differentiation. *Biochemical and biophysical research communications*, 358(3), pp.813–818.
- Ikeda, K., Maretich, P. & Kajimura, S., 2018. The Common and Distinct Features of Brown and Beige Adipocytes. *Trends in endocrinology and metabolism: TEM*, 29(3), pp.191–200.
- Irving, M.D. et al., 2008. Chondrodysplasia punctata: a clinical diagnostic and radiological review. *Clinical dysmorphology*, 17(4), pp.229–241.
- Kaiser, V.B. & Semple, C.A., 2017. When TADs go bad: chromatin structure and nuclear organisation in human disease. *F1000Research*, 6, p.314.
- Kent, W.J. et al., 2002. The human genome browser at UCSC. *Genome research*, 12(6), pp.996–1006.
- Khan, T. et al., 2009. Metabolic dysregulation and adipose tissue fibrosis: role of collagen VI. *Molecular and cellular biology*, 29(6), pp.1575–1591.
- Kirschning, C.J. et al., 1997. Similar organization of the lipopolysaccharide-binding protein (LBP) and phospholipid transfer protein (PLTP) genes suggests a common gene family of lipid-binding proteins. *Genomics*, 46(3), pp.416–425.
- Komori, T., 2019. Regulation of Proliferation, Differentiation and Functions of Osteoblasts by Runx2. *International journal of molecular sciences*, 20(7), p.1694.
- Krakov, D. & Rimoin, D.L., 2010. The skeletal dysplasias. *Genetics in medicine : official journal of the American College of Medical Genetics*, 12(6), pp.327–341.
- Krijger, P.H.L. & de Laat, W., 2016. Regulation of disease-associated gene expression in the 3D genome. *Nature Reviews Molecular Cell Biology*, 17(12), pp.771–782.
- Krueger, F., 2012. Trim Galore. Available at: https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/.
- Kuchta, K. et al., 2016. FAM46 proteins are novel eukaryotic non-canonical poly(A) polymerases. *Nucleic Acids Research*, 44(8), pp.3534–3548.
- Lagali, P.S. et al., 2002. Identification and characterization of C6orf37, a novel candidate human retinal disease gene on chromosome 6q14. *Biochemical and biophysical research communications*, 293(1), pp.356–365.
- Langmead, B. & Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. *Nature methods*, 9(4), pp.357–359.
- Larsen, T.M. et al., 2010. The Diet, Obesity and Genes (Diogenes) Dietary Study in eight European countries - a comprehensive design for long-term intervention. *Obesity reviews : an official journal of the International Association for the Study of Obesity*, 11(1), pp.76–91.
- Lass, A. et al., 2011. Lipolysis - a highly regulated multi-enzyme complex mediates the catabolism of cellular fat stores. *Progress in lipid research*, 50(1), pp.14–27.
- Lemos, D.R. et al., 2012. Functionally convergent white adipogenic progenitors of different lineages participate in a diffused system supporting tissue regeneration. *STEM CELLS*, 30(6), pp.1152–1162.
- Li, S.-N. & Wu, J.-F., 2020. TGF- β /SMAD signaling regulation of mesenchymal stem cells in adipocyte commitment. *Stem cell research & therapy*, 11(1), pp.41–10.

- Lieberman-Aiden, E. et al., 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science (New York, N.Y.)*, 326(5950), pp.289–293.
- Liu, S. et al., 2018. Effect of Long-Term Exercise Training on lncRNAs Expression in the Vascular Injury of Insulin Resistance. *Journal of cardiovascular translational research*, 11(6), pp.459–469.
- Liu, T.M. et al., 2007. Identification of common pathways mediating differentiation of bone marrow- and adipose tissue-derived human mesenchymal stem cells into three mesenchymal lineages. *STEM CELLS*, 25(3), pp.750–760.
- Longo, M. et al., 2019. Adipose Tissue Dysfunction as Determinant of Obesity-Associated Metabolic Complications. *International journal of molecular sciences*, 20(9), p.2358.
- Loos, R.J.F. & Kilpeläinen, T.O., 2018. Genes that make you fat, but keep you healthy. *Journal of internal medicine*, 284(5), pp.450–463.
- Lowe, C.E., O'Rahilly, S. & Rochford, J.J., 2011. Adipogenesis at a glance. *Journal of cell science*, 124(Pt 16), pp.2681–2686.
- Lutz, T.A., 2020. An Overview of Rodent Models of Obesity and Type 2 Diabetes. *Methods in molecular biology (Clifton, N.J.)*, 2128(2948), pp.11–24.
- Magoč, T. & Salzberg, S.L., 2011. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics (Oxford, England)*, 27(21), pp.2957–2963.
- Mariani, L. et al., 2017. Identification of Human Lineage-Specific Transcriptional Coregulators Enabled by a Glossary of Binding Modules and Tunable Genomic Backgrounds. *Cell systems*, 5(6), p.654.
- Marks, J., Debnam, E.S. & Unwin, R.J., 2010. Phosphate homeostasis and the renal-gastrointestinal axis. *American journal of physiology. Renal physiology*, 299(2), pp.F285–96.
- Marvyn, P.M. et al., 2016. Data on oxygen consumption rate, respiratory exchange ratio, and movement in C57BL/6J female mice on the third day of consuming a high-fat diet. *Data in brief*, 7, pp.472–475.
- Massagué, J. & Wotton, D., 2000. Transcriptional control by the TGF-beta/Smad signaling system. *The EMBO journal*, 19(8), pp.1745–1754.
- McCord, R.P., Kaplan, N. & Giorgetti, L., 2020. Chromosome Conformation Capture and Beyond: Toward an Integrative View of Chromosome Structure and Function. *Molecular cell*, 77(4), pp.688–708.
- McGowan, S.J. et al., 2013. MIG: Multi-Image Genome viewer. *Bioinformatics (Oxford, England)*, 29(19), pp.2477–2478.
- Mehlem, A. et al., 2013. Imaging of neutral lipids by oil red O for analyzing the metabolic status in health and disease. *Nature protocols*, 8(6), pp.1149–1154.
- Mi, H. et al., 2019. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Research*, 47(D1), pp.D419–D426.
- Moreno-Navarrete, J.M. et al., 2013. A role for adipocyte-derived lipopolysaccharide-binding protein in inflammation- and obesity-associated adipose tissue dysfunction. *Diabetologia*, 56(11), pp.2524–2537.

- Mroczek, S. et al., 2017. The non-canonical poly(A) polymerase FAM46C acts as an onco-suppressor in multiple myeloma. *Nature communications*, 8(1), pp.619–17.
- Nakachi, Y. et al., 2008. Identification of novel PPARgamma target genes by integrated analysis of ChIP-on-chip and microarray expression data during adipocyte differentiation. *Biochemical and biophysical research communications*, 372(2), pp.362–366.
- Naumova, N. et al., 2012. Analysis of long-range chromatin interactions using Chromosome Conformation Capture. *Methods (San Diego, Calif.)*, 58(3), pp.192–203.
- Neri, G. et al., 2013. *Simpson-Golabi-Behmel syndrome: an X-linked encephalo-trophoschisis syndrome*. 1988, John Wiley & Sons, Ltd.
- Newgard, C.B. et al., 2009. A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance. *Cell metabolism*, 9(4), pp.311–326.
- Niesler, C.U., Siddle, K. & Prins, J.B., 1998. Human preadipocytes display a depot-specific susceptibility to apoptosis. *Diabetes*, 47(8), pp.1365–1368.
- Nieto, M.A. et al., 2016. EMT: 2016. *Cell*, 166(1), pp.21–45.
- Nishikawa, S. et al., 2007. Involvement of sex, strain and age factors in high fat diet-induced obesity in C57BL/6J and BALB/cA mice. *Experimental animals*, 56(4), pp.263–272.
- Oikonomopoulou, K. et al., 2012. Interactions between coagulation and complement—their role in inflammation. *Seminars in immunopathology*, 34(1), pp.151–165.
- Oishi, Y. et al., 2005. Krüppel-like transcription factor KLF5 is a key regulator of adipocyte differentiation. *Cell metabolism*, 1(1), pp.27–39.
- Otto, G.P. et al., 2016. Clinical Chemistry Reference Intervals for C57BL/6J, C57BL/6N, and C3HeB/FeJ Mice (*Mus musculus*). *Journal of the American Association for Laboratory Animal Science : JAALAS*, 55(4), pp.375–386.
- Packard, G.C. & Boardman, T.J., 1999. The use of percentages and size-specific indices to normalize physiological data for variation in body size: wasted time, wasted effort? *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, 122(1), pp.37–44. Available at: <https://www.sciencedirect.com/science/article/pii/S1095643398101708>.
- Palou, A. & Bonet, M.L., 2013. Challenges in obesity research. *Nutricion hospitalaria*, 28 Suppl 5, pp.144–153.
- Paulo, E. & Wang, B., 2019. Towards a Better Understanding of Beige Adipocyte Plasticity. *Cells*, 8(12), p.1552.
- Pessentheiner, A.R., Ducasa, G.M. & Gordts, P.L.S.M., 2020. Proteoglycans in Obesity-Associated Metabolic Dysfunction and Meta-Inflammation. *Frontiers in immunology*, 11, p.769.
- Petrus, P. et al., 2018. Transforming Growth Factor-β3 Regulates Adipocyte Number in Subcutaneous White Adipose Tissue. *Cell reports*, 25(3), pp.551–560.e5.
- Pope, B.D. et al., 2014. Topologically associating domains are stable units of replication-timing regulation. *Nature*, 515(7527), pp.402–405.

- Poznanski, W.J., Waheed, I. & Van, R., 1973. Human fat cell precursors. Morphologic and metabolic differentiation in culture. *Laboratory investigation; a journal of technical methods and pathology*, 29(5), pp.570–576.
- Quach, J.M. et al., 2011. Zinc finger protein 467 is a novel regulator of osteoblast and adipocyte commitment. *The Journal of biological chemistry*, 286(6), pp.4186–4198.
- Quadro, L. et al., 2003. Understanding the physiological role of retinol-binding protein in vitamin A metabolism using transgenic and knockout mouse models. *Molecular aspects of medicine*, 24(6), pp.421–430.
- Richard, D. & Picard, F., 2011. Brown fat biology and thermogenesis. *Frontiers in bioscience (Landmark edition)*, 16(1), pp.1233–1260.
- Roncari, D.A., Lau, D.C. & Kindler, S., 1981. Exaggerated replication in culture of adipocyte precursors from massively obese persons. *Metabolism: clinical and experimental*, 30(5), pp.425–427.
- Rondinone, C.M., 2006. Adipocyte-derived hormones, cytokines, and mediators. *Endocrine*, 29(1), pp.81–90.
- Rosen, E.D. & MacDougald, O.A., 2006. Adipocyte differentiation from the inside out. *Nature Reviews Molecular Cell Biology*, 7(12), pp.885–896.
- Rosenbaum, M. et al., 1997. Effects of weight change on plasma leptin concentrations and energy expenditure. *The Journal of clinical endocrinology and metabolism*, 82(11), pp.3647–3654.
- Ross, S.E. et al., 2000. Inhibition of adipogenesis by Wnt signaling. *Science (New York, N.Y.)*, 289(5481), pp.950–953.
- Ruiz-Ojeda, F.J. et al., 2019. Extracellular Matrix Remodeling of Adipose Tissue in Obesity and Metabolic Diseases. *International journal of molecular sciences*, 20(19), p.4888.
- Rumin, J. et al., 2015. The use of fluorescent Nile red and BODIPY for lipid measurement in microalgae. *Biotechnology for biofuels*, 8(1), pp.42–16.
- Sadie-Van Gijsen, H., 2019. Adipocyte biology: It is time to upgrade to a new model. *Journal of cellular physiology*, 234(3), pp.2399–2425.
- Saely, C.H., Geiger, K. & Drexel, H., 2012. Brown versus white adipose tissue: a mini-review. *Gerontology*, 58(1), pp.15–23.
- Sanchez-Gurmaches, J. & Guertin, D.A., 2014. Adipocyte lineages: tracing back the origins of fat. *Biochimica et biophysica acta*, 1842(3), pp.340–351.
- Santos, I. et al., 2017. Prevalence of personal weight control attempts in adults: a systematic review and meta-analysis. *Obesity reviews : an official journal of the International Association for the Study of Obesity*, 18(1), pp.32–50.
- Saraç, F. & Saygılı, F., 2014. Causes of High Bone Alkaline Phosphatase. *Biotechnology & Biotechnological Equipment*, 21(2), pp.194–197. Available at: <https://www.tandfonline.com/doi/abs/10.1080/13102818.2007.10817444>.
- Scherer, P.E. et al., 1995. A novel serum protein similar to C1q, produced exclusively in adipocytes. *The Journal of biological chemistry*, 270(45), pp.26746–26749.
- Schmitt, A.D. et al., 2016. A Compendium of Chromatin Contact Maps Reveals Spatially Active Regions in the Human Genome. *Cell reports*, 17(8), pp.2042–2059.

- Schoettl, T., Fischer, I.P. & Ussar, S., 2018. Heterogeneity of adipose tissue in development and metabolic function. *The Journal of experimental biology*, 221(Pt Suppl 1).
- Schwartz, M.W. et al., 2017. Obesity Pathogenesis: An Endocrine Society Scientific Statement. *Endocrine reviews*, 38(4), pp.267–296.
- Scott, M.A. et al., 2011. Current methods of adipogenic differentiation of mesenchymal stem cells. *Stem cells and development*, 20(10), pp.1793–1804.
- Sharma, U., Pal, D. & Prasad, R., 2014. Alkaline phosphatase: an overview. *Indian journal of clinical biochemistry : IJCB*, 29(3), pp.269–278.
- Shen, Y. et al., 2012. A map of the cis-regulatory sequences in the mouse genome. *Nature*, 488(7409), pp.116–120.
- Shungin, D. et al., 2015. New genetic loci link adipose and insulin biology to body fat distribution. *Nature*, 518(7538), pp.187–196.
- Siersbæk, R. & Mandrup, S., 2011. Transcriptional networks controlling adipocyte differentiation. *Cold Spring Harbor symposia on quantitative biology*, 76(0), pp.247–255.
- Smemo, S. et al., 2014. Obesity-associated variants within FTO form long-range functional connections with IRX3. *Nature*, 507(7492), pp.371–375.
- Smith, J.L., 1908. On the simultaneous staining of neutral fat and fatty acid by oxazine dyes. *The Journal of Pathology and Bacteriology*, 12(1), pp.1–4.
- Soler-Oliva, M.E. et al., 2017. Analysis of the relationship between coexpression domains and chromatin 3D organization. *PLoS computational biology*, 13(9), p.e1005708.
- Soubere Mahamoud, Y. et al., 2016. Additive Effects of Millimeter Waves and 2-Deoxyglucose Co-Exposure on the Human Keratinocyte Transcriptome. A. T. Slominski, ed. *PLOS ONE*, 11(8), p.e0160810.
- Spalding, K.L. et al., 2008. Dynamics of fat cell turnover in humans. *Nature*, 453(7196), pp.783–787.
- Spiegelman, B.M. & Flier, J.S., 2001. Obesity and the regulation of energy balance. *Cell*, 104(4), pp.531–543.
- Spitz, F. & Furlong, E.E.M., 2012. Transcription factors: from enhancer binding to developmental control. *Nature reviews. Genetics*, 13(9), pp.613–626.
- Stepanenko, A.A. & Heng, H.H., 2017. Transient and stable vector transfection: Pitfalls, off-target effects, artifacts. *Mutation research*, 773, pp.91–103.
- Stottmann, R. & Beier, D., 2014. ENU Mutagenesis in the Mouse. *Current protocols in human genetics*, 82(1), pp.15.4.1–10.
- Tabula Muris Consortium et al., 2018. Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. *Nature*, 562(7727), pp.367–372.
- Tan, C.K. et al., 2012. Getting “Smad” about obesity and diabetes. *Nutrition & diabetes*, 2(3), pp.e29–e29.
- Tan, C.K. et al., 2011. Smad3 deficiency in mice protects against insulin resistance and obesity induced by a high-fat diet. *Diabetes*, 60(2), pp.464–476.

- Teitelbaum, S.L., 2000. Bone resorption by osteoclasts. *Science (New York, N.Y.)*, 289(5484), pp.1504–1508.
- Tolhuis, B. et al., 2002. Looping and interaction between hypersensitive sites in the active beta-globin locus. *Molecular cell*, 10(6), pp.1453–1465.
- Tong, Q. et al., 2000. Function of GATA transcription factors in preadipocyte-adipocyte transition. *Science (New York, N.Y.)*, 290(5489), pp.134–138.
- Tsompana, M. & Buck, M.J., 2014. Chromatin accessibility: a window into the genome. *Epigenetics & chromatin*, 7(1), pp.33–16.
- van Baak, M.A. & Mariman, E.C.M., 2019. Mechanisms of weight regain after weight loss - the role of adipose tissue. *Nature reviews. Endocrinology*, 15(5), pp.274–287.
- Wabitsch, M. et al., 2001. Characterization of a human preadipocyte cell strain with high capacity for adipose differentiation. *International journal of obesity and related metabolic disorders : journal of the International Association for the Study of Obesity*, 25(1), pp.8–15.
- Wang, E.A. et al., 1993. Bone morphogenetic protein-2 causes commitment and differentiation in C3H10T1/2 and 3T3 cells. *Growth factors (Chur, Switzerland)*, 9(1), pp.57–71.
- Wang, Y. & Sul, H.S., 2009. Pref-1 regulates mesenchymal cell commitment and differentiation through Sox9. *Cell metabolism*, 9(3), pp.287–302.
- Wang, Y. et al., 2018. The 3D Genome Browser: a web-based browser for visualizing 3D genome organization and long-range chromatin interactions. *Genome biology*, 19(1), pp.151–12.
- Wang, Y.-X. et al., 2003. Peroxisome-proliferator-activated receptor delta activates fat metabolism to prevent obesity. *Cell*, 113(2), pp.159–170.
- Warkocki, Z. et al., 2018. Terminal nucleotidyl transferases (TENTs) in mammalian RNA metabolism. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 373(1762), p.20180162.
- Wasserman, W.W. & Sandelin, A., 2004. Applied bioinformatics for the identification of regulatory elements. *Nature reviews. Genetics*, 5(4), pp.276–287.
- Watanabe, T. et al., 2018. Fam46a regulates BMP-dependent pre-placodal ectoderm differentiation in *Xenopus*. *Development (Cambridge, England)*, 145(20), p.dev166710.
- WHO, 2009. Global health risks: mortality and burden of disease attributable to selected major risks. *WHO Press*. Available at: https://www.who.int/healthinfo/global_burden_disease/GlobalHealthRisks_report_full.pdf?ua=1&ua=1.
- WHO, 2020. Obesity and overweight. *World Health Organization*. Available at: <https://www.who.int/en/news-room/fact-sheets/detail/obesity-and-overweight> [Accessed April 1, 2021].
- Yamauchi, T. et al., 2002. Adiponectin stimulates glucose utilization and fatty-acid oxidation by activating AMP-activated protein kinase. *Nature medicine*, 8(11), pp.1288–1295.
- Yamauchi, T. et al., 2003. Cloning of adiponectin receptors that mediate antidiabetic metabolic effects. *Nature*, 423(6941), pp.762–769.

- Yamauchi, T. et al., 2007. Targeted disruption of AdipoR1 and AdipoR2 causes abrogation of adiponectin binding and metabolic actions. *Nature medicine*, 13(3), pp.332–339.
- Young, P., Arch, J.R. & Ashwell, M., 1984. Brown adipose tissue in the parametrial fat pad of the mouse. *FEBS letters*, 167(1), pp.10–14.
- Zhang, W. et al., 2018. The TEA domain family transcription factor TEAD4 represses murine adipogenesis by recruiting the cofactors VGLL4 and CtBP2 into a transcriptional complex. *The Journal of biological chemistry*, 293(44), pp.17119–17134.
- Zhang, Y. et al., 1994. Positional cloning of the mouse obese gene and its human homologue. *Nature*, 372(6505), pp.425–432.
- Zhang, Y.E., 2018. Mechanistic insight into contextual TGF- β signaling. *Current opinion in cell biology*, 51, pp.1–7.
- Zhao, C. et al., 2018. Knockdown of ubiquitin D inhibits adipogenesis during the differentiation of porcine intramuscular and subcutaneous preadipocytes. *Cell proliferation*, 51(2), p.e12401.
- Zhao, L. et al., 2009. Comparison of multipotent differentiation potentials of murine primary bone marrow stromal cells and mesenchymal stem cell line C3H10T1/2. *Calcified tissue international*, 84(1), pp.56–64.
- Zhou, H.Y. et al., 2014. A Sox2 distal enhancer cluster regulates embryonic stem cell differentiation potential. *Genes & Development*, 28(24), pp.2699–2711.
- Zhou, Y. et al., 2019. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nature communications*, 10(1), pp.1523–10.
- Zizola, C.F., Schwartz, G.J. & Vogel, S., 2008. Cellular retinol-binding protein type III is a PPARgamma target gene and plays a role in lipid metabolism. *American journal of physiology. Endocrinology and metabolism*, 295(6), pp.E1358–68.



CURRICULUM VITAE

ANN-KRISTIN HOV

MOLECULAR BIOLOGIST

Ruelle du Lapin-Vert 2 - 1005 Lausanne, Switzerland

+41 79 307 06 66

annkristin.hov@gmail.com

Norwegian // Permit B

WORK & RESEARCH EXPERIENCE

2016 - 2020

DOCTORAL ASSISTANT, EPFL, LAUSANNE, SWITZERLAND

2017 - 2019

TEACHING ASSISTANT (300+ HOURS), EPFL, LAUSANNE, SWITZERLAND

M.Sc. course "Biotechnology laboratory" & B.Sc. course "Integrated laboratory in Life Sciences I"

JAN. - AUG. 2016

MASTER THESIS PROJECT, UNIVERSITY OF SUSSEX, BRIGHTON, UK

2013 - 2015

VISITING SCIENTIST, EMBL, HEIDELBERG, GERMANY

FEB. - MAY 2013

BACHELOR THESIS PROJECT, NEWCASTLE UNIVERSITY, NEWCASTLE, UK

EDUCATION

2016 - PRESENT

PHD MOLECULAR LIFE SCIENCES

EPFL, LAUSANNE, SWITZERLAND

Thesis: "The role and transcriptional regulation of FAM46A in adipocyte differentiation"

2015 - 2016

MSC GENETIC MANIPULATION AND MOLECULAR CELL BIOLOGY

UNIVERSITY OF SUSSEX, BRIGHTON, UK

Thesis: "Understanding Asymmetric Division and the Potential Role of Greatwall Kinase in Glioblastoma Multiforme Stem Cells"

2011 - 2013

BSC (HONS) BIOMEDICAL SCIENCES

NEWCASTLE UNIVERSITY, NEWCASTLE, UK

Thesis: "Detecting Interactions between Escherichia coli Divisome Proteins Using a Chemical Cross-Linking Method"

SKILLS & COMPETENCES

Languages

English (Native), French (Excellent), Norwegian (Native)

<u>Software</u>	MS Office, GraphPad Prism, Adobe Illustrator, Affinity Designer, R, ImageJ/Fiji, QuPath, Unix, SLIMS (ELN)
<u>Genetics</u>	SerialCloner, SnapGene, ApE, IGV, UCSC Genome Browser, IPA, NCBI Resources, Clustal Omega, HOMER, fastQC, FLASH, bowtie

LABORATORY TECHNIQUES

<u>Cell culture</u>	P1/P2 bacterial and mammalian culture, transfection (electroporation, lipofectamine, PEI), primary cell isolation (APCs, BM-MSCs)
<u>Molecular</u>	DNA/RNA extraction, SDS-PAGE, Western blotting, PCR, qPCR, protein purification, ELISA, CRISPR-Cas9 gene editing, gene cloning (Gibson, TA/TOPO), plasmid purification, light/fluorescence microscopy, gene knockdown (shRNA), gene overexpression
<u>Genomics</u>	Chromosome conformation capture (4C-seq, CaptureC), preparation of sequencing libraries for Illumina sequencing

PUBLICATIONS

Banzhaf M, Yau HC, Verheul J, Lodge A, Kritikos G, Mateus A, Cordier B, **Hov AK**, Stein F, Wartel M, Pazos M, Solovyova AS, Breukink E, van Teeffelen S, Savitski MM, den Blaauwen T, Typas A, Vollmer W. (2020)

Outer membrane lipoprotein Nlpl scaffolds peptidoglycan hydrolases within multi-enzyme complexes in Escherichia coli. EMBO J. 39(5):e102246.

Cho SH, Szewczyk J, Pesavento C, Zietek M, Banzhaf M, Roszczenko P, Asmar A, Laloux G, **Hov AK**, Leverrier P, Van der Henst C, Vertommen D, Typas A, Collet JF. (2014) Detecting Envelope Stress by Monitoring β -Barrel Assembly. Cell 159(7):1652-1664.

EXTRACURRICULARS AND INTERESTS

2019 - PRESENT

PRESIDENT AND IMPROVISATIONAL COACH FOR THE CATALYST

LAUSANNE, SWITZERLAND

The Catalyst is a science entertainment collective, passionate about communicating science and inspiring scientists to become better communicators. The organisation provides weekly improvisational classes available to scientists and other science-enthusiasts. As a coach for these weekly sessions at EPFL and CERN, my duties included preparing lesson plans with particular themes, accommodating students with varying levels of experience and creating a safe and fun environment for all.

Interests Cooking, theatre, hiking, scuba diving (PADI diver lvl 2), travel