LUND UNIVERSITY

Functional genomics approaches to dissect causal DNA variations in obesity and type 2 diabetes

Huang, Mi

2023

Document Version:
Publisher's PDF, also known as Version of record

Link to publication

Citation for published version (APA):
Huang, M. (2023). *Functional genomics approaches to dissect causal DNA variations in obesity and type 2 diabetes*. [Doctoral Thesis (compilation), Department of Clinical Sciences, Malmö]. Lund University, Faculty of Medicine.

Total number of authors:
1

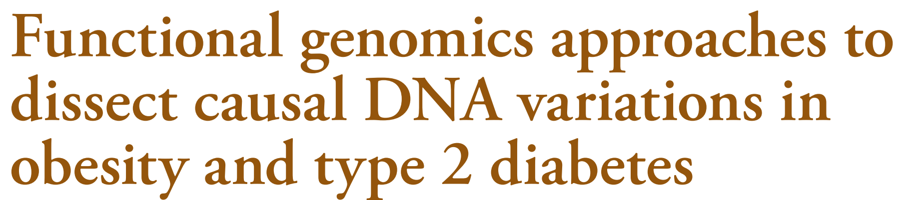
General rights
Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00



Functional genomics approaches to dissect causal DNA variations in obesity and type 2 diabetes

MI HUANG
DEPARTMENT OF CLINICAL SCIENCES | FACULTY OF MEDICINE | LUND UNIVERSITY

Functional genomics approaches to dissect causal DNA variations in obesity and type 2 diabetes

Mi Huang
黄 密

DOCTORAL DISSERTATION

By due permission of the Faculty of Medicine, Lund University, Sweden.
To be defended at Aulan, Clinical Research Centre, Jan Waldenströms gata 35

SE-21428 Malmö, Sweden

09:00, 19 June 2023

Faculty opponent
Professor Mikael Rydén, Karolinska Institutet, Stockholm

Organization LUND UNIVERSITY **Author** Mi Huang (黄密)	**Document name** DOCTORAL DISSERTATION **Date of issue** 19 June 2023 Sponsoring organization

Title and subtitle
Functional genomics approaches to dissect causal DNA variations in obesity and type 2 diabetes

Abstract

GWAS studies have identified hundreds of loci that associate with obesity and type 2 diabetes. However, the functional relevance of the genetic variants, and their impact on disease or traits, is largely unknown. To address those questions, we applied functional genomics approaches to study genetic variants pertinent to obesity and type 2 diabetes-related traits.

In paper I, we demonstrate that rs67785913 is a causal cis-eQTL for *MTIF3* expression. We also show that MTIF3 regulates mitochondrial respiration and endogenous fatty acid oxidation in adipocytes, probably as a consequence of MTIF3 affecting mitochondrial oxidative phosphorylation complex assembly. MTIF3 deficiency also leads to higher retention of triglycerides after glucose restriction challenges.

In paper II, we show that rs8192678 confers allele-specific causal effects on white adipocyte differentiation. The T allele presents a dose-dependent effect on increased lipogenesis and mitochondrial function, and on the expression of genes involved in adipogenesis, lipid catabolism, lipogenesis, and lipolysis. The T allele also confers higher levels of *PPARGC1A* mRNA and PGC-1α protein, as well as faster turnover and higher activity of PGC-1α.

In paper III, we experimentally document that rs10071329 is a cis-eQTL for *PPARGC1B* expression, and influences brown adipocyte lipolysis and mitochondrial function. In a Tanzanian cohort lookup, rs10071329 shows a trend of association for BMI, and an association with mid-upper arm circumference.

In paper IV, we show that allele substitution at rs2289669 and rs8065082 does not affect *SLC47A1* expression or splicing in HepG2 cells. Moreover, among MetGen cohort participants, we observed that no common genetic variants in *SLC47A1* were associated with metformin-mediated efficacy on glycemia reduction.

The findings presented in this thesis provide biological mechanistic insights underlying epidemiological observations and may aid in developing precision medicine for obesity and type 2 diabetes.

Key words: GWAS, obesity, type 2 diabetes, CRISPR/Cas9, precision medicine, adipocytes, functional genomics

Classification system and/or index terms (if any)		
Supplementary bibliographical information		**Language** English
ISSN and key title **1652-8220** Lund University, Faculty of Medicine Doctoral Dissertation Series 2023:88		**ISBN** **978-91-8021-428-5**
Recipient's notes	**Number of pages 158**	Price
	Security classification	

I, the undersigned, being the copyright owner of the abstract of the above-mentioned dissertation, hereby grant to all reference sources permission to publish and disseminate the abstract of the above-mentioned dissertation.

Signature Date 2023-05-03

Functional genomics approaches to dissect causal DNA variations in obesity and type 2 diabetes

Mi Huang
黄 密

LUND
UNIVERSITY

Cover photo by Mi Huang (created with Midjourney).

Copyright pp 1-74 (Mi Huang)

Paper 1 © by the Authors (Open Access)

Paper 2 © by the Authors (Open Access)

Paper 3 © by the Authors (Manuscript unpublished)

Paper 4 © by the Authors (Open Access)

Faculty of Medicine
Department of Clinical Sciences, Malmö
ISBN 978-91-8021-428-5
ISSN 1652-8220

Printed in Sweden by Media-Tryck, Lund University
Lund 2023

Media-Tryck is a Nordic Swan Ecolabel
certified provider of printed material.
Read more about our environmental
work at www.mediatryck.lu.se

To my family and Jane

不忘初心，方得始终

Table of Contents

Papers included in this thesis **9**

Papers not included in this thesis **10**

Abbreviations **11**

Popular Science Summary **13**

Populärvetenskaplig Sammanfattning **15**

科普总结 **18**

Acknowledgments **20**

Chapter 1 - Introduction **24**

- Obesity 24
- Type 2 diabetes 25
- The role of adipose cells in obesity and T2DM 26
- GWAS in obesity and T2DM 27
- From association to causality: the challenges 28
- From association to causality: the experimental approaches 29
- rs8192678 and *PPARGC1A* 31
- rs1885988 and *MTIF3* 31
- rs10071329 and *PPARGC1B* 32
- rs2289669 and *SLC47A1* 33

Chapter 2 - Methods **34**

- Cell lines 34
- Luciferase reporter assays for causal variant prioritization 34
- sgRNA and HDR donor template design for CRISPR/Cas9-mediated allele editing 36
- Limiting dilution cloning for clonal screening 36
- RT-qPCR 37
- Mitochondrial respiration profiling using Seahorse analyzer 38
- SDS-PAGE and Western Blotting 40

Blue Native PAGE and immunoblotting 41
Oil red O staining 41
Chapter 3 - Aims of the Thesis 42
Paper I 42
Paper II 43
Paper III 43
Paper IV 44
Chapter 4 - Results and Discussions 45
Paper I 45
Paper II 49
Paper III 53
Paper IV 56
Overall summary 59
Future perspectives 61
References 63

Papers included in this thesis

I. **Mi Huang**, Daniel Coral, Hamidreza Ardalani, Peter Spégel, Alham Saadat, Melina Claussnitzer, Hindrik Mulder, Paul W Franks, Sebastian Kalamajski. (2023) Identification of a weight loss-associated causal eQTL in *MTIF3* and the effects of MTIF3 deficiency on human adipocyte function. *eLife* **12**: e84168 https://doi.org/10.7554/eLife.84168.

II. **Mi Huang**, Melina Claussnitzer, Alham Saadat, Daniel Coral, Sebastian Kalamajski, Paul W Franks. (2023) Engineered allele substitution at *PPARGC1A* rs8192678 alters human white adipocyte differentiation, lipogenesis, and PGC-1α content and turnover. Accepted Manuscript. *Diabetologia.*

III. **Mi Huang**, Rashmi B Prasad, Daniel Coral, Hindrik Mulder, Paul W Franks, Sebastian Kalamajski. Single-allele human genetic variation (rs10071329) correlates with adiposity-related traits, modulates *PPARGC1B* expression, and alters brown adipocyte function. (Manuscript submitted)

IV. Sebastian Kalamajski, **Mi Huang**, Jonathan Dalla-Riva, Maria Keller, Adem Y Dawed, Ola Hansson, Ewan R Pearson, MetGen Plus Consortium, Hindrik Mulder, Paul W Franks. (2022) Genomic editing of metformin efficacy-associated genetic variants in *SLC47A1* does not alter *SLC47A1* expression. *Human Molecular Genetics*, 31, **4**: 491–498, https://doi.org/10.1093/hmg/ddab266.

Papers not included in this thesis

I. Sevda Gheibi, Luis Rodrigo Cataldo, Alexander Hamilton, **Mi Huang**, Sebastian Kalamajski, Malin Fex, Hindrik Mulder. (2023) Reduced Expression Level of Protein Phosphatase PPM1E Serves to Maintain Insulin Secretion in Type 2 Diabetes. *Diabetes*, 72 (**4**): 455–466. https://doi.org/10.2337/db22-0472

II. **Mi Huang**, Daniel Coral, Peter Spégel, Alham Saadat, Melina Claussnitzer, Hindrik Mulder, Paul W Franks, Sebastian Kalamajski. CRISPR editing shows rs8192678 polymorphism differentially regulates human brown adipocytes mitochondrial function under cold exposure through regulating PGC-1α stability. (Manuscript)

Abbreviations

BMI	body mass index
T2DM	type 2 diabetes mellitus
GWAS	genome-wide association studies
WAT	white adipose tissue
BAT	brown adipose tissue
UCP1	uncoupling protein 1
SNP	single nuclear polymorphism
LD	linkage disequilibrium
CRISPR/Cas9	Clustered Regularly Interspaced Palindromic Repeats (CRISPR) and its associated protein 9
crRNA	CRISPR RNA
tracrRNA	trans activating CRISPR RNA
sgRNA	single guide RNA
PAM	protospacer adjacent motif
DSB	DNA double strand break
HDR	homology directed repair
NHEJ	non homologous end joining
MPRAs	massively parallel reporter assays

PGC-1α	peroxisome proliferator activated receptor γ coactivator 1α
MTIF3	mitochondrial translation initiation factor 3 protein
OXPHOS	oxidative phosphorylation
PGC-1β	peroxisome proliferator activated receptor γ coactivator 1β
MATE1	multidrug and toxin extrusion 1
hWAs	human white preadipocyte cell line
hBAs	human brown preadipocyte cell line
hTERT	human telomerase reverse transcriptase
hWAs-iCas9	human white pre-adipocyte inducible cas9 expressing cell line
RNP	ribonucleoprotein
RT-qPCR	reverse transcription quantitative real time PCR
OCR	oxygen consumption rate
ECAR	extracellular acidification rate
FCCP	carbonyl cyanide 4 (trifluoromethoxy) phenylhydrazone
NE	norepinephrine
SDS-PAGE	sodium dodecylsulfate polyacrylamide gel electrophoresis
SDS	sodium dodecylsulfate
PVDF	hydrophilic polyvinylidene fluoride
BN-PAGE	blue native polyacrylamide gel electrophoresis
EMSA	electrophoretic mobility shift assay

Popular Science Summary

Obesity and type 2 diabetes mellitus (T2DM) are increasingly common health issues that affect people all over the world. These conditions are caused by a combination of different factors, including our DNA and the environment we live in. The way our DNA interacts with our lifestyle and surroundings can impact our risk of developing obesity and T2DM. We know that overeating, lack of exercise, and other lifestyle factors can contribute to these conditions, but it is also important to understand how variations in our DNA may cause obesity and T2DM. By study of human DNA mutations, i.e. small differences in DNA found in different individuals, we can gain a better understanding of how these diseases develop. By use of this knowledge, we may eventually develop more effective prevention and treatment strategies, ultimately helping people live healthier lives.

How do scientists identify which DNA mutations contribute to disease? Researchers use a technique called genome-wide association studies (GWAS) to explore which mutations could be involved in causing obesity and T2DM. Through GWAS, scientists have found hundreds of mutations that are connected to either the development or treatment efficacy of obesity and T2DM.

Although we now have associated many mutations with obesity and T2DM, we still rarely know exactly how they work or why they affect some people more than others. Scientists have found that many of the mutations are found in the parts of the DNA that may control how certain genes are turned on (up) or off (down). However, we still do not fully understand which genes are affected or how this regulation works in different organs or different environments. These gaps in our knowledge make it difficult to use the findings from GWAS to develop new treatments for obesity and T2DM.

Scientists are therefore developing new ways to study DNA mutations. One promising approach is called functional genomics, which helps researchers investigate how particular mutations affect cells and living organisms. One of the most promising functional genomics approaches is a technology called CRISPR/Cas9 genomic editing. This cutting-edge technique allows researchers to make precise changes to the genome of a living cell and study the effects of those changes.

How does the CRISPR/Cas9 machinery work? CRISPR/Cas9 can be likened to a tiny pair of scissors that can cut our DNA at a specific location. Scientists guide the

scissors to the right location using a small piece of barcode (called guide RNA). Once the DNA is cut, the cell will attempt to repair it, and can be tricked into introducing new, specific DNA mutations if we provide it with a specific DNA template. This allows scientists to edit the DNA, e.g. by switching the genetic code from A to G.

In our research, we used the CRISPR/Cas9 technique to investigate the genetic basis of body weight regulation. We wanted to identify which specific DNA mutations are responsible for differences in body weight change and fat metabolism, and how they affect the cells involved.

Our work led to several exciting discoveries. We found that a mutation in our DNA called rs67785913 can affect the production of a gene called *MTIF3*. Curious to discover the function of this gene, we studied how it works in fat cells (adipocytes). We found that *MTIF3* is important for controlling how the cells use and store fat, and it accomplishes this by helping the cells' energy factories called mitochondria to work efficiently during fat combustion.

In our second study, we looked at a specific DNA mutation known as rs8192678. We found it to be involved in the development of white fat cells and their metabolism. The mutation itself influenced the activity of a gene called *PPARGC1A*, which then affected the production of a protein called PGC-1α; it is important for the function and production of several other proteins.

In our third study, we focused on a specific DNA mutation called rs10071329 and how it influences the function of fat-burning brown fat cells. We found that it affected the production of a gene called *PPARGC1B*, which is responsible for the proper functioning of brown fat cell powerhouse – the mitochondria – and their response to norepinephrine – a hormone that helps burn fat.

In our fourth paper, we focused on a specific DNA mutation (called rs2289669) that has been linked with the effectiveness of a commonly prescribed drug called metformin, which helps lower blood glucose in people with T2DM. We found that this mutation and its tightly linked mutation (called rs8065082) do not have any impact on a gene called *SLC47A1* in liver cells. While these results are negative, they further show the value of combining association studies with experimental science to validate which DNA mutations are relevant for T2DM and its treatment.

In conclusion, we used a state-of-the-art technique, CRISPR/Cas9, to unveil the specific DNA mutations and genes responsible for regulating body weight and T2DM, and our work provides a valuable workflow for future studies of other DNA mutations. By studying how our genes control body weight and metabolism, we hope to eventually be able to devise personalized plans to help people stay healthy, and to develop new drugs that help people lose or maintain weight. Our research has shown how complex the interplay between DNA mutations and lifestyle can be, but we also hope it will help us find new ways to help people live healthier lives.

Populärvetenskaplig Sammanfattning

Fetma och typ 2-diabetes (T2DM) är allt vanligare hälsoproblem som påverkar människor över hela världen. Dessa tillstånd orsakas av en kombination av olika faktorer, inklusive vårt DNA och den miljö vi lever i. Sättet som vårt DNA interagerar med vår livsstil och omgivning kan påverka risken att utveckla fetma och T2DM. Vi vet att överätande, brist på motion och andra livsstilsfaktorer kan bidra till dessa tillstånd, men det är också viktigt att förstå hur vårt DNA bidrar till att orsaka fetma och T2DM. Genom att studera människans DNA mutationer, d.v.s. små skillnader i DNA som är unika för olika individer, kan vi få en bättre förståelse för hur dessa sjukdomar utvecklas. Genom att använda denna kunskap hoppas vi senare kunna föreslå mer effektiva förebyggande- och behandlingsstrategier, vilket i slutändan hjälper människor att leva hälsosammare liv.

Hur upptäcker forskare DNA mutationer som bidrar till sjukdomar? Forskare använder en teknik som kallas för genomtäckande associationssstudier (GWAS) för att utforska vilka mutationer som kan vara involverade i orsakande av fetma och T2DM. Genom GWAS har forskare funnit hundratals mutationer som är kopplade till antingen utveckling eller behandlingseffektivitet av dessa sjukdomar.

Även om många mutationer har kopplats till fetma och T2DM, vet vi ofta inte hur de fungerar eller varför de påverkar vissa människor mer än andra. Forskare har funnit att många av mutationerna finns i de delar av DNA som kontrollerar hur vissa gener aktiveras eller inaktiveras. Men vi förstår fortfarande inte fullt ut vilka gener som påverkas eller hur denna reglering fungerar i olika organ eller i olika miljöer. Dessa kunksapsluckor gör det svårt att använda GWAS-resultaten för att utveckla nya behandlingar mot fetma och T2DM.

Därför utvecklar forskare nya sätt att studera DNA mutationer. Ett lovande tillvägagångssätt kallas funktionell genomik, vilket hjälper forskare att undersöka hur specifika mutationer påverkar celler och levande organismer. En av de mest lovande metoderna inom funktionell genomik är en teknik som kallas CRISPR/Cas9-genomredigering. Denna toppmoderna teknik gör det möjligt för forskare att göra precisa förändringar i genomet hos en levande cell och studera effekterna av dessa förändringar.

Hur fungerar CRISPR/Cas9? CRISPR/Cas9 kan liknas vid en liten sax som kan klippa vårt DNA på en specifik plats. Forskare styr saxarna till rätt plats med hjälp av en "streckkod" (kallad guide-RNA). När DNA:t har klippts försöker cellen att

reparera det, och kan då luras att introducera nya, specifika DNA-mutationer, förutsatt att vi tillhandahåller den en specifik DNA-mall. Detta gör det möjligt för forskare att redigera DNA:t, t.ex. genom att byta genetisk kod från A till G.

I vår forskning använde vi CRISPR/Cas9-tekniken för att undersöka den genetiska grunden för reglering av kroppsvikt. Vi ville identifiera vilka specifika DNA-mutationer som är ansvariga för skillnader i förändringar av kroppsvikt och fettmetabolism, och hur de påverkar celler som är involverade i dessa process.

Vårt arbete ledde till flera spännande upptäckter. Vi fann att en mutation i vårt DNA kallad rs67785913 kan påverka produktionen av en gen som kallas *MTIF3*. För att upptäcka funktionen av denna gen, studerade vi hur den påverkar fettceller (adipocyter). Vi fann att *MTIF3* är viktig för kontroll av cellernas användning och lagring av fett, och att *MTIF3* åstadkommer detta genom att hjälpa cellernas energifabriker, så kallade mitokondrier, att arbeta effektivt när fettförbränning sker.

Det andra studiet fokuserade på en specifik DNA mutation betecknad rs8192678. Vi fann att den är involverad i utvecklingen av vita fettceller och deras ämnesomsättning. Själva mutationen påverkade aktiviteten av en gen kallad *PPARGC1A*, vilket i sin tur påverkade produktionen av ett protein kallat PGC-1α, som är viktigt för funktionen och produktionen av flera andra proteiner.

I vår tredje studie fokuserade vi på en specifik DNA mutation som kallas rs10071329 och hur den påverkar funktionen av fettförbrännande bruna fettceller. Vi fann att mutationen påverkade produktionen av en gen som kallas *PPARGC1B*, som är ansvarig för den korrekta funktionen av de bruna fettcellernas kraftverk - mitokondrierna - och deras respons på noradrenalin - ett hormon som hjälper till att förbränna fett.

I vår fjärde artikel fokuserade vi på en specifik DNA mutation (kallad rs2289669) som har kopplats till effektiviteten av en vanligt föreskriven medicin kallad metformin, som hjälper till att sänka blodsockret hos personer med T2DM. Vi fann att denna mutation och dess närliggande mutation (kallad rs8065082) inte har någon påverkan på en gen kallad *SLC47A1* i leverceller. Även om dessa resultat är negativa visar de ytterligare värdet av att kombinera associationsstudier med experimentell forskning för att validera vilka DNA mutationer som är relevanta för T2DM och dess behandling.

För att summera, i vår forskning använde vi en toppmodern teknik, CRISPR/Cas9, för att kartlägga de specifika DNA-mutationerna och generna som är ansvariga för att reglera kroppsvikt och T2DM. Vårt arbete demonstrerar även ett arbetsflöde som kan utnyttjas för framtida studier av andra DNA-mutationer. Genom att studera hur våra gener kontrollerar kroppsvikt och metabolism hoppas vi så småningom kunna utveckla individanpassade strategier som hjälper människor att hålla sig friska, samt utveckla nya läkemedel som hjälper människor att gå ner i vikt eller behålla sin vikt. Vår forskning har visat hur komplex samverkan mellan DNA mutationer och livsstil

kan vara, men vi hoppas också att den kommer att hjälpa oss att hitta nya sätt att hjälpa människor att leva hälsosammare liv.

科普总结

最近几十年来，肥胖和 2 型糖尿病的发病率在全球范围内急剧上升，严重威胁着人们的生命健康。导致这些疾病产生的原因是多方面的，其中包括我们的基因和所处的生活环境。众所周知，过度饮食、缺乏运动和其他不健康的生活方式等这些因素会增加患肥胖和糖尿病的风险，我们的基因在这些疾病的发生和发展过程中也扮演着非常重要的角色。科学家们通过研究人类基因中的突变，也就是不同个体之间基因中微小的差异能够帮助我们更好的理解肥胖和糖尿病，从而提出更有效的预防和治疗策略，最终帮助人们过上更健康的生活。

那么，科学家是如何来研究到底是哪些基因突变会导致疾病的呢？通常来讲，研究人员会利用一种称为全基因组关联研究（GWAS）的技术来寻找这些可能致病的基因突变。到目前为止，科学家们已经通过GWAS的研究发现了数百个与肥胖和 2 型糖尿病的发生或治疗疗效相关的基因突变。

然而，科学家们发现的这些基因突变，它们的具体功能我们目前所知甚少，我们也并不清楚为什么这些基因突变对某些人的影响比其他人更大。有意思的是，科学家们还发现，大多数的基因突变都不存在于基因中编码蛋白质的区域，这就意味着它们的功能可能是控制某些基因的开启（上调）或关闭（下调）。即便如此，我们仍然不了解这些基因突变是调控哪些基因以及在什么条件下它们才发挥基因调节的作用。这些知识鸿沟使得我们难以利用 GWAS 的研究结果来开发针对肥胖和 2 型糖尿病的新治疗方法。

科学家尝试开发新的方法来研究基因突变的脚步从未停止。近年发现的 CRISPR/Cas9 基因组编辑的技术，在这一研究领域展现出了极大的优势。利用这种前沿技术，研究人员可以对人类细胞的基因组进行精确的修改，并研究这些改变对细胞生命活动产生的影响。

那么，CRISPR/Cas9 这个工具是如何工作的呢？我们可以把 CRISPR/Cas9 比作一把微小的分子剪刀，它可以在特定的位置切割我们的基因组。科学家们首先使用一个小的条形码（称为向导RNA）将这个剪刀引导到基因组中想要进行编辑的位置， 这把分子剪刀就会将基因切割，紧接着细胞将会启动相应的机制来修复这个切口。在这一阶段，如果我们提供给细胞特定的基因修复

模板，那么细胞就有可能直接利用这个模板来修复断裂的基因，从而实现精准的基因编码的替换。例如，在特定的基因座上，将基因编码从 A 变为 G。

在我的博士课题中，我们使用了 CRISPR/Cas9 的技术深入研究了一些调控人体体重的基因突变。通过这些研究，我们希望找到那些在调节体重变化和脂肪代谢中真正起作用的基因突变，以及阐明它们是如何来影响我们的细胞和组织的功能。

在这些研究中，我们有一些振奋人心的发现。在我们的第一个课题中，我们发现一个叫做 rs67785913 的基因突变，它可能会影响基因 MTIF3 生成对应的蛋白质。接着，我们也研究了 MTIF3 蛋白质在脂肪细胞中的基因功能。我们发现 MTIF3 可以帮助协调细胞的能量工厂 - 线粒体在脂肪燃烧过程中高效工作来控制细胞如何利用和储存脂肪。

在我们的第二项研究中，我们研究了一个叫做 rs8192678 的基因突变。我们发现它参与了白色脂肪细胞的生成和细胞代谢。该基因突变影响了一个叫做 *PPARGC1A* 的基因的活性，从而影响了一个名为 PGC-1α 的蛋白质的产生，该蛋白质对细胞代谢过程中其他的一些关键蛋白质的功能和产生都具有极其重要的影响。

在我们的第三项研究中，我们研究了 rs10071329 的基因突变对棕色脂肪细胞的功能的影响。棕色脂肪细胞是一种可以"燃烧脂肪"的细胞，我们发现这个基因突变可以调控 *PPARGC1B* 的基因产物的合成，从而调控棕色脂肪细胞能量工厂- 线粒体的正常功能以及对去甲肾上腺素介导的"脂肪燃烧"的效率。

在我们的第四篇论文中，我们研究了与一种常用的降血糖的处方药 - 二甲双胍的有效性相关的基因突变，rs2289669。我们发现这个基因突变及其紧密联系的突变（称为 rs8065082）对肝细胞中的一个名为 *SLC47A1* 的基因没有任何影响。这些阴性结果，进一步展示了基础实验在验证二型糖尿病相关的基因突变中的必要性。

总而言之，我们使用了先进的 CRISPR/Cas9 基因编辑技术，揭示了特定的基因突变在调节体重和 2 型糖尿病中的作用机制，提供了具有指导价值的工作思路和流程。通过研究我们的基因突变如何控制体重和新陈代谢，我们希望最终能够制定个性化的方案来帮助人们保持健康，并开发新药来帮助人们减肥或维持体重。我们的研究也同时显示了基因突变和环境之间存在着的复杂的相互作用，也希望通过我们的这些研究来帮助人们来寻找一种更健康的生活方式。

Acknowledgments

This thesis would not have been possible without the invaluable contributions of many people surrounding me in these five and a half year's journeys. I am deeply grateful for the assistance, advice, and encouragement provided by my supervisors, colleagues, and friends. Their dedication, expertise, and generosity have made a significant impact on this work and on my personal and professional growth. I am also thankful to my family and my girlfriend for their unwavering support. Thank you all for being a part of this amazing experience.

I extend my sincerest thanks to **Paul Franks**, who has been my main supervisor throughout the first three and a half years of my Ph.D. studies. His trust in my abilities, guidance, and support has been invaluable in shaping me into the researcher I am today. I am deeply grateful for the opportunity he gave me to pursue my Ph.D. under his supervision. I would also like to express my gratitude for the excellent opportunities he provided me with, such as the chance to visit the **Broad Institute of MIT and Harvard**, which greatly inspired me and enhanced my research experience. **Paul**, thank you for your dedication to helping me achieve my academic and personal goals and for being a role model in the scientific community. Your mentorship has had a significant impact on my growth as a researcher.

I would also like to express my gratitude for **Sebastian Kalamajski's** exceptional qualities as a supervisor, including his patience, helpfulness, and willingness to answer my questions. **Sebastian** has been my main supervisor during the last two years of my Ph.D. studies and my co-supervisor at the beginning of my academic journey, providing me with hands-on experiments training and outstanding mentorship. His passion for science and abundance of innovative ideas have been instrumental in motivating my research journey. **Sebastian's** patience in answering my naive questions and his helpfulness in guiding me through challenging tasks have been invaluable in my academic and personal growth. I consider myself fortunate to have had **Sebastian** as my supervisor and friend, and without his guidance and support, completing all my projects would have been a daunting task. I must also express my thanks for **Sebastian Kalamajski's** sense of humor and ability to make us laugh, which made my time in Sweden even more enjoyable. **Sebastian**, thank you for being an excellent supervisor, and a great friend.

I am also grateful to my co-supervisor **Hindrik Mulder** for providing me with an excellent opportunity to work with his talented research team. His guidance and

support, along with the collaboration with other group members, have been essential in enhancing my research experience. Working in **Hindrik Mulder's** team has also been a lot of fun. From Christmas dinner competitions to occasional retreats, the collaborative and supportive atmosphere within the team has made my Ph.D. journey an unforgettable experience.

My heartfelt thanks go to **Melina Claussnitzer** for giving me the opportunity to work in her lab in **Boston**. **Melina's** passion for science and her ability to navigate complex challenges inspired me greatly. She demonstrated intelligence and humility in her work, and I learned a lot from her during my time in her team. I appreciate her support during this experience and for being a positive influence on my academic journey.

I want to express my great thanks to **Giuseppe** (**Nick**) **Giordano** for taking on the role of our group leader when **Paul** left, and for his excellent management of all administrative tasks for us students. His support and assistance have been crucial in enabling me to focus on my research. I appreciate his willingness to help whenever I have asked for it. Thank you, **Nick**.

I really appreciate all the help from our lab manager, **Laila Jacobsson**, for her tireless efforts in taking care of our lab-related issues and ensuring that we have everything we need for our experiments. She has been fundamental in helping me to order the lab materials I require and has always been there to help whenever I need it. I am grateful for her contributions to the lab and for making our work environment efficient and organized.

I want to give special thanks to **Pernilla Siming** for her exceptional support and dedication during my study. Her help with administrative tasks, visa applications, flight bookings, and so on has been invaluable. **Pernilla**, thank you for going above and beyond to make our research possible.

I am incredibly thankful for the support and guidance of my colleagues and friends in the **Molecular Metabolism Unit**: **Malin Fex**, **Sevda Gheibi**, **Rodrigo Cataldo**, **Leonardo de Moura Alvorcem**, **Sergey Hladkou**, **Fiona Roberts**, **Tania Singh**, **Claire Lyons**, **Gaurav Verma**, **Mahmoud Al-Majdoub**, **Neelanjan Vishnu**, **Elaine Cowan**, **Alexander Hamilton**, **Joao Paulo da Cunha**, **Klinsmann Carolo Dos Santos.** I want to express my deepest gratitude for the opportunity to work with each of you. Your intelligence, creativity, and kindness have not only made our research more successful but also made it a joy to come to work every day. It is truly a pleasure and an honor to be a part of this team.

I feel privileged to have been a part of the **GAME Unit** and want to thank my colleagues and friends for making it an enjoyable and rewarding experience: **Daniel Coral, Hugo Fitipaldi**, **Naeimeh Atabaki Pasdar**, **Hugo Pomares M**, **Pascal**

Mutie, **Marketa Sjögren**, **Maria Keller, Yan Chen**, **Tibor V Varga**, **Simon Timpka**, **Alaitz Poveda**.

I also want to express my sincere thanks to my friends and colleagues in **Claussnitzer's lab** at **Broad Institute**: **Anlu Chen**, **Alham Saadat**, **Garrett Garborcauskas**, **Samantha Laber**, **Virtu Calabuig Navarro**, and **Sophie Strobel.** Their kindness and generosity have made my time in the lab and in Boston all the more enjoyable. I am also delighted for the excellent discussions and insights that we have shared together. It has been a pleasure to work alongside such intelligent and thoughtful individuals, and I am honored to have been working in this team.

Lots of thanks to my co-authors: **Hamidreza Ardalani** and **Peter Spégel**, for their excellent contributions to our papers; also to my half-time examiners: **Nils Wierup** and **Camilla Schéele**, for their constructive suggestions.

My time in Sweden was made unforgettable by my Chinese friends, and I would like to extend my deepest appreciation to them: **Yanni Li**, **Naiqi Zhang**, **Ming Sun**, **Rui Wu**, **Daowei Yang**, **Wuqing Huang**, **Zhiyi Ding**, **Feifei Du**, **Yang Liu**, **Guoqiao Zheng**, **Yishan Liu, Mengyu Pan (Ammy), Yaxiong Deng**, **Jian Li**, **Chen Liu**, **Suxun Pan**, **Xingchen Li**, **Xue Bao**, **Songjiang Yin**, **Chuanyan Wu**, **Jingkai Zhou** and **all in CRC**. I have truly enjoyed every moment spent with you, and your warmth and kindness have made me feel at home. I wish you all success in your future endeavors.

My thanks also go to senior scientists in **CRC**: **Yang De Marinis**, **Cheng Luan**, **Yingying Ye**, **Yan Borné**, **Jiangming Sun**, **Enming Zhang**, **Jianguang Ji**, and **Xiao Wang**. Thank you very much for all the conversations and suggestions. I really learned a lot from you.

I am grateful to Professor **Xinzhou Yang** in China for his encouragement and support throughout my Ph.D. journey.

I would like to take this opportunity to express my gratitude to **Mark** and **Mairead** for their generous hosting during our stay in the USA. Their passion for life and positive attitude have left a lasting impression on me.

Thanks to **Ronnie**, our cuttest cat for the accompany and for giving us so much fun.

Last but not least, I would like to thank my girlfriend **Jingxue Pan (Jane)** for her unwavering love and support throughout my time in Sweden. I feel incredibly lucky to have you in my life, Jane. Your presence has added so much joy, happiness, and color to my everyday experiences. Your unconditional love and encouragement have been a source of strength and motivation for me, especially during challenging times.

Many thanks to the funding bodies for the support of this thesis work: China Scholarship Council (201708420158); the European Commission (ERC-2015-CoG-

681742), and Swedish Research Council, LUDC-IRC and The Albert Påhlsson Foundation.

最后我想感谢我的父母，姐姐，姐夫和我的小外甥，感谢你们一直以来对我的支持, 理解和关心，感谢你们对家庭的付出。你们的支持和理解，让我有信心面对一切挑战，有你们做我最可靠的后盾，我相信我可以度过人生中的各种难关。

Chapter 1 - Introduction

Obesity

Overweight and obesity are major risk factors for diabetes, cancer, heart disease, and stroke, and are leading causes of death and disability [1]. Body mass index (BMI), calculated as the ratio of weight (in kilograms) to square of height (in metres), is commonly used to define categories of body corpulence. For adults, a person with a BMI ≥ 25-29.9 kg/m^2 is considered overweight, and a BMI ≥ 30 kg/m^2 is considered obese.

In 2016, the World Health Organization (WHO) reported that worldwide 39% of adults were overweight, and 13% were obese [2]. Strikingly, the incidence of obesity has tripled in recent years, a health problem that continues to worsen [3,4].

The growing number of people with overweight or obesity is overwhelming healthcare systems. People with obesity often require more medical attention and specialized care, and experience more interactions with healthcare services, and are hospitalized and under surgery more frequently than people with low BMIs [5,6]. One study estimated the economic impact of obesity and overweight will be between 2.19% and 3.3% of gross domestic product (GDP) in 161 countries by 2060 [7].

Efforts have been made to reveal the mechanisms of overweight and obesity. There are key drivers of overweight and obesity: **Environmental factors**: factors for positive energy balance and weight gain include increased eating and consumption, especially of high-calorie foods, reduced physical activity, use of medications with adverse effects of weight gain, and insufficient sleep [8-11]. Many people who have overweight or obesity will have energy homeostasis (i.e. they will not be losing or gaining weight) and they may have developed overweight or obesity in childhood or adolescence. **Genetic factors:** families-based studies, some involving twins have shown that BMI is highly heritable, with an estimated heritability between 40% and 70% [12]. In addition, 11 rare genetic forms of obesity have been observed, such as genetic deficiency in the leptin gene and melanocortin-4 receptors, with the hypothalamus and play a role in regulating energy balance [13,14].

Current therapies for obesity and overweight management are lifestyle intervention, pharmacotherapy, and bariatric surgery [11]. **Lifestyle interventions**, which typically consist of diet intervention, physical activity, and behavioral therapy, are a standard approach for weight loss in people with overweight or obesity [15]. Examples of

lifestyle interventions can be found in two ground-breaking randomized control trials: the Diabetes Prevention Program (DPP) [16] and the Action for Health in Diabetes (Look AHEAD) trial [17]. People with obesity typically lose about 8 kg (about 8% of their initial weight) with intensive lifestyle interventions using the Guidelines for the Management of Overweight and Obesity in Adults as a framework [18]. **Pharmacotherapy** is often used in conjunction with lifestyle interventions to achieve weight loss and to prevent weight regain [19,20]. There are 7 medications so far that can be used for chronic weight management in adults, namely: semaglutide, bupropion-naltrexone, liraglutide, orlistat, phentermine-topiramate, tirzepatide, and setmelanotide. However, concerns about side effects and the common occurrence of weight regain after stopping the medication may discourage some patients from using it [11]. **Bariatric surgery,** which is a weight management option for patients with a BMI of 40 or higher, or a BMI between 35 and 39 in the presence of an additional health condition. There are three main types of bariatric surgery: laparoscopic adjustable gastric banding, vertical-sleeve gastrectomy, and Roux-en-Y procedures. Although the surgeries are more effective in reducing body weight than lifestyle changes and medications, they are also riskier and more expensive [21,22].

Nevertheless, patients with lifestyle interventions and treatments often show a very wide interindividual response to bodyweight loss and maintenance [23-26]. Accumulating evidence has shown these inconsistent results might be mediated by genetics [26,27], which highlights the importance to study the biological mechanisms of conception: how genetics and environmental factors interplay affects body weight control.

Type 2 diabetes mellitus

Type 2 diabetes mellitus (T2DM) is a chronic metabolic disorder characterised by impaired insulin secretion and insulin resistance resulting in hyperglycaemia. It has become a global health threat; by 2040, the estimated prevalence will be 642 million people worldwide, with about 98% with T2DM [28]. This type of diabetes usually develops slowly over many years, and is often accompanied by obesity, lack of physical activity, and an unhealthy diet [29].

It is generally accepted that T2DM is a polygenic disease resulting from a complex interaction between environmental factors and genetic components [30]. Environmental risk factors that influence the development of T2DM include obesity, sedentary lifestyle, low or high birth weight and stress. Other dietary factors and toxins may also play a role [31]. These environmental factors undoubtedly affect the onset of T2DM, but they do not impact everyone in the same way.

Epidemiological studies have shown that the estimated heritability of T2DM ranges from 20% to 80% in different populations [32,33]. The lifetime risk of developing T2DM is 40% for people with one parent with T2DM and 70% if both parents are affected. An abundance of genes associated with T2DM have been uncovered through linkage investigations, candidate gene studies, and genome-wide association studies; these include *PPARG*, *TCF7L2*, *IRS1*, *IRS2*, *FTO, MTNR1B,* and others. However, how these genes interact with each other, and with the environmental factors to cause T2DM is still not well understood.

Although the etiology and disease mechanisms are not fully understood, there are several treatment options for T2DM. Glucose-lowering drug therapy, along with maintaining a healthy lifestyle, is an essential part of managing T2DM. Among these drugs, metformin remains one frontline therapy for people with T2DM in many countries.

Metformin primarily acts in the liver to reduce the expression of gluconeogenic enzymes and inhibit mitochondrial respiration to decrease hepatic glucose output [34,35]. It also improves insulin sensitivity of peripheral tissues and stimulates GLP-1 secretion [36]. Furthermore, metformin nominally reduces HbA1c concentration by about 1–2%, and rarely causes hypoglycaemia; it has modest beneficial effects on blood pressure and lipid profile [37]. Because of these biological actions, the benefits of metformin are not limited to T2DM, but also apply to the prevention of cardiovascular diseases, obesity, etc [37]. However, metformin is not effective for all the people with T2DM, nearly around 30% of patients do not respond to the drug [38], and ~20 to 30% of patients experience intolerable side effects [39]. This inefficacy and side effects might be partly explained by genetics [40].

The role of adipose cells in obesity and T2DM

In humans, normal fat mass ranges from 9 to 18% in males and from 14 to 28% in females; while people with obesity have fat mass exceeding 22% of body weight in males and 32% in females [41]. Adipose tissue is highly heterogenous and has diverse functions in regulating human metabolic homeostasis, including insulin sensitivity [42], glucose and lipid metabolism, inflammatory responses [43], as well as non-shivering thermogenesis [44].

There are distinguished adipose tissue types: white adipose tissue (WAT) and brown adipose tissue (BAT) [45]. The primary role of WAT is as an energy reservoir to regulate energy homeostasis by storing excess lipids, and secreting paracrine factors to regulate other metabolic tissues, e.g., leptin [46,47]. By exposing the adipocytes to a sustained excessive energy supply, their capacity for lipid storage can be pushed to their limits [47]. Consequently, the ectopic fat may end up stored in other organs, which are essential for glucose homeostases, such as the muscle, pancreas (islets),

and liver. This phenomenon is referred to as "lipotoxicity" and is the cause of whole-body insulin resistance [47,48]. Unhealthy WAT expansion occurs when people are under prolonged positive energy balance conditions. To compensate for the increased need for lipid reserves, white adipocytes increase in size and number, which may cause stress in the fat cells [47]. Many adverse effects are associated with unhealthy WAT expansion, including inflammation, impaired adipokine secretion, and mitochondrial dysfunction. Targeting each of the listed outcomes may be helpful in the development of treatment for obesity [45].

Unlike the role of white adipose tissue in energy storage, BAT largely oxidizes fatty acids without generation of ATP. This uncoupling is achieved through uncoupling protein 1 (*UCP1*) which depolarizes the inner mitochondrial membrane and generates heat [49]. BAT is found predominantly in rodents and infant humans, where it's located around the scapula area [50]. BAT may contribute 60% of "non-shivering" thermogenesis in small mammals, allowing them to survive in cold environments [51]. In human adults, recent studies using the 18-fluorodeoxyglucose (^{18}FDG) imaging analysis also provided clear evidence of the existence of BAT, and the estimated volume is typically from 50 to 150 ml. However, the true volume of BAT in humans might be underestimated due to the technical limitations of ^{18}FDG PET/CT; for example, it is very challenging to measure regions smaller than 1 cm^3. [51]. Brown fat mass and metabolic activity have been inversely correlated with body-mass index in humans [52,53]. Enhancing the activity of BAT has been proposed as a strategy for the prevention and treatment of obesity and obesity-related metabolic disorders [54].

GWAS in obesity and T2DM

The completion of the Human Genome and International HapMap Projects created unique opportunities for studying genetic associations with selected diseases and traits in a "genome-wide" fashion. GWAS was at that point propelled, and rapidly became a powerful tool in finding multiple genetic variants in diseases with individually small effects. The single nuclear polymorphism (SNP) found in GWAS could have additionally identified some novel genetic insights into obesity. However, the hope for personalized prevention and treatment of obesity by using the identified variants is still far away because most genetic variants connected to BMI and obesity have not been proven to be biologically or clinically relevant for obesity treatment or prevention [55].

To facilitate the clinical use of the genetic discoveries through GWAS, attempts have been made by testing the variants and lifestyle interactions on body weight loss and regain in randomized control trials [56]. For example, in DPP and Look AHEAD trials, researchers found the minor G allele of *MTIF3* rs1885988 was consistently

associated with greater weight loss following lifestyle intervention over 4 years, indicating the significant SNP and treatment interaction [56].

Regarding T2DM, there have been over 400 genetic variants identified by GWAS associated with T2DM in European ancestry participants [57] and 242 loci associated with glycemic traits [58]. A large portion of the identified loci seems to be relevant for β-cell function. Indeed, as many SNPs associated with T2DM and glycemic traits are enriched in open chromatin sites in islets, suggests that their potential regulatory functions are for insulin production or secretion. These genetic findings also provided an opportunity for pharmacogenetic studies, which may help generating precision therapeutic options for T2DM. There is a growing body of evidence from cohort studies and clinical trials supporting the use of genotype to predict the response to diabetes medication [59].

From association to causality: the challenges

The information provided by GWAS helps link a genetic variant to a particular disease but because it only identifies the genomic region that contains the causal variant, it does not clearly define the functional DNA sequence. Identifying disease-causal variants underlying GWAS associations remains challenging due to several reasons:

The existence of linkage disequilibrium (LD) When two or more genetic variants (or alleles) are physically close to each other on a chromosome, they tend to be inherited together more often than expected by chance, which is called Linkage disequilibrium (LD) [60]. On average, there is an occurrence of SNPs every 300 base pairs [61]. Consequently, a SNP identified in GWAS (tag SNP) is often just a representative of a genomic region (the whole LD block) that harbors the causal variant. The causal variant might be the tag variant itself or other variants in the haploblock.

Most GWAS signals reside in non-coding regions More than 90% of variants found in GWAS are in non-coding regions, making it difficult to identify which gene or genes are affected by the genetic variation. These non-coding variants are commonly enriched in predicted regulatory regions, suggesting that their effects on affecting disease risk occur through regulating gene expression, splicing or mRNA stability [62]. Once a specific polymorphism is associated with a disease, it is usually annotated by naming the gene in closest proximity. However, owing to the complex nature of eukaryotic transcriptional regulations, this doesn´t implicate the nearest gene as the actual target gene [63].

The role of the non-coding variant may be cell type specific Many functional elements in non-coding regions are tissue- and cell-type specific [64]. Assaying the

correct cell types that are causally related to the phenotype is essential for the annotation of GWAS SNPs. Integrating gene expression and chromatin state annotation into SNP enrichment assays is a helpful approach for prioritizing cell types [64,65]. However, the involvement of multiple cell types in the mediation of many traits through SNPs adds complexity to the dissection of disease-causal variants. The use of single-cell approaches may aid in the elucidation of the role of disease-associated variants in specific cell subpopulations [65].

From association to causality: the experimental approaches

Mapping causal DNA variants from GWAS studies and identifying their target genes, and furthermore, linking the causal variants and target genes to molecular, cellular, and physiological phenotypes is extremely important for the translation of GWAS results into clinical practice, with the goal of advancing precision medicine [66]. While numerous statistical methods and genomic functional annotations have been applied to prioritize causal variants (termed fine-mapping) and their target genes [67,68], the crucial evidence, validating these causal variants and their targets in a laboratory setting, is largely lacking. To dissect the causal variants, genome editing in the native genome provides perhaps the best evidence for causality or lack thereof. Thanks to the development of genome editing techniques, we now have numerous tools for precise genome editing. Among them, CRISPR/Cas9 genome editing system has shown great potential [66].

CRISPR/Cas9 is an abbreviation of Clustered Regularly Interspaced Palindromic Repeats (CRISPR) and its associated protein 9, and was initially discovered as an adaptive immune system in bacteria and archaea [69]. Later, this system was utilized for modifying genomes in eukaryotic cells and extensively employed in biomedical research and clinical applications [70]. The CRISPR/Cas9 system consists of two components, the Cas9 nuclease and a guide RNA. The Cas9 nuclease is the effector, which cuts the target double-strand DNA, while the guide RNA connects (targets) to the Cas9 enzyme to a specific DNA sequence. The guide RNA can be either a complex of crispr RNA (crRNA) and trans-activating crispr RNA (tracrRNA), or a chimeric single-guide RNA (sgRNA). Guide RNA has a 20-nucleotide sequence (spacer sequence) that can complementarily match with the DNA target (protospacer) in the genome. In front of the protospacer DNA sequence, it requires a motif 5′-NGG-3′ (the protospacer adjacent motif, called PAM) for Cas9 binding [70].

Once a DNA double stand break (DSB) has been introduced by the Cas9 nuclease, the cells start to repair the DSB using two mechanisms: either homology-directed repair (HDR) or non-homologous end joining (NHEJ) [71]. NHEJ typically results in

random mutations with the creation of small insertions or deletions at the location where the breakage occurred, and this repair system predominates in most cases [72]. Alternatively, when provided a donor DNA template, the HDR mechanism may be triggered, resulting in an error-free repair that can introduce precise genetic modifications [72]. These two distinguished pathways are used by researchers for generating different genomic engineering outcomes. The former is normally used for creating a general gene knockout or DNA element deletion, while the latter is used for precise genome editing, e.g., allele switching or DNA fragment insertion. Notably, when applying the CRISPR/Cas9 system for single nucleotide substitution, as the two repair mechanisms are competing for the DSB repair, usually a mixed proportion of precisely edited cells and cells with random mutations is found [73].

Using CRISPR/Cas9 genomic editing system, researchers have successfully identified some causal variants observed in GWAS (summarized elsewhere [66]). The methods used include the introduction of indels nearby the causal variants, genomic deletion of the causal variants, as well as the allele substitution of the causal variant. A pioneering study using CRISPR/Cas9 to generate allele switch allowed identifying causal variant in *FTO* locus was conducted in 2015 [74]. In that study, Claussnitzer et al. applied CRISPR/Cas9 in primary pre-adipocyte cells, successfully made an rs1421085 T-to-C single-nucleotide switch, and confirmed its causal role in regulating *IRX3* and *IRX5* expression which affected adipocyte thermogenesis [74]. This revealed the potential of using CRISPR/Cas9-mediated allele editing in dissecting GWAS causal variants.

Nevertheless, the application of CRISPR/Cas9 in precise genomic editing is limited by the low efficiency of HDR. In fact, the CRISPR/Cas9 mediated HDR efficiency depends on many factors such as cell type or the sequence and accessibility of the edited locus. Although some methodologies have been developed for improving HDR efficiency [75], their efficiency is often not satisfactory. Given the hundreds of SNPs found for one specific trait and the colocalized variants in the same haplotype block, it is extremely resource-demaning to edit all the variants and evaluate their causality using the CRISPR/Cas9-mediated allele switch approach. Therefore, the common initial step for dissecting the causal variants is to perform fine mapping using either publicly available annotation tools or prioritize variants in an experimental setting using reporter assays.

It is usually straightforward to test the causal role of a coding DNA variation, as the coding variation may disrupt the protein structure and further affect the protein function. For the non-coding genetic variants, a luciferase reporter assay may be applied for prioritizing potential causal variants, assaying both the risk and protective alleles of the candidate SNPs. The basic workflow for the cell culture-based luciferase reporter assay is that several hundred base pairs surrounding the SNPs are cloned into a plasmid carrying a promoter and a reporter gene. They are then transfected into a relevant cell type, and the transcriptional activity of the fragments is read out. This assay is scalable; so rather than testing reporter

constructs one-by-one in cell culture contexts, several research groups have developed massively parallel reporter assays (MPRAs), in which thousands of variants can be tested in a single experiment [76].

rs8192678 and *PPARGC1A*

rs8192678 (C/T, Gly482Ser) is a missense mutation in the *PPARGC1A* gene, and it is the only common polymorphism in its haploblock. The minor T allele frequency is 5% in African populations, and ranges from 26 to 44% in American, East Asian, European, and South Asian populations. Epidemiology studies have shown the T allele is extensively associated with T2DM, insulin resistance, and obesity [77]. Interestingly, the T allele also associates with a better response to certain lifestyle interventions on metabolic outcomes [78-80].

rs8192678 C/T encodes peroxisome proliferator-activated receptor γ coactivator 1-α (PGC-1α) Gly482 and Ser482, respectively. PGC-1α protein is a transcription coactivator, which interacts with numerous transcription factors to regulate the expression of genes involved in oxidative metabolism and mitochondrial function [81,82]. The rs8192678 T allele has shown an association with lower *PPARGC1A* mRNA expression in muscle and islets [83,84]. When PGC-1α was episomally overexpressed in HepG2 cells and Ins-1 cells, the T allele-encoded Ser482 degraded faster than the C allele-encoded 482Gly protein [85,86]. The above studies, in epidemiological and experimental biology, suggest that the rs8192678 polymorphism may affect PGC-1α protein abundance, which in turn may contribute to the observed phenotypes. However, direct evidence that the biological mechanisms linking the Gly482Ser PGC-1α variant with cellular phenotypes in metabolic disorders are still missing.

rs1885988 and *MTIF3*

rs1885988 (A/G) is a non-coding genetic variation, and the G allele is the minor allele, with a frequency of 19% in the European population and 15% in the East Asia population. This variant has shown significant association with lifestyle intervention-mediated weight change in the world's largest randomized controlled weight loss trials (Diabetes Prevention Program [DPP] and Look AHEAD). Researchers found that the G allele carriers were slightly more prone to weight gain in the control arm, yet achieved significantly greater weight loss at 12 months post-randomization and retained lost weight longer (18-36 months) than A allele carriers [56]. Elsewhere, the same locus has been associated with greater and more sustained weight loss following bariatric surgery [87]. These observational studies give a hint

that genetic variations in rs1885988 haplotype may play a causal role in lifestyle intervention-mediated weight regulation. However, there are some questions that remain to be answered: which genetic variation(s) is causal for the observed phenotype? Which gene(s) is the target of the variant(s)? What is the function of the target gene? What is the biological mechanism linking the gene to the weight change outcome?

MTIF3 is the nearest gene to rs1885988, and other genetic variants within the *MTIF3* gene locus have also been associated with body weight regulation outcomes. *MTIF3* encodes the Mitochondrial Translation Initiation Factor 3 protein [88]. Functional studies in cell and mouse models have shown the important role of MTIF3 in regulating mitochondrial protein synthesis, and oxidative phosphorylation OXPHOS complexes assembly in heart and skeletal muscle [89, 90]. In the context of body weight regulation, these studies suggest that *MTIF3* may influence obesity predisposition and weight loss potential by modulating mitochondrial function. *MTIF3* may play a key role in adipose tissue metabolic homeostasis, as adipocyte mitochondria are involved in lipid metabolism, thermogenesis, and regulation of whole-body energy homeostasis [91,92]. However, studies for validation of the causal role of genetic variant candidates and genes in the weight regulation-related cellular models are needed.

rs10071329 and *PPARGC1B*

rs10071329 is a biallelic (A/G) variant, the minor allele frequency is 1% globally, but higher in African populations, reaching 7-8% in ESN (Esan in Nigeria) and YRI (Yoruba) subpopulations (according to population genetics data in 1000 Genomes Project Phase 3). This genetic variation has been associated with subcutaneous adiposity change after lifestyle intervention [93]. But whether the rs10071329 is a causal variation that regulates the observed phenotype is still unknown.

rs10071329 resides downstream of *PPARGC1B*, whose protein coding product, peroxisome proliferator-activated receptor γ coactivator 1β (PGC-1β), is established as an effector in brown adipocyte lipid metabolism and energy metabolism [94 95]. Importantly, rs10071329 is the only variant in its haplotype block. These findings together suggest that rs10071329 may play a regulatory role in regulating *PPARGC1B* expression, thus affecting body weight regulation. However, this hypothesis needs to be experimentally examined in the relevant cellular model and in vivo models.

rs2289669 and *SLC47A1*

Metformin is the first-line drug for glycaemic control in type 2 diabetes. However, patients´ responses to metformin vary widely [96], which motivates researchers to search for the genetic determinants of this phenotype. There is one genetic variant, rs2289669, that has been consistently associated with an HbA1c-lowering effect after metformin treatment, with the A allele being more beneficial [97-101]. Moreover, rs8065082, which is in tight linkage disequilibrium with rs2289669, has been shown to mediate the risk of T2DM development after metformin treatment, with the minor T allele being more beneficial [102]. In contrast, the rs2289669-mediated metformin HbA1c-lowering effect was not supported by analyses within the MetGen consortium[103]. To explain these contradictory findings, direct evidence is needed that links the causal SNP candidates to their effector genes.

rs2289669 resides in an intron of the *SLC47A1* gene, which encodes a transmembrane cationic transporter MATE1 (multidrug and toxin extrusion 1). MATE1 modulates metformin efflux from the liver and kidney [104-109]. Putting the above evidence together, it is reasonable to hypothesize that the intronic genetic variants at rs2289669, or closely linked variants, could affect *SLC47A1* gene expression or splicing, ultimately determining metformin efflux rate from the liver and kidney, thus affecting the metformin efficacy in lowering HbA1c. However, this hypothesis needs to be experimentally tested.

Chapter 2 – Methods

The experimental procedures of each assay are described in the papers included in the thesis. In this chapter, I will focus on explaining the principles of the key methods, and on describing the key resources used in the projects.

Cell lines

Human white pre-adipocyte cell line (hWAs) and brown pre-adipocyte cell line (hBAs) were kindly shared by Professor Yu-Hua Tseng (Joslin Diabetes Center, Boston, USA). The cells of origin were from a female subject aged 56, with a BMI of 30.8 [110]. The cells were immortalized by introducing human telomerase reverse transcriptase (hTERT) into primary white pre-adipocyte cells isolated from the subcutaneous fat pad, or into primary brown pre-adipocyte cells from deep neck fat. Both cell lines were used in this thesis for genomic editing and for functional assays to identify causal genetic variation. hWAs cell line was used in **paper I** and **II** and hBAs cell line was used in **paper III**.

The hWAs-iCas9 cell line was established in our lab by integrating the doxycycline-inducible Cas9 cassette into the hWAs genome. This cell model was used for elucidating the function of *MTIF3* in human white adipocytes in **paper I**.

HepG2 cell line was purchased from Sigma-Aldrich. This cell line was originally isolated from a hepatocellular carcinoma of a 15-year-old, white, male youth [111]. Here, it was used for genomic editing and assessment of hypothetical causal genetic variation in **paper IV**.

Luciferase reporter assays for causal variant prioritization

A luciferase reporter assay typically requires an expression plasmid that carries luciferase-encoding cDNA (reporter) and a proximal regulatory element. In **paper I**, for prioritization of causal variants, hypothetical regulatory elements were derived from genomic DNA sequences that surround genetic variants in the haplotype block of interest. Thus-constructed plasmids are then transfected into cells, whereby the

regulatory elements recruit cellular transcriptional machinery to express the luciferase reporter gene. The luciferase expression level in the cells is dependent on the regulatory effects of the DNA element inserted (e.g. enhancer), and can be quantified with a bioluminescence assay. Different luciferase expression levels suggest different functional impacts of the interrogated regulatory elements, and potentially also of the associated genetic variants [112]. An illustration of how a luciferase reporter assay works is shown in Figure 1.

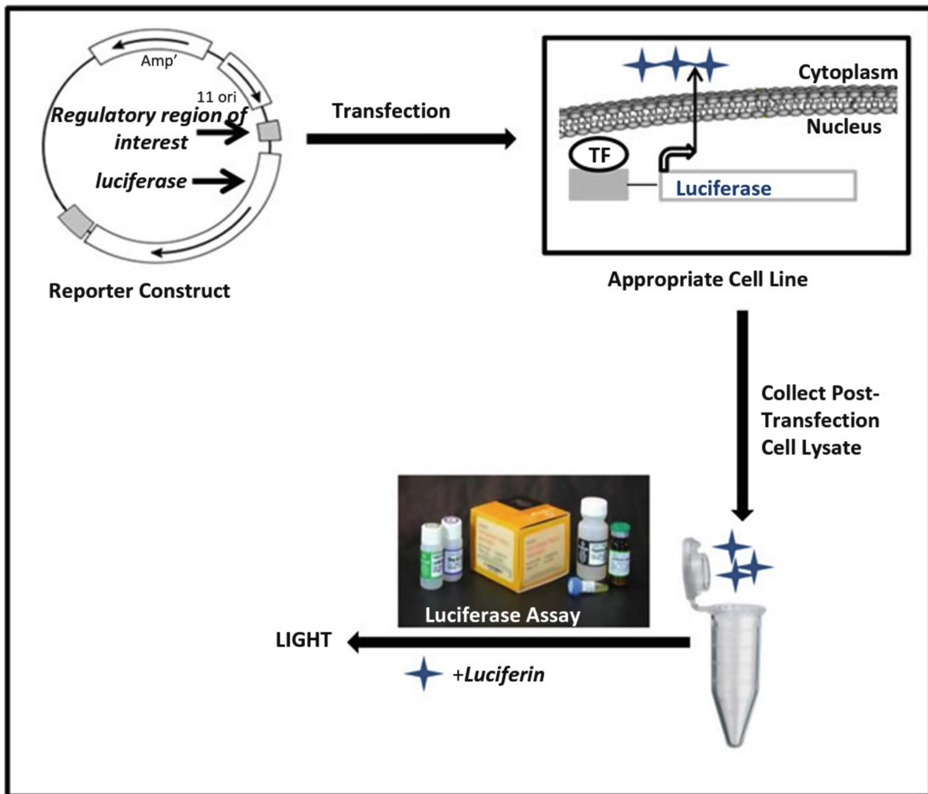


Figure 1. An illustration of a luciferase reporter assay, the figure was obtained from [112].

sgRNA and HDR donor template design for CRISPR/Cas9-mediated allele editing

To simplify the genomic editing process and to improve its efficiency, we used ribonucleoprotein (RNP) complexes [113] instead of CRISPR/Cas9 plasmids in our experiments. To design the editing assays, we first used the UCSC genome browser (https://genome-euro.ucsc.edu) to retrieve the human reference genome sequences surrounding the genetic variants of interest. The sequences were parsed into Primer3 web tool (https://primer3.ut.ee/) to design primers for PCR amplification and later Sanger sequencing. The sgRNAs and HDR donor templates used for CRISPR/Cas9 allele substitution were designed manually by, in general, finding the shortest distance between the possible CRISPR/Cas9 cut sites and the alleles to be edited.

Limiting dilution cloning for clonal screening

After the initial CRISPR editing reaction, the cell population is a mixture of scarlessly allele-edited cells and cells with random mutations. To isolate the properly edited cells, we, therefore, utilized single-cell cloning. The illustration of the workflow is as below:

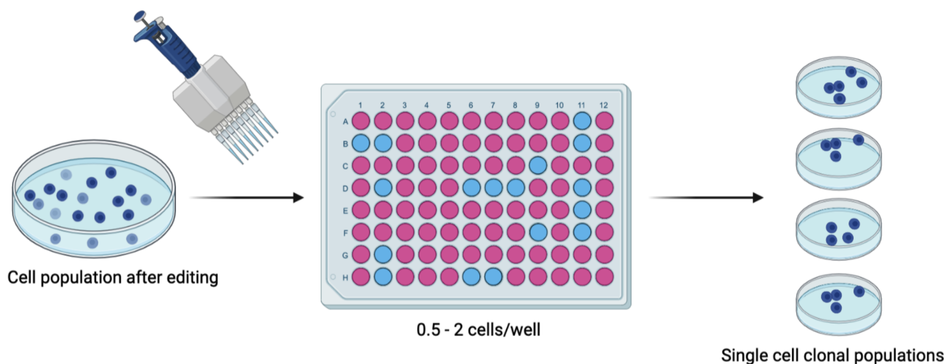


Figure 2. The workflow of limiting dilution cloning. The CRISPR-edited cell population is diluted to 5 - 20 cells/ml medium and seeded into 96-well plates at 100 µl/well. The cells in the plate are then allowed to proliferate for 3 weeks and apparent single clonal populations are genotyped. (This illustration is made with Biorender, https://app.biorender.com/).

RT-qPCR

To quantify gene expression in our allele-edited cells, in all four papers, we used reverse transcription quantitative real-time PCR (RT-qPCR). The RT-qPCR assay includes two procedures. First, mRNA extracted from samples is reverse-transcribed into cDNA by reverse transcriptase; second, the synthesized cDNA is used as a template for PCR amplification, in the presence of either dyes or probes for quantification of the amplification products. In our study, we used the Taqman probe-based method for the RT-qPCR assay. The reason for using Taqman assays is, compared with the dye-based method (SYBR-Green), higher accuracy and sensitivity, and the optional multiple-target detection.

The principle for Taqman assay is described in Figure 3 [114]. A Taqman probe is a short DNA oligonucleotide complementary to the cDNA of the PCR-amplified target, carrying a fluorescent reporter (reporter) and a quencher at opposite ends. An intact probe does not emit fluorescence, as the reporter and the quencher remain in proximity. However, during an ongoing real-time PCR reaction, the probe that binds to its target is cleaved, whereby the quencher inhibition of the fluorescent reporter is released, increasing the fluorescence in the sample.

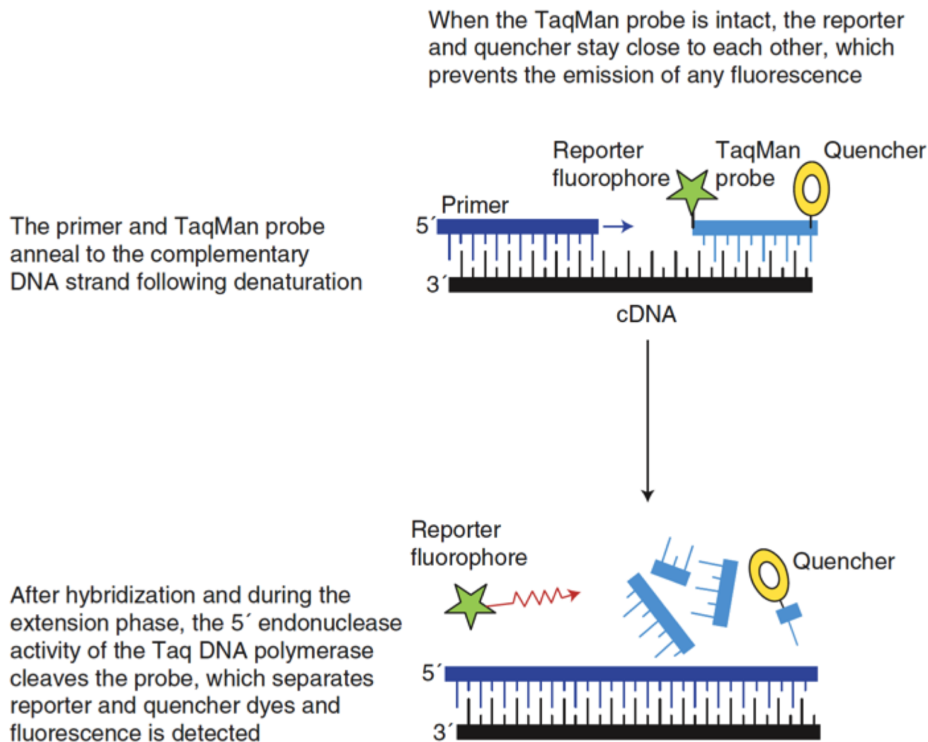


Figure 3. The principle of Taqman assay, the figure is from [114].

Mitochondrial respiration profiling using Seahorse analyzer

Mitochondria play a crucial role in regulating adipocyte glucose and lipid metabolism, and their dysfunction has been demonstrated to contribute to the development of obesity and T2DM [115]. In **papers I**, **II,** and **III**, we found that the genetic variants studied could regulate the expression of genes involved in mitochondrial function. To measure cellular respiration in our allele-edited cells, we employed a Seahorse XF analyzer (Agilent Technologies, Santa Clara, California, USA) which is a sensitive instrument for directly monitoring cellular metabolic functions. Seahorse analyzer provides label-free measurements of the oxygen consumption rate (OCR) and the extracellular acidification rate (ECAR) in a microchamber with about 2 µl medium 200 µm above a monolayer of cells, all in a microplate. OCR measurements trace the mitochondrial oxidative phosphorylation, as oxygen is consumed as the final electron acceptor in aerobic metabolism. ECAR

represents the glycolytic (lactate or lactic acid production) and non-glycolytic (CO_2 production through the TCA cycle) acidification, but mainly reflects glycolytic metabolism.

During a Seahorse assay, different drugs or substrates are added to allow evaluation of the different aspects of mitochondrial respiration. For example, to perform a mitochondrial stress test, oligomycin, carbonyl cyanide 4 (trifluoromethoxy) phenylhydrazone (FCCP), rotenone, and antimycin A are sequentially added during the assay. Oligomycin inhibits ATP synthase (complex V) resulting in decreased OCR, which reveals how much of the respiration is driving ATP synthesis in the cells. Here, the residual OCR can be attributed to proton leak maintaining minimal Electron transport chain (ETC) and non-mitochondrial respiration. FCCP, an uncoupler of oxidative phosphorylation, disrupts the proton gradient across the mitochondrial membrane, showing maximal oxygen consumption capacity. Rotenone and antimycin A inhibit Oxidative phosphorylation (OXHPOS) complex I and III respectively, completely stopping ETC activity. Taken together, by extracting the OCR data from the respiration trace, we can calculate ATP production OCR, maximal respiration OCR, proton leak OCR, and non-mitochondrial respiration OCR accordingly (an illustration is shown in Figure 4).

Other than the above-mentioned types of drugs can be used for other specific purposes. For instance, to study brown fat thermogenesis, we used norepinephrine (NE) to activate UCP1 (**paper III**), and to study the effects of high glucose on white adipocyte respiration, we used glucose as an additive during Seahorse assays (**Paper II**).

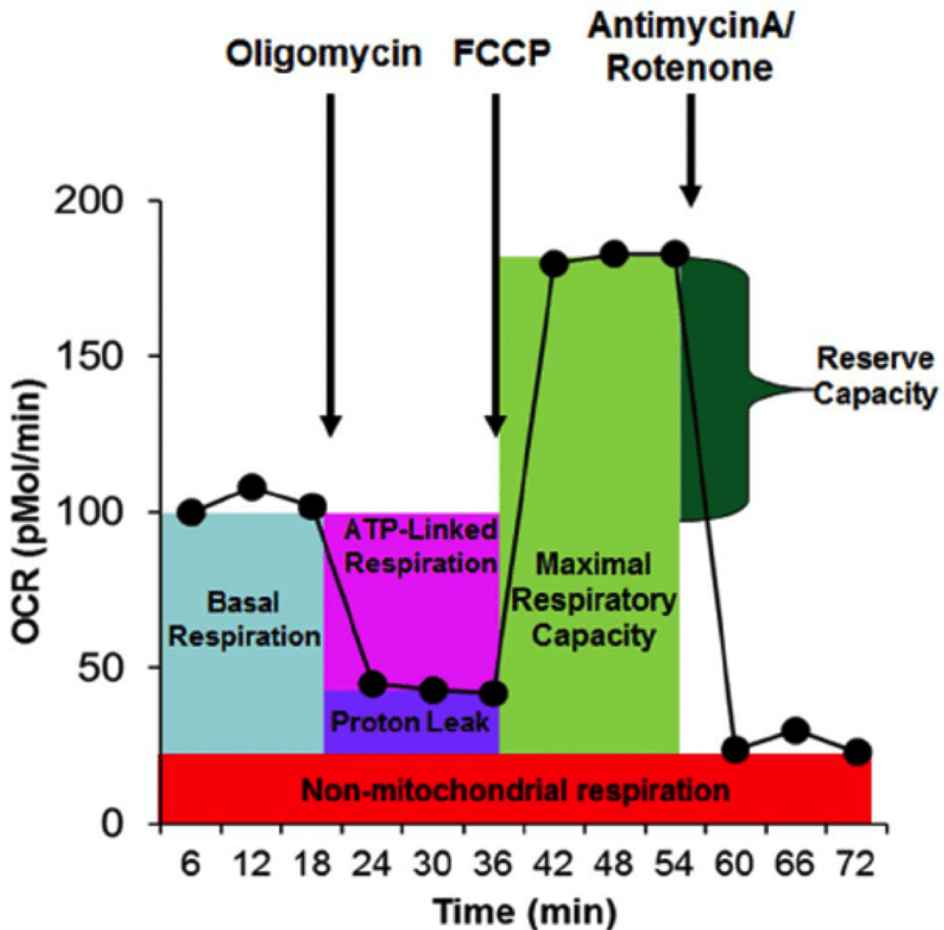


Figure 4. An illustration showing how the different respiration components are extracted from a Seahorse assay-derived OCR trace. The figure is from [116].

SDS-PAGE and Western Blotting

For protein expression detection, SDS-PAGE followed by immunoblotting (also called Western Blotting) has been extensively used since it was first described in 1981 [117]. In the thesis, this technique was used in **papers I**, **II,** and **III** for quantification of target proteins.

SDS-PAGE is used to separate denatured protein mixtures. As the protein samples are prepared with SDS, the protein molecules become negatively charged. When the samples are electrophoresed on polyacrylamide gels the proteins migrate toward the

positively charged electrode. Because of the size differences of the proteins, the light proteins will travel faster than the heavy ones, resulting in a size-dependent separation of the proteins in the gel.

The proteins on the gel can be electrophoretically transferred to a hydrophilic polyvinylidene fluoride (PVDF) membrane, which captures the proteins through hydrophobic interactions. There are different types of transfer methods, including tank transfer (wet transfer) and rapid-transfer systems. Both methods were used in this thesis, the choice is based on the expression level and the size of the target protein. After transfer, the membrane can be incubated with milk or BSA-blocking buffer to avoid non-specific binding of the antibodies. The membrane is then incubated with specific primary and secondary antibodies, and visualized for protein quantification (see details in the methods parts in each paper).

Blue Native PAGE and immunoblotting

In **paper I**, we applied Blue native polyacrylamide gel electrophoresis (BN-PAGE) to study the function of MTIF3 in mitochondrial OXPHOS complex assembly in adipocytes. BN-PAGE was developed in the early 1990s by Schägger and Von Jagow [118] and is based on the use of mild neutral detergents to extract OXPHOS complexes, followed by electrophoretic separation of the complexes. Electrophoresis is performed in native buffer conditions to maintain the interactions between the subunits of each complex.

After BN-PAGE, to quantify the amount of each OXPHOS complex, the proteins can either be stained with Coomassie Blue or transferred to a PVDF membrane for immunodetection using antibodies against specific complex subunits. Yet another option is to use in-gel activity assays, which employ substrates for catalytic staining of different complexes (e.g. diaminobenzidine for CIII and CIV) [119]. In **paper I**, we used immunoblotting to visualize the assembly of OXPHOS complexes (a detailed method description can be found in the paper).

Oil red O staining

Oil red O is an inexpensive fat-soluble dye commonly used to stain lipids in tissues or cells [120]. Oil red O is dissolved in alcohol, but when applied to biological material it migrates from the solution to the cellular fat because of its greater solubility in lipids. In this thesis, we used Oil red O to stain the differentiated adipocytes in **papers I**, **II,** and **III** (detailed staining protocols can be found in the papers).

Chapter 3 – Aims of the Thesis

The major aim of this thesis was to apply novel experimental approaches to pinpoint the genetic variants that underpin obesity and T2DM. We also aimed to investigate the biological mechanisms through which genetic variation interacts with environmental factors. Knowledge of how such interactions contribute to the development of obesity and T2DM can provide support for genotype-based precision medicine and prevention. To accomplish that, we used various functional genomics approaches, including genome editing (CRISPR/Cas9) of human adipocyte and hepatocyte cell lines. Below, the aims of each paper are briefly described:

Paper I

The main aim of this study was to identify the molecular basis behind the epidemiological findings wherein the rs1885988 variant was associated with body weight change outcome during specific lifestyle interventions. The specific aims were:

1. To map the causal genetic variation in the rs1885988 haplotype block and determine its target gene in human white adipocytes.
2. To study the role of MTIF3 in white adipocyte biology using a gene knockout cell model.
3. To investigate the biological mechanism through which MTIF3 affects adipocyte function under physiological or glucose-restricted conditions (lifestyle mimetics).
4. To establish the molecular mechanism that links causal genetic variation with lifestyle intervention-mediated weight loss outcome in previously published observational studies.

Paper II

Genetic variation at rs8192678 has been extensively associated with obesity and T2DM in various cohorts, but the functional basis at the molecular level is largely missing. In this study, we aimed to gain more understanding of how rs8192678 affects the adipocyte phenotype in an experimental setting. The specific aims were:

1. To apply CRISPR/Cas9 tool kit to perform allele editing at rs8192678 in a human white pre-adipocyte cell line.
2. Using the rs8192678 allele-edited cell lines to investigate the functional consequences caused by the allele switch in human white adipose cells.
3. To study the biological mechanisms of how the rs8192678 variant affects PGC-1α protein activity in white adipose cells.
4. To establish the causal role of rs8192678 in regulating the metabolic activity in human white adipocytes.
5. To elucidate the mechanisms behind the association between rs8192678 and various metabolic outcomes relating to obesity and T2DM.

Paper III

Motivated by a previous epidemiological finding that the *PPARGC1B* rs10071329 variant is associated with adiposity in humans, we aimed to test if rs10071329 is a causal variant that affects adipocyte biology. The specific aims were:

1. To conduct a genotype-based cohort study to replicate the previously published association between rs10071329 and adiposity.
2. To generate human brown isogenic pre-adipocyte cell lines for rs10071329 using CRISPR/Cas9.
3. To assess whether rs10071329 causally affects *PPARGC1B* expression in differentiating brown adipocytes, using the allele-edited cells.
4. To differentiate the allele-edited cells, and evaluate the brown adipocyte-specific functional consequences caused by the rs10071329 allele switch. This aims to explore the molecular mechanisms that underlie the association between rs10071329 and adiposity.

Paper IV

In this study, we aimed to validate if the rs2289669 variant, or the closely linked genetic variants, causally affect *SLC47A1* expression, which could shed light on studies that demonstrated an association of rs2289669 with the glucose-lowering efficacy of metformin. The specific aims were:

1. To perform genomic editing using CRISPR/Cas9 in a human liver cancer cell line (HepG2) thereby generating cell lines homozygous for the different alleles at rs2289669 and the closely linked rs8065082.
2. To use the allele-edited cell lines to study if rs2289669 or rs8065082 affect *SLC47A1* expression or splicing.
3. To provide a mechanistic explanation of the previously observed association between rs2289669 or rs8065082 and metformin efficacy.

Chapter 4 – Results and Discussions

Paper I

Previous studies have robustly associated genetic variants at the *MTIF3* (Mitochondrial Translational Initiation Factor 3) locus with obesity and body weight regulation in humans [56]. However, before genetic variation can become a reliable toolkit in precision medicine, the question of genetic causality needs to be answered. In this study, we applied methods like luciferase reporter assay and CRISPR/Cas9 genomic editing in a human white adipocyte cell line to investigate whether the reported variants are truly causal for the observed associations, and whether they target *MTIF3* expression. We also investigated the biological mechanisms that potentially underlie these associations.

rs67785913 is a causal regulatory variant for *MTIF3* expression

Firstly, through lookups in the GTEx database, we found that individuals who carry the C allele at the rs1885988 locus have increased levels of *MTIF3* expression in subcutaneous fat. And we also found *MTIF3* has a high level of expression in adipose tissue. This finding suggests that genetic variants within the rs1885988 haplotype block may regulate *MTIF3* expression in adipose tissue. To prioritize putative causal variants experimentally, we segmented all tightly rs1885988-linked (r^2>0.8) SNPs, and their surrounding genomic sequences, into 12 DNA fragments. We then inserted these fragments into luciferase reporter plasmids, and measured their impact on luciferase expression in hWAs adipocytes. We found that the genomic fragment containing rs67785913 had the highest capability to increase luciferase transcription, when compared with other fragments and with minimal promoter-driven luciferase expression. This data indicated the potential transcription-enhancing role of the rs67785913 locus. The luciferase assays, due to their plasmid-based design, do not reveal any specific gene targets of the discovered enhancer sequences. Also, at that point, it was unclear whether the enhancement of target gene expression could be dependent on the different alleles at rs67785913.

We, therefore, employed the CRISPR/Cas9 toolkit to substitute the prominent CT allele at rs67785913 with the alternative CTCT allele in the hWAs cell line. Because differential *MTIF3* and *GTF3A* gene expression are both associated with rs67785913 in adipose tissue and muscle (in the GTEx database), we examined the

expression of both genes in the allele-edited cells before and after adipogenic differentiation. In all experimental conditions, the rs67785913 CTCT/CTCT genotype conferred a higher level of *MTIF3* expression, and unchanged *GTF3A* expression, when compared with the rs67785913 CT/CT genotype. Taken together, our data demonstrated that the rs67785913 variant is probably the causal eQTL for *MTIF3* expression. This finding could explain the correlation found between the tightly linked rs1885988 and *MTIF3* expression in adipose tissue.

Generating inducible Cas9-expressing pre-adipocyte cell line (hWas-iCas9)

Once we established that rs67785913 confers differential *MTIF3* expression, we then attempted to examine the biological consequences caused by the rs67785913 allele substitution. However, after differentiation, we found that the allele-edited cells exhibited large intra-group variation in adipogenesis and mitochondrial content, probably due to clonal variation introduced during single-cell cloning of allele-edited cells. This hampered further use of these cells in functional studies of adipocyte metabolism, as large variations in baseline parameters (e.g. differentiation capacity, triglyceride, and mitochondrial content) would confound the results. To further study the function of *MTIF3* in adipocytes, we, therefore, sought to establish a cell line that allowed us to generate gene knockouts at any stage during adipocyte differentiation, without the necessity of single-cell cloning. We thus established a doxycycline-inducible Cas9-expressing human pre-adipocyte cell line (hWAs-iCas9), in which endogenously expressed Cas9 can form complexes with synthetic single guide RNAs (sgRNAs) that can be transfected into cells to generate gene knockouts. This cell model allowed us to generate consistently efficient *MTIF3* knockouts in mature adipocytes, but even more importantly it provides a relatively accessible method for generating other gene knockouts. Another advantage of this model is that it circumvents the limitations of using pre-adipocyte cell lines where cell passage-dependent altered differentiation capacity [121] and differential genetic effects [122] could easily introduce confounders. The established hWAs-iCas9 cell model may therefore be very useful for the adipocyte research community, as it allows e.g. metabolism studies that require similar baseline parameters, such as triglyceride or mitochondrial content.

MTIF3 knockout disrupts mitochondrial DNA-encoding gene and protein expression, mitochondrial content, as well as mitochondrial OXPHOS assembly in hWAs-iCas9 adipocytes

MTIF3 encodes Mitochondrial Translation Initiation Factor 3, and is important for the expression of proteins in mitochondria. MTIF3 facilitates the initiation of complex formation on mitochondrial 55S ribosomes in the presence of MTIF2, fMet-tRNA, and poly(A, U, G) [123]. In humans, there are 13 proteins encoded by the

mitochondrial genome, all forming parts of the OXPHOS complex I, III, IV, and V [124]. Loss of *MTIF3* can decrease the assembly of OXPHOS complexes in muscle and heart in the mouse [89], and decrease the translation of *ATP6* mRNA in the HepG2 cells [90]. In adipose tissue, mitochondria play an essential role not only by making ATP but also by triggering cellular signaling pathways by OXPHOS respiration products [125]. However, the impact of MTIF3 on white adipocyte metabolism has not been investigated, but is of interest because of the association between the *MTIF3* locus and body weight regulation. In our study, using the hWAs-iCas9 cell model, we examined the effects of *MTIF3* ablation on differentiated hWAs adipocytes. Firstly, compared with the wildtype control cells, we found that the *MTIF3* knockout cells had lower expression of COX II (a subunit of OXPHOS complex IV), ND2 (a subunit of OXPHOS complex I), and trending lower expression of CYTB (a subunit of OXPHOS complex III). We did not find any changes in the expression of ATP8 (a subunit of OXPHOS complex V). Secondly, using RT-qPCR, we observed that several mitochondrial DNA-encoding genes' expression was significantly altered in *MTIF3* knockout cells. Specifically, we found increased expression of *MT-ND1* and *MT-ND2*, a trending increase of *MT-ND4*, and lower expression of *MT-ND3*, and *MT-CO3*. This imbalanced mitochondrial gene expression could reflect a feedback regulation caused by mitochondrial dysfunction after the loss of MTIF3. In addition, we also found that mitochondrial content was significantly reduced in the *MTIF3* knockout adipocytes. Taken together, our data suggested that *MTIF3* knockout disrupts mitochondrial DNA-encoding gene and protein expression, as well as mitochondrial biogenesis.

Furthermore, we examined if MTIF3 plays a role in mitochondrial OXPHOS complex assembly in adipocytes. We found that *MTIF3* knockouts assembled significantly less complex III2/IV2 and IV1, but, interestingly, assembled more complex II. We also found several undefined bands in the *MTIF3* knockouts, possibly suggesting complex disassembly or degradation.

Taken together, our data showed that MTIF3 deficiency in adipocytes results in disrupted mitochondrial gene and protein expression, as well as disassembly of mitochondrial respiration OXPHOS complex, along with decreased mitochondrial content. These data reveal the molecular mechanisms of how *MTIF3* affects mitochondrial respiration in adipocytes.

MTIF3 knockout affects mitochondrial respiration and endogenous fatty acid oxidation in hWAs-iCas9 adipocytes

To study the functional consequences caused by the disrupted mitochondrial OXPHOS complex expression and assembly, we employed a Seahorse analyzer to assess the oxygen consumption rate (OCR) in *MTIF3* knockout adipocytes. As expected, we found that *MTIF3* knockout cells exhibited significantly lower basal

respiration OCR, ATP production OCR and a trending decrease in maximal respiration OCR.

Moreover, we also found that endogenous fatty acid oxidation decreased significantly, which was evaluated by treating the cells with etomoxir in the Seahorse assay. Etomoxir is an inhibitor of carnitine palmitoyltransferase (CPT)-1 and diglyceride acyltransferase (DGAT) activity in mitochondria, which are two key enzymes involved in fatty acid oxidation in mammalian cells. Therefore the decreased OCR found in etomoxir-treated adipocytes in the Seahorse assay reflects the mitochondrial fatty acid oxidation respiration.

Furthermore, we detected the expression of genes related to mitochondrial function, fatty acid oxidation, lipolysis, and lipid catabolism. We found that *MTIF3* knockout cells had decreased expression of the mitochondria-related *MT-CO1*, and the fatty acid oxidation-related *ACADM* and *ACAT1*.

Collectively, our data shows that MTIF3 loss in adipocytes can lead to truncated mitochondrial function, and decreased fatty acid oxidation and ATP production in human white adipose tissue. These consequences have frequently been associated with obesity [126,127], and while other factors could also underpin these phenotypes, our findings suggest that low MTIF3 content, which could result from specific genetic variation, might be a contributing cause to the development of obesity.

MTIF3 knockout affects hWAs-iCas9 adipocyte triglyceride content after glucose restriction challenge

Mitochondrial function in white adipocytes has been inversely correlated with weight change [128]; we therefore hypothesized that *MTIF3* expression in adipocytes could affect a dietary intervention outcome. We tested this hypothesis *in vitro* by exposing hypertrophic adipocytes to glucose restriction (which limits glucose usage by the cells) and evaluated its effects on triglyceride content. Compared with 25 mM glucose medium, we observed that different levels of glucose restriction led to decreased triglyceride content in *MTIF3* knockout and control cells. Interestingly, when cultured in 1 mM glucose medium for 3 days, the control cells showed a more extensive triglyceride content decrease than the *MTIF3* knockout cells ($p=0.053$, $n=4$). These data might suggest that MTIF3-deficient adipocytes have limited capacity for lipid metabolism under glucose restriction.

In summary, in this study, using different functional genomics approaches, we established a causal link between rs67785913 and *MTIF3* expression, as well as showing the molecular mechanisms through which MTIF3 affects mitochondrial function and lipid metabolism in a human adipocyte cell model. Our results thus help establish the link between common genetic variation in *MTIF3* and the mechanisms related to the regulation of human body weight. Additionally, our data may also imply that individuals possessing cis-eQTLs leading to higher *MTIF3*

expression may achieve greater benefits from lifestyle interventions aimed at body weight management.

Key findings:

- rs67785913 is a causal genetic variant in the rs1885988 haplotype block, with the CTCT allele conferring higher *MTIF3* expression in white adipocytes.
- The newly established hWAs-iCas9 cell line allows an efficient generation of *MTIF3* knockouts in mature adipocytes, and could be useful for the generation of various gene knockouts in other projects.
- MTIF3 plays a vital role in regulating mitochondrial respiration, fatty acid oxidation, and lipid metabolism in human white adipocytes.
- MTIF3 ablation in adipocytes hampers triglyceride content reduction during glucose restriction, which might be explained by decreased fatty acid oxidation.
- Mechanistically, MTIF3 controls mitochondrial gene and protein expression, as well as affects mitochondrial OXPHOS complex assembly.
- The findings may support the notion that people carrying the rs67785913 CTCT allele may have higher *MTIF3* expression, and could therefore benefit more from lifestyle interventions aiming at body weight loss.

Paper II

PGC-1α protein is encoded by the *PPARGC1A* gene, which plays a central regulatory role in energy metabolism and mitochondrial function [81]. Genetic variation in *PPARGC1A* has been extensively examined in association with metabolism-related traits [77]. Among the variants, a common polymorphism, rs8192678 (C/T, Gly482Ser), showed a robust signal. However, the direct allele-specific functional effects of rs8192678 are still unknown; in this study, therefore, we tried to fill in this knowledge gap. We applied the CRISPR/Cas9 genome editing tool to generate rs8192678 C/C, C/T, and T/T genotypes in an isogenic human pre-adipocyte white adipose tissue (hWAs) cell line, and conducted functional studies to examine rs8192678 allele-dependent effects on white adipocyte differentiation and mitochondrial function.

CRISPR/Cas9-mediated rs8192678 editing in hWAs pre-adipocyte cell line

rs8192678 is a missense variant in *PPARGC1A*, the C allele encodes Gly and the T allele encodes Ser. Interestingly, this variation is the only common polymorphism in its haplotype block. In this study, motivated by numerous epidemiological studies on the association between rs8192678 and metabolic disorders-related traits [77,93,129-131], we dedicated our efforts to dissect the functional role of rs8192678 in white adipocytes. Since the hWAs cells are heterozygous at rs8192678, we designed sgRNAs and donor templates for CRISPR/Cas9 editing to establish rs8192678 C/C, T/T, and C/T clonal cell lines. After clone screening, we obtained eight C/C clonal populations, eight T/T clonal populations, and six C/T clonal populations, as confirmed by Sanger sequencing. These edited clones were used in further functional assessments.

The rs8192678 polymorphism regulates hWAs differentiation, lipid accumulation, and lipogenesis

Next, we studied the effects of rs8192678 on adipocytic phenotypes including white adipocyte differentiation, lipid accumulation, and lipogenesis. We observed that rs8192678 T/T and C/T cells had higher differentiation capacity than C/C cells, as assessed by BODIPY staining. Consistently, the triacylglycerol content was markedly higher in T/T and C/T cells than in C/C cells.

When evaluating lipogenesis in the allele-edited cells, we found that T/T cells had higher lipogenesis ($p=0.05$) than C/C cells at basal condition. Interestingly, after insulin stimulation, lipogenesis increased in all three genotypes and the T allele showed an apparent additive dose-dependent effect.

Moreover, to further confirm these phenotypes, we detected the expression of gene markers involved in white adipocyte metabolism. The white adipocyte adipogenesis markers *PPARG*, *ADIPOQ, CEBPA,* and *SLC2A4* were significantly more expressed in T/T vs C/C cells. T/T cells also showed higher expression of *ACACB* (a lipid catabolism marker gene), *FASN, SCD, SREBF1,* and *FABP4* (lipogenesis marker genes), and *ABHD5* and *PNPLA2* (lipolysis marker genes) than C/C cells. Collectively, our data showed that rs8192678 affects white adipocyte differentiation and lipid metabolism.

The role of *PPARGC1A* in adipocyte differentiation remains inconclusive despite several studies. In humans, *PPARGC1A* expression was shown to continuously increase during subcutaneous adipocyte differentiation, and *PPARGC1A* knockdown was detrimental for human brown adipocyte differentiation [132,133]. In mice, although *Ppargc1a* deficiency led to reduced body fat mass despite being fed with a normal or high-fat diet [134], the white pre-adipocyte cell differentiation was not affected when cultured *ex vivo* [135]. Similarly, *Ppargc1a* knockout did not affect mouse brown adipocyte differentiation, but it decreased thermogenic gene

expression [136]. Taking our results into account *PPARGC1A* appears to regulate adipocyte differentiation in a species- and tissue-dependent manner. Our study may also raise concern that when seeking to determine the functional role of human genetic polymorphisms, researchers may not readily draw conclusions from results obtained from animal genetic knockout models.

The rs8192678 polymorphism affects mitochondrial respiration during white adipocyte differentiation

Previous studies have shown that PGC-1α regulates mitochondrial biogenesis [137]. Given the crucial role of mitochondria in adipocyte differentiation [138], we tested the effects of rs8192678 on mitochondrial respiration during adipogenic differentiation in the allele-edited cells. Using the Seahorse extracellular flux analyzer, we found that T/T and C/T adipocytes had higher basal respiration and ATP production OCR than C/C adipocytes. T/T adipocytes also showed higher maximal respiration capacity than C/C and C/T adipocytes.

To further confirm the effects of the rs8192678 variant on mitochondrial function, we measured the mRNA expression of genes involved in mitochondrial function and found significantly higher transcript levels of *TOMM20*, *MT-CO2,* and *CS* in T/T cells than in C/C cells. Taken together, our data suggest that the rs8192678 polymorphism regulates white adipocyte differentiation by influencing mitochondrial oxidative respiration. The improved mitochondrial function in T/T cells may also explain the associations where the T allele carriers appear to benefit more from various weight loss interventions like energy-restricted diets [139], bariatric surgery [80], and acarbose treatment [78].

The rs8192678 polymorphism regulates *PPARGC1A* mRNA expression, PGC-1α protein content, and degradation, as well as PPARγ transcriptional activity

Next, to understand the molecular mechanisms underlying the effect of rs8192678 variation on adipogenesis, we examined cell proliferation, *PPARGC1A* mRNA stability, PGC-1α protein expression, and degradation. We found that rs8192678 does not affect pre-adipocyte differentiation. We also found that at the end of adipogenic differentiation, the T/T cells had more *PPARGC1A* mRNA than C/C cells, but the endogenous *PPARGC1A* mRNA degradation rate was the same as in C/C cells. In terms of protein degradation, we found that the PGC-1α protein contents in T/T adipocytes decreased rapidly after 1 h cycloheximide treatment (that blocks protein translation) and remained at a low level after 2 h, but were largely unchanged in C/C adipocytes at all treatment timepoints. This result is consistent with previous findings in HepG2 and Ins-1 cells [85,86]. Interestingly, we also found that PGC-1α content in C/C cells is significantly lower than in T/T cells at steady state, which is coherent with the *PPARGC1A* mRNA expression difference.

PGC-1α is a coactivator of PPARγ (Peroxisome proliferator-activated receptor γ) which is a master regulator of adipocyte differentiation [140]. Because we found that T/T cells had higher PGC-1α content, we reasoned that their higher differentiation capacity might be due to an increased PPARγ transcription activity. To test this hypothesis, we assessed endogenous PPARγ transcriptional activity by transfecting the PPRE–luciferase reporter plasmid into C/C and T/T cells. We found that T/T cells had higher PPRE–luciferase expression, indicating a higher PPARγ transcriptional activity. We further validated these findings by checking PPARγ downstream target gene expression using RT-qPCR. As expected, *LPL*, *FABP4*, *CEBPA*, *ADIPOQ,* and *FASN* were significantly more expressed in T/T than in C/C cells. Collectively, the data suggests that during adipogenic differentiation, the T allele confers higher PPARγ transcriptional activity than the C allele.

In summary, we established the following mechanisms of rs8192678 allele-dependent biological processes. rs8192678 482Ser (T allele) confers a higher level of *PPARGC1A* mRNA and PGC-1α content, a faster turnover of PGC-1α protein, and higher PPARγ transcriptional activity during white adipocyte differentiation. Judging from the current results, the higher differentiation capacity in the rs8192678 T/T cells may well be explained by the higher PGC-1α content modulating PPARγ activity, but the faster PGC-1α protein turnover rate may also play a role. As a coactivator in energy metabolism, the half-life of PGC-1α is short [141] and tightly controlled by 20S– and ubiquitin–proteasome-mediated degradation [142,143]. This might suggest the need for tight regulation of PGC-1α, to quickly respond to shifting metabolic demands, and to decrease the interference with shifting metabolic pathways (e.g. adipogenesis vs lipolysis) [140].

Our experimental data shows that the rs8192678 genetic variation can cause different white adipocyte phenotypes, and may provide some explanation on why T allele carriers had higher body fat mass and BMI, and excessive weight gain, as well as better response to certain lifestyle interventions [77,93,129-131,144]. However, our data fails to explain why the T allele carriers have a higher risk for type 2 diabetes. To answer this question, more data are needed on how rs8192678 affects other tissues as well as whole-body metabolism.

Key findings:

- We established rs8192678 allele-edited human white pre-adipocyte cell lines on isogenic background, and studied the allele-dependent effects on adipocyte-specific biological processes.
- rs8192678 can regulate human white adipocyte differentiation, with the T allele conferring higher adipogenesis capacity.
- In differentiating adipocytes, cells carrying the rs8192678 T allele showed higher *PPARGC1A* mRNA and PGC-1α protein content, as well as more dynamic and higher PPARγ transcriptional activity.

- Our findings may provide guidance for genotype-based precision medicine in obesity treatment, but the proof-of-concept remains to be shown in genotype-based clinical trials.

Paper III

Genetic variation in *PPARGC1B* has previously been associated with adiposity, but the experimental validation is largely lacking. rs10071329 is a non-coding variant in the *PPARGC1B* locus and has been linked with lifestyle intervention-mediated subcutaneous adiposity [28]. However, if this variant is the actual cause of differential *PPARGC1B* expression, and how this variant affects adiposity, is yet to be answered. In the current study, we used CRISPR/Cas9 to scarlessly edit the alleles of rs10071329 in the human brown adipocyte cell line (hBAs). Using the generated homozygous A/A and G/G cell lines, we examined *PPARGC1B* expression and mitochondrial function, in order to establish the rs10071329 allele-dependent effects on brown adipocyte biology.

Epigenetics lookups for rs10071329

The rs10071329 SNP has no common variants in tight linkage disequilibrium (r^2>0.8). When checking for chromatin features at rs10071329 in WashU Epigenome Browser (http://epigenomegateway.wustl.edu/browser/), we observed that the rs10071329 locus resides downstream of *PPARGC1B* and upstream of the *PDE6A* gene, and shows high DNase I hypersensitivity, H3K27Ac, and H3K4me1 signals in human adipocytes and subcutaneous fat. These features drove us to hypothesize that rs10071329 may be a causal eQTL for gene expression regulation in human adipose tissue.

Establishment of rs10071329 allele-edited human brown adipocyte cell lines (hBAs)

Next, we performed rs10071329 allele editing in a human brown pre-adipocyte cell line hBAs using CRISPR/Cas9 toolkit. The hBAs rs10071329 A/A homozygous cells were electroporated with sgRNA/Cas9 RNP complex and DNA donor templates to accomplish a genotype switch, either from A/A to G/G, or mock A/A to A/A. After the editing, the single-cell clones were expanded through limiting cloning, and the homozygous scarlessly edited cells were identified by Sanger sequencing. We obtained 3 correctly edited A/A and 3 G/G clonal cell lines at the rs10071329 locus, out of 95 and 90 screened single-cell clones, respectively. These cell lines were validated for their capacity to differentiate into brown adipocytes,

and for their lipolytic response to norepinephrine, and were further used in our functional studies.

rs10071329 locus regulates *PPARGC1B* expression in hBAs

To examine whether the rs10071329 editing affects potential target gene expression, we differentiated A/A and G/G hBAs cells to mature adipocytes and detected *PPARGC1B* and *PDE6A* gene expression by RT-qPCR. At all examined differentiation days 7, 12, and 28, the G/G cells showed consistently higher expression of *PPARGC1B*, and the expression of *PDE6A* was below the limit of detection in both A/A and G/G cells. Our results experimentally validate that rs10071329 is an eQTL for *PPARGC1B* expression in human brown adipocytes.

Electrophoretic mobility shift assay (EMSA) on rs10071329-derived DNA probes

To gain further insight into how the rs10071329 variant affects *PPARGC1B* expression, we performed an EMSA assay to examine the binding ability of nuclear proteins to genomic sequence-based DNA probes with rs10071329 A or G alleles, respectively. We observed that the G allele probes bind a larger quantity of at least one nuclear protein than the A allele probes. This result confirms the differential role of rs10071329 alleles on nuclear protein binding, which may further support its role as a regulatory variant. However, the identity of the binding proteins remains to be determined.

rs10071329 regulates mitochondrial respiration and mitochondrial gene expression in hBAs adipocytes

The role of the target gene of rs10071329, *PPARGC1B,* in regulating mitochondrial function in brown adipocytes has been previously established. Adipocyte-specific *Ppargc1b* knockout mice showed disrupted brown adipocyte lipid metabolism, thermogenesis, adrenergic stimulated lipolysis, as well as interactions between mitochondria and lipid droplets [94,95]. In addition, the *Ppargc1b*-deficient brown adipocytes, which were isolated from mice exposed to cold (21°C), showed a clear whitening phenotype, with large triglyceride-filled vacuoles and reduced mitochondrial gene and *Ucp1* expression [145]. In this project, we focused on mitochondrial function as the functional consequence caused by rs10071329 allele substitution. In human brown adipocytes, mitochondria are the main source of ATP, but when activated by e.g. norepinephrine or cold exposure, they shift to adaptive thermogenesis mediated by uncoupling protein-1 (UCP1) [36]. We found a higher basal mitochondrial respiration in G/G than in A/A cells. After the addition of norepinephrine, which activates thermogenesis in brown adipocytes [37], both genotypes increased their respiration, but G/G cells responded more than A/A cells. The G/G cells also had a higher fold change of norepinephrine-stimulated-to-basal

OCR. Lastly, the proton leak was higher in G/G cells, but the ATP production was comparable between the genotypes.

Furthermore, we also examined the effect of rs10071329 on the expression of mitochondrial-encoded genes. The expression of *MT-ND1* and *MT-ND2* trended higher, and *MT-ND3* and *MT-CO3* were significantly increased in differentiated G/G cells.

Taken together, the higher *PPARGC1B* expression in rs10071329 G/G brown adipocytes correlated with improved mitochondrial gene expression, mitochondrial function, and response to norepinephrine. Our findings are thus largely consistent with previous results in *Ppargc1b*-knockout mice [94,95,146]. Furthermore, the improved mitochondrial respiration and uncoupled, norepinephrine-induced, respiration in rs10071329 G/G cells, are also supported by the previously reported improved mitochondrial respiration after *PPARGC1B* overexpression [147,148].

rs10071329 regulates norepinephrine-stimulated lipolysis in hBAs adipocytes

To further assess the effect of the rs10071329 genotype on hBAs, we also examined their lipolytic activity (glycerol release) in the presence of norepinephrine, which is normally secreted during cold exposure or exercise. We measured the fold change of norepinephrine-stimulated-to-basal lipolysis and observed that G/G cells had markedly higher glycerol release than A/A cells.

In summary, we have experimentally demonstrated that rs10071329 is a causal eQTL in human brown adipocytes, with the G allele conferring higher *PPARGC1B* expression, as well as improved mitochondrial function and response to norepinephrine. This suggests that the development of drugs that manipulate *PPARGC1B* levels, to improve beige/brown adipocyte and mitochondrial function, could be desirable in treating T2DM and obesity. Conceivably though, one could also leverage the gain-of-function effect of genetic variants in *PPARGC1B* carried by specific individuals: previous observational data, and now the experimental validation, suggest that rs10071329 G allele carriers could benefit from response to lifestyle interventions to improve weight loss or attenuate weight regain.

Key findings:

- We found rs10071329 is showing a trend association ($p = 0.055$) for BMI and a significant association with mid-upper arm circumference ($p = 0.008$) in a cohort of 132 women from Tanzania.
- We generated isogenic rs10071329 G/G and A/A homozygous brown adipocyte cell lines.
- We validated rs10071329 as a *cis*-eQTL, with the G allele conferring enhanced *PPARGC1B* expression in brown adipocytes.

- Brown adipocytes with the rs10071329 G/G genotype have improved mitochondrial function and response to norepinephrine.
- We have established the causal role of rs10071329 in human brown adipocyte function, which provides molecular insights into the previously reported associations with adiposity.
- Our findings may help the development of precision medicine for obesity treatment based on this genotype.

Paper IV

Human multidrug and toxin extrusion 1 (MATE1) is encoded by the *SLC47A1* gene and plays an important role in metformin transport in the liver and kidney. Therefore, hypothetically, any DNA polymorphisms that affect *SLC47A1* expression may affect metformin efficacy by regulating the cellular metformin flux. In support of this hypothesis, previous pharmacogenetics studies have associated a *SLC47A1* intron genetic variant, rs2289669, with metformin-dependent HbA1c decrease. However, if this SNP is causal, or how this genetic variant affects the putative gene target (*SLC47A1*) was still unknown. In this paper, we sought to answer the above questions using various approaches including *in silico* analysis, *in vitro* CRISPR/Cas9 genome engineering as well as locus-wide meta-analysis from the MetGen cohort.

In silico functional annotation identifies rs2289669 and rs8065082 as putative causal variants

Starting with the tag SNP rs2289669, we used the PredictSNP2 algorithm and manually analyzed chromatin features in the liver, using WashU EpiGenome Browser (v52.5.2), to map potential causal variants in the tag SNP haplotype block. This *in silico* functional annotation revealed two variants, rs2289669, and rs8065082, that could play gene regulatory roles, as rs2289669 resides in an *Alu* motif and rs8065082 may interact with several transcription factors like JUND1, HNF4G, and FOXA1. Whether these variants are responsible for the cis-regulation of *SLC47A1* had to be experimentally validated.

rs2289669 and rs8065082 do not regulate *SLC47A1* expression or splicing in HepG2 cells

To experimentally validate the role of predicted causal variants, we applied CRISPR/Cas9 to substitute the alleles at rs2289669 and rs8065082 in the human liver HepG2 cell model. As HepG2 cells are heterozygous at rs2289669 (G/A) and rs8065082 (C/T), we generated homozygous cell clonal populations at rs2289669 (G/G vs A/A), and rs8065082 (C/C vs T/T). We then examined if these two variants are functional by quantifying *SLC47A1* gene expression and by assessing splicing. Unexpectedly, we did not observe any *SLC47A1* expression or splicing differences between the two genotypes at either rs2289669 or rs8065082. These results showed that the predicted putative causal variants were not regulatory for *SLC47A1* in HepG2 cells. The *SLC47A1* gene expression data, however, showed a rather large intra-group variance among clones of the same genotype, and it could have obscured small differences in gene expression. (The variance probably originated from the single cell cloning process that was necessary for obtaining scarless allele-edited cells.) To further expand on our findings, we therefore used CRISPR/Cas9 to make small deletions at both rs2289669 and rs8065082 loci.

SLC47A1 expression in HepG2 cells carrying small deletions around rs2289669 and rs8065082

Using the CRISPR/Cas9-mediated deletion approach, we successfully deleted 44 bp surrounding rs2289669 and 36–37 bp surrounding rs8065082 locus with over 85% editing efficiency. The high editing efficiency, unlike the allele editing, allowed us to perform the downstream functional assays without the need for single cell cloning. The results were consistent with what we found in the allele-edited cells, i.e. small deletions around rs2289669 and rs8065082 did not have any effects on *SLC47A1* gene expression or splicing. Importantly, compared with the results obtained from allele-edited cells, the data from the deletion mutants had low intra-group variance, and supported the conclusion of null effects on gene expression by rs2289669 and rs8065082.

MetGen analysis of SNPs in *SLC47A1* and association with metformin glycemia-lowering effect

Based on our experimental findings, we next re-examined the associations between genetic variants in *SLC47A1* and the metformin glycemia-lowering effect in the MetGen cohort. The MetGen cohort includes over 10 000 participants, allowing a high-powered association study. The analysis revealed that all common genetic variants in the *SLC47A1* locus were not significantly associated with metformin-mediated lowering of glycemia. Particularly, the previously reported SNPs rs2289669 and rs8065082 showed no significant association, with P values of 0.58 and 0.83, respectively.

To summarize, in this study, we experimentally documented that the SNPs rs2289669 and the tightly linked rs8065082 are not causal variants for *SLC47A1* gene regulation, and, using a statistical approach, we also showed that common genetic variants in the *SLC47A1* locus do not associate with metformin efficacy in reducing glycemia. The present study does not exclude the possibility that *SLC47A1* plays a role in regulating metformin response. However, it is improbable that the previously reported genetic variations in the *SLC47A1* region have any functional impact on metformin action in the liver. Our findings suggest that the net effect of metformin on glucose regulation is likely to be multi-dimensional and influenced not only by the cellular flux rate of metformin but also by factors such as target protein abundance, differential organ distribution, individual organ decline, and possibly even blood pressure and cardiovascular health. Studies have also indicated that gut cells and microbiota play a significant role in regulating metformin release [149-152]. Other transporters like OCT1 [153] or proteins like GLUT2 [154] in the liver or kidneys may also impact the efficacy of metformin. Our study suggests researchers to focus on other loci for functional investigation of the metformin glucose-lowering efficacy.

Key findings:

- We generated rs2289669 and rs8065082 allele-edited isogenic HepG2 cell lines.
- We showed that rs2289669 and the tightly-linked variant rs8065082 do not affect *SLC47A1* expression or splicing in HepG2 cells.
- We also found there were no genetic variants in *SLC47A1* that were associated with metformin-mediated efficacy on glycemia reduction in MetGen participants.
- Our results raise a concern that the “candidate gene” approach might not be a good strategy for mapping causal genes.

Overall summary and discussion

Genome-wide association studies have connected thousands of human genetic variants to diseases and trait-relevant phenotypes [155]. However, the biological impact of those variants is rarely experimentally validated, which hinders the usage of those findings. The ultimate goal of our research is to advance precision medicine, identify novel druggable targets, and to design gene therapies for specific diseases. To accelerate the translation of GWAS results to clinical practice, it is desirable to identify the causal variants, their target genes and tissues, and importantly, the molecular basis that causally links the variant to certain phenotypes at cellular and organismal levels [62].

In this thesis, we used functional genomics approaches to dissect genetic variation that underpins obesity and type 2 diabetes-related traits. These two chronic metabolic disorders affect hundreds of millions of people around the world and have strong genetic components. In **paper I**, we identified rs67785913 as a cis-eQTL for *MTIF3* and studied the impact of *MTIF3* in regulating lipid metabolism in human adipocytes. Our data reveals the important role of *MTIF3* in the maintenance of mitochondrial function, which may potentially explain why people with rs67785913 CTCT allele are associated with better responses to weight loss intervention. In **paper II**, we provided the first experimental evidence for the allele-specific causal effects of the rs8192678 variant on human white adipocyte differentiation. The C allele confers lower levels of *PPARGC1A* mRNA and PGC-1α protein, as well as disrupted dynamics of PGC-1α turnover and activity, with downstream effects on cellular differentiation and mitochondrial function. Our experimental data may provide an explanation of why rs8192678 T allele carriers have higher body fat mass and BMI, and excessive weight gain [93,129-131,144,156], but fails to explain the T allele-associated increased risk of type 2 diabetes. However, since our investigation only focused on human white adipocytes, we must be aware that the effects of the T allele in other tissues, including the liver, muscle, and brain, may also contribute to metabolic dysfunction. In **paper III**, we documented that rs10071329 is a cis-eQTL, with the G allele conferring enhanced *PPARGC1B* expression, which consequently improves mitochondrial function and response to norepinephrine in brown adipocytes. These data may provide an explanation for the observed associations between rs10071329 and adiposity [93].

The findings presented in this thesis may have clinical significance in the future. Firstly, the identified eQTL variants may prove useful in the development of genotype-based precision medicine for obesity treatment. Secondly, studying the function of the genes whose regulation is affected by the variants may help to identify potential druggable targets for the development of anti-obesity drugs. Lastly, knowing the genetic variants that contribute to causing disease may help risk allele carriers in evaluating their lifestyles and entice them to actively prevent obesity.

Besides the obesity-related research, in **paper IV** we investigated the putative causal variant rs2289669 in T2DM treatment-related traits. Through our functional studies, we found that rs2289669-containing haploblock does not appear to regulate *SLC47A1* expression or splicing, and therefore it may not contain causal variants that explain its previously reported association with glucose-lowering effects of metformin.

The functional genomics approaches used in all four projects may provide a framework for dissecting GWAS findings, especially the use of CRISPR/Cas9 mediated allele editing. The discovery of the CRISPR/Cas9 system has provided an excellent tool in genetics research [66]. Supplying the cells with a DNA donor template, along with CRISPR/Cas9, will result in some cells using the template to repair the DNA strand break through HDR, resulting in the generation of DNA knock-in cells [157]. This mechanism allows us to substitute single alleles of putative causal GWAS-associated SNPs and to study their effects on gene expression or protein function. However, one should bear in mind that the HDR efficiency mediated by CRISPR/Cas9 is often low, and generating homozygous without DNA-scarring edited clones can thus become very time-consuming. This highlights the necessity of prioritization of putative causal variants for CRISPR/Cas9 allele editing.

As most of the GWAS-associated SNPs reside in non-coding genomic regions, the locations of the causal variants are often found in promoters or enhancers. The activity of these two types of functional DNA elements is highly cell type-specific [158,159], and many disease-associated variants may regulate target gene expression and cellular functions in a cell type-specific manner. While the obesity-associated variants may be functional in different organs, in our studies we mainly focused on human adipose cells **(paper I**, **II,** and **III)**. It should be pointed out though that further work on the function of these variants in other cell types is warranted.

In **paper IV**, we studied the genetic variants in the *SLC47A1* gene in liver cells, based on enriched expression of this gene in the liver (and kidney). Although the present study does not completely rule out a role for *SLC47A1* in the complex network that probably controls the metformin response, it is unlikely that the *SLC47A1* variants reported to date play a functional role in liver-mediated metformin action.

In summary, in the post-GWAS era, we are facing challenges to investigate the function of the disease-associated genetic variation. In my thesis, I tried to fill this knowledge gap by applying different functional genomics approaches to dissect genetic variants correlated with obesity and T2DM. In my papers, I present a genetic variant-to-function experimental workflow, and I have identified several causal genetic variants that could contribute to obesity and T2DM. Our methodologies and findings may provide valuable information on further functional genomics research and the development of precision medicine.

Future perspectives

In the current thesis, we applied functional genomics to human cell lines to identify causal genetic variations that may contribute to obesity and type 2 diabetes. Future investigations based on our findings could be extended as follows:

In **paper I**, we linked the genetic variant rs67785913 to *MTIF3* expression, and illustrated the functional basis of MTIF3 in adipocyte metabolism. However, all studies were only from a cell model cultured in a dish. How the variation affects human metabolism at the organismal level is still unknown. To further support the translation of our findings to drug discovery, precision medicine, and developing personalized therapy against obesity, future studies should aim to understand the effects of rs67785913 on human metabolism through a genotype-based recall study (clinical trial). To reveal the adipose-specific role of MTIF3, a conditional knockout mouse study might also be helpful. Furthermore, rs67785913 may play a regulatory role in other tissues, such as muscle or liver, and future functional genomics studies could be performed to address this question.

In **paper II**, we found that rs8192678 causally affects PGC-1α protein abundance and stability, resulting in altered adipogenic differentiation in white adipocytes. Although these findings can help explain several epidemiology findings, such as T allele carriers having higher body fat mass, BMI, and excessive weight gain, the long-term *in vivo* effects of this allele on other variables, such as insulin sensitivity, remains to be investigated. Future studies may focus on the effect of rs8192678 on other tissues, and on gaining a fuller understanding of how rs8192678 affects whole-body metabolism. Furthermore, in the current study, we did not perform gene-by-environment interaction experiments in similarly differentiated cells with similar triacylglycerol content, and were unable to investigate some of the epidemiological findings, e.g. insulin sensitivity, experimentally. Future genotype-based recall studies, focusing on different environmental exposures, may be useful in finding the environmental factors that will aid the development of genotype-guided precision medicine.

Paper III shows that rs10071329 is a *cis*-eQTL for *PPARGC1B* expression. This genetic variant, and yet unidentified eQTLs at *PPARGC1B,* could eventually prove useful in genotype-based precision medicine for obesity treatment. As we only assayed the effects of rs10071329 on brown adipocytes, studies to reveal its function in other tissues or cell types may also be needed. Furthermore, a genotype-based cohort study will be helpful in providing further evidence for this concept.

In **paper IV**, although we found no causal relationship between the genetic variants, future studies may focus on dissecting other variants which may affect the metformin-mediated glucose-lowering effects. Also, variants in other presumed metformin transporters could be investigated.

Lastly, future functional genomics approaches, including methods used in this thesis, may also be improved by integrating high throughput CRISPR screening to prioritize candidate gene targets, and by MPRA assays to prioritize functional genetic variants. Additionally, CRISPRi and CRISPRa screening also provides new tools for finding regulatory sites in the genome, which may help to facilitate functional genomics research.

References

1 Must, A. *et al.* The disease burden associated with overweight and obesity. *JAMA* **282**, 1523-1529, doi:10.1001/jama.282.16.1523 (1999).

2 WHO. World Health Organisation: obesity and overweight. (2017).

3 Collaboration, N. C. D. R. F. Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128.9 million children, adolescents, and adults. *Lancet* **390**, 2627-2642, doi:10.1016/S0140-6736(17)32129-3 (2017).

4 Bluher, M. Obesity: global epidemiology and pathogenesis. *Nat. Rev. Endocrinol.* **15**, 288-298, doi:10.1038/s41574-019-0176-8 (2019).

5 Cecchini, M. Use of healthcare services and expenditure in the US in 2025: The effect of obesity and morbid obesity. *PLoS One* **13**, e0206703, doi:10.1371/journal.pone.0206703 (2018).

6 Wolfenstetter, S. B. *et al.* Body weight changes and outpatient medical care utilisation: Results of the MONICA/KORA cohorts S3/F3 and S4/F4. *Psychosoc. Med.* **9**, Doc09, doi:10.3205/psm000087 (2012).

7 Okunogbe, A., Nugent, R., Spencer, G., Ralston, J. & Wilding, J. Economic impacts of overweight and obesity: current and future estimates for eight countries. *BMJ Glob Health* **6**, doi:10.1136/bmjgh-2021-006351 (2021).

8 Hall, K. D., Guo, J., Dore, M. & Chow, C. C. The progressive increase of food waste in America and its environmental impact. *PLoS One* **4**, e7940, doi:10.1371/journal.pone.0007940 (2009).

9 Popkin, B. M. & Hawkes, C. Sweetening of the global diet, particularly beverages: patterns, trends, and policy responses. *Lancet Diabetes Endocrinol* **4**, 174-186, doi:10.1016/S2213-8587(15)00419-2 (2016).

10 Church, T. S. *et al.* Trends over 5 decades in U.S. occupation-related physical activity and their associations with obesity. *PLoS One* **6**, e19657, doi:10.1371/journal.pone.0019657 (2011).

11 Heymsfield, S. B. & Wadden, T. A. Mechanisms, Pathophysiology, and Management of Obesity. *N. Engl. J. Med.* **376**, 254-266, doi:10.1056/NEJMra1514009 (2017).

12 Bray, M. S. *et al.* NIH working group report-using genomic information to guide weight management: From universal to precision treatment. *Obesity (Silver Spring)* **24**, 14-22, doi:10.1002/oby.21381 (2016).

13 Pigeyre, M., Yazdi, F. T., Kaur, Y. & Meyre, D. Recent progress in genetics, epigenetics and metagenomics unveils the pathophysiology of

human obesity. *Clin. Sci. (Lond.)* **130**, 943-986, doi:10.1042/CS20160136 (2016).

14 van der Klaauw, A. A. & Farooqi, I. S. The hunger genes: pathways to obesity. *Cell* **161**, 119-132, doi:10.1016/j.cell.2015.03.008 (2015).

15 Jensen, M. D. *et al.* 2013 AHA/ACC/TOS guideline for the management of overweight and obesity in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and The Obesity Society. *J. Am. Coll. Cardiol.* **63**, 2985-3023, doi:10.1016/j.jacc.2013.11.004 (2014).

16 Diabetes Prevention Program Research, G. The Diabetes Prevention Program (DPP): description of lifestyle intervention. *Diabetes Care* **25**, 2165-2171, doi:10.2337/diacare.25.12.2165 (2002).

17 Look, A. R. G. *et al.* The Look AHEAD study: a description of the lifestyle intervention and the evidence supporting it. *Obesity (Silver Spring)* **14**, 737-752, doi:10.1038/oby.2006.84 (2006).

18 Webb, V. L. & Wadden, T. A. Intensive Lifestyle Intervention for Obesity: Principles, Practices, and Results. *Gastroenterology* **152**, 1752-1764, doi:10.1053/j.gastro.2017.01.045 (2017).

19 Apovian, C. M. *et al.* Pharmacological management of obesity: an endocrine Society clinical practice guideline. *J. Clin. Endocrinol. Metab.* **100**, 342-362, doi:10.1210/jc.2014-3415 (2015).

20 Yanovski, S. Z. & Yanovski, J. A. Long-term drug treatment for obesity: a systematic and clinical review. *JAMA* **311**, 74-86, doi:10.1001/jama.2013.281361 (2014).

21 Schauer, P. R., Mingrone, G., Ikramuddin, S. & Wolfe, B. Clinical Outcomes of Metabolic Surgery: Efficacy of Glycemic Control, Weight Loss, and Remission of Diabetes. *Diabetes Care* **39**, 902-911, doi:10.2337/dc16-0382 (2016).

22 Puzziferri, N. *et al.* Long-term follow-up after bariatric surgery: a systematic review. *JAMA* **312**, 934-942, doi:10.1001/jama.2014.10706 (2014).

23 Holzapfel, C. *et al.* Association of a MTNR1B gene variant with fasting glucose and HOMA-B in children and adolescents with high BMI-SDS. *Eur. J. Endocrinol.* **164**, 205-212, doi:10.1530/EJE-10-0588 (2011).

24 Moleres, A. *et al.* Obesity susceptibility loci on body mass index and weight loss in Spanish adolescents after a lifestyle intervention. *J. Pediatr.* **161**, 466-470 e462, doi:10.1016/j.jpeds.2012.04.004 (2012).

25 Reinehr, T., Wolters, B., Roth, C. L. & Hinney, A. FTO gene: association to weight regain after lifestyle intervention in overweight children. *Horm. Res. Paediatr.* **81**, 391-396, doi:10.1159/000358328 (2014).

26 Heitkamp, M. *et al.* Obesity Genes and Weight Loss During Lifestyle Intervention in Children With Obesity. *JAMA Pediatr* **175**, e205142, doi:10.1001/jamapediatrics.2020.5142 (2021).

27 Huang, T. & Hu, F. B. Gene-environment interactions and obesity: recent developments and future directions. *BMC Med. Genomics* **8 Suppl 1**, S2, doi:10.1186/1755-8794-8-S1-S2 (2015).

28 Green, A. *et al.* Type 1 diabetes in 2017: global estimates of incident and prevalent cases in children and adults. *Diabetologia* **64**, 2741-2750, doi:10.1007/s00125-021-05571-8 (2021).

29 Chatterjee, S., Khunti, K. & Davies, M. J. Type 2 diabetes. *Lancet* **389**, 2239-2251, doi:10.1016/S0140-6736(17)30058-2 (2017).

30 Cornelis, M. C. & Hu, F. B. Gene-environment interactions in the development of type 2 diabetes: recent progress and continuing challenges. *Annu. Rev. Nutr.* **32**, 245-259, doi:10.1146/annurev-nutr-071811-150648 (2012).

31 Adeghate, E., Schattner, P. & Dunn, E. An update on the etiology and epidemiology of diabetes mellitus. *Ann. N. Y. Acad. Sci.* **1084**, 1-29, doi:10.1196/annals.1372.029 (2006).

32 Meigs, J. B., Cupples, L. A. & Wilson, P. W. Parental transmission of type 2 diabetes: the Framingham Offspring Study. *Diabetes* **49**, 2201-2207, doi:10.2337/diabetes.49.12.2201 (2000).

33 Poulsen, P., Kyvik, K. O., Vaag, A. & Beck-Nielsen, H. Heritability of type II (non-insulin-dependent) diabetes mellitus and abnormal glucose tolerance--a population-based twin study. *Diabetologia* **42**, 139-145, doi:10.1007/s001250051131 (1999).

34 Shaw, R. J. *et al.* The kinase LKB1 mediates glucose homeostasis in liver and therapeutic effects of metformin. *Science* **310**, 1642-1646, doi:10.1126/science.1120781 (2005).

35 El-Mir, M. Y. *et al.* Dimethylbiguanide inhibits cell respiration via an indirect effect targeted on the respiratory chain complex I. *J. Biol. Chem.* **275**, 223-228, doi:10.1074/jbc.275.1.223 (2000).

36 DeFronzo, R. A. *et al.* Once-daily delayed-release metformin lowers plasma glucose and enhances fasting and postprandial GLP-1 and PYY: results from two randomised trials. *Diabetologia* **59**, 1645-1654, doi:10.1007/s00125-016-3992-6 (2016).

37 Pernicova, I. & Korbonits, M. Metformin--mode of action and clinical implications for diabetes and cancer. *Nat. Rev. Endocrinol.* **10**, 143-156, doi:10.1038/nrendo.2013.256 (2014).

38 Cook, M. N., Girman, C. J., Stein, P. P. & Alexander, C. M. Initial monotherapy with either metformin or sulphonylureas often fails to achieve or maintain current glycaemic goals in patients with Type 2 diabetes in UK primary care. *Diabet. Med.* **24**, 350-358, doi:10.1111/j.1464-5491.2007.02078.x (2007).

39 Kahn, S. E. *et al.* Glycemic durability of rosiglitazone, metformin, or glyburide monotherapy. *N. Engl. J. Med.* **355**, 2427-2443, doi:10.1056/NEJMoa066224 (2006).

40 Garcia-Calzon, S. *et al.* Epigenetic markers associated with metformin response and intolerance in drug-naive patients with type 2 diabetes. *Sci. Transl. Med.* **12**, doi:10.1126/scitranslmed.aaz1803 (2020).
41 Hausman, D. B., DiGirolamo, M., Bartness, T. J., Hausman, G. J. & Martin, R. J. The biology of white adipocyte proliferation. *Obes. Rev.* **2**, 239-254, doi:10.1046/j.1467-789x.2001.00042.x (2001).
42 Saltiel, A. R. & Kahn, C. R. Insulin signalling and the regulation of glucose and lipid metabolism. *Nature* **414**, 799-806, doi:10.1038/414799a (2001).
43 Tilg, H. & Moschen, A. R. Adipocytokines: mediators linking adipose tissue, inflammation and immunity. *Nat. Rev. Immunol.* **6**, 772-783, doi:10.1038/nri1937 (2006).
44 Cannon, B. & Nedergaard, J. Brown adipose tissue: function and physiological significance. *Physiol. Rev.* **84**, 277-359, doi:10.1152/physrev.00015.2003 (2004).
45 Longo, M. *et al.* Adipose Tissue Dysfunction as Determinant of Obesity-Associated Metabolic Complications. *Int. J. Mol. Sci.* **20**, doi:10.3390/ijms20092358 (2019).
46 Scherer, P. E. Adipose tissue: from lipid storage compartment to endocrine organ. *Diabetes* **55**, 1537-1545, doi:10.2337/db06-0263 (2006).
47 Reilly, S. M. & Saltiel, A. R. Adapting to obesity with adipose tissue inflammation. *Nat. Rev. Endocrinol.* **13**, 633-643, doi:10.1038/nrendo.2017.90 (2017).
48 Rutkowski, J. M., Stern, J. H. & Scherer, P. E. The cell biology of fat expansion. *J. Cell Biol.* **208**, 501-512, doi:10.1083/jcb.201409063 (2015).
49 Ouellet, V. *et al.* Brown adipose tissue oxidative metabolism contributes to energy expenditure during acute cold exposure in humans. *J. Clin. Invest.* **122**, 545-552, doi:10.1172/JCI60433 (2012).
50 Kwok, K. H., Lam, K. S. & Xu, A. Heterogeneity of white adipose tissue: molecular basis and clinical implications. *Exp. Mol. Med.* **48**, e215, doi:10.1038/emm.2016.5 (2016).
51 Carpentier, A. C. *et al.* Brown Adipose Tissue Energy Metabolism in Humans. *Front. Endocrinol. (Lausanne)* **9**, 447, doi:10.3389/fendo.2018.00447 (2018).
52 Leitner, B. P. *et al.* Mapping of human brown adipose tissue in lean and obese young men. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 8649-8654, doi:10.1073/pnas.1705287114 (2017).
53 Cypess, A. M. *et al.* Identification and importance of brown adipose tissue in adult humans. *N. Engl. J. Med.* **360**, 1509-1517, doi:10.1056/NEJMoa0810780 (2009).
54 Saito, M. *et al.* High incidence of metabolically active brown adipose tissue in healthy adult humans: effects of cold exposure and adiposity. *Diabetes* **58**, 1526-1531, doi:10.2337/db09-0530 (2009).
55 Muller, M. J. *et al.* The case of GWAS of obesity: does body weight control play by the rules? *Int. J. Obes. (Lond.)* **42**, 1395-1405, doi:10.1038/s41366-018-0081-6 (2018).

56 Papandonatos, G. D. *et al.* Genetic Predisposition to Weight Loss and Regain With Lifestyle Intervention: Analyses From the Diabetes Prevention Program and the Look AHEAD Randomized Controlled Trials. *Diabetes* **64**, 4312-4321, doi:10.2337/db15-0441 (2015).

57 Mahajan, A. *et al.* Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. *Nat. Genet.* **50**, 1505-1513, doi:10.1038/s41588-018-0241-6 (2018).

58 Chen, J. *et al.* The trans-ancestral genomic architecture of glycemic traits. *Nat. Genet.* **53**, 840-860, doi:10.1038/s41588-021-00852-9 (2021).

59 Maruthur, N. M. *et al.* The pharmacogenetics of type 2 diabetes: a systematic review. *Diabetes Care* **37**, 876-886, doi:10.2337/dc13-1276 (2014).

60 Slatkin, M. Linkage disequilibrium--understanding the evolutionary past and mapping the medical future. *Nat Rev Genet* **9**, 477-485, doi:10.1038/nrg2361 (2008).

61 Billings, L. K. & Florez, J. C. The genetics of type 2 diabetes: what have we learned from GWAS? *Ann. N. Y. Acad. Sci.* **1212**, 59-77, doi:10.1111/j.1749-6632.2010.05838.x (2010).

62 Gallagher, M. D. & Chen-Plotkin, A. S. The Post-GWAS Era: From Association to Function. *Am. J. Hum. Genet.* **102**, 717-730, doi:10.1016/j.ajhg.2018.04.002 (2018).

63 Edwards, S. L., Beesley, J., French, J. D. & Dunning, A. M. Beyond GWASs: illuminating the dark road from association to function. *Am. J. Hum. Genet.* **93**, 779-797, doi:10.1016/j.ajhg.2013.10.012 (2013).

64 Cano-Gamez, E. & Trynka, G. From GWAS to Function: Using Functional Genomics to Identify the Mechanisms Underlying Complex Diseases. *Front Genet* **11**, 424, doi:10.3389/fgene.2020.00424 (2020).

65 French, J. D. & Edwards, S. L. The Role of Noncoding Variants in Heritable Disease. *Trends Genet.* **36**, 880-891, doi:10.1016/j.tig.2020.07.004 (2020).

66 Rao, S., Yao, Y. & Bauer, D. E. Editing GWAS: experimental approaches to dissect and exploit disease-associated genetic variation. *Genome Med.* **13**, 41, doi:10.1186/s13073-021-00857-3 (2021).

67 Schaid, D. J., Chen, W. & Larson, N. B. From genome-wide associations to candidate causal variants by statistical fine-mapping. *Nat Rev Genet* **19**, 491-504, doi:10.1038/s41576-018-0016-z (2018).

68 Broekema, R. V., Bakker, O. B. & Jonkers, I. H. A practical view of fine-mapping and gene prioritization in the post-genome-wide association era. *Open Biol* **10**, 190221, doi:10.1098/rsob.190221 (2020).

69 Garneau, J. E. *et al.* The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* **468**, 67-71, doi:10.1038/nature09523 (2010).

70 Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819-823, doi:10.1126/science.1231143 (2013).

71 Ranjha, L., Howard, S. M. & Cejka, P. Main steps in DNA double-strand break repair: an introduction to homologous recombination and related

processes. *Chromosoma* **127**, 187-214, doi:10.1007/s00412-017-0658-1 (2018).

72 Shen, M. W. *et al.* Predictable and precise template-free CRISPR editing of pathogenic variants. *Nature* **563**, 646-651, doi:10.1038/s41586-018-0686-x (2018).

73 Adams, K. F. *et al.* Overweight, obesity, and mortality in a large prospective cohort of persons 50 to 71 years old. *N. Engl. J. Med.* **355**, 763-778, doi:10.1056/NEJMoa055643 (2006).

74 Claussnitzer, M. *et al.* FTO Obesity Variant Circuitry and Adipocyte Browning in Humans. *N. Engl. J. Med.* **373**, 895-907, doi:10.1056/NEJMoa1502214 (2015).

75 Liu, M. *et al.* Methodologies for Improving HDR Efficiency. *Front Genet* **9**, 691, doi:10.3389/fgene.2018.00691 (2018).

76 Klein, J. C. *et al.* A systematic evaluation of the design and context dependencies of massively parallel reporter assays. *Nat Methods* **17**, 1083-1091, doi:10.1038/s41592-020-0965-y (2020).

77 Vandenbeek, R., Khan, N. P. & Estall, J. L. Linking Metabolic Disease With the PGC-1α Gly482Ser Polymorphism. *Endocrinology* **159**, 853-865, doi:10.1210/en.2017-00872 (2018).

78 Andrulionyte, L., Zacharova, J., Chiasson, J. L., Laakso, M. & Group, S.-N. S. Common polymorphisms of the PPAR-gamma2 (Pro12Ala) and PGC-1alpha (Gly482Ser) genes are associated with the conversion from impaired glucose tolerance to type 2 diabetes in the STOP-NIDDM trial. *Diabetologia* **47**, 2176-2184, doi:10.1007/s00125-004-1577-2 (2004).

79 Ambye, L. *et al.* Studies of the Gly482Ser polymorphism of the peroxisome proliferator-activated receptor gamma coactivator 1alpha (PGC-1alpha) gene in Danish subjects with the metabolic syndrome. *Diabetes Res. Clin. Pract.* **67**, 175-179, doi:10.1016/j.diabres.2004.06.013 (2005).

80 Geloneze, S. R. *et al.* PGC1α gene Gly482Ser polymorphism predicts improved metabolic, inflammatory and vascular outcomes following bariatric surgery. *Int. J. Obes. (Lond.)* **36**, 363-368, doi:10.1038/ijo.2011.176 (2012).

81 Finck, B. N. & Kelly, D. P. PGC-1 coactivators: inducible regulators of energy metabolism in health and disease. *J. Clin. Invest.* **116**, 615-622, doi:10.1172/JCI27794 (2006).

82 Austin, S. & St-Pierre, J. PGC1α and mitochondrial metabolism – emerging concepts and relevance in ageing and neurodegenerative disorders. *J. Cell Sci.* **125**, 4963-4971, doi:10.1242/jcs.113662 (2012).

83 Ling, C. *et al.* Epigenetic regulation of PPARGC1A in human type 2 diabetic islets and effect on insulin secretion. *Diabetologia* **51**, 615-622, doi:10.1007/s00125-007-0916-5 (2008).

84 Ling, C. *et al.* Multiple environmental and genetic factors influence skeletal muscle PGC-1alpha and PGC-1beta gene expression in twins. *J. Clin. Invest.* **114**, 1518-1526, doi:10.1172/jci21889 (2004).

85 Besse-Patin, A. *et al.* Estrogen Signals Through Peroxisome Proliferator-Activated Receptor-γ Coactivator 1α to Reduce Oxidative Damage Associated With Diet-Induced Fatty Liver Disease. *Gastroenterology* **152**, 243-256, doi:10.1053/j.gastro.2016.09.017 (2017).
86 Galipeau, M. *et al.* 299-OR: Molecular and Physiological Consequences of the Diabetes-Related PGC1A Gly482Ser Polymorphism. *Diabetes* **71**, doi:10.2337/db22-299-OR (2022).
87 Rasmussen-Torvik, L. J. *et al.* rs4771122 Predicts Multiple Measures of Long-Term Weight Loss After Bariatric Surgery. *Obes. Surg.* **25**, 2225-2229, doi:10.1007/s11695-015-1872-7 (2015).
88 Kuzmenko, A. *et al.* Mitochondrial translation initiation machinery: conservation and diversification. *Biochimie* **100**, 132-140, doi:10.1016/j.biochi.2013.07.024 (2014).
89 Rudler, D. L. *et al.* Fidelity of translation initiation is required for coordinated respiratory complex assembly. *Sci Adv* **5**, eaay2118, doi:10.1126/sciadv.aay2118 (2019).
90 Chicherin, I. V. *et al.* Initiation Factor 3 is Dispensable For Mitochondrial Translation in Cultured Human Cells. *Sci. Rep.* **10**, 7110, doi:10.1038/s41598-020-64139-5 (2020).
91 Gregoire, F. M., Smas, C. M. & Sul, H. S. Understanding adipocyte differentiation. *Physiol. Rev.* **78**, 783-809, doi:10.1152/physrev.1998.78.3.783 (1998).
92 Boudina, S. & Graham, T. E. Mitochondrial function/dysfunction in white adipose tissue. *Exp. Physiol.* **99**, 1168-1178, doi:10.1113/expphysiol.2014.081414 (2014).
93 Franks, P. W. *et al.* Common variation at PPARGC1A/B and change in body composition and metabolic traits following preventive interventions: the Diabetes Prevention Program. *Diabetologia* **57**, 485-490, doi:10.1007/s00125-013-3133-4 (2014).
94 Sonoda, J., Mehl, I. R., Chong, L. W., Nofsinger, R. R. & Evans, R. M. PGC-1beta controls mitochondrial metabolism to modulate circadian activity, adaptive thermogenesis, and hepatic steatosis. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 5223-5228, doi:10.1073/pnas.0611623104 (2007).
95 Funda, J. *et al.* Adipose tissue-specific ablation of PGC-1beta impairs thermogenesis in brown fat. *Dis Model Mech* **15**, doi:10.1242/dmm.049223 (2022).
96 Zolk, O. Disposition of metformin: variability due to polymorphisms of organic cation transporters. *Ann Med* **44**, 119-129, doi:10.3109/07853890.2010.549144 (2012).
97 Nies, A. T., Damme, K., Kruck, S., Schaeffeler, E. & Schwab, M. Structure and function of multidrug and toxin extrusion proteins (MATEs) and their relevance to drug therapy and personalized medicine. *Arch Toxicol* **90**, 1555-1584, doi:10.1007/s00204-016-1728-5 (2016).

98 Stocker, S. L. *et al.* The effect of novel promoter variants in MATE1 and MATE2 on the pharmacokinetics and pharmacodynamics of metformin. *Clin Pharmacol Ther* **93**, 186-194, doi:10.1038/clpt.2012.210 (2013).

99 Becker, M. L. *et al.* Genetic variation in the multidrug and toxin extrusion 1 transporter protein influences the glucose-lowering effect of metformin in patients with diabetes: a preliminary study. *Diabetes* **58**, 745-749, doi:10.2337/db08-1028 (2009).

100 Becker, M. L. *et al.* Interaction between polymorphisms in the OCT1 and MATE1 transporter and metformin response. *Pharmacogenet Genomics* **20**, 38-44, doi:10.1097/FPC.0b013e328333bb11 (2010).

101 Tkac, I. *et al.* Pharmacogenomic association between a variant in SLC47A1 gene and therapeutic response to metformin in type 2 diabetes. *Diabetes Obes Metab* **15**, 189-191, doi:10.1111/j.1463-1326.2012.01691.x (2013).

102 Jablonski, K. A. *et al.* Common variants in 40 genes assessed for diabetes incidence and response to metformin and lifestyle intervention in the diabetes prevention program. *Diabetes* **59**, 2672-2681, doi:10.2337/db10-0543 (2010).

103 Dujic, T. *et al.* Variants in Pharmacokinetic Transporters and Glycemic Response to Metformin: A Metgen Meta-Analysis. *Clin Pharmacol Ther* **101**, 763-772, doi:10.1002/cpt.567 (2017).

104 Hibma, J. E. *et al.* The Effect of Famotidine, a MATE1-Selective Inhibitor, on the Pharmacokinetics and Pharmacodynamics of Metformin. *Clin Pharmacokinet* **55**, 711-721, doi:10.1007/s40262-015-0346-3 (2016).

105 Kusuhara, H. *et al.* Effects of a MATE protein inhibitor, pyrimethamine, on the renal elimination of metformin at oral microdose and at therapeutic dose in healthy subjects. *Clin Pharmacol Ther* **89**, 837-844, doi:10.1038/clpt.2011.36 (2011).

106 Masuda, S. *et al.* Identification and functional characterization of a new human kidney-specific H+/organic cation antiporter, kidney-specific multidrug and toxin extrusion 2. *J Am Soc Nephrol* **17**, 2127-2135, doi:10.1681/ASN.2006030205 (2006).

107 Otsuka, M. *et al.* A human transporter protein that mediates the final excretion step for toxic organic cations. *Proc Natl Acad Sci U S A* **102**, 17923-17928, doi:10.1073/pnas.0506483102 (2005).

108 Toyama, K. *et al.* Loss of multidrug and toxin extrusion 1 (MATE1) is associated with metformin-induced lactic acidosis. *Br J Pharmacol* **166**, 1183-1191, doi:10.1111/j.1476-5381.2012.01853.x (2012).

109 Tsuda, M. *et al.* Targeted disruption of the multidrug and toxin extrusion 1 (mate1) gene in mice reduces renal secretion of metformin. *Mol Pharmacol* **75**, 1280-1286, doi:10.1124/mol.109.056242 (2009).

110 Shamsi, F. & Tseng, Y. H. Protocols for Generation of Immortalized Human Brown and White Preadipocyte Cell Lines. *Methods Mol. Biol.* **1566**, 77-85, doi:10.1007/978-1-4939-6820-6_8 (2017).

111 Aden, D. P., Fogel, A., Plotkin, S., Damjanov, I. & Knowles, B. B. Controlled synthesis of HBsAg in a differentiated human liver carcinoma-derived cell line. *Nature* **282**, 615-616, doi:10.1038/282615a0 (1979).

112 Nair, A. K. & Baier, L. J. Using Luciferase Reporter Assays to Identify Functional Variants at Disease-Associated Loci. *Methods Mol. Biol.* **1706**, 303-319, doi:10.1007/978-1-4939-7471-9_17 (2018).

113 Vakulskas, C. A. *et al.* A high-fidelity Cas9 mutant delivered as a ribonucleoprotein complex enables efficient gene editing in human hematopoietic stem and progenitor cells. *Nat. Med.* **24**, 1216-1224, doi:10.1038/s41591-018-0137-0 (2018).

114 Arya, M. *et al.* Basic principles of real-time quantitative PCR. *Expert Rev. Mol. Diagn.* **5**, 209-219, doi:10.1586/14737159.5.2.209 (2005).

115 Bournat, J. C. & Brown, C. W. Mitochondrial dysfunction in obesity. *Curr. Opin. Endocrinol. Diabetes Obes.* **17**, 446-452, doi:10.1097/MED.0b013e32833c3026 (2010).

116 Rose, S. *et al.* Oxidative stress induces mitochondrial dysfunction in a subset of autism lymphoblastoid cell lines in a well-matched case control cohort. *PLoS One* **9**, e85436, doi:10.1371/journal.pone.0085436 (2014).

117 Burnette, W. N. "Western blotting": electrophoretic transfer of proteins from sodium dodecyl sulfate--polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. *Anal. Biochem.* **112**, 195-203, doi:10.1016/0003-2697(81)90281-5 (1981).

118 Schagger, H. & von Jagow, G. Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. *Anal. Biochem.* **199**, 223-231, doi:10.1016/0003-2697(91)90094-a (1991).

119 Fernandez-Vizarra, E. & Zeviani, M. Blue-Native Electrophoresis to Study the OXPHOS Complexes. *Methods Mol. Biol.* **2192**, 287-311, doi:10.1007/978-1-0716-0834-0_20 (2021).

120 Kraus, N. A. *et al.* Quantitative assessment of adipocyte differentiation in cell culture. *Adipocyte* **5**, 351-358, doi:10.1080/21623945.2016.1240137 (2016).

121 Poulos, S. P., Dodson, M. V. & Hausman, G. J. Cell line models for differentiation: preadipocytes and adipocytes. *Exp. Biol. Med. (Maywood)* **235**, 1185-1193, doi:10.1258/ebm.2010.010063 (2010).

122 Kamble, P. G. *et al.* Proof-of-concept for CRISPR/Cas9 gene editing in human preadipocytes: Deletion of FKBP5 and PPARG and effects on adipocyte differentiation and metabolism. *Sci. Rep.* **10**, 10565, doi:10.1038/s41598-020-67293-y (2020).

123 Bhargava, K. & Spremulli, L. L. Role of the N- and C-terminal extensions on the activity of mammalian mitochondrial translational initiation factor 3. *Nucleic Acids Res.* **33**, 7011-7018, doi:10.1093/nar/gki1007 (2005).

124 Kuhlbrandt, W. Structure and function of mitochondrial membrane protein complexes. *BMC Biol.* **13**, 89, doi:10.1186/s12915-015-0201-x (2015).

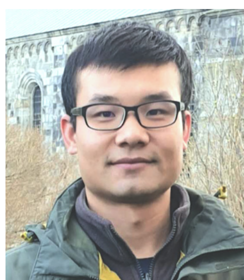
125 Maechler, P. & Wollheim, C. B. Mitochondrial function in normal and diabetic beta-cells. *Nature* **414**, 807-812, doi:10.1038/414807a (2001).
126 Kaaman, M. *et al.* Strong association between mitochondrial DNA copy number and lipogenesis in human white adipose tissue. *Diabetologia* **50**, 2526-2533, doi:10.1007/s00125-007-0818-6 (2007).
127 Heinonen, S., Jokinen, R., Rissanen, A. & Pietilainen, K. H. White adipose tissue mitochondrial metabolism in health and in obesity. *Obes. Rev.* **21**, e12958, doi:10.1111/obr.12958 (2020).
128 Jokinen, R. *et al.* Adipose tissue mitochondrial capacity associates with long-term weight loss success. *Int. J. Obes. (Lond.)* **42**, 817-825, doi:10.1038/ijo.2017.299 (2018).
129 Ha, C. D., Cho, J. K., Han, T., Lee, S. H. & Kang, H. S. Relationship of PGC-1α gene polymorphism with insulin resistance syndrome in Korean children. *Asia Pac. J. Public Health* **27**, Np544-551, doi:10.1177/1010539513477685 (2015).
130 Weng, S. W. *et al.* Gly482Ser polymorphism in the peroxisome proliferator-activated receptor gamma coactivator-1alpha gene is associated with oxidative stress and abdominal obesity. *Metabolism* **59**, 581-586, doi:10.1016/j.metabol.2009.08.021 (2010).
131 Vazquez-Del Mercado, M. *et al.* The 482Ser of PPARGC1A and 12Pro of PPARG2 Alleles Are Associated with Reduction of Metabolic Risk Factors Even Obesity in a Mexican-Mestizo Population. *Biomed Res Int* **2015**, 285491, doi:10.1155/2015/285491 (2015).
132 Huang, P. I. *et al.* PGC-1alpha mediates differentiation of mesenchymal stem cells to brown adipose cells. *J Atheroscler Thromb* **18**, 966-980, doi:10.5551/jat.7401 (2011).
133 Semple, R. K. *et al.* Expression of the thermogenic nuclear hormone receptor coactivator PGC-1alpha is reduced in the adipose tissue of morbidly obese subjects. *Int. J. Obes. Relat. Metab. Disord.* **28**, 176-179, doi:10.1038/sj.ijo.0802482 (2004).
134 Lin, J. *et al.* Defects in adaptive energy metabolism with CNS-linked hyperactivity in PGC-1alpha null mice. *Cell* **119**, 121-135, doi:10.1016/j.cell.2004.09.013 (2004).
135 Kleiner, S. *et al.* Development of insulin resistance in mice lacking PGC-1alpha in adipose tissues. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 9635-9640, doi:10.1073/pnas.1207287109 (2012).
136 Uldry, M. *et al.* Complementary action of the PGC-1 coactivators in mitochondrial biogenesis and brown fat differentiation. *Cell Metab.* **3**, 333-341, doi:10.1016/j.cmet.2006.04.002 (2006).
137 Puigserver, P. & Spiegelman, B. M. Peroxisome proliferator-activated receptor-gamma coactivator 1 alpha (PGC-1 alpha): transcriptional coactivator and metabolic regulator. *Endocr. Rev.* **24**, 78-90, doi:10.1210/er.2002-0012 (2003).
138 De Pauw, A., Tejerina, S., Raes, M., Keijer, J. & Arnould, T. Mitochondrial (dys)function in adipocyte (de)differentiation and systemic metabolic

alterations. *Am. J. Pathol.* **175**, 927-939, doi:10.2353/ajpath.2009.081155 (2009).
139 Goyenechea, E., Crujeiras, A. B., Abete, I., Parra, D. & Martínez, J. A. Enhanced short-term improvement of insulin response to a low-caloric diet in obese carriers the Gly482Ser variant of the PGC-1alpha gene. *Diabetes Res. Clin. Pract.* **82**, 190-196, doi:10.1016/j.diabres.2008.08.011 (2008).
140 Liang, H. & Ward, W. F. PGC-1alpha: a key regulator of energy metabolism. *Adv. Physiol. Educ.* **30**, 145-151, doi:10.1152/advan.00052.2006 (2006).
141 Sano, M. *et al.* Intramolecular control of protein stability, subnuclear compartmentalization, and coactivator function of peroxisome proliferator-activated receptor gamma coactivator 1alpha. *J. Biol. Chem.* **282**, 25970-25980, doi:10.1074/jbc.M703634200 (2007).
142 Trausch-Azar, J., Leone, T. C., Kelly, D. P. & Schwartz, A. L. Ubiquitin proteasome-dependent degradation of the transcriptional coactivator PGC-1{alpha} via the N-terminal pathway. *J. Biol. Chem.* **285**, 40192-40200, doi:10.1074/jbc.M110.131615 (2010).
143 Adamovich, Y. *et al.* The protein level of PGC-1α, a key metabolic regulator, is controlled by NADH-NQO1. *Mol. Cell. Biol.* **33**, 2603-2613, doi:10.1128/mcb.01672-12 (2013).
144 Deeb, S. S. & Brunzell, J. D. The role of the PGC1alpha Gly482Ser polymorphism in weight gain due to intensive diabetes therapy. *PPAR Res* **2009**, 649286, doi:10.1155/2009/649286 (2009).
145 Enguix, N. *et al.* Mice lacking PGC-1beta in adipose tissues reveal a dissociation between mitochondrial dysfunction and insulin resistance. *Mol Metab* **2**, 215-226, doi:10.1016/j.molmet.2013.05.004 (2013).
146 Lelliott, C. J. *et al.* Ablation of PGC-1beta results in defective mitochondrial activity, thermogenesis, hepatic function, and cardiac performance. *PLoS Biol* **4**, e369, doi:10.1371/journal.pbio.0040369 (2006).
147 Meirhaeghe, A. *et al.* Characterization of the human, mouse and rat PGC1 beta (peroxisome-proliferator-activated receptor-gamma co-activator 1 beta) gene in vitro and in vivo. *Biochem J* **373**, 155-165, doi:10.1042/BJ20030200 (2003).
148 St-Pierre, J. *et al.* Bioenergetic analysis of peroxisome proliferator-activated receptor gamma coactivators 1alpha and 1beta (PGC-1alpha and PGC-1beta) in muscle cells. *J Biol Chem* **278**, 26597-26603, doi:10.1074/jbc.M301850200 (2003).
149 Duca, F. A. *et al.* Metformin activates a duodenal Ampk-dependent pathway to lower hepatic glucose production in rats. *Nat Med* **21**, 506-511, doi:10.1038/nm.3787 (2015).
150 Napolitano, A. *et al.* Novel gut-based pharmacology of metformin in patients with type 2 diabetes mellitus. *PLoS One* **9**, e100778, doi:10.1371/journal.pone.0100778 (2014).
151 Buse, J. B. *et al.* The Primary Glucose-Lowering Effect of Metformin Resides in the Gut, Not the Circulation: Results From Short-term

Pharmacokinetic and 12-Week Dose-Ranging Studies. *Diabetes Care* **39**, 198-205, doi:10.2337/dc15-0488 (2016).
152 Forslund, K. *et al.* Disentangling type 2 diabetes and metformin treatment signatures in the human gut microbiota. *Nature* **528**, 262-266, doi:10.1038/nature15766 (2015).
153 Graham, G. G. *et al.* Clinical pharmacokinetics of metformin. *Clin Pharmacokinet* **50**, 81-98, doi:10.2165/11534750-000000000-00000 (2011).
154 Zhou, K. *et al.* Variation in the glucose transporter gene SLC2A2 is associated with glycemic response to metformin. *Nat Genet* **48**, 1055-1059, doi:10.1038/ng.3632 (2016).
155 Visscher, P. M. *et al.* 10 Years of GWAS Discovery: Biology, Function, and Translation. *Am. J. Hum. Genet.* **101**, 5-22, doi:10.1016/j.ajhg.2017.06.005 (2017).
156 Franks, P. W. *et al.* PPARGC1A coding variation may initiate impaired NEFA clearance during glucose challenge. *Diabetologia* **50**, 569-573, doi:10.1007/s00125-006-0580-1 (2007).
157 Wang, J. Y. & Doudna, J. A. CRISPR technology: A decade of genome editing is only the beginning. *Science* **379**, eadd8643, doi:10.1126/science.add8643 (2023).
158 Schoenfelder, S. & Fraser, P. Long-range enhancer-promoter contacts in gene expression control. *Nat Rev Genet* **20**, 437-455, doi:10.1038/s41576-019-0128-0 (2019).
159 Nott, A. *et al.* Brain cell type-specific enhancer-promoter interactome maps and disease-risk association. *Science* **366**, 1134-1139, doi:10.1126/science.aay0793 (2019).

Printed by Media-Tryck, Lund 2023 NORDIC SWAN ECOLABEL 3041 0903

About the Author



Mi Huang obtained his Bachelor of Science degree in Pharmacology and Master of Science degree in Pharmaceutical Chemistry from South-Central University for Nationalities in China. He completed his Ph.D. at Lund University Diabetes Centre, where he conducted research in the Genetic and Molecular Epidemiology unit. Mi Huang's doctoral thesis centered on utilizing novel experimental techniques to pinpoint genetic variations that contribute to obesity and type 2 diabetes. Additionally, he investigated the biological mechanisms underlying how these genetic variations interact with environmental factors to influence the development of these diseases.

LUND UNIVERSITY

FACULTY OF MEDICINE

Department of Clinical Sciences

Lund University, Faculty of Medicine
Doctoral Dissertation Series 2023:88
ISBN 978-91-8021-428-5
ISSN 1652-8220

9 789180 214285