Technische Universität München
TUM School of Medicine and Health

Extreme phenotype recovery in human adipocytes

Sophie Madlen Strobel

Vollständiger Abdruck der von der TUM School of Medicine and Health der Technischen Universität München zur Erlangung eines

Doctor of Philosophy (Ph.D.)

genehmigten Dissertation.

Vorsitz: Prof. Dr. Dr. Stefan Engelhardt

Betreuer*in: Prof. Dr. Johann J. Hauner

Prüfer*innen der Dissertation:

1. Prof. Dr. Marc Martignoni
2. Prof. Dr. Melina Claussnitzer
3. Prof. Dr. Matthias Blüher

Die Dissertation wurde am 03.07.2023 bei der TUM School of Medicine and Health der Technischen Universität München eingereicht und durch die TUM School of Medicine and Health am 25.11.2023 angenommen.

Acknowledgement

I am grateful and would like to express my appreciation to the individuals who have played a significant role in the successful completion of my dissertation.

First and foremost, I am deeply grateful to my thesis advisor, Prof. Hans Hauner, my second supervisor, Prof. Marc Martignoni, and my mentor, Prof. Melina Claussnitzer. Their support, expertise, and commitment have been instrumental in guiding and shaping my scientific work. Thanks to their guidance, I achieved results that make an important scientific contribution to the field of genetics. I am truly grateful for the time and resources they have dedicated to my academic development. Their understanding and assistance in striking a balance between my medical studies and experimental work have been truly invaluable.

I would also like to express my sincere thanks to my colleagues in the laboratory. Their tireless efforts, expertise, and collaborative spirit have played a pivotal role in the successful outcome of my dissertation. In particular, I am deeply grateful to Samantha Laber and Alham Sadaat for their support, insightful scientific discussions, and everything they taught me.

An incredible opportunity in my PhD journey was the work at the Broad Institute of MIT and Harvard. I am tremendously thankful for the invaluable experience and the inspiring atmosphere I was fortunate enough to encounter during my time there. The insights gained and the collaborations with internationally renowned researchers have significantly broadened my scientific horizons and opened up new perspectives. I extend my sincere gratitude to Prof. Melina Claussnitzer for granting me this exceptional opportunity.

The skills acquired through my scientific work and the inspirations I have gained have further deepened my passion for research. I am firmly committed to continuing this pursuit and applying my knowledge in the field of pediatrics to contribute to the well-being of young people.

Lastly, I would like to express my deepest appreciation to my family. Their unwavering support and encouragement have been my driving force, enabling me to overcome challenges and consistently perform at my best.

To all those who have accompanied me and provided support throughout this journey, I extend my heartfelt thanks

Table of contents

1 Aim of the work and summary

Over the last decade a plethora of genetic loci influencing susceptibility to human common complex disease have been identified. The next challenge is to identify the underlying mechanisms driving disease risk for each of these individual trait- and disease-associated genetic loci and to understand patterns of cellular programs that are shared between loci of a particular disease, and across diseases. Development of novel technologies and approaches to characterize and recover disease-relevant and cell-type specific phenotypes at scale will be powerful in elucidating mechanistic insights of genetic signals and to generate hypotheses about therapeutic interventions.

In this thesis I present and discuss three manuscripts, which I have co-led. The first manuscript describes the systematic characterization of mechanisms of disease-associated genetic and polygenic variation using a high-content imaging based phenotypic profiling assay, namely LipocyteProfiler. In the second manuscript, I report the identification of cellular and morphological consequences of the *2q24.3* metabolic risk locus in adipose derived mesenchymal stem cells. Using LipocyteProfiler, I characterize the underlying mechanism of this locus on adipocyte-specific actin-cytoskeleton remodelling and consequently adipocyte differentiation. Finally, the third manuscript elucidates cellular programs underlying the *18q21.33* locus association characterized by a lipodystrophy-like phenotype in visceral and subcutaneous adipocytes by integrating image-based profiles and transcriptomic data. Lastly, I summarize my work on a CRISPR/Cas9-based tool to engineer genetic variants *in vitro* to establish causality between genetic variation and molecular and cellular phenotypic effects. While this work is not at a point to compile a manuscript, it contributes to the evaluation of CRISPR-based genetic perturbation schemes in the context of extreme phenotype recovery.

Together, the data presented in this thesis demonstrate novel frameworks for the dissection of pathophysiology of common complex diseases at scale and may support the development of novel therapeutic targets for personalized medicine.

2 Ziel der Arbeit und Zusammenfassung

Im letzten Jahrzehnt wurde eine Vielzahl von Genloki identifiziert, die die Anfälligkeit für Volkskrankheiten des Menschen beeinflussen. Die Herausforderung besteht nun darin, die zugrunde liegenden Mechanismen zu identifizieren, die dem Krankheitsrisiko für jeden dieser merkmals- und krankheits-assoziierten genetischen Loki zugrunde liegen und Muster zellulärer Programme zu verstehen, die mehreren Loki einer bestimmten Krankheit unterliegen. Die Entwicklung neuer Technologien und Ansätze zur skalierbaren Charakterisierung von krankheitsrelevanten und zelltypspezifischen Phänotypen wird dazu beitragen, mechanistische Einblicke in genetische Signale zu bekommen und Hypothesen über therapeutische Interventionen zu generieren.

In dieser Dissertation präsentiere und diskutiere ich drei Manuskripte, an deren Entstehung ich maßgeblich mitgearbeitet habe. Das erste Manuskript beschreibt die systematische Charakterisierung von Mechanismen, die krankheits-assoziierter genetischer und polygenetischer Variation zugrunde liegen. Hierzu habe ich eine auf High-Content Imaging basierende phänotypische Profiling Methode, die wir LipocyteProfiler genannt haben, mitentwickelt und angewandt. Im zweiten Manuskript identifiziere ich zelluläre und morphologische Konsequenzen des metabolischen Risiko-Lokus *2q24.3* in aus Fettgewebe stammenden mesenchymalen Stammzellen. Unter Verwendung von LipocyteProfiler charakterisiere ich die zugrunde liegenden Mechanismen dieses Lokus auf die Adipozyten-spezifische Aktin-Zytoskelett-Umstrukturierung und die daraus resultierenden Folgen für die Differenzierung von Adipozyten. Das dritte Manuskript beschäftigt sich mit den zellulären Programmen, die der Assoziation des *18q21.33*-Lokus mit Lipodystrophie-ähnlichen Phänotypen in viszeralen und subkutanen Adipozyten zugrunde liegen, unter Einbindung von High Content Imaging Profilierungs und Transkriptom Daten-Analysen. Schließlich fasse ich meine Arbeit an einer auf CRISPR/Cas9 basierenden Methode zusammen, um genetische Varianten *in-vitro* zu modifizieren, um Kausalität zwischen genetischen Veränderungen und molekularen und zellulären phänotypischen Effekten herzustellen. Obwohl diese Arbeit noch nicht an dem Punkt ist, um daraus ein Manuskript zu erstellen, trägt sie zur Bewertung von CRISPR-basierten genetischen Modifikation im Kontext der Entstehung von extremen Phänotypen bei.

Gemeinsam bieten die in dieser Dissertation präsentierten Daten ein neues Gerüst für die skalierbare Analysen der Pathophysiologie komplexer Volkskrankheiten und können die Entwicklung neuer therapeutischer Interventionen für Personalisierte Medizin unterstützen.

3 Introduction

3.1 Genetic variation influences susceptibility to disease

3.1.1 Human geneticists have identified thousands of genetic loci associated with hundreds of diseases and traits

In 2003 researchers of the international human genome sequencing consortium published the complete genetic code in humans[1]. Advances of high-throughput sequencing technologies and efforts of the International HapMap project to characterize haplotype structures and common patterns of genetic variation inheritance[2–4] paved the way to systematically study and elucidate the contribution of genetics to human traits and diseases.

Since then, genomes of hundreds of thousands of individuals have been genotyped and/or sequenced, identifying common variations of nucleotides between individuals with the same traits or diseases. These single base-pair (bp) DNA variations among individuals, called single-nucleotide polymorphism (SNP), occur on average every 300bp in the genome[5]. Many of those SNPs lie within so-called haplotypes and are inherited together. More specifically, haplotypes are genomic regions that are more likely to preserve linear structure during recombination of meiosis. Therefore, nearby SNPs of a haplotype are more likely to stay together during meiosis and to be inherited together. This likelihood can be quantified using linkage disequilibrium (LD) and depends on recombination rates. As a consequence, SNPs with high LD are more frequently inherited together than SNPs in low LD[6]. By defining common patterns of inheritance of genetic variation using LD structures, scientists are able to interrogate SNPs at genome-scale, by genotyping one SNP of a haplotype, namely tag-SNP, and imputing other SNPs using LD structures. Using this approach enabled genome-wide association studies (GWAS) of large cohorts, in which disease or traits within a population were found to be associated with genome-wide genetic variation, and led to the discovery of genetic variants associated with human disease and traits (Figure 1).

Comprehensive genome-wide-association studies have since discovered thousands of loci associated with hundreds of human traits and diseases[7]. It is now known that genetic variants tend to be spread across the whole genome and have pleiotropic effects, i.e. one variant influencing more than one trait in a context-specific manner, including different cell types, regulation of different genes, action under specific conditions, altogether resulting in a variety of intermediate phenotypes converging in a complex disease phenotype. Interestingly, over 90% of those trait- and disease-associated SNPs are mapped to the non-coding genome[8,9], suggesting susceptibility to common complex diseases is due to alterations in gene regulatory networks, rather than changes of protein-coding sequences. Further, common variants often lie within complex haplotypes with a lot of variants in high LD making it computationally and experimentally challenging to identify the causal variant[10]. Taken together, the characterization of underlying mechanisms leading from genetic variation to disease phenotype remains an obstacle.

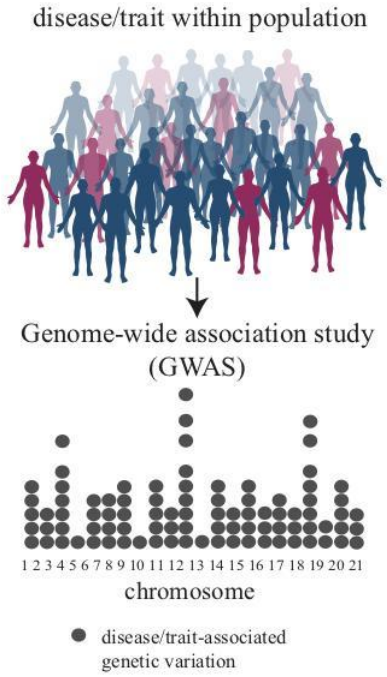


Figure 1: GWAS of large cohorts led to the discovery of genetic variation associated with human disease and traits.

3.1.2 Characterizing the underlying mechanisms of disease-associated genetic variation is challenging

First attempts to dissect genetic variants associated with disease were focused on rare variants (< 0.5% minor allele frequency (MAF)) with large phenotypic effects mediated through the regulation of a single gene leading to deleterious effects on protein function[11]. So-called monogenic diseases occur within families and are identified using linkage analysis, pedigrees and nowadays whole exome sequencing or whole genome sequencing, followed by functional *in vitro* and *in vivo* assays[12].

Other than monogenic diseases, also known as Mendelian disorders, caused by typically one rare mutation with large effects, common diseases, such as diabetes, obesity, cardiovascular diseases, psychiatric illnesses and inflammatory diseases, are associated with common variants with

higher allele frequency (> 5% MAF) within populations[12]. Each common variant usually has only small to modest phenotypic effects, however many common variants are thought to influence human disease and in combination lead to complex functional consequences[13–15].

Initially, the understanding of genetic inheritability on common complex disease was that multiple genetic variants - polygenic - with small effects in concert would explain genetic susceptibility for disease[16]. However, unexpectedly most genome-wide significant hits identified by GWAS could not explain total vulnerability for common disease, and the question arose what contributes to the missing genetic signal. Over the years, this so-called "missing heritability"[14,17,18] mystery has been addressed and is expected to be mostly resolved with increasing large-scale sequencing studies in diverse ancestries enabled by advances in sequencing technologies. It was also observed that rare variants could have modest rather than deleterious effects, based on the notion that rare variants have likely arisen after migration out of Africa and as such shared among closely related geographical and ethnical groups[19]. With emerging population-based sequencing studies the genetics community has realized that the genetic architecture of complex diseases and traits can be explained by a continuous spectrum of variant allele frequencies and associated effects, which together support a primary role of common variants in the pathogenesis of common complex diseases and traits. This expansion of knowledge of genetic architecture and genetic contribution to common complex diseases changed the historical understanding of two distinct models (monogenic and polygenic) to a spectrum of allele frequencies and effect sizes influencing common diseases[11] (Figure 2).

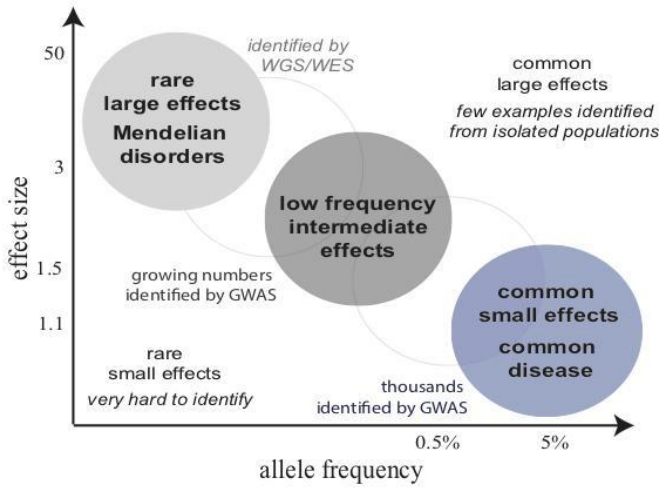


Figure 2: Spectrum of allelic frequency and effect size contributing to disease suceptibility.
Figure adapted from Manolio *et al.*

In parallel, advances in understanding the structure and function of the human genome were used to unravel genetic association with human disease. For example, novel approaches to study epigenetic regulation of genetic information in context of cell type specificity and environmental conditions[20,21], conformation and physical contacts across the genome[22], and functional relationship between genetic variation and gene expression across cell types and environmental conditions[23,24] were applied to identify mechanism of genotype-phenotype association.

Another fundamental discovery was made in 2012, when scientists discovered and engineered natural defense tools of prokaryotic organisms to edit or perturb the genetic code of humans[25]. The discovery of CRISPR genome editing technologies and its growing applications and scaling approaches emerged as a new dimension for characterizing the function of genetic variants and epigenetic regulation, and accelerated advances in the field of functional genomics.

In parallel, read-outs were developed to assess functional consequences of natural genetic variation or introduced genetic manipulation on intermediate (eg. transcriptome, proteome) and complex phenotypes. However, technologies to systematically interrogate genetic variation are currently sparse.

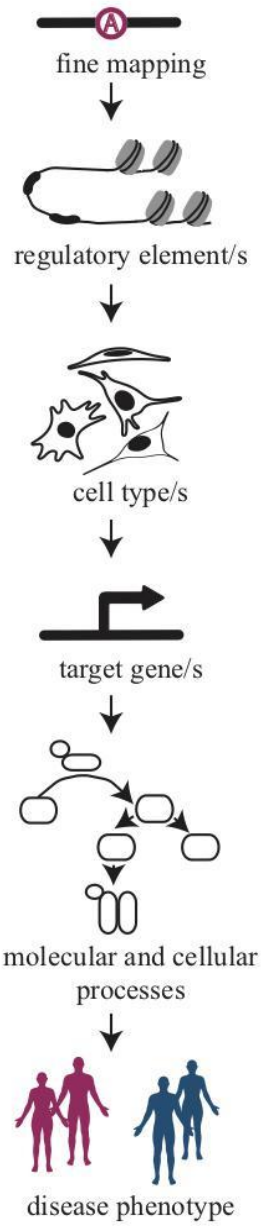


Figure 3: Variant-to-function concept

3.1.3 Single locus dissection studies showcase the variant-to-function concept

Over the last decades thousands of associations of common genetic variation with diseases have been identified through GWAS leveraging large biobanks. However, advances in understanding how these variants are linked to disease, by step-by-step characterization of their mechanisms involving identification of cell-type/s and context of action, mediated target gene/genes, effects on molecular, cellular and physiological processes, is challenging

and therefore has not proceeded as fast as identifying those variants[12] (Figure 3).

Studies of dissecting human disease-associated genetic variants are time consuming and experimentally exhausting, hence scientists focused on dissecting one locus at a time. Successful single-locus dissection studies, like the dissection of the *FTO* locus[26] or the characterization of common complex variants affecting red blood cells[27], are pioneering studies and showcased the idea of variant-to-function approaches in the field of functional genomics. These studies also demonstrated that genetic variants can regulate target genes in far distance including target genes without obvious connection or previously thought relationship to pathophysiology of disease[26].

To path from a full understanding of genetic variation and characterization of mechanistic insights to clinical application and ultimately genomic medicine, efforts of dissecting one-locus-at-the-time are ineffective and insufficient to keep up with the speed of genome-wide variant discovery. Hence, novel approaches are needed to systematically dissect genome-wide genetic risk loci across diseases and different ethnicities/populations[12]. The ongoing rise of new scalable technologies and data integration of complex readouts assessing epigenomics, transcriptomics, proteomics and morphological and cellular signatures, will help to accelerate functional genomics and translation of genomics into clinical practice.

3.2 Integration of phenotypic profiling technologies accelerate functional characterization of genetic variation

3.2.1 Multiparametric image-based phenotypic profiling quantifies complex phenotypes

To investigate and quantify complex disease-associated phenotypes, phenotypic profiling technologies are powerful tools. Phenotypic profiling assays cover the full spectrum of phenotypic read-outs, e.g. transcriptomic, proteomic, metabolomic and image-based technologies[28].

Image-based profiling provides high information content that is currently several orders of magnitude higher than other phenotypic profiling tools[28]. While observations by-eye can capture one or few cellular characteristics and are often biased from *a priori* knowledge, image-based

profiling technologies simultaneously score thousands of features of fluorescence-labeled cellular components in a relatively unbiased way[29] .

One example of an established high-content imaging method is Cell Painting[30]. Cell Painting multiplexes six fluorescent dyes for generic cellular organelles to reveal large amounts of morphological signatures, while maintaining compatibility with standard microscopy systems, and using relatively inexpensive easy-to-use fluorescence dyes (instead of high-cost antibodies), making it generalizable, inexpensive and broadly applicable to perform high-throughput experiments[30].

Biological information gained by thousands of precise measurements elucidate a complex phenotype that might represent the convergence of genetic, transcriptional, and proteomic states[31,32]. Due to the quantity of information obtained, appropriate image analysis pipelines are needed to process and generate intuitive data outputs which are relatively easy to understand. One example of such pipelines is the open source software Cell Profiler[29]. Cell Profiler uses a supervised machine learning algorithm to extract thousands of features from fluorescence microscope images. Each feature represents the quantification, obtained by different measurements, of either cell size, shape, fluorescence stain intensity, object granularity, or pixel patterning of different cellular sub-compartments or organelles. The features in concert provide information-rich cellular and morphological profiles[29] (Figure 4).

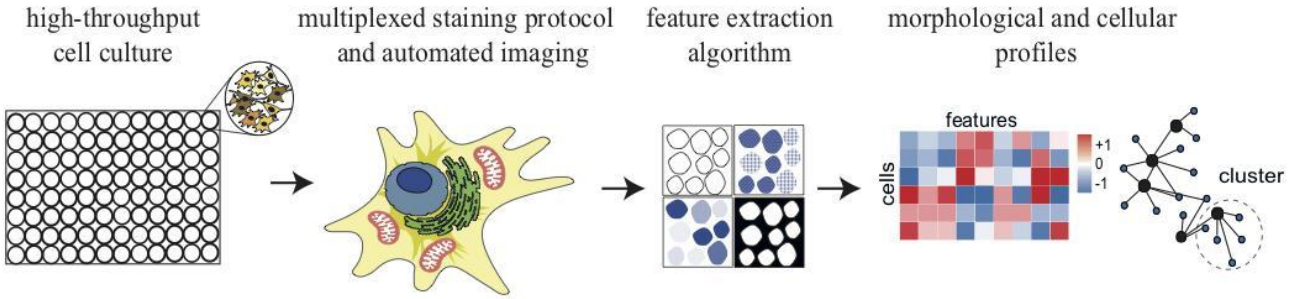


Figure 4: Schematic of a multiparametric high-content imaging assay

Since the development of multiparametric high-content imaging assays, scientists used image-based phenotypic profiling or screening to illuminate mechanism of action of small molecules, ascertain compound toxicity, generate compound fingerprints, and to annotated gene function in gene expression screens or CRISPR/Cas9 gene perturbation screens[33–38].

However, disease-aware customised high-content imaging assays for specific disease-relevant cell models, e.g. adipocytes, to interrogate genetic association signals of common complex disease, such as metabolic traits, are currently missing.

3.2.2 Disease-relevant high-fidelity cellular model systems are necessary to study genetic association with disease

Although morphological profiling enables unconstrained discovery, interpretation of profiles to obtain biological meaningful results is challenging and stays often at an abstract and descriptive level, making it hard to decipher mechanistic underpinnings of genetic variation. To gain biological meaningful insights and create appropriate disease models enabling studying pathophysiology, disease-relevant cell assays, including disease-relevant cell-types and conditions, are needed.

Immortalised cell-lines, although limitless available and therefore mostly used to study cell biology, are not necessarily the perfect model system due to several limitations such as: (1) they harbor complex genetic aberrations, and are often not diploid; (2) they are adapted to standardized unphysiological culture conditions; and (3) they are often derived from tissue of diseased individuals, hence have little to no similarity to their original derived tissue[39].

In recent years, efforts in reprogramming and differentiation of isogenic pluripotent stem cells (iPSC) have been made to obtain more suitable cell models as well as to increase genetic heterogeneity in cell models that can be leveraged to study natural genetic variation and how this contributes to cell behaviour and disease[40–42]. Although, iPSCs are a great resource some things need to be considered when using iPSCs: (1) reprogramming and culturing potentially introduce unwanted epigenetic and/or genetic variability that provide noise and may confound downstream analysis[43,44], (2) differentiation protocols can be inefficient, and (3) high experience and detailed experimental planning is needed to use them successfully[39]. The Human Induced Pluripotent Stem Cells Initiative (HipSci, http://www.hipsci.org/), as one example, started to build a large biobank of comprehensively characterized isogenic pluripotent stem cells (iPSC) from healthy and diseased donors for the research community to study population-based genetic variation. Indeed, it has been reported that common genetic variation drives molecular heterogeneity in human iPSCs and quantitative assays of cell morphology demonstrated a donor contribution in the range of 8%-23%[42]. Identifying extrinsic versus intrinsic drivers of variation in cell behaviour

in human iPSC lines further revealed that genetic variation contributes to phenotypic variation[45]. This suggests that genetically-driven intrinsic differences are maintained in these cells and that they can be used to study inter-individual genetic variation.

A potentially more biologically relevant alternative to iPSCs are primary cells derived from disease-relevant tissue from healthy and unhealthy donors. For example it was previously shown that human adipose-tissue derived mesenchymal stem cells can be differentiated into fully functional adipocytes and therefore have been used to study adipocyte biology[46,47] and adipose-related metabolic disease phenotypes[48,49].

With the creation of large biobanks and the development of high-throughput technologies for epi-/genetic screens and phenotypic profiling powerful tools are in place and need to be systematically applied in appropriate disease models to comprehensively characterize genetic variation on population scale. The integration of these different data is challenging, but impactful when elucidating underlying mechanisms of disease-associated genetic variation and further identifying personalized drugs.

3.3 Metabolic diseases are influenced by complex genetic variation

3.3.1 Metabolic traits are highly heritable

The prevalence of metabolic diseases like type 2 diabetes (T2D) and obesity are rising worldwide especially in developing countries and are a threat for public health. Metabolic diseases are interrelated and often described in context of the metabolic syndrome, which summarises the collective manifestation of metabolic traits, such as dyslipidemia, hypertension and elevated glucose, and associate with increased risk for cardiovascular diseases and common cancer resulting in higher morbidity and mortality[50–52].

Metabolic disease/traits are heritable and GWAS of clinically phenotyped populations for metabolic disease/traits including T2D, fasting glucose, waist-to-hip ratio and body mass index have identified hundreds of genetic loci[11]. Together, these genetic risk loci harbor thousands of associated variants that point to a diverse set of tissues of action and biological mechanisms, including adipocytes and their precursors[26,53–56]. However, the majority of loci have not been conclusively linked to their effector genes and biological mechanisms [57].

Overall heritability of T2D, for example, is estimated to be 25-80%[58,59]. T2D is a heterogeneous disease associated with various metabolic traits (e.g. hyperglycemia, insulin resistance, adverse waist-to-hip ratio, hyperlipidemia), and known to be influenced by genetic, environmental, and behavioural factors; to implicate various organs (e.g. adipose tissue, skeletal muscle, liver and pancreas); to lead to various metabolic dysregulation impacting whole body physiology; converging in a complex disease phenotype[60–63]. Therefore, T2D is no longer thought to be only a disease of hyperglycemia, but rather a syndrome of multiple metabolic disturbances[60,64].

Although monogenic forms of T2D, like maturity-onset diabetes of the young (MODY), occur within affected families and map to rare variants with strong phenotypic effects, the majority of the genetic inheritability is explained by common variants that lie in intronic or intergenic regions often mapping to regulatory elements[11]. Due to their characteristics genetic variants associated with type 2 diabetes are thought to be cell-type and context-specific, and often have pleiotropic effects. To unravel the complexity of association of genetic variation and disease phenotype, researchers developed frameworks using epigenomic and transcriptomic data to score loci according to their tissue of action[62] or cluster loci according to their phenotypic presentation[63]. However, as eluted to above the functional consequences of associated genetic variants are poorly understood.

3.3.2 Lipodystrophy-like phenotypes demonstrate complexity of metabolic disease

The clinical presentation of T2D is often complex involving the collective manifestation of metabolic traits. Causal mechanisms and their interactions of metabolic traits are not fully understood, complicating prevention and management of disease with established therapeutic options.

Lipodystrophy-like phenotypes, for example, demonstrate heterogeneity of metabolic traits associated with T2D in context of underlying mechanism and phenotypic presentation. Although T2D is often associated with obesity, there is growing evidence that some patients with normal weight occasionally are characterized by metabolic abnormalities. These patients are described in literature as metabolically obese normal-weight (MONW) individuals[65]. GWAS for insulin resistance, known as key mediator of elevated metabolic and cardiovascular risk[66,67] , identified loci that are linked to a lipodystrophy-like phenotype with insulin resistance, elevated waist-to-hip ratio, normal body mass index (BMI) with central adiposity and peripheral fat

depletion[68]. Lipodystrophy syndromes are known as a monogenic disease and molecular mechanisms of affected genes have been previously described. However, some of the lipodystrophic features are also present in general populations suggesting that the spectrum of disease susceptibility, from low to high polygenic risk, is linked to a spectrum of phenotypic effects, ranging from minor lipodystrophic clinical representations to a severe clinical phenotype, like familial partial lipodystrophy type 1 (FPLD1) or Köbberling-type lipodystrophy[68,69]. However, a comprehensive understanding of the underlying mechanistic regulation of polygenic risk for lipodystrophy-like phenotypes and mediated genes are missing.

3.3.3 Polygenic risk scores quantify individual genetic liability of metabolic disease

Historically, genetic testing in the clinic to predict or diagnose an individual's risk to a particular disease was employed in monogenic disease that ran in families and where the effect size was relatively large[70]. More recently, the realisation that thousands of genetic loci influence an individual's susceptibility to develop a common complex disease led to the concept of generating a combined score, called polygenic risk score, to predict genetic vulnerability of an individual[71].

Polygenic risk, a cumulative risk of millions of genetic common variants associated with disease, is scored by combining genetic information of all single variants associated with a particular disease into a number representing an individual's polygenic risk score (Figure 5). Today, such scores have been generated for various common diseases, like obesity[72], coronary artery disease[71] and type 2 diabetes[73]. It has been shown that individuals at the tail ends of polygenic risk scores (top and bottom 2.5%) differ 10-times in their prevalence to T2D[73]. Additionally, the effect of

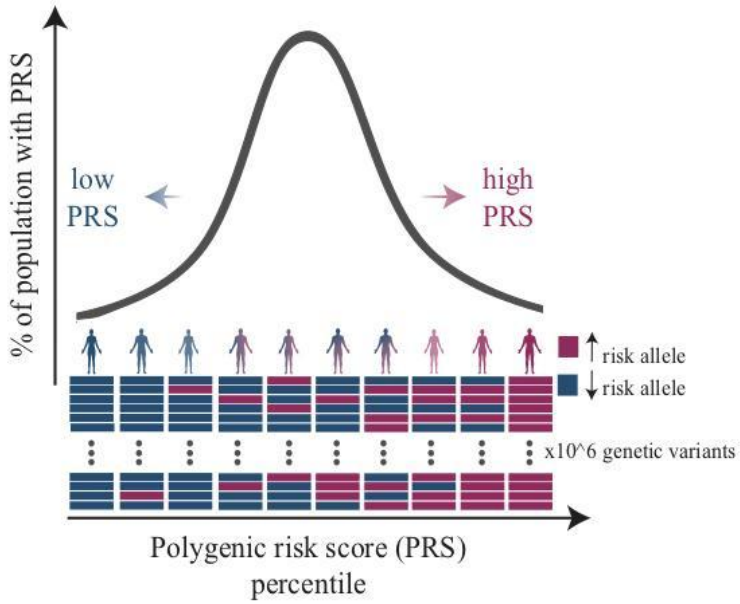


Figure 5: Polygenic risk scores quantify cumulative risk of genetic common variants associated with disease

polygenic scores for obesity for example, can be similar to rare, monogenic forms of this disease[71,72], and a high polygenic score is a strong risk factor for severe obesity and associated diseases. However, the molecular mechanisms underlying differential polygenic scores are not characterized yet.

Polygenic risk scores (PRS) quantify disease predisposition and have the potential, especially in combination with other clinical data to advance prevention, progression, and clinical management of disease. Although, PRSs are promising, they are not routinely used in clinic, due to several obstacles: (1) PRS are constructed based on genomic data of Europeans, and show reduced performance in other ethnicities[74,75]; (2) type 2 diabetes is influenced by genetic and non-genetic factors (eg. lifestyle, environment), hence conclusions only based on PRS without considering non-genetic influences could lead to misleading risk prediction; (3) established clinical parameters and biomarkers already perform quite well in predicting type 2 diabetes raising the question what additional benefit PRS could add[61]; (4) type 2 diabetes is a heterogeneous disease, involving many cellular processes, which are not fully understood[64].
To tackle the latter, partitioned - process specific - polygenic scores were generated by clustering genetic variants due to biological processes[62]. Clinically informed, genetically anchored scores, account for genetic and phenotypic heterogeneity of metabolic disease like T2D and are likely to be more informative for optimizing clinical management (eg. selecting a process-tailored drug) and monitoring disease progression[61].
The complexity of metabolic phenotypes and their interrelationship showcase the challenge of elucidating genetic association with metabolic disease. Leveraging genetically-driven differences in cell behaviour in relevant cell types (e.g. adipocytes and hepatocytes for T2D traits) using PRSs together with scalable disease-relevant profiling assays therefore have the potential for unravelling underlying molecular mechanisms of polygenic disease.

3.3.4 Genome engineering is necessary to establish causality of disease mechanisms

Over the last decade, genome engineering technologies have been developed and optimised to perform targeted genome modifications. The toolbox of CRISPR-based genome engineering technologies has been expanded with versatile approaches to study genetic variation at scale and to perform precise genome editing enabling to establish causality of disease mechanisms.

Genetic screens are powerful assays to study genetic variation at scale using CRISPR or RNA interference (RNAi) technologies[76]. Over the last decade various applications have been developed to study gene function (eg. CRISPR knockout screens[77,78]) and interactions (eg. Perturb-seq[79]), functional consequences of SNPs (eg. base editor screens[80]), and genetic regulatory elements (eg. CRISPR interference, CRISPR activation screens[81,82]) in pooled or arrayed experimental settings combined with high-throughput phenotypic read-out technologies (eg. single cell transcriptomics and epigenomics, high-content imaging)[12,83].

To ascertain causality between a genetic variant and disease phenotype, genome engineering approaches are necessary. CRISPR-based technologies are powerful tools to edit disease-associated genetic variation. Today, many approaches exist using either classic CRISPR/Cas9-mediated homology-directed repair pathways or novel technologies like base editor[84–86] or prime editor[87]. However, currently available approaches mostly have low editing efficiency[88], often require time consuming single clonal expansion steps, potentially introduce transcriptome-wide off-target edits[89], and/or are hampered by limited targetability of much of the genome[84,87]. To leverage the power of genetic engineering tools for systematically characterizing functional consequences of identified genetic variants (coding and non-coding), it is necessary to optimise for tools that are broadly applicable, that edit accurately (only change a predetermined mutation) and are scalable by reducing experimental time and costs.

Together, phenotypic profiling in combination with CRISPR engineering approaches pave the way for recovering extreme phenotypes and establishing causality between genetic modifications and disease pathophysiology.

In this thesis, I present and discuss three manuscripts aiming to create frameworks for dissecting disease-associated genetic variation at scale. First, I introduce a framework for systematically characterizing underlying mechanisms of disease-associated genetic and polygenic variation using high-content imaging, called LipocyteProfiler. In the second manuscript, I apply LipocyteProfiler on the *2q24.3* metabolic risk locus and identify cellular and morphological consequences of the *2q24.3* metabolic risk locus on actin cytoskeleton remodeling and

consequently adipocyte differentiation in human adipose derived mesenchymal stem cells. In the third manuscript, I elucidate cellular programs of the *18q21.33* locus association with a lipodystrophy-like phenotype in visceral and subcutaneous adipocytes by integrating image-based profiling and transcriptional data. Finally, I briefly introduce my work to develop and test an experimental method, CRISPR-SAVE, to engineer genetic variation in a human adipocyte model system and stress its challenges and limitations that I discovered while developing the approach.

4 Summary of publications

4.1 Discovering cellular programs of intrinsic and extrinsic drivers of metabolic traits using LipocyteProfiler

This manuscript describes the development and application of a multiparametric phenotypic profiling assay that I co-developed to characterize extrinsic and intrinsic drivers of metabolic disease in lipid-filled cell types, called LipocyteProfiler. This work was co-led by Samantha Laber and myself.

Rationale: Thus far, image-based phenotypic profiling assays have been limited to capture generic morphological structures of cellular organelles and cellular processes, such as cell proliferation and growth[30]. Leveraging high-information content obtained from image-based phenotypic profiling to discover cell- and/or disease-specific processes requires customized assays, which could be used to link genetic variation to disease-relevant cellular programs at scale.

Lipid droplets are storage organelles and involved in whole body metabolism and energy homeostasis. They are found in all cell types, either as part of physiological processes or under pathophysiological conditions[90]. Dynamics of lipid droplet formation are highly relevant processes for the manifestation and progression of metabolic disease. Due to their morphological characteristics, lipid droplets and their dynamic changes can be captured and studied using image-based profiling.

In this manuscript, we introduce LipocyteProfiler, a metabolic disease-oriented phenotypic profiling system for lipid-accumulating cells that can be used to generate high-throughput disease-focused read-outs. LipocyteProfiler is a modified version of CellPainting[30] by incorporating BODIPY, a dye that stains neutral lipids, to measure morphological and cellular profiles of lipid-related features, and thus captures lipocyte-relevant phenotypes in addition to generic morphological profiles.

Results and Contribution: Briefly, we validated LipocyteProfiler by testing its ability to capture phenotypes of interest: (1) We showed its ability to identify known lipocyte specific biology (e.g. lipid accumulation capacity during differentiation); (2) we knocked-out genes known to alter adipogenesis or adipocyte function using a CRISPR/Cas9 and assessed morphological and cellular signatures; (3) we integrated gene expression of RNA-seq data with lipocyte profiles of corresponding AMSCs and identified biological meaningful feature-gene connections; and (4) we characterized cellular programs of drug perturbations. Next, we applied LipocyteProfiler to subcutaneous and visceral human adipocytes and demonstrated its ability to identify adipose depot-specific morphological and cellular signatures. We then used our assay to discover novel functional insights of genetic association to polygenic risk of metabolic traits, including insulin resistance, waist-to-hip ratio and the polygenic contribution to lipodystrophy. In summary, we demonstrated that LipocyteProfiler presents a comprehensive framework to interrogate genetic signals and environmental effects associated with complex human disease in an unbiased way.

I was mainly responsible for the following parts of the manuscript: (1) I differentiated 65 adipose derived mesenchymal stem cells (AMSCs) of the Munich Obesity BioBank (MOBB). (2) I carried out a customised staining protocol in AMSCs, namely LipocytePainting, at four time points of adipocyte differentiation. (3) I generated a Cas9 stable cell line of an established white adipocyte line (hWAT)[91] and performed a pilot CRISPR/Cas9 knockout screen of 42 genes, known to be involved in adipocyte function, using LipocyteProfiler as a read-out. (3) I conceived the analytical LipocyteProfiler framework and performed all computations on high-content imaging data and RNA-seq data. (4) I discussed all results with my co-authors and contributed equally to the visualization of the results and writing of the manuscript.

4.2 A non-coding variant linked to metabolic obesity with normal weight affects actin remodelling in subcutaneous adipocyte

This manuscript reports the functional characterization of the link between the *2q24.3* locus and a lipodystrophy-like phenotype, characterized by increased risk of type 2 diabetes and decreased risk of adiposity related traits.

Rationale: When I joined the project, my co-authors had identified the causal variant (rs6712203), the putative target gene (*COBLL1*) and identified a function of *COBLL1* in cytoskeletal remodeling during adipocyte differentiation, subsequently resulting in impaired insulin sensitivity. To further link functional consequences of the *2q24.3* locus on morphological and cellular phenotypes in AMSCs, I used LipocyteProfiler for image-based phenotypic profiling at this locus.

Results and Contribution: My contribution consisted of linking the risk haplotype at this locus to its phenotypic consequences using LipocyteProfiler, that I had co-developed (see 4.1). More specifically, I performed LipocytePainting at 3-4 time points of differentiation in a number of settings: (1) using siRNA for *COBLL1* knockdown in pre-adipocytes (knockdown was performed by Samantha Laber), followed by 14 days of differentiation to assess morphological consequences on actin cytoskeleton remodeling and differentiation capacity in subcutaneous AMSCs following *COBLL1* knockdown; (2) using LipocytePainting data that I had generated in 4.1 by stratifying samples based on rs6712203 genotype to assess consequences on morphological and cellular signatures in visceral and subcutaneous adipocytes at four time points of adipogenesis; (3) using mouse AMSCs from *Cobll1*-KO animals, isolated by Debora Sobreira from the University of Chicago, which were differentiated for 10 days.

Finally, I compared morphological profiles driven by genetic variation of rs6712203 with siRNA-mediated targeted gene perturbation of *COBLL1* and identified that the risk haplotype matches a profile of perturbed *COBLL1* expression by altering actin cytoskeleton remodeling during adipocyte differentiation, and subsequently lipid accumulation capacity in mature subcutaneous adipocytes. I did not observe any phenotypic effects in visceral adipocytes, suggesting that underlying mechanisms of *2q24.3* locus are specific for adipose depot. In the murine model I found that *Cobll1* knockout in mice affects actin cytoskeleton remodeling and

lipid accumulation during *in vitro* adipocyte differentiation, mimicking our observations in human adipocytes. I discussed and interpreted my results with my co-authors and was responsible for visualization of the results, compiling of the figures and writing the first draft of the image-based profiling part of the manuscript.

4.3 A non-coding metabolic risk variant, rs12454712, associates with context-dependent adipocyte molecular and cellular programs in humans

This manuscript describes the functional dissection of the *18q21.33* pleiotropic risk locus, associated with a lipodystrophy-like phenotype, into mediating mechanisms. This work was co-led by Samantha Laber and myself.

Rationale: With the advent of GWAS hundreds of disease-associated genetic loci have been identified that map to more than one disease or trait, act in different cell-types and/or under specific conditions, and converge on a complex phenotype. Type 2 diabetes is a heterogeneous disease associated with hundreds of genetic loci with pleiotropic effects making it challenging to characterize mechanistic underpinnings of genetic risk loci by linking genetic variation to phenotypes.
In this study, we elucidate functional consequences of the *18q21.33* metabolic risk locus associated with a lipodystrophy-like phenotype in visceral and subcutaneous adipocytes by integrating transcriptional and image-based profiling.

Results and Contribution: Briefly, we identified the pleiotropic characteristics of *18q21.33* illuminating different cellular programs in visceral and subcutaneous differentiated adipocytes mainly mediated through three target genes *BCL2*, *KDSR* and *VPS4B*. By phenotypic profiling, using LipocyteProfiler, we characterized depot- and context-specific cellular pathways of apoptosis and altered thermogenesis activity in subcutaneous and visceral adipocytes, respectively, which are cellular phenotypes relevant for peripheral subcutaneous fat depletion and central adiposity on the organismal level. Finally, we identified similarities in morphological and cellular signatures between genetically-driven signals of *18q21.33* locus and polygenic risk for waist-to-hip ratio adjusted for BMI (WHRadjBMI; the morphological and cellular profile for

WHRadjBMI polygenic risk was generated in 4.1) in subcutaneous mature adipocytes. Together these results highlight the complexity of pleiotropic loci and their association with heterogeneous disease phenotypes and provide an approach for resolving their mechanistic underpinnings.

My contribution to this study consisted of the following: (1) I analysed high-content imaging data using the framework I developed in 4.1 and RNA-seq data we generated in 4.1 stratified based on rs12454712 genotype in subcutaneous and visceral AMSCs at 4 time point of differentiation. (3) I performed LipocytePainting in subcutaneous AMSCs following siRNA-mediated *BCL2* knockdown in preadipocytes, which was performed by Samantha Laber, and 14 days of differentiation at four time points and analysed the data using LipocyteProfiler. (4) I discussed the results with my co-authors and contributed equally to the visualization and writing of the manuscript.

4.4. CRISPR-SAVE, a *piggyBac*-mediated CRISPR/Cas9 genome editing approach

In this project I developed a *piggyBac*-mediated CRISPR/Cas9 genome editing tool, namely CRISPR-SAVE, which potentially enables scalable accurate variant editing in future studies. I tested the approach in an established cell model of immortalised white adipose-derived mesenchymal cells (hWAT)[91] using prioritised variants of the *2q24.3* metabolic risk locus that we functionally characterized in 4.2.

Rationale and Contribution: To establish causality between genetic variants and morphological and cellular profiles identified by LipocyteProfiler, we aimed to develop a tool to engineer genetic variants recovering genetically-driven phenotypes. The approach builds on a method described in [92] and performs genome modifications in four steps: (1) Ribonucleoprotein complex (RNP), Cas9 complexed with a guide RNA (gRNA), and a single stranded DNA (ssDNA) repair template comprising homology arms (~600bp) and a selection cassette flanked by *piggyBac* transposase-specific inverted repeat sequences are delivered to the cell. (2) Cells that have integrated the selection cassette are selected by antibiotic-based and FACS-assisted double positive selection. (3) After enrichment of cells that underwent genomic modifications, a *piggyBac* transposase removes the selection cassette traceless. (4) FACS-assisted negative

selection enriches transposon-negative edited cells (Figure 6). Together, CRISPR-SAVE holds promise to ameliorate one challenge of genome editing, the need for single clonal expansion of correctly edited cells, and further allows for higher efficiencies, reduced hands-on (cloning-free) and faster (no single-cloning required) generation of targeted genome modifications.

I optimised CRISPR-SAVE in hWAT and was able to (1) increase mono-allelic on-target integration efficiency of the selection cassette by optimising the delivery format and/or design of Cas9, gRNA, and DNA repair template. (2) Using RNP instead of Cas9/gRNA encoded on plasmid and long ssDNA instead of dsDNA repair template, I was able to decrease off-target and random integration events. (3) I improved *piggyBac*-mediated transposon removal in terms of experimental turn around time and cytotoxicity. I demonstrated its ability to enrich mono-allelic edited cells within ~5-6 weeks in a pooled format. To comprehensively characterize edited cell pools and corresponding single cell expansion derived clones I used Sanger and Next-Generation Sequencing technologies revealing vulnerabilities of the reported approach. More specifically, I observed genetic modification at the transposon excision site and hypothesised that during mono-allelic *piggyBac*-mediated excision of the selection cassette homologous recombination mechanisms between wild-type and edited DNA strand occur resulting in genomic aberrations at transposon removal site.

In this project, I led the conceptual work of all optimizations and the experimental work. However, further optimization is necessary to investigate causal drivers of the observed genomic modifications e.g. generating bi-allelically modified cells which might resolve genetic modifications at the transposon removal region.

5 Discussion

The manuscripts presented in this dissertation introduce a framework - LipocyteProfiler - for characterizing associations of genetic variation with metabolic disease by phenotypic profiling of disease-relevant cell models and integration of genetic/polygenic, epigenetic, transcriptional and image-based multiparametric data in lipocytes. Lipocytes are cells that accumulate lipids either under physiological conditions, like adipocytes, hepatocytes, macrophages/foam cells and glial

cells[90,93–96], or under pathophysiological conditions, like vascular smooth muscle cells, skeletal muscle cells, renal podocytes and cancer cells[97–101]. We functionally validated our assay using adipocytes and hepatocytes which are highly specialized cells to store excess energy in the form of lipid droplets and are associated with metabolic disease. We assessed and perturbed their capacity to accumulate lipids; and used the assay to identify morphological and cellular signatures of cellular programs of marker gene expression or drug treatment. Furthermore, we used LipocyteProfiler to discover novel mechanistic underpinnings of polygenic effects on metabolic traits like insulin resistance and lipodystrophy, and used it to functionally dissect two metabolic risk loci, *2q24.3* and *18q21.33,* associated with lipodystrophy-like phenotype.

Briefly, in the second manuscript I applied LipocyteProfiler on *2q24.3* metabolic risk locus associated with increased risk for T2D and decreased risk for adiposity. I identified morphological and cellular consequences on actin cytoskeleton remodeling during adipogenesis, and subsequently lipid storage resulting in a complex phenotype of disturbed insulin sensitivity and fat distribution. The study supports the hypothesis that adipocyte dysfunction and impaired adipocyte tissue expandability influences insulin resistance[68,102,103]. Notably, the *2q24.3* metabolic risk locus has been previously associated with a lipodystrophy-like phenotype[68] and identified as one of the top scoring loci among 20 T2D risk loci in the lipodystrophy cluster described in Udler *et al.*[62], inferring strong contribution to a lipodystrophic-like phenotype mediated through adipose tissue. Previous studies have characterized functional consequences of some identified metabolic risk loci associated with MONW (eg. *FAM13A* locus was linked to fat distribution and adipocyte differentiation[104]) revealing distinct underlying cellular programs of the observed phenotype[102,105]. In the presented manuscript (4.2), we linked the *2q24.3* lipodystrophy risk locus to its functional variant, its cell type and context specific effect, its regulatory element and effector gene, and characterized its cellular program comprising the perturbation of actin cytoskeleton remodeling during subcutaneous adipocyte differentiation resulting in distrubed lipid metabolism and insulin sensitivity in mature subcutaneous adipocytes. Indeed, actin cytoskeleton remodeling has been described to play a key role in adipocyte differentiation and function such as regulating lipid droplet formation and insulin-stimulated *GLUT4* vesicles trafficking and fusion[106]. Although the identified variant rs6712203 scored highest across all *2q24.3* haplotype variants in our sequence based predictive models, other variants of this locus

are predicted to affect regulatory activity as well. Therefore further studies, including CRISPR-based genome engineering approaches like CRISPR-SAVE (described in 4.4), are necessary to test the hypothesis that multiple variants of the *2q24.3* risk locus could act in concert regulating the identified target gene *COBLL1* along with other effector genes like *GRB14*. Together, this manuscript demonstrates a step-by-step dissection of a common genetic locus integrating epigenetic, transcriptional, morphological and cellular signatures.

In the third manuscript, we applied LipocyteProfiler on *18q21.33* risk locus associated with a lipodystrophy-like phenotype, and illuminate mechanistic insights underlying its pleiotropic effects in different cell-types and conditions resulting in a complex phenotype. In light of the pleiotropic effects, phenotypic informed clustering of T2D-associated genetic variants showed that some variants show strong weighting in more than one cluster, suggesting that these are involved in more than one mechanistic process[62]. Extensive pleiotropic effects have been described across loci (90% of loci associated with multiple traits), SNPs, and gene sets underlying genetic associations with human traits[107]. In the third reported manuscript (4.3), we highlighted the complexity of determining underlying mechanisms of liability of disease by identifying pleiotropic effects of the *18q21.33* metabolic risk locus mediated through at least three cell types (muscle, visceral and subcutaneous adipose tissue), three target genes (BCL2, KDSR, and VPS4B), and distinct cellular programs (eg. apoptosis, oxidative phosphorylation). Pleiotropy is thought to support identifying reasons for comorbidity between traits, characterize underlying shared genetic mechanisms, and may aid in establishing the direction of causality between traits[107]. For example *BCL2*-inhibitors (eg. Venetoclax) are established drugs for the treatment of chronic lymphocytic leukemia and lymphoma by inhibiting tumor cell proliferation and inducing apoptosis. Intriguingly, a reported drug side effect of Venetoclax is hyperglycemia (occured in 16% of patients (5% showed severe hyperglycemia) in a 1-year follow-up clinical trial[108]), which is in line with the described association of the *18q21.33* locus with impaired insulin sensitivity mediated through *BCL2* as effector gene in subcutaneous adipocytes. Using image-based phenotypic profiling we identified shared morphological and cellular signatures between the *18q21.33* metabolic risk locus and polygenic risk of adverse body fat distribution in subcutaneous adipocytes pointing to mitochondrial-related cellular programs involved in the regulation of apoptotic processes. Indeed, it has been previously reported that impaired

mitochondrial function influences insulin resistance[49,109] and that apoptotic processes might be involved in adverse body fat distribution[110]. Further, it is known that visceral adipose tissue stores fat when subcutaneous adipose tissue cannot accommodate excess fat because of its limited expandability which is associated with impaired insulin sensitivity and risk of developing type 2 diabetes and cardiovascular disease[111]. Based on our findings we hypothesized that the *18q21.33* metabolic risk locus and polygenic risk of increased waist-to-hip ratio in subcutaneous adipocytes modify similar pathways resulting in limited expandability of peripheral fat storage capacity and central fat accumulation which is linked to adverse metabolic effects.

The three reported studies showcase a framework that can be used to study functional genetics of single loci and polygenic signals, however due to the small sample size the reported findings should be further validated in future studies.

The introduced framework and the two single locus applications demonstrate, consistent with literature, that disease-relevant high-confidence cellular models are necessary in discovering underlying mechanisms of disease risk. LipocyteProfiling of primary AMSCs revealed that genetically-driven disease-relevant cellular processes are specific to adipose depot, differentiation state, and conditions. Comparably to the natural genetic variation that was reported for donor-derived iPSCs ranging from 5 to 46%[42], we observed inter-individual genetic variation of 3 to 41% and low variability explained by extrinsic, experimentally introduced confounders (eg. batch-to-batch variation explained on average 6.02% and cell plating density on average 3.75% of total variability). Batch-to-batch variation is wide-spread in high-throughput methodologies and poses a major challenge[112]. Since undetection can lead to misinterpretation and false conclusion, identifying and accounting for batch effects that result from undesired technical variation (eg. changes in laboratory conditions, handling variability between experimentalists) rather than representing a meaningful biological result is an important preliminary step[112,113]. In the reported LipocyteProfiler framework, we applied quality control steps before downstream data processing (eg. removed images on well edges and images with low cell number), normalised data, applied feature selection steps (e.g. removed features that are known to be noisy and generally unreliable), and quantified batch effects by differentiating three control cell lines in parallel. To account for variable feature-specific contributions of batch, sex,

age, and BMI to overall feature variance, we corrected for those covariables in our analyses. Based on these steps we gained confidence that LipocyteProfiler can be used to study the effect of genetic contributions to morphological and cellular programs.

However, it will be necessary in the future to substantially increase sample size to capture the full range of allele frequencies, to further standardize, automate strategies, and use suitable controls to keep technical confounders as low as possible, while power to detect contribution of natural genetic variability increases. Notably, power calculation of genome-wide functional studies *in vitro* could differ from previous calculations of similar studies in whole tissue, due to the heterogeneity of cell states in whole tissue. Cells cultured and differentiated *in vitro* are enriched for cell-type and cell-state of interest, whereas whole tissue is composed of various cell-types of different differentiation states, which introduces noise and therefore reduces power to detect disease-relevant phenotypes[39]. The reported studies demonstrated the ability to detect subtle phenotypes with small sample sizes in disease-relevant cell models, suggesting that relatively small sample size -- if selected thoroughly for example by sampling from the tail ends of PRS distributions -- might be sufficient to perform *in vitro* functional genetic studies using multiparametric phenotypic profiling assays in disease-relevant models. While our work serves as proof-of-principle that genetic variants affect morphological and cellular phenotypes, the majority of the signals we detected did not survive multiple testing schemes, stressing the need to (1) substantially increase sample size when testing more than one genetic risk locus for phenotype associations and (2) combine phenotypic profiling assays such as LipocyteProfiler with CRISPR-based engineering approaches.

CRISPR-based genome engineering tools are necessary to establish causality between genetic variants and genetically-driven disease phenotypes. These tools need to be accurate, versatilly applicable, and scalable. The reported approach, CRISPR-SAVE, introduces strategies to overcome challenges such as increasing editing efficiency, decreasing undesired modification (off-targets and/or random integration) and improving scalability by reducing time and cost consuming experimental procedures. However, it also demonstrated pitfalls limiting precise predetermined genome modifications requiring further optimisation efforts. Nevertheless, the combination of phenotypic profiling approaches like LipcoteProfiler and CRISPR-based genome

engineering tools hold great promise to characterize functional consequences of genetic variants and to recover genetically-driven phenotypes.

Pooled experimental designs rather than arrayed assays can process many samples in parallel, reducing experimental costs and time, and have the ability to culture cells under the same environmental conditions, reducing batch effects and increasing statistical power. However, pooled assays for population-scale genetic studies face different experimental challenges, like heterogeneity of growth rate and differentiation capacity between donors, and analytical challenges like deconvolution of individual genotype-phenotype relationship[83,114]. Studies, namely *village-in-dish*[114] and *pooled optical screens*[83], introduced frameworks of pooled designs to elucidate genetic association with complex phenotypes. In future, combining appropriate large-scale experimental and analytical approaches will be necessary to characterize genotype-phenotype association at genome scale in disease-relevant cell models in a well-controlled, experimental feasible and inexpensive way.

Etiology of metabolic disease is multifactorial, including intrinsic and extrinsic factors. Some genetically-driven phenotypic effects may be conditional on environmental stimulation. For example, we showed that morphological and cellular signatures of different polygenic risk for metabolic traits or locus-specific effects of *2q24.3* and *18q21.33* were specific to adipose depot and differentiation state. Additionally, risk allele carriers of *18q21.33* show different responses to free fatty acid treatment in visceral adipocytes. Similar observation was previously described when studying expression quantitative trait loci (eQTLs)[115]. In this context, it was shown that some eQTLs are context-specific and can only be observed under specific environmental simulations, and therefore are called dynamic eQTLs[116]. These findings demonstrate that besides disease relevant cell models, potentially disease-causing environmental conditions should be included in study designs to capture the full bandwidth of multifactorial, modifiable disease phenotypes.

Genetic association signals of complex traits are spread across the whole genome and are in close proximity to genes without obvious connection to disease. This implies that pleiotropy is a wide-spread feature of pathophysiology of common complex diseases and that gene expression is affected globally, including genes that are previously thought to have no obvious influence on disease. Further, it is known that these variants lie more likely in open chromatin of cell-type

specific regulatory elements rather than ubiquitously regulating elements[11]. The omnigenic model[16] describes the hypothesis that, due to the architecture of common genetic variants and associated gene regulation, complex indirect networks of genes contribute to total heritability and converge in core pathways directly affecting disease phenotypes. This implies that mapping cell-specific regulatory networks will be essential to fully understand human disease biology.
In the reported manuscripts, we demonstrated that Lipocyte profiling captures cellular networks of complex phenotypes, which may be explained by the fact that image-based profiling captures a magnitude of subtle morphological and cellular changes downstream of complex transcriptional and protein-protein interaction networks. In the future, systematic integration of "omics" data, including epigenomics, transcriptomics, proteomics, and metabolomics, with high-content imaging may further improve to capture disease-relevant cell-specific regulatory networks associated with genetic variation.

Polygenic risk scores have the potential to provide clinical benefit by quantifying disease predisposition, predicting disease progression, propensity to develop disease complications, and response to pharmacological and behavioral interventions[61]. Process-specific PRSs quantify polygenic risk of individuals belonging to phenotypic-informed subclusters of T2D. It was shown that one third of individuals scored highest for at least one cluster, and most of them (75%) scored lower in other clusters[62]. Therefore, understanding of mechanistic insights of process-specific PRSs by characterizing the influence of a high polygenic risk score on molecular and cellular pathways of a specific process, as demonstrated in LipocyteProfiler manuscript, is impactful and may accelerate clinic application. Individuals predominantly assigned to one disease-driving process could particularly benefit, in combination with established biomarkers, from process-tailored drugs and improvements of tailored monitoring of progression and complications. Expanding the phenotypic profiling approaches, by integrating more phenotypic data (eg. metabolomics) and generating phenotypic profiles under various disease-relevant environmental conditions (eg. metabolic challenges, drug treatments) may further illuminate intrinsic and extrinsic drivers of disease-relevant cellular programs. Further, most GWAS have been performed in Europeans, most large biobanks consist of samples of European ancestry, and polygenic scores have been calculated from European sequences[74,75,117]. However, it was shown that accuracy in predicting polygenic risk in other ethnicities is reduced, due to different allele frequencies and strength of correlation with disease[74,75]. This implies that

genetically mediated cellular programs may differ between different populations and ethnicities. Therefore, calculation of polygenic risk scores and characterization of genetically-driven mechanistic insights should be assessed across populations and ethnicities to develop novel therapy strategies, including pharmacological and behavioral intervention, for everyone.

In summary, multiparametric genotypic/phenotypic profiling of large diverse cohorts of disease-relevant cellular and environmental models, with standardized well-controlled experimental settings, will provide an unique possibility to characterize genetic association with cellular function in an unbiased genome-wide manner. Such technologies, if combined with large biobanks of primary and iPS derived cells and scalable CRISPR perturbation schemes, will advance the field of functional genomics and pharmacogenomics making it possible to develop tailored therapeutic hypotheses for individual patients with common complex and heterogeneous disease.

6 List of publications

Laber S*, **Strobel SM***, Mercader JM, Dashti H, Dos Santos F, Kubitz P, Jackson M, Ainbinder A, Honecker J, Agrawal S, Garborcauskas G, Stirling DR, Leong A, Figueroa K, Sinnott-Armstrong N, Kost-Alimova M, Deodato G, Harney A, Way GP, Saadat A, Harken S, Reibe-Pal S, Ebert H, Zhang Y, Calabuig-Navarro V, McGonagle E, Stefek A, Dupuis J, Cimini B, Hauner H, Udler M, Carpenter A, Florez JC, Lindgren CM, Jacobs S, Claussnitzer M#. Discovering cellular programs of intrinsic and extrinsic drivers of metabolic traits using LipocyteProfiler. *Cell Genomics, 2023. */#Denotes shared, equal contribution*

Glunk V*, Laber S*, Sinnott-Armstrong N*, Sobreira DR*, **Strobel SM***, Batista TM, ... & Claussnitzer M#. A non-coding variant linked to metabolic obesity with normal weight affects actin remodelling in subcutaneous adipocytes. *Nature Metabolism*, 2023. **/#Denotes shared, equal contribution*

Laber S*, **Strobel SM***, Dashti M, Mercader JM, Agrawal S, Ainbinder A, Honecker J, Garborcauskas G, Stirling DR, Saadat A, Kost-Alimova M, Deodato G, Harney A, Way GP, Harken S, Sinnott-Armstrong N, Ma A, Björkegren JLM, Lindgren C, Florez JC, Jacobs S, Cimini B, Engreitz J, Khera A, Udler M, Carpenter A, Hauner H, Claussnitzer M#. A non-coding metabolic risk variant, rs12454712, associates with context-dependent adipocyte molecular and cellular programs in humans. *Manuscript in preparation*. **/#Denotes shared, equal contribution*

Hansen GT*, Sobreira DR*, Weber ZT, Thornburg AG, Aneas I, Zhang L, Sakabe NJ, Joslin AC, Haddad GA, **Strobel SM**, Laber S, Li Y, Claussnitzer M, Battaglino R#, Nobrega M#. Genetics of sexually dimorphic adipose distribution in humans. *Nature Genetics*, *2023*. **/#Denotes shared, equal contribution*

Emont MP, Jacobs C, Essene AL, Pant D, Tenen D, Colleluori G, Di Vincenzo A, Jørgensen AM, Dashti H, Stefek A, McGonagle E, **Strobel SM**, … & Rosen ED#. A single-cell atlas of human and mouse white adipose tissue. *Nature*, *2022.*/#Denotes shared, equal contribution*

Honecker J*, Ruschke S*, Seeliger C, Laber S, **Strobel SM**, Pröll P, Nellaker C, Lindgren CM, Kulozik U, Ecker J, Karampinos DC, Claussnitzer M#, Hauner, H#. Transcriptome and fatty-acid signatures of adipocyte hypertrophy and its non-invasive MR-based characterization in human adipose tissue. *EBioMedicine*, *2022.*/#Denotes shared, equal contribution*

Laber S*, Forcisi S*, Bentley L, Petzold J, Moritz F, Smirnov K, Al Sadat L, Williamson I, **Strobel SM**, Agnew T, Sengupta S, Nicol T, Honecker J, Mianne J, Teboul L, Lindgren CM, Bickmore W, Hauner H, Schmitt-Kopplin P, Claussnitzer M#, Cox RD#. Linking the FTO obesity rs1421085 variant circuitry to cellular, metabolic and organismal phenotypes *in vivo*. *Science Advances, 2021. */#Denotes shared, equal contribution*

Wang L*, Sinnott-Armstrong NA*, Wagschal A, Wark AR, Camporez JP, Perry RJ, Ji F, Sohn Y, Oh J, Wu S, Chery J, NematiMoud B, Saadat A, Dankel SN, Mellgren G, Sri Priyanka Tallapragada D, **Strobel SM**, Lee M-J, Tewhey R, Sabeti PC, Schaefer A, Petri A, Kauppinen S, Chung RT, Soukas A, Avruch J, Fried S, Hauner H, Sadreyev RI, Shulman GI, Claussnitzer M#, Näär AM#. A thrifty microRNA linking human positive selection and metabolic disorders. *Cell, 2020. */#Denotes shared, equal contribution*

7 References

1. Collins, F. S., Green, E. D., Guttmacher, A. E., Guyer, M. S. & US National Human Genome Research Institute. A vision for the future of genomics research. *Nature* **422**, 835–847 (2003).
2. Consortium, T. H. R. & the Haplotype Reference Consortium. A reference panel of 64,976 haplotypes for genotype imputation. *Nature Genetics* vol. 48 1279–1283 (2016).
3. 1000 Genomes Project Consortium *et al.* A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).

4. International HapMap Consortium. The International HapMap Project. *Nature* **426**, 789–796 (2003).
5. Nelson, M. R. *et al.* Large-scale validation of single nucleotide polymorphisms in gene regions. *Genome Res.* **14**, 1664–1668 (2004).
6. Slatkin, M. Linkage disequilibrium--understanding the evolutionary past and mapping the medical future. *Nat. Rev. Genet.* **9**, 477–485 (2008).
7. Buniello, A. *et al.* The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res.* **47**, D1005–D1012 (2019).
8. Maurano, M. T. *et al.* Systematic localization of common disease-associated variation in regulatory DNA. *Science* **337**, 1190–1195 (2012).
9. Cookson, W., Liang, L., Abecasis, G., Moffatt, M. & Lathrop, M. Mapping complex disease traits with global gene expression. *Nat. Rev. Genet.* **10**, 184–194 (2009).
10. 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 (2018).
11. Barroso, I. & McCarthy, M. I. The Genetic Basis of Metabolic Disease. *Cell* **177**, 146–161 (2019).
12. Claussnitzer, M. *et al.* A brief history of human disease genetics. *Nature* **577**, 179–189 (2020).
13. King, R. A., Rotter, J. I. & Motulsky, A. G. *The Genetic Basis of Common Diseases*. (Oxford University Press, 2002).
14. Gibson, G. Rare and common variants: twenty arguments. *Nat. Rev. Genet.* **13**, 135–145 (2012).
15. Hindorff, L. A., Gillanders, E. M. & Manolio, T. A. Genetic architecture of cancer and other

complex diseases: lessons learned and future directions. *Carcinogenesis* **32**, 945–954 (2011).

16. Boyle, E. A., Li, Y. I. & Pritchard, J. K. An Expanded View of Complex Traits: From Polygenic to Omnigenic. *Cell* vol. 169 1177–1186 (2017).
17. Manolio, T. A. *et al.* Finding the missing heritability of complex diseases. *Nature* **461**, 747–753 (2009).
18. Golan, D., Lander, E. S. & Rosset, S. Measuring missing heritability: inferring the contribution of common variants. *Proc. Natl. Acad. Sci. U. S. A.* **111**, E5272–81 (2014).
19. Mahajan, A. *et al.* Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. *Nat. Genet.* **46**, 234–244 (2014).
20. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).
21. Roadmap Epigenomics Consortium *et al.* Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317–330 (2015).
22. Kempfer, R. & Pombo, A. Methods for mapping 3D chromosome architecture. *Nat. Rev. Genet.* **21**, 207–226 (2020).
23. Consortium, G. & GTEx Consortium. Genetic effects on gene expression across human tissues. *Nature* vol. 550 204–213 (2017).
24. Regev, A. *et al.* Science forum: the human cell atlas. *Elife* **6**, e27041 (2017).
25. Jinek, M. *et al.* A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* **337**, 816–821 (2012).
26. Claussnitzer, M. *et al.* FTO Obesity Variant Circuitry and Adipocyte Browning in Humans. *N. Engl. J. Med.* **373**, 895–907 (2015).
27. Sankaran, V. G. *et al.* Human fetal hemoglobin expression is regulated by the

developmental stage-specific repressor BCL11A. *Science* **322**, 1839–1842 (2008).

28. Feng, Y., Mitchison, T. J., Bender, A., Young, D. W. & Tallarico, J. A. Multi-parameter phenotypic profiling: using cellular effects to characterize small-molecule compounds. *Nat. Rev. Drug Discov.* **8**, 567–578 (2009).
29. Carpenter, A. E. *et al.* CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol.* **7**, R100 (2006).
30. Bray, M.-A. *et al.* Cell Painting, a high-content image-based assay for morphological profiling using multiplexed fluorescent dyes. *Nat. Protoc.* **11**, 1757–1774 (2016).
31. Caicedo, J. C., Singh, S. & Carpenter, A. E. Applications in image-based profiling of perturbations. *Curr. Opin. Biotechnol.* **39**, 134–142 (2016).
32. Meijering, E., Carpenter, A. E., Peng, H., Hamprecht, F. A. & Olivo-Marin, J.-C. Imagining the future of bioimage analysis. *Nat. Biotechnol.* **34**, 1250–1255 (2016).
33. Kepiro, M., Varkuti, B. H. & Davis, R. L. High Content, Phenotypic Assays and Screens for Compounds Modulating Cellular Processes in Primary Neurons. *Methods Enzymol.* **610**, 219–250 (2018).
34. Scheeder, C., Heigwer, F. & Boutros, M. Machine learning and image-based profiling in drug discovery. *Curr Opin Syst Biol* **10**, 43–52 (2018).
35. Simm, J. *et al.* Repurposing High-Throughput Image Assays Enables Biological Activity Prediction for Drug Discovery. *Cell Chem Biol* **25**, 611–618.e3 (2018).
36. Wawer, M. J. *et al.* Toward performance-diverse small-molecule libraries for cell-based phenotypic screening using multiplexed high-dimensional profiling. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 10911–10916 (2014).
37. Rohban, M. H. *et al.* Systematic morphological profiling of human gene and allele function via Cell Painting. *Elife* **6**, (2017).

38. Groot, R., Lüthi, J., Lindsay, H., Holtackers, R. & Pelkmans, L. Large‐scale image‐based profiling of single‐cell phenotypes in arrayed CRISPR‐Cas9 gene perturbation screens. *Molecular Systems Biology* vol. 14 (2018).

39. Warren, C. R. & Cowan, C. A. Humanity in a Dish: Population Genetics with iPSCs. *Trends Cell Biol.* **28**, 46–57 (2018).

40. Kumar, S., Blangero, J. & Curran, J. E. Induced Pluripotent Stem Cells in Disease Modeling and Gene Identification. *Methods Mol. Biol.* **1706**, 17–38 (2018).

41. Pashos, E. E. *et al.* Large, Diverse Population Cohorts of hiPSCs and Derived Hepatocyte-like Cells Reveal Functional Genetic Variation at Blood Lipid-Associated Loci. *Cell Stem Cell* **20**, 558–570.e10 (2017).

42. Kilpinen, H. *et al.* Common genetic variation drives molecular heterogeneity in human iPSCs. *Nature* **546**, 370–375 (2017).

43. Kanchan, K. *et al.* Genomic integrity of human induced pluripotent stem cells across nine studies in the NHLBI NextGen program. *Stem Cell Res.* **46**, 101803 (2020).

44. Merkle, F. T. *et al.* Human pluripotent stem cells recurrently acquire and expand dominant negative P53 mutations. *Nature* **545**, 229–233 (2017).

45. Vigilante, A. *et al.* Identifying Extrinsic versus Intrinsic Drivers of Variation in Cell Behavior in Human iPSC Lines from Healthy Donors. *Cell Rep.* **26**, 2078–2087.e3 (2019).

46. Ehrlund, A. *et al.* Transcriptional Dynamics During Human Adipogenesis and Its Link to Adipose Morphology and Distribution. *Diabetes* **66**, 218–230 (2017).

47. Dicker, A. *et al.* Functional studies of mesenchymal stem cells derived from adult human adipose tissue. *Experimental Cell Research* vol. 308 283–290 (2005).

48. Böhm, A. *et al.* Metabolic signatures of cultured human adipocytes from metabolically healthy versus unhealthy obese individuals. *PLoS One* **9**, e93148 (2014).

49. Böhm, A. *et al.* Increased mitochondrial respiration of adipocytes from metabolically unhealthy obese compared to healthy obese individuals. *Sci. Rep.* **10**, 12407 (2020).
50. Groop, L. Genetics of the metabolic syndrome. *Br. J. Nutr.* **83 Suppl 1**, S39–48 (2000).
51. Micucci, C., Valli, D., Matacchione, G. & Catalano, A. Current perspectives between metabolic syndrome and cancer. *Oncotarget* **7**, 38959–38972 (2016).
52. Shi, T. H., Wang, B. & Natarajan, S. The Influence of Metabolic Syndrome in Predicting Mortality Risk Among US Adults: Importance of Metabolic Syndrome Even in Adults With Normal Weight. *Prev. Chronic Dis.* **17**, E36 (2020).
53. Claussnitzer, M. *et al.* Leveraging cross-species transcription factor binding site patterns: from diabetes risk loci to disease mechanisms. *Cell* **156**, 343–358 (2014).
54. Sinnott-Armstrong, N. *et al.* A regulatory variant at 3q21.1 confers an increased pleiotropic risk for hyperglycemia and altered bone mineral density. *Cell Metab.* **33**, 615–628.e13 (2021).
55. Small, K. S. *et al.* Regulatory variants at KLF14 influence type 2 diabetes risk via a female-specific effect on adipocyte size and body composition. *Nat. Genet.* **50**, 572–580 (2018).
56. Shungin, D. *et al.* New genetic loci link adipose and insulin biology to body fat distribution. *Nature* **518**, 187–196 (2015).
57. Dimas, A. S. *et al.* Impact of type 2 diabetes susceptibility variants on quantitative glycemic traits reveals mechanistic heterogeneity. *Diabetes* **63**, 2158–2171 (2014).
58. Avery, A. R. & Duncan, G. E. Heritability of Type 2 Diabetes in the Washington State Twin Registry. *Twin Res. Hum. Genet.* **22**, 95–98 (2019).
59. Prasad, R. B. & Groop, L. Genetics of type 2 diabetes-pitfalls and possibilities. *Genes* **6**, 87–123 (2015).

60. Ahlqvist, E. *et al.* Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables. *Lancet Diabetes Endocrinol* **6**, 361–369 (2018).

61. Udler, M. S., McCarthy, M. I., Florez, J. C. & Mahajan, A. Genetic Risk Scores for Diabetes Diagnosis and Precision Medicine. *Endocr. Rev.* **40**, 1500–1520 (2019).

62. Udler, M. S. *et al.* Type 2 diabetes genetic loci informed by multi-trait associations point to disease mechanisms and subtypes: A soft clustering analysis. *PLoS Med.* **15**, e1002654 (2018).

63. Torres, J. M. *et al.* A Multi-omic Integrative Scheme Characterizes Tissues of Action at Loci Associated with Type 2 Diabetes. *Am. J. Hum. Genet.* **107**, 1011–1028 (2020).

64. Elksnis, A., Martinell, M., Eriksson, O. & Espes, D. Heterogeneity of Metabolic Defects in Type 2 Diabetes and Its Relation to Reactive Oxygen Species and Alterations in Beta-Cell Mass. *Front. Physiol.* **10**, 107 (2019).

65. Stefan, N., Schick, F. & Häring, H.-U. Causes, Characteristics, and Consequences of Metabolically Unhealthy Normal Weight in Humans. *Cell Metab.* **26**, 292–300 (2017).

66. Ginsberg, H. N. Insulin resistance and cardiovascular disease. *J. Clin. Invest.* **106**, 453–458 (2000).

67. Shulman, G. I. Ectopic fat in insulin resistance, dyslipidemia, and cardiometabolic disease. *The New England journal of medicine* vol. 371 2237–2238 (2014).

68. Lotta, L. A. *et al.* Integrative genomic analysis implicates limited peripheral adipose storage capacity in the pathogenesis of human insulin resistance. *Nat. Genet.* **49**, 17–26 (2017).

69. Guillín-Amarelle, C. *et al.* Type 1 familial partial lipodystrophy: understanding the Köbberling syndrome. *Endocrine* **54**, 411–421 (2016).

70. Alford, R. L., Rossiter, B. J. & Caskey, C. T. DNA diagnosis in monogenic diseases. *Int. J.*

Technol. Assess. Health Care **10**, 628–643 (1994).

71. Khera, A. V. *et al.* Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat. Genet.* **50**, 1219–1224 (2018).
72. Khera, A. V. *et al.* Polygenic Prediction of Weight and Obesity Trajectories from Birth to Adulthood. *Cell* **177**, 587–596.e9 (2019).
73. 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 (2018).
74. Martin, A. R. *et al.* Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations. *Am. J. Hum. Genet.* **100**, 635–649 (2017).
75. Martin, A. R. *et al.* Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat. Genet.* **51**, 584–591 (2019).
76. Schuster, A. *et al.* RNAi/CRISPR Screens: from a Pool to a Valid Hit. *Trends Biotechnol.* **37**, 38–55 (2019).
77. Shalem, O. *et al.* Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* **343**, 84–87 (2014).
78. Wang, T., Wei, J. J., Sabatini, D. M. & Lander, E. S. Genetic screens in human cells using the CRISPR-Cas9 system. *Science* **343**, 80–84 (2014).
79. Dixit, A. *et al.* Perturb-Seq: Dissecting Molecular Circuits with Scalable Single-Cell RNA Profiling of Pooled Genetic Screens. *Cell* **167**, 1853–1866.e17 (2016).
80. Hanna, R. E. *et al.* Massively parallel assessment of human variants with base editor screens. *Cell* **184**, 1064–1080.e20 (2021).
81. Gilbert, L. A. *et al.* Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell* **159**, 647–661 (2014).

82. Fulco, C. P. *et al.* Activity-by-contact model of enhancer-promoter regulation from thousands of CRISPR perturbations. *Nat. Genet.* **51**, 1664–1669 (2019).

83. Feldman, D. *et al.* Optical Pooled Screens in Human Cells. *Cell* **179**, 787–799.e17 (2019).

84. Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A. & Liu, D. R. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **533**, 420–424 (2016).

85. Gaudelli, N. M. *et al.* Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature* **551**, 464–471 (2017).

86. Rees, H. A. & Liu, D. R. Base editing: precision chemistry on the genome and transcriptome of living cells. *Nat. Rev. Genet.* **19**, 770–788 (2018).

87. Anzalone, A. V. *et al.* Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* **576**, 149–157 (2019).

88. Steyer, B. *et al.* Scarless Genome Editing of Human Pluripotent Stem Cells via Transient Puromycin Selection. *Stem Cell Reports* **10**, 642–654 (2018).

89. Grünewald, J. *et al.* Transcriptome-wide off-target RNA editing induced by CRISPR-guided DNA base editors. *Nature* **569**, 433–437 (2019).

90. Olzmann, J. A. & Carvalho, P. Dynamics and functions of lipid droplets. *Nat. Rev. Mol. Cell Biol.* **20**, 137–155 (2019).

91. Xue, R. *et al.* Clonal analyses and gene profiling identify genetic biomarkers of the thermogenic potential of human brown and white preadipocytes. *Nat. Med.* **21**, 760–768 (2015).

92. Yusa, K. Seamless genome editing in human pluripotent stem cells using custom endonuclease–based gene targeting and the piggyBac transposon. *Nat. Protoc.* **8**, 2061–2078 (2013).

93. Liu, L. *et al.* Glial lipid droplets and ROS induced by mitochondrial defects promote neurodegeneration. *Cell* **160**, 177–190 (2015).

94. Wang, H., Quiroga, A. D. & Lehner, R. Analysis of Lipid Droplets in Hepatocytes. *Methods in Cell Biology* 107–127 (2013) doi:10.1016/b978-0-12-408051-5.00007-3.

95. Grandl, M. & Schmitz, G. Fluorescent high-content imaging allows the discrimination and quantitation of E-LDL-induced lipid droplets and Ox-LDL-generated phospholipidosis in human macrophages. *Cytometry Part A: The Journal of the International Society for Advancement of Cytometry* **77**, 231–242 (2010).

96. Robichaud, S. *et al.* Identification of novel lipid droplet factors that regulate lipophagy and cholesterol efflux in macrophage foam cells. *Autophagy* 1–19 (2021).

97. Hershey, B. J., Vazzana, R., Joppi, D. L. & Havas, K. M. Lipid Droplets Define a Sub-Population of Breast Cancer Stem Cells. *J. Clin. Med. Res.* **9**, (2019).

98. Cruz, A. L. S., Barreto, E. de A., Fazolini, N. P. B., Viola, J. P. B. & Bozza, P. T. Lipid droplets: platforms with multiple functions in cancer hallmarks. *Cell Death Dis.* **11**, 105 (2020).

99. Wang, Z. *et al.* Regulation of renal lipid metabolism, lipid accumulation, and glomerulosclerosis in FVBdb/db mice with type 2 diabetes. *Diabetes* **54**, 2328–2335 (2005).

100. Weinert, S. *et al.* The lysosomal transfer of LDL/cholesterol from macrophages into vascular smooth muscle cells induces their phenotypic alteration. *Cardiovasc. Res.* **97**, 544–552 (2013).

101. Prats, C. *et al.* Decrease in intramuscular lipid droplets and translocation of HSL in response to muscle contraction and epinephrine. *J. Lipid Res.* **47**, 2392–2399 (2006).

102. Loos, R. J. F. & Kilpeläinen, T. O. Genes that make you fat, but keep you healthy. *J. Intern. Med.* **284**, 450–463 (2018).

103. Stears, A., O'Rahilly, S., Semple, R. K. & Savage, D. B. Metabolic insights from extreme human insulin resistance phenotypes. *Best Pract. Res. Clin. Endocrinol. Metab.* **26**, 145–157 (2012).

104. Fathzadeh, M. *et al.* FAM13A affects body fat distribution and adipocyte function. *Nat. Commun.* **11**, 1465 (2020).

105. Kilpeläinen, T. O. *et al.* Genetic variation near IRS1 associates with reduced adiposity and an impaired metabolic profile. *Nat. Genet.* **43**, 753–760 (2011).

106. Kim, J. I. *et al.* During Adipocyte Remodeling, Lipid Droplet Configurations Regulate Insulin Sensitivity through F-Actin and G-Actin Reorganization. *Mol. Cell. Biol.* **39**, (2019).

107. Watanabe, K. *et al.* A global overview of pleiotropy and genetic architecture in complex traits. *Nat. Genet.* **51**, 1339–1348 (2019).

108. Roberts, A. W. *et al.* Targeting BCL2 with Venetoclax in Relapsed Chronic Lymphocytic Leukemia. *N. Engl. J. Med.* **374**, 311–322 (2016).

109. Montgomery, M. K. & Turner, N. Mitochondrial dysfunction and insulin resistance: an update. *Endocr Connect* **4**, R1–R15 (2015).

110. Loh, N. Y. *et al.* RSPO3 impacts body fat distribution and regulates adipose cell biology in vitro. *Nat. Commun.* **11**, 2797 (2020).

111. Smith, U. Abdominal obesity: a marker of ectopic fat accumulation. *J. Clin. Invest.* **125**, 1790–1792 (2015).

112. Leek, J. T. *et al.* Tackling the widespread and critical impact of batch effects in high-throughput data. *Nat. Rev. Genet.* **11**, 733–739 (2010).

113. Caicedo, J. C. *et al.* Data-analysis strategies for image-based cell profiling. *Nat. Methods* **14**, 849–863 (2017).

114. Mitchell, J. M., Nemesh, J., Ghosh, S. & Handsaker, R. E. Mapping genetic effects on

cellular phenotypes with‘ cell villages’. *bioRxiv* (2020).

115. Umans, B. D., Battle, A. & Gilad, Y. Where Are the Disease-Associated eQTLs? *Trends Genet.* **37**, 109–124 (2021).
116. Ward, M. C., Banovich, N. E., Sarkar, A., Stephens, M. & Gilad, Y. Dynamic effects of genetic variation on gene expression revealed following hypoxic stress in cardiomyocytes. *Elife* **10**, (2021).
117. Manrai, A. K. *et al.* Genetic Misdiagnoses and the Potential for Health Disparities. *N. Engl. J. Med.* **375**, 655–665 (2016).

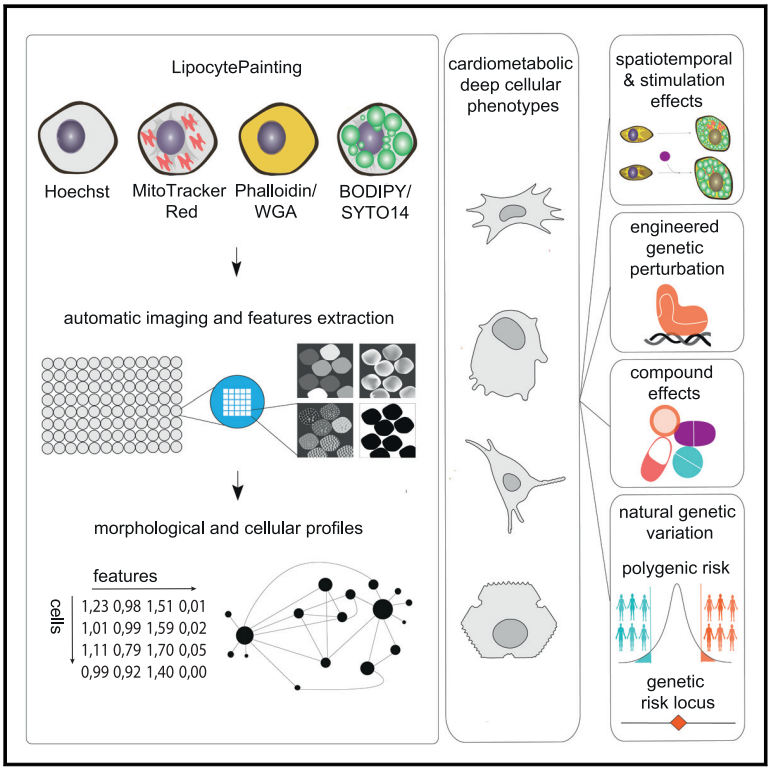
8 Manuscripts

Cell Genomics

Technology

Discovering cellular programs of intrinsic and extrinsic drivers of metabolic traits using LipocyteProfiler

Graphical abstract



Highlights

- LipocyteProfiler, a metabolic-disease-oriented image-based phenotypic profiling tool
- LipocyteProfiler links genetic variation, genes, and compounds to cellular programs
- Polygenic metabolic disease risk dictates context-specific cellular programs
- LipocyteProfiler allows us to link variants to function (V2F) without a priori hypotheses

Authors

Samantha Laber, Sophie Strobel, Josep M. Mercader, ..., Cecilia Lindgren, Suzanne B.R. Jacobs, Melina Claussnitzer

Correspondence

melina@broadinstitute.org

In brief

Laber and Strobel et al. introduce LipocyteProfiler, a metabolic-disease-oriented image-based deep phenotypic profiling tool, which allows one to systematically link engineered and natural genetic variation as well as chemical perturbations to cell-state-dependent cellular programs in lipocytes. Using LipocyteProfiler, they discover known and novel drivers of cellular programs related to human health and disease.

Laber et al., 2023, Cell Genomics *3*, 100346
July 12, 2023 © 2023
https://doi.org/10.1016/j.xgen.2023.100346

CellPress

Please cite this article in press as: Laber et al., Discovering cellular programs of intrinsic and extrinsic drivers of metabolic traits using LipocyteProfiler, Cell Genomics (2023), https://doi.org/10.1016/j.xgen.2023.100346

Cell Genomics

CellPress
OPEN ACCESS

Technology

Discovering cellular programs of intrinsic and extrinsic drivers of metabolic traits using LipocyteProfiler

Samantha Laber,[1,2,3,4,14] Sophie Strobel,[1,5,14] Josep M. Mercader,[1,6,7] Hesam Dashti,[1,7,13] Felipe R.C. dos Santos,[1,13] Phil Kubitz,[1,5,13] Maya Jackson,[1,13] Alina Ainbinder,[1] Julius Honecker,[8] Saaket Agrawal,[1] Garrett Garborcauskas,[1] David R. Stirling,[1] Aaron Leong,[1,6,7] Katherine Figueroa,[1,6] Nasa Sinnott-Armstrong,[1,9] Maria Kost-Alimova,[1] Giacomo Deodato,[1] Alycen Harney,[1] Gregory P. Way,[1] Alham Saadat,[1] Sierra Harken,[1] Saskia Reibe-Pal,[2] Hannah Ebert,[10] Yixin Zhang,[13] Virtu Calabuig-Navarro,[1,10] Elizabeth McGonagle,[1] Adam Stefek,[1] Josée Dupuis,[12] Beth A. Cimini,[1] Hans Hauner,[5,8,11] Miriam S. Udler,[1,6,7] Anne E. Carpenter,[1] Jose C. Florez,[1,6,7] Cecilia Lindgren,[1,2,3] Q1 Suzanne B.R. Jacobs,[1,6] and Melina Claussnitzer[1,4,6,7,13,15,*]

[1]Programs in Metabolism and Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA
[2]Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, University of Oxford, Oxford OX3 7FZ, UK
[3]Wellcome Centre for Human Genetics, University of Oxford, Oxford OX3 7BN, UK
[4]Massachusetts General Hospital, Harvard Medical School, Boston, MA 02114, USA
[5]Institute of Nutritional Medicine, School of Medicine, Technical University of Munich, 85354 Freising-Weihenstephan, Germany
[6]Diabetes Unit and Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA 02114, USA
[7]Department of Medicine, Harvard Medical School, Boston, MA 02114, USA
[8]Else Kröner-Fresenius-Centre for Nutritional Medicine, School of Life Sciences, Technical University of Munich, 85354 Freising-Weihenstephan, Germany
Q9 [9]Department of Genetics, Stanford University, San Francisco, CA, USA
[10]Institute of Nutritional Science, University Hohenheim, 70599 Stuttgart, Germany
[11]German Center for Diabetes Research (DZD), 85764 Neuherberg, Germany
[12]Department of Biostatistics, Boston University School of Public Health, Boston, MA 02118, USA
[13]The Novo Nordisk Foundation Center for Genomic Mechanisms of Disease, Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA
[14]These authors contributed equally
[15]Lead contact
*Correspondence: melina@broadinstitute.org
https://doi.org/10.1016/j.xgen.2023.100346

SUMMARY

A primary obstacle in translating genetic associations with disease into therapeutic strategies is elucidating the cellular programs affected by genetic risk variants and effector genes. Here, we introduce LipocyteProfiler, a cardiometabolic-disease-oriented high-content image-based profiling tool that enables evaluation of thousands of morphological and cellular profiles that can be systematically linked to genes and genetic variants relevant to cardiometabolic disease. We show that LipocyteProfiler allows surveillance of diverse cellular programs by generating rich context- and process-specific cellular profiles across hepatocyte and adipocyte cell-state transitions. We use LipocyteProfiler to identify known and novel cellular mechanisms altered by polygenic risk of metabolic disease, including insulin resistance, fat distribution, and the polygenic contribution to lipodystrophy. LipocyteProfiler paves the way for large-scale forward and reverse deep phenotypic profiling in lipocytes and provides a framework for the unbiased identification of causal relationships between genetic variants and cellular programs relevant to human disease.

Q3 Q4 Q7 Q2 INTRODUCTION

With the rise of human genome sequencing data, the number of genetic variants known to be associated with human diseases has increased substantially; however, elucidating the pathogenic mechanisms through which genetic variants impact disease remains limiting. Phenotypic profiling is a powerful tool to systematically discover external and internal regulators of biological processes in cellular systems in an unbiased manner.[1–4] High-content imaging is an established multi-parametric approach that captures and quantifies biological processes from microscopy images, yielding a rich set of morphological and cellular profiles.[5] To date, image-based profiling has been used in small-molecule screens to identify compound fingerprints, ascertain compound toxicity, and predict compound assay activity[6–9] and in gene expression screens to annotate gene function.[10] In all cases, the basic strategy is to match the profile of a given sample based on similarity to morphological profiles of previously annotated samples.

Please cite this article in press as: Laber et al., Discovering cellular programs of intrinsic and extrinsic drivers of metabolic traits using LipocyteProfiler, Cell Genomics (2023), https://doi.org/10.1016/j.xgen.2023.100346

CellPress
OPEN ACCESS

Cell Genomics
Technology

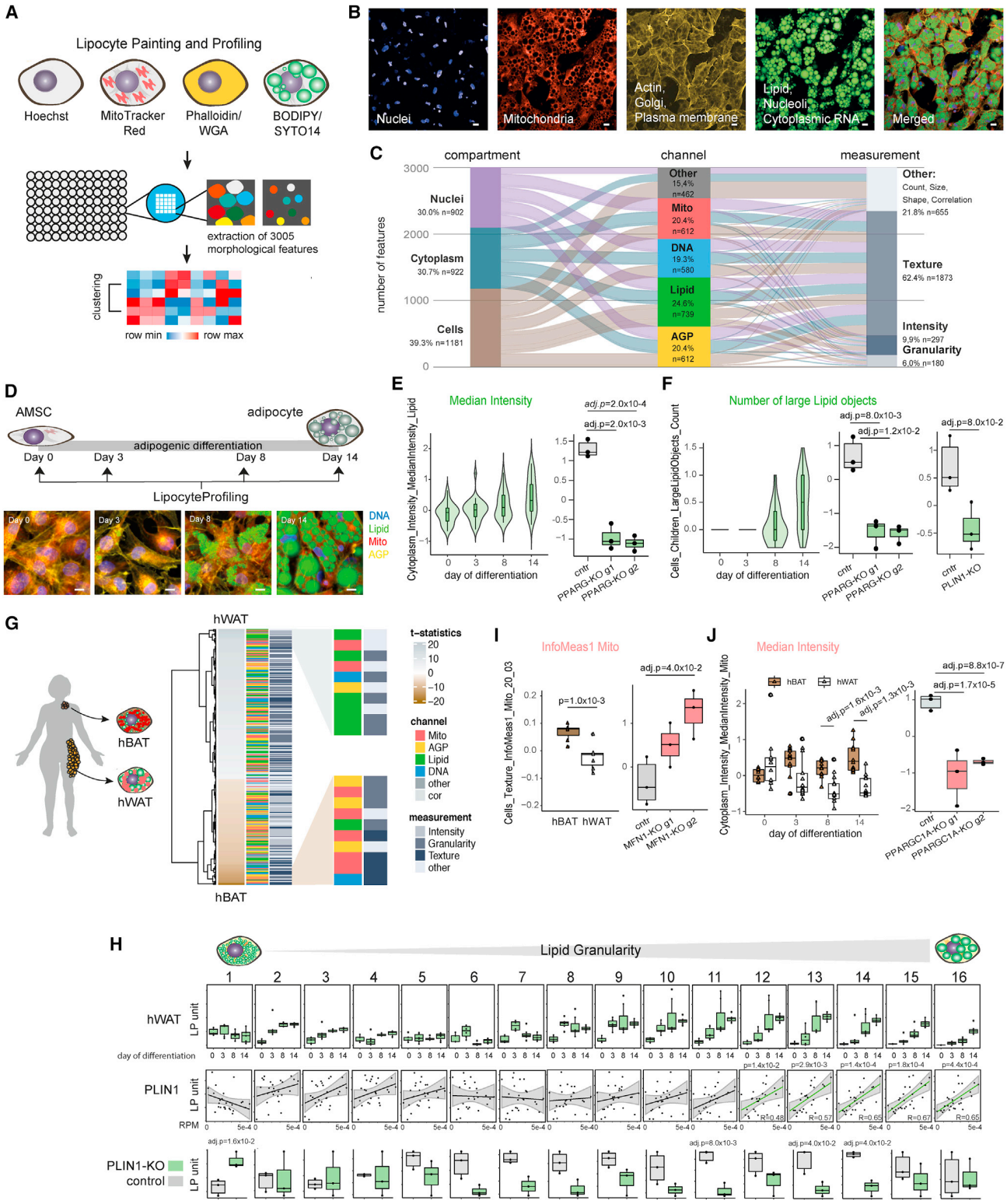


Figure 1. LipocyteProfiler creates rich morphological and cellular profiles in adipocytes that are informative for known cellular functions

(A) Schematic of LipocyteProfiler, which is a high-content imaging assay that multiplexes six fluorescent stains imaged in four channels in conjunction with an automated image-analysis pipeline to generate rich morphological and cellular profiles in lipid-storing cell types (lipocytes), such as adipocytes during differentiation.

(legend continued on next page)

Please cite this article in press as: Laber et al., Discovering cellular programs of intrinsic and extrinsic drivers of metabolic traits using LipocyteProfiler, Cell Genomics (2023), https://doi.org/10.1016/j.xgen.2023.100346

CellPress
OPEN ACCESS

In metabolism, lipid droplets represent a relevant feature that is amenable to image-based profiling. Lipid droplets are storage organelles central to both whole-body metabolism and energy homeostasis. These droplets are highly dynamic and found in all cell types.[11] They are functional in either cellular homeostasis in lipid-accumulating cells (lipocytes), such as adipocytes, hepatocytes, macrophages/foam cells, and glial cells,[11–15] or in pathophysiological processes in cells including vascular smooth muscle cells, skeletal muscle cells, renal podocytes, and cancer cells.[16–20] Changes in lipid-droplet dynamics such as the number and size of lipid droplets and overall lipid content are associated with the progression of numerous metabolic diseases including type 2 diabetes (T2D), obesity, and non-alcoholic fatty liver disease.[21]

Here, we introduce LipocyteProfiler, a metabolic-disease-oriented phenotypic profiling system for lipid-accumulating cells bridging the gap between high-throughput generalizable assays and low-throughput, highly customized, disease-focused readouts. LipocyteProfiler is an adaptation of Cell Painting[5,22–24] that incorporates BODIPY to measure dynamic features of lipid droplets and thus captures lipocyte-relevant phenotypes in addition to generic morphological profiles.

Design

Elucidating cellular programs that underlie the association of genetic variants, regulatory elements, and genes with diseases largely remains a non-systematic, labor- and cost-intensive endeavor that is biased toward hypotheses drawn from a priori knowledge. High-content imaging captures and quantifies numerous distinct biological processes from microscopy images in an unbiased manner, yielding a rich set of morphological and cellular profiles.[5]

Thus far the phenotypic data ascertained from scalable morphological profiling assays has been limited to features informative for the generic organelles of the cell.[5,22,23] This includes structural information about nuclei, endoplasmic reticulum, cytoskeleton, and mitochondria, or generic processes, such as cell growth or proliferation. It is currently unknown how gene and compound effects translate to changes in specific cellular pathways and processes. Image-based deep cellular phenotypic profiling tools provide a plethora of quantitative features. However, narrowing down these high-dimensional data matrices to a set of the most informative features that drive cellular processes requires differentiating first-order from second-order relationships among the features. As such, there is a pressing need to develop foundational technologies that allow systematic linking of genetic variation to disease-relevant cellular programs at scale and in a broadly accessible way.

To quantitatively map dynamic and context-dependent morphological and cellular signatures in lipocytes as well as discover intrinsic and extrinsic drivers of cellular programs, we developed a high-content image-based profiling approach called LipocyteProfiler (Figure 1A). LipocyteProfiler is an unbiased high-throughput profiling assay that generates rich generic and lipocyte-specific cellular profiles from six multiplexed fluorescent dyes imaged in four channels (Figure 1B) in conjunction with an automated image-analysis pipeline (see STAR Methods and Method S1).

LipocyteProfiler extracts 3,005 morphological and cellular features that map to three cellular compartments (Cell, Cytoplasm, and Nucleus) across four organelles, namely nucleus (Hoechst), *Mito* (MitoTracker red, which stains mitochondria), *AGP* (actin, Golgi, plasma membrane; stained with phalloidin [F-actin cytoskeleton] and wheat germ agglutinin [Golgi and plasma membranes]), and *Lipid* (BODIPY, which stains neutral lipids, multiplexed with SYTO14, which stains nucleoli and cytoplasmic RNA) (Figures 1C and S1A). Within each compartment and channel, features quantify morphological changes based on four different measurement classes: *Intensity*, *Granularity*, *Texture*, and *Others* (Figures 1C and S1A).

(B) Representative microscopy image of fully differentiated adipocytes for four individual channels and a merged representation across channels. Scale bars, 10 μm.

(C) LipocyteProfiler extracts 3,005 morphological and cellular features that map to three cellular compartments and across four channels using four measurement classes.

(D) Schematic of LipocyteProfiling in differentiating hWAT at four time points of adipocyte differentiation (days 0, 3, 8, 14). Representative images of AMSCs stained using LipocytePainting at four time points of differentiation (days 0, 3, 8, 14). Scale bars, 10 μm.

(E) Cytoplasm *MedianIntensity Lipid*, a measurement of lipid content within a cell, significantly increases with adipogenic differentiation and decreases following CRISPR-Cas9-mediated knockdown of *PPARG* in differentiated white adipocytes. Data are shown for two guides used (g1 and g2), and y axis shows LP units (normalized LipocyteProfiling [LP] values across three batches, see STAR Methods).

(F) *Number of large Lipid objects* informative for large lipid droplets are absent in the progenitor state (day 0) and in early differentiation (day 3) and progressively increase in later stages of differentiation (days 8 and 14). *Number of large Lipid objects* is reduced following CRISPR-Cas9-mediated knockout (KO) of *PPARG* (data are shown for two guides used [g1 and g2]) and *PLIN1*, at day 14 of differentiation. y axis shows LP units (normalized LP values across three batches, see STAR Methods).

(G) Morphological profiles of white (hWAT) and brown (hBAT) adipocytes at day 14 of differentiation differ significantly across all feature classes (FDR < 0.1%). Features are clustered based on effect size. Features with the highest effect size in hWAT and hBAT adipocytes are lipid- and mitochondria-related, respectively. Graph shows zoom-in for top ten features with largest effect sizes in hWAT (top panel) and hBAT (bottom panel).

(H) *Lipid* Granularity measures, as spectra of 16 lipid-droplet size measures, show size-specific changes in hWAT and hBAT during differentiation. See also Figure S1H. Granularity features informative for larger lipid droplets (*Lipid Granularity* 10–16) correlate positively with *PLIN1* gene expression and are reduced in *PLIN1*-KO adipocytes. See also Figures S1I and S1J (PLIN2, FASN-KO). y axis shows autoscaled LP units (normalized LP values across three batches, seeSTAR Methods).

(I) Brown adipocytes (hBAT) show higher *Mito_Texture_InfoMeas1*, a measure of spatial relationship between specific intensity values, compared with white adipocytes (hWAT). CRISPR-Cas9-mediated knockout of *MFN1*, a mitochondrial fusion gene, changes *Mito_Texture_InfoMeas1* (data shown for two guides used [g1 and g2]). y axis shows LP units (normalized LP values across three batches [hBAT/hWAT] or normalized across CRISPR-KO data, see STAR Methods).

(J) *Mito_MedianIntensity* is higher in brown (hBAT) compared with white (hWAT) adipocytes throughout differentiation and decreased after CRISPR-Cas9-mediated knockout of *PPARGC1A* in hWAT. y axis shows LP units (normalized LP values across three batches, see STAR Methods).

Please cite this article in press as: Laber et al., Discovering cellular programs of intrinsic and extrinsic drivers of metabolic traits using LipocyteProfiler, Cell Genomics (2023), https://doi.org/10.1016/j.xgen.2023.100346

More specifically, *Intensity* features are a collection of features that measure pixel intensities across an image using various measurement types such as *MedianIntensity*, *MaxIntensity*, and *RadialDistribution* of Intensity. Texture features describe the complexity or homogeneity within an image of a compartment using a plethora of different quantification approaches including *Entropy*, *AngularSecondMoment*, and *Variance* measurements. *Granularity* features are informative for a spectrum of different structural elements (sizes 1–16) that fit into an image. For example, *Granularity Lipid* features are indicative for small (sizes 1–5), medium (sizes 6–10), and large (sizes 11–16) lipid droplets. The fourth measurement class, referred to as *Other* features, is a composition of different measurements that quantify *shape*, *size*, and *count* of cells as well as correlations between different channel intensities (e.g., between *Mito* and *Lipid*). In concert, these features build rich lipocyte-specific cellular profiles that enable elucidation of cellular programs that link genetic loci and variation to human disease.

To nominate LipocyteProfiler core features that drive cellular processes we applied an information theoretic algorithm, which reduces the 3,005 LipocyteProfiler features, based on mutual information (MI) between features, and prioritizes first-order interactions indicative of direct interactions between features[25–27] (Figure S1B). By representing features as nodes of a graph and MI-based calculated interactions between features as weighted edges, we constructed an MI network representing interactions between the features. After ranking the nodes based on their degree of connectivity (number of edges), we defined LipocyteProfiler core features as those features among the 75% upper quantile of the ranked nodes and the lower 25% percentile of the average MI. The MI-based feature reductions allowed us to nominate approximately one-third of the features as core features (Figures S1B and S1C) that can be used to identify intrinsic and extrinsic drivers of phenotypic changes in a concise way.

We demonstrate that our LipocyteProfiler tool can identify diverse cardiometabolic-disease-relevant cellular mechanisms by generating context-, process-, and allele-specific morphological and cellular profiles. We prototyped LipocyteProfiler in adipocytes and hepatocytes, which are highly specialized cells that store excess energy in the form of lipid droplets and have key roles in cardiometabolic disease. First, we demonstrate that LipocyteProfiler can identify meaningful changes in feature profiles (1) during adipocyte differentiation, (2) across white and brown adipocyte lineages, and (3) following genetic and drug perturbations. Next, we correlated LipocyteProfiler features with transcriptomic data from RNA sequencing (RNA-seq) to link gene sets with morphological and cellular features, capturing a broad range of cellular activity in differentiating adipocytes. We then applied LipocyteProfiler to connect polygenic risk scores for type 2 diabetes (T2D)-related traits to cellular phenotypes and discover novel trait-specific cellular mechanisms underlying polygenic risk. Finally, we used our method to uncover cellular traits under the genetic control of an individual genetic risk locus, demonstrated for the *2p23.3* metabolic risk locus at *DNMT3A*.[28]

RESULTS

LipocyteProfiler generates meaningful morphological and cellular profiles in differentiating adipocytes

To test the ability of LipocyteProfiler to extract biologically meaningful high-dimensional representations of morphological and cellular programs, we used our metabolic-disease-oriented image-based profiling tool to detect (1) changes associated with adipocyte differentiation, (2) differences between white and brown adipocytes, and (3) phenotypic effects of directed gene perturbation using CRISPR-Cas9 to knock out key regulators of adipocyte function.

First, we applied LipocyteProfiler to a model of adipocyte differentiation using an established white adipocyte line (hWAT),[29] which undergoes phenotypic changes from fibroblast-shaped to spherical lipid-filled cells during differentiation (days 0, 3, 8, and 14; Figure 1D). We mapped the phenotypic signature of progressive lipid accumulation and cytoskeletal remodeling during adipocyte differentiation in hWAT using tractable *Lipid* and *AGP* features. We show that cytoplasmic intensity of *Lipid*, a proxy of overall lipid content within a cell, increased with adipogenic differentiation (Figure 1E). In addition, *large Lipid objects* (large lipid droplets) were absent in the progenitor state (day 0) and in early differentiation (day 3), and the number of these objects increased in later stages of differentiation (Figure 1F). We confirmed that CRISPR-Cas9-directed perturbation of *PPARG*, the master regulator of adipogenesis, decreases the overall *Lipid Intensity* in differentiated white adipocytes (guide 1 adjusted p [adj.p] = 2.0×10^{-3}, guide 2 adj.p = 2.0×10^{-4}; Figure 1E). Furthermore, *large Lipid objects* present at day 14 of differentiation were reduced when we perturbed regulators of lipid accumulation, *PPARG* (guide 1 adj.p = 8.0×10^{-3}, guide 2 adj.p = 1×10^{-2}) and *PLIN1* (adj.p = 8.0×10^{-2}), a key regulator of lipid-droplet homeostasis (Figure 1F). These data demonstrate that LipocyteProfiler detects expected changes in lipid dynamics associated with adipocyte differentiation. Another cellular change that occurs during adipocyte differentiation is a drastic reorganization of the actin cytoskeleton, which transitions from well-defined stress fibers in pre-adipocytes to relatively thick cortical actin lining composed of patches of punctate F-actin at the inner surface of the plasma membrane in fully differentiated adipocytes[30] (Figure S1D). This cytoskeletal remodeling is stimulated by insulin and essential for GLUT4 translocation into the membrane to facilitate insulin-responsive glucose uptake in the cell.[30] Concordantly, we found that CRISPR-Cas9-mediated disruption of the insulin receptor (*INSR*) and insulin receptor substrate 1 (*IRS1*) in pre-adipocytes altered *AGP Texture* features (describing the smoothness of a given stain) in mature adipocytes at day 14 of differentiation (Figure S1E). Specifically, *INSR* and *IRS* knockout reduced variation of cytoplasmic *AGP* stain intensities most significantly near the plasma membrane (Cytoplasm_RadialDisribution_RadialCV_AGP_4_of_4, *IRS1* guide 1 adj.p = 8.0×10^{-3}, guide 2 adj.p = 8.0×10^{-2}; *INSR* guide 1 adj.p = 8.0×10^{-2}, guide 2 adj.p = 2.0×10^{-2}), indicative of less punctuated *AGP*, which is in line with less cortical actin in *INSR*- and *IRS*-knockout cells.

Next, we used brown and white adipocyte model systems to elucidate mitochondrial and lipid-related informational content. Intrinsic differences distinguishing white and brown adipocytes

Please cite this article in press as: Laber et al., Discovering cellular programs of intrinsic and extrinsic drivers of metabolic traits using LipocyteProfiler, Cell Genomics (2023), https://doi.org/10.1016/j.xgen.2023.100346

CellPress
OPEN ACCESS

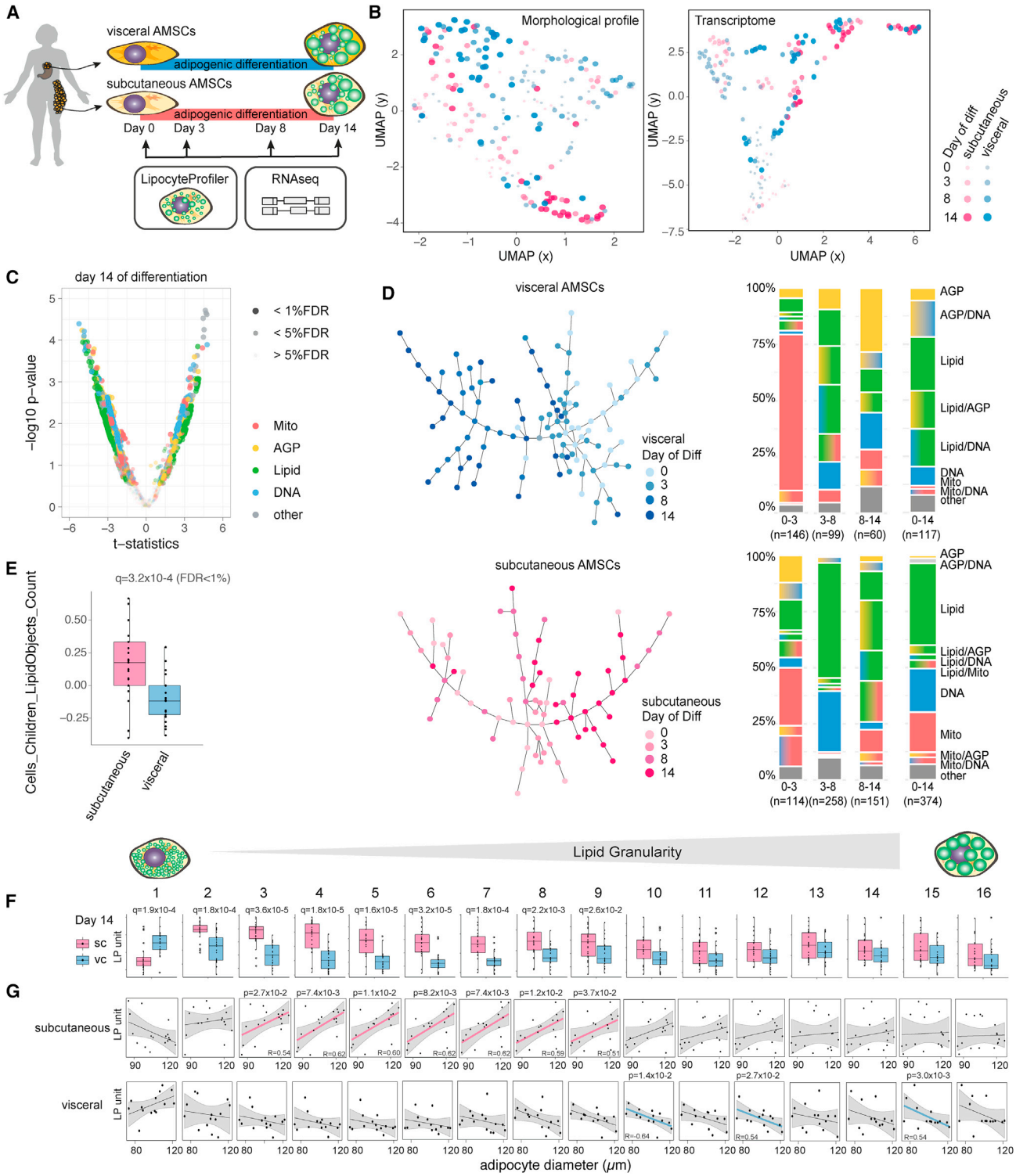


Figure 2. LipocyteProfiler identifies distinct depot-specific morphological and cellular signatures associated with differentiation trajectories in both visceral and subcutaneous AMSCs

(A) Human AMSCs isolated from subcutaneous and visceral adipose depots were differentiated for 14 days, and LipocyteProfiler and RNA-seq profiling were performed throughout adipocyte differentiation (days 0, 3, 8, and 14).

(legend continued on next page)

Please cite this article in press as: Laber et al., Discovering cellular programs of intrinsic and extrinsic drivers of metabolic traits using LipocyteProfiler, Cell Genomics (2023), https://doi.org/10.1016/j.xgen.2023.100346

are known to be predominantly driven by differences in mitochondrial number and activity that translate into differential lipid accumulation.[31] Using an established brown adipocyte line derived from human neck fat (hBAT) from the same individual as for the hWAT line, we showed that morphological profiles from differentiated hWAT and hBAT differ significantly in every channel and feature category (Figure 1G). *Lipid Granularity* measures, a class of metrics that capture the typical sizes of bright spots for a stain, predominated among those increased in hWAT. During adipocyte differentiation, lipid droplets typically increase first in number and then enlarge and fuse to form larger lipid droplets over the course of maturation.[32] We observed that the number of small and medium-sized lipid droplets (*Lipid Granularity* measures 1–9) present in early differentiating hWAT saturate in early stages of differentiation (Figure 1H). Larger lipid droplets (*Lipid Granularity* measures 10–16) increase in terminal differentiation, indicating that lipid droplets form in early differentiation and grow thereafter, a process that is reflected in *Lipid Granularity* measures and *Lipid objects* count (Figures 1H and S1F). Consistent with the notion that adipocytes from brown adipose have smaller lipid droplets, we found that during differentiation, hBAT adipocytes accumulate fewer medium-sized and large lipid droplets as seen by lower values across the spectra of granularity (Figures 1H and S1G). Intuitively, LipocyteProfiler-derived size estimates showed that white hWAT are larger than brown hBAT adipocytes after 14 days of adipogenic differentiation as cells become lipid laden (*Cells_AreaShape_Area* $p = 5.1 \times 10^{-5}$; Figure S1H). To test whether lipid-droplet-associated perilipins can be linked to lipid-droplet sizes, we correlated *Lipid Granularity* measures with mRNA expression levels of *PLIN1*, which is specifically expressed in adipocytes where it directs the formation of large lipid droplets[33,34] and *PLIN2*, the only constitutively and ubiquitously expressed lipid-droplet protein that is associated with a range of lipid droplets in diverse cell types.[35,36] We observed that mRNA expression levels of *PLIN1* positively correlated with the *Lipid Granularity* features informative for larger spot sizes (*Lipid Granularity* measures 12–16) (Figure 1H). *PLIN2* correlated best with *Lipid Granularity* measures of smaller and larger spectra (Figure S1I). Accordingly, when we knocked out *PLIN1* and *FASN*, genes involved in lipid-droplet dynamics and lipid metabolism, we observed a size-specific reduction of *Lipid Granularity* (Figures 1H and S1J), suggesting that *Lipid Granularity* features are a suitable output measure of lipid-droplet size spectra and an indicator of adipocyte differentiation.

Consistent with the relevance of mitochondria for brown adipocyte function, mitochondrial measures were among the features that increased the most in hBAT (Figure 1G), particularly the *Texture* feature *Cells_Texture_InfoMeas1_Mito* ($p = 1.0 \times 10^{-3}$), which describes the overall information content based on the smoothness of a given stain. Perturbation of *MFN1*, a mitochondrial fusion gene, increased *Cells_Texture_InfoMeas1_Mito* in hWAT adipocytes ($adj.p = 4.0 \times 10^{-2}$; Figure 1I), suggesting that the higher values of this measurement in differentiated hBAT could be indicative of higher mitochondrial fission in hBAT compared with hWAT. This finding is consistent with brown adipocytes elevating mitochondrial thermogenesis by increasing mitochondrial fission.[37] hBAT adipocytes are further characterized by increased *Mito Intensity* compared with hWAT adipocytes throughout differentiation, with the most substantial increase in the fully differentiated state (median, day 8 $adj.p = 1.6 \times 10^{-3}$, day 14 $adj.p = 1.3 \times 10^{-3}$; Figure 1J), demonstrating that LipocyteProfiler can identify known cellular programs that distinguish different adipocyte lineages. Indeed, when we perturbed *PPARGC1A*, the master regulator of mitochondrial biogenesis and thermogenesis in adipocytes, using CRISPR-Cas9-mediated knockout in hWAT, mitochondrial intensity decreased (guide 1 $adj.p = 1.7 \times 10^{-5}$, guide 2 $adj.p = 8.8 \times 10^{-7}$; Figure 1J). We also confirmed similar morphological and cellular differences between brown hBAT adipocytes and another established white adipocyte cell line, differentiated SGBS (Simpson-Golabi-Behmel syndrome) cells (Figure S2A). Taken together, our data demonstrate that LipocyteProfiler can generate rich sets of morphological and cellular features that correlate with cellular function.

LipocyteProfiler identifies distinct depot-specific signatures associated with differentiation trajectories in visceral and subcutaneous adipocytes

We next used LipocyteProfiler to distinguish phenotypes of primary human adipose-derived mesenchymal stem cells (AMSCs) derived from the two main adipose tissue depots in the body, namely subcutaneous and visceral, across the course of differentiation (Figure 2A). We differentiated subcutaneous and visceral AMSCs and generated morphological profiles at days 0, 3, 8, and 14 using LipocyteProfiler and validated successful differentiation in both depots by an increase of adipogenesis marker genes (*LIPE*, *PPARG*, *PLIN1*, *GLUT4*) (Figure S2B). Concomitantly, we used RNA-seq to profile the transcriptome

(B) LipocyteProfiler and transcriptome profiles show time-course-specific signatures revealing a differentiation trajectory, but only LipocyteProfiler additionally resolves adipose-depot-specific signatures.
(C) Subcutaneous and visceral AMSCs at terminal differentiation (day 14) have distinct morphological and cellular profiles with differences that are spread across all channels. See also Figure S2C (volcano plot reporting the $-\log_{10}$ p value and the effect comparing subcutaneous and visceral adipocytes, t test).
(D) Sample progression discovery analysis (SPD). Proportions of subgroups of features characterizing differentiation differ between subcutaneous and visceral adipocytes and dynamically change over the course of differentiation. In both depots, *Mito* features drive differentiation predominantly in the early phase of differentiation (days 0–3) whereas *Lipid* features predominate in the terminal phases (days 8–14). See also Figure S2D for SPD of hWAT and SGBS.
(E) The number of lipid droplets is higher in subcutaneous AMSCs than in visceral AMSCs at terminal differentiation. y axis shows LP units (normalized LP values across eight batches, see STAR Methods).
(F) Mature subcutaneous AMSCs have larger intracellular lipid droplets compared with visceral AMSCs at day 14 of differentiation (*Lipid Granularity*). y axis shows autoscaled LP units (normalized LP values across eight batches, see STAR Methods).
(G) *Lipid Granularity* from subcutaneous AMSCs at day 14 of differentiation correlates positively with floating mature adipocyte diameter but shows an inverse relationship for visceral adipose tissue, suggesting distinct cellular mechanisms that lead to adipose tissue hypertrophy in these two depots. y axis shows autoscaled LP units (normalized LP values across eight batches; x axis, histology adipocytes diameter [μm], see STAR Methods).

Please cite this article in press as: Laber et al., Discovering cellular programs of intrinsic and extrinsic drivers of metabolic traits using LipocyteProfiler, Cell Genomics (2023), https://doi.org/10.1016/j.xgen.2023.100346

on the same differentiation time points. We observed that both the morphological and transcriptomic profiles show time-course-specific signatures revealing a differentiation trajectory; however, only morphological profiles generated by LipocyteProfiler also resolved adipose-depot-specific signatures throughout differentiation (Figure 2B). At day 14 of differentiation, morphological differences between subcutaneous and visceral adipocytes were spread across a large number of features in all feature classes (Figures 2C and S2C).

To discover patterns associated with progression through adipocyte differentiation in each depot, we performed a sample progression discovery analysis (SPD).[38] SPD clusters samples to reveal their underlying progression and simultaneously identifies subsets of features that show the same progression pattern and are illustrative of differentiation. We discovered that subsets of features distinguish the differentiation patterns of subcutaneous and visceral adipocytes and that most dominant feature classes were dynamically changing over the time course of differentiation (Figure 2D). In visceral adipocytes, the early phase of differentiation was predominantly associated with mitochondrial features, whereas terminal phases of differentiation were primarily associated with changes in lipid-related features (Figure 2D). In subcutaneous adipocytes, we observed that the feature classes (actin cytoskeleton, lipid, mitochondrial, and nucleic acid) were more evenly involved through adipogenesis and that the contribution of *Lipid* features started in early phases of differentiation, consistent with an earlier initiation of lipid accumulation in subcutaneous compared with visceral adipocytes (Figure 2D). To demonstrate that LipocyteProfiler captures consistent morphological patterns across white adipocyte models, we applied SPD to differentiate white adipocytes and show that SGBS cells, the immortalized subcutaneous hWAT line, and subcutaneous AMSCs progress similarly throughout adipocyte differentiation (Figure S2D). More specifically, we show that among the feature classes, *Lipid* features contribute the most to dynamical changes during adipocyte differentiation. SGBS cells begin the differentiation process by initiating the formation of lipid droplets (percentage of *Lipid* features between day 0 and day 3), which grow in the later stages (contribution of *AGP* features between days 8 and 14). These dynamic changes are similar to morphological changes of subcutaneous adipocytes and indicate that LipocyteProfiler can capture cell-specific morphological characteristics.

We next compared lipid-related signatures in mature AMSCs and observed that subcutaneous AMSCs had more lipid droplets than visceral AMSCs (*Cells_LipidObject_count*, Figure 2E, $q = 3.2 \times 10^{-4}$, false discovery rate [FDR] < 1%). More specifically, mature subcutaneous AMSCs showed significantly higher *Lipid Granularity* of small to medium-sized lipid objects, whereas visceral adipocytes showed higher *Lipid Granularity* of very small lipid objects, suggesting that mature subcutaneous AMSCs have larger intracellular lipid droplets compared with visceral AMSCs, which present higher abundance of very small lipid droplets (Figure 2F). These apparent intrinsic differences in differentiation capacity and lipid accumulation between subcutaneous and visceral AMSCs are consistent with previously described distinctions between AMSC depot properties across differentiation.[39] Our data suggest that LipocyteProfiler can facilitate identification of distinct lineage differences and programs of cellular differentiation.

Lastly, to assess the *in vivo* relevance of morphological features of *in vitro* differentiated adipocytes, we correlated *Lipid Granularity* features of adipocytes at day 14 of differentiation with diameter estimates of tissue-derived mature adipocytes from the same individual (see STAR Methods). We showed that changes in *Lipid Granularity* of *in vitro* differentiated female subcutaneous adipocytes correlated significantly with the mean diameter of mature adipocytes (Figure 2G). More specifically, medium-size granularity measures increased with larger *in vivo* size estimates, suggesting that *in vivo* adipocyte size is reflected by medium-sized lipid droplets in subcutaneous adipocytes that have been differentiated *in vitro*. Strikingly, we found the opposite effect between correlation of visceral *Lipid Granularity* and diameter estimates from mature adipocytes, suggesting that subcutaneous and visceral adipose tissues differ in cellular programs that govern depot-specific adipose tissue expansion, which may account for different depot-specific susceptibility to metabolic diseases. Indeed, white adipose depots have been reported to differ in their respective mechanisms of fat mass expansion under metabolic challenges, with subcutaneous adipose tissue being more capable of hyperplasia whereas visceral adipose tissue expands mainly via hypertrophy.[40]

LipocyteProfiler features reflect transcriptional states in adipocytes

To identify relevant processes that manifest in morphological and cellular features and to identify pathways of a given set of features, we next used a linear mixed model to link the expression of 52,170 genes derived from RNA-seq with each of the image-based LipocyteProfiler features in subcutaneous adipocytes at day 14 of differentiation across 26 individuals (Figure 3A and STAR Methods). We found 20,296 non-redundant significant feature-gene connections that were composed of 7,012 genes and 669 features (FDR < 0.01%, Figure 3B and Table S1A; FDR < 0.1%: 44,736 non-redundant feature-gene connections, 10,931 genes and 869 features, Figure S3A and Table S1B), and mapped across all channels (Figure 3A). Although features from every channel had significant gene correlations, *Lipid* features showed the highest number of gene connections compared with any other channel. This suggests that lipid-droplet structure, localization, and dynamics in adipocytes most closely represent the transcriptional state of the differentiated cell (Figure 3B). Pathway enrichment analyses of lists of genes connected to a feature at FDR < 0.01% add support to the idea that genes that correlated with a particular feature are biologically meaningful. For example, *Mito Granularity* associated with genes that are enriched for pathways such as the tricarboxylic acid cycle (TCA), which oxidizes acetyl-coenzyme A in mitochondria,[41] and lipid and lipoprotein metabolism and triglyceride biosynthesis, which are known to involve mitochondrial processes (pathway enrichment analysis FDR < 5%). This connection between lipid and mitochondrial processes was also detected in feature-gene associations for *Lipid Intensity*, which associates with genes significantly enriched in oxidative phosphorylation (OXPHOS) and β-oxidation (WikiPathway 368 and WikiPathway 143), and *Lipid Granularity* in the cytoplasm,

Please cite this article in press as: Laber et al., Discovering cellular programs of intrinsic and extrinsic drivers of metabolic traits using LipocyteProfiler, Cell Genomics (2023), https://doi.org/10.1016/j.xgen.2023.100346

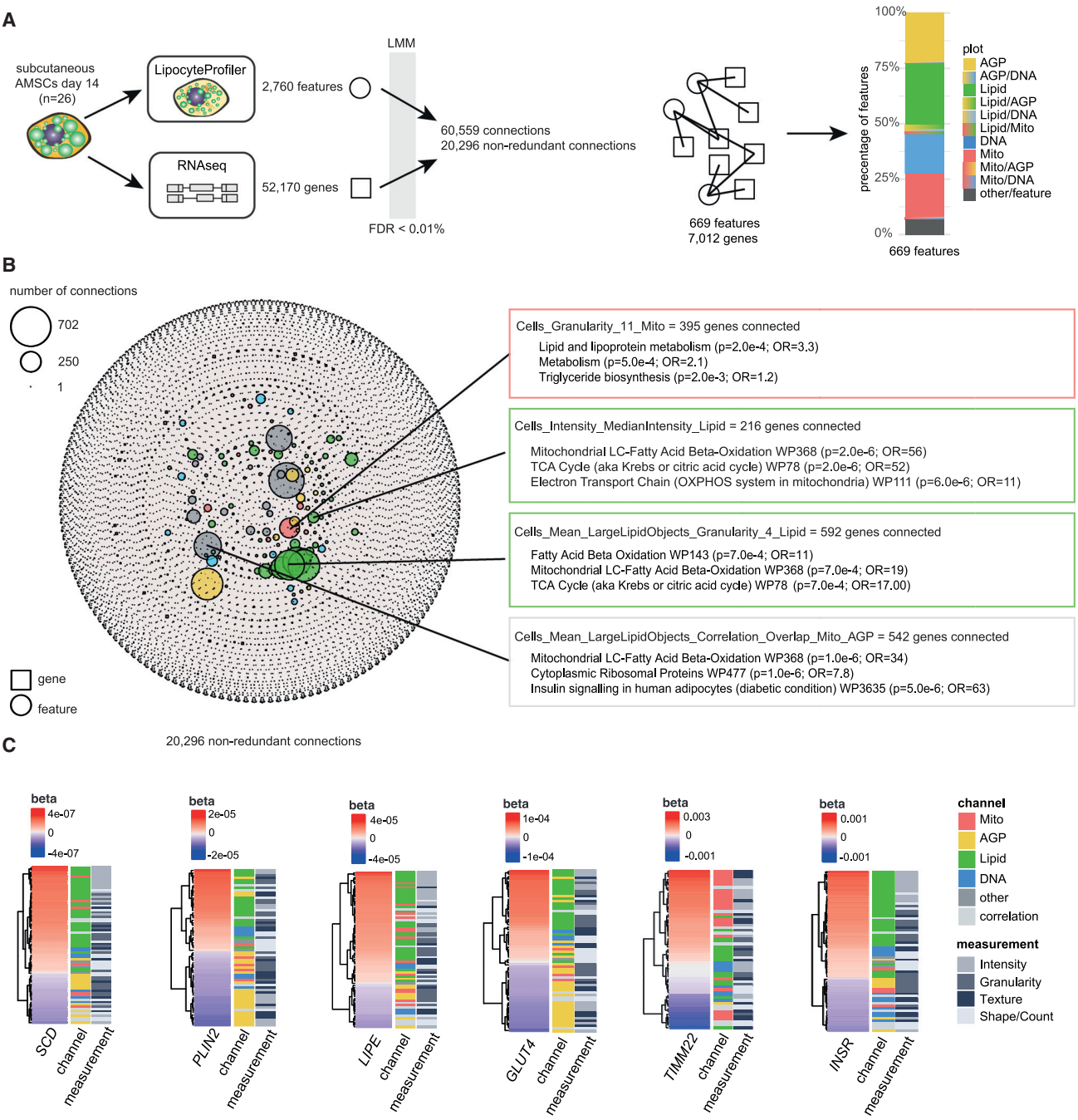


Figure 3. Correlations between morphological and transcriptional profiles

(A) Linear mixed model (LMM) was applied to correlate 2,760 morphological features derived from LipocyteProfiler with 52,170 transcripts derived from RNA-seq in matched samples of subcutaneous AMSCs at terminal differentiation (day 14). With FDR < 0.01%, we discover 20,296 non-redundant connections that map to 669 morphological features and 7,012 genes.

(B) Network of transcript-LipocyteProfiler feature correlations (significant connections FDR < 0.01%). Genes correlated with individual LipocyteProfiler features are enriched for relevant pathways (FDR < 5%). Node size is determined by number of connections. See also Figure S3A for a network with a significance level threshold of FDR < 0.1%.

(C) LipocyteProfiler signatures of adipocyte marker genes *SCD*, *PLIN2*, *LIPE*, *GLUT4*, *TIMM22*, and *INSR* recapitulate their known cellular function. Features are clustered based on beta of linear regression.

Please cite this article in press as: Laber et al., Discovering cellular programs of intrinsic and extrinsic drivers of metabolic traits using LipocyteProfiler, Cell Genomics (2023), https://doi.org/10.1016/j.xgen.2023.100346

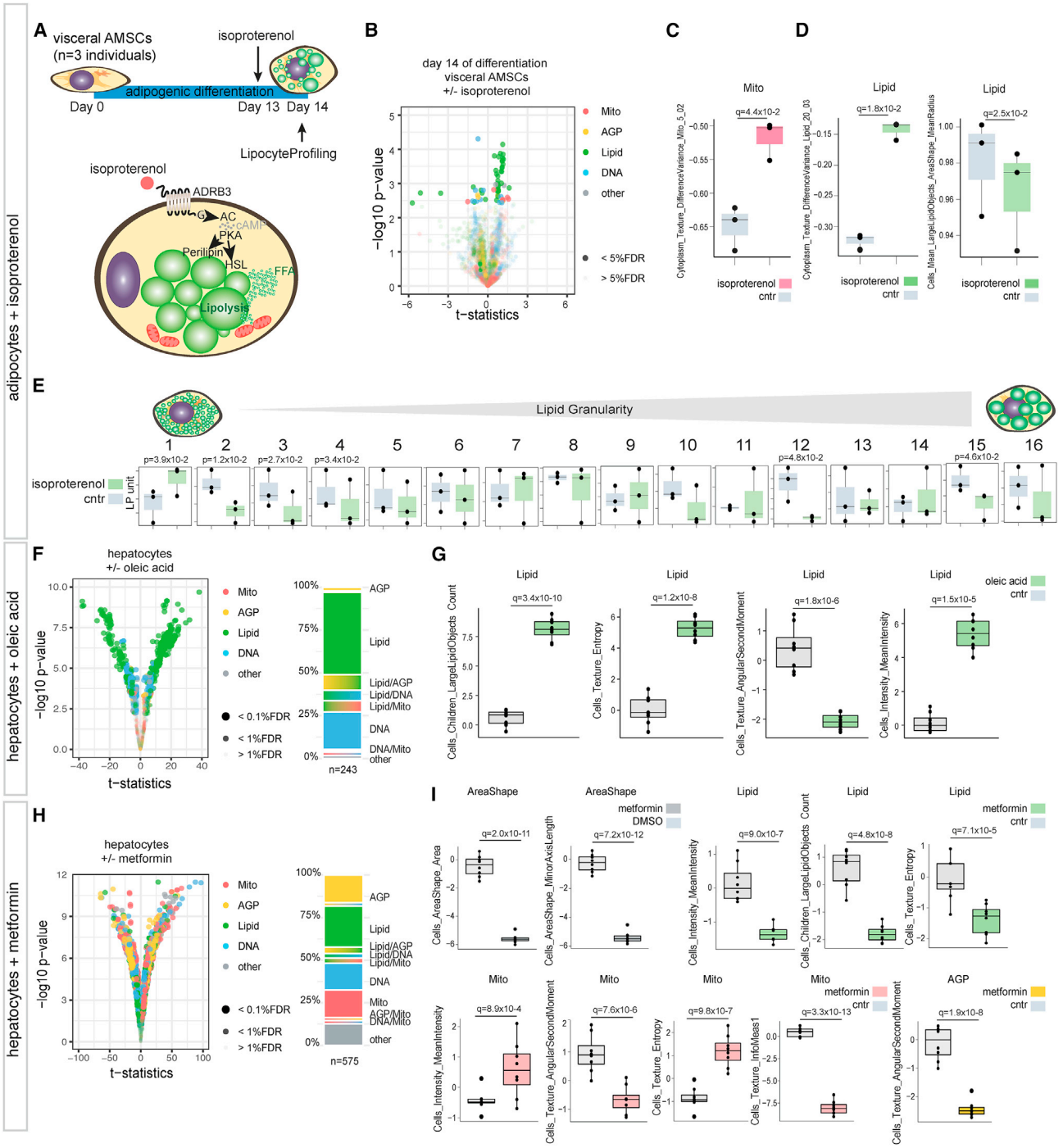


Figure 4. LipocyteProfiler identifies molecular mechanisms of drug stimulations in adipocytes and hepatocytes

(A) LipocyteProfiler was performed in visceral AMSCs (n = 3) treated with the β-adrenergic receptor agonist isoproterenol for 24 h.

(B) Isoproterenol treatment results in changes of lipid-related and mitochondrial traits in visceral AMSCs at day 14 of differentiation. See also Figure S3B (volcano plot reporting the $-\log_{10}$ p value and the effect comparing isoproterenol-treated cells and DMSO-treated cells, t test).

(C and D) Isoproterenol treatment of visceral AMSCs increase *Mito* and *Lipid TextureDifferenceVariance* while decreasing the respective *LargeLipidObject* mean radius features. y axis shows LP units (normalized LP values across eight batches, see STAR Methods).

(E) Isoproterenol treatment reduces lipid-droplet sizes measured via lipid granularity. y axis shows autoscaled LP units (normalized LP values across eight batches, see STAR Methods).

(F) Oleic acid treatment in PHH results in changes of lipid-related features.

(legend continued on next page)

Please cite this article in press as: Laber et al., Discovering cellular programs of intrinsic and extrinsic drivers of metabolic traits using LipocyteProfiler, Cell Genomics (2023), https://doi.org/10.1016/j.xgen.2023.100346

which associates with genes involved in fatty acid oxidation, mitochondrial long-chain fatty acid β-oxidation, and the TCA cycle. Finally, correlation features, which capture the overlap between lipid droplets, mitochondria, and AGP, were enriched for cytoplasmic ribosomal proteins, genes involved in mitochondrial long-chain fatty acid β-oxidation, and genes involved in insulin signaling in human adipocytes (Figures 3B and S3A; Table S1).

In addition to examining genes connected to feature groups, we also explored morphological features connected to specific genes. We found that morphological signatures of *SCD*, *PLIN2*, *LIPE*, *GLUT4*, *TIMM22*, and *INSR* revealed their known cellular functions (Figure 3C and Table S2). For example, the expression of *TIMM22*, a mitochondrial membrane gene, was most strongly correlated with *Mito Texture*. Expression of the insulin receptor (*INSR*) most strongly correlated with *Lipid Intensity* features indicative of lipid accumulation. *PLIN2* and *GLUT4* showed the highest positive and negative correlations with *Lipid* and *AGP* features, respectively. Together, these data show that mechanistic information gained from LipocyteProfiler features is not limited to generic cellular organelles but reflects the transcriptional state of the cell and can be deployed to gain relevant mechanistic insights.

LipocyteProfiler identifies cellular processes affected by drug perturbations in adipocytes and hepatocytes

To investigate whether LipocyteProfiler can identify effects of drug perturbations on cellular profiles, we first compared subcutaneous and visceral adipocytes that had been stimulated with the β-adrenergic agonist isoproterenol (Figure 4A). Isoproterenol is known to induce lipolysis and increase mitochondrial energy dissipation.[42] We observed that visceral adipocytes responded to isoproterenol treatment by changes in *Lipid* and *Mito* features (Figure 4B; Tables S3A and S3B). More specifically, we observed that isoproterenol-treated visceral adipocytes were characterized by differences in mitochondrial *Texture* (Difference Variance, $q = 4.4 \times 10^{-2}$), indicative of a less-smooth appearance of mitochondrial staining compared with DMSO-treated controls (Figure 4C). This suggests that isoproterenol treatment results in more hyperpolarized and fragmented mitochondria, which is a reported mechanism of norepinephrine-stimulated browning in adipocytes.[37] Isoproterenol-treated visceral adipocytes are further characterized by increased *Lipid* Difference Variance ($q = 1.8 \times 10^{-2}$) and decreased area of *large Lipid objects*, i.e., decreased mean radius and area of large lipid droplets ($q = 2.5 \times 10^{-2}$) (Figure 4D) as well as decreased *Lipid Granularity* across the full granularity size spectra, particularly at the smallest lipid-droplet sizes (Figure 4E). In fact, LipocyteProfiler core features highlight the importance of decreased lipid-droplet size and lipid intensity (Table S3B). This pattern suggests less overall lipid content in isoproterenol-treated lipolytic visceral adipocytes. Finally, the phenotypic response following isoproterenol treatment was predominant in visceral adipocytes, as we did not observe a significant effect (FDR < 5%) in subcutaneous adipocytes (Figure S3B). Indeed, adrenergic induced lipolysis is observed to be higher in visceral than subcutaneous in overweight and obese individuals.[43,44]

To test LipocyteProfiler in cell types beyond adipocytes, we assayed the effects of oleic acid and metformin in primary human hepatocytes (PHH). Consistent with the finding that free fatty acid treatment induces lipid-droplet accumulation in PHH,[45] our results showed that treatment of PHH with oleic acid yielded predominantly *Lipid* feature changes in the cell (Figure 4F and Table S4), with a morphological profile indicative of increased lipid-droplet number (*LargeLipidObjects_Count*, $q = 3.4 \times 10^{-10}$) and overall lipid content (*Cells_MeanIntensity_Lipid*, $q = 1.5 \times 10^{-5}$) as well as differences in *Texture* (*Cells_Texture_Entropy_Lipid*, $q = 1.2 \times 10^{-8}$; *Cells_Texture_AngularSecondMoment_Lipid*, $q = 1.8 \times 10^{-6}$; Figure 4G). By contrast, treatment of PHH with metformin caused morphological and cellular changes that were spread across all channels (Figure 4H and Table S5), with a profile suggestive of smaller cells (*Cells_AreaShape_Area*, $q = 2.0 \times 10^{-11}$; *Cells_AreaShape_MinorAxisLength*, $q = 7.2 \times 10^{-12}$) with increased mitochondrial membrane potential (*Cells_MeanIntensity_Mito*, $q = 8.9 \times 10^{-4}$), and mitochondrial heterogeneity (*Cells_Texture_AngularSecondMoment_Mito*, $q = 7.6 \times 10^{-6}$; *Cells_Texture_Entropy_Mito*, $q = 9.8 \times 10^{-7}$; *Cells_Texture_InfoMeas1_Mito*, $q = 3.3 \times 10^{-13}$). Additionally, we observed reduced lipid content (*Cells_MeanIntensity_Lipid*, $q = 9.0 \times 10^{-7}$), reduced lipid-droplet number (*LargeLipidObjects_Count*, $q = 4.8 \times 10^{-8}$), and differences in *Texture* (*Cells_Texture_Entropy_Lipid*, $q = 7.1 \times 10^{-5}$) (Figure 4I). This concerted effect of metformin on mitochondrial structure and function as well as lipid-related features is consistent with a less uniform appearance of the cytoskeleton, Golgi, and plasma membrane in metformin-treated hepatocytes compared with control (*Cells_Texture_AngularSecondMoment_AGP*, $q = 1.9 \times 10^{-8}$, Figure 4I). Indeed, prolonged treatment with high doses of metformin leads to mitochondrial uncoupling, resulting in mitochondrial hyperpolarization and diminished lipid accumulation in PHH.[45–47] Together, these data demonstrate that morphological and cellular profiles of drug perturbation in lipocytes yield cellular signatures reflecting known biology and drug action in a single concerted snapshot of cell behavior.

Polygenic risk effects for insulin resistance affects lipid degradation in differentiated visceral adipocytes

Next, we used LipocyteProfiler to discover cellular programs of metabolic polygenic risk in adipocytes. For systematic profiling of AMSCs in the context of natural genetic variation (Table S6), we first assessed the effect of both technical and biological variance on LipocyteProfiler features. To obtain a measure of batch-to-batch variance associated with our experimental setup, we differentiated hWAT, hBAT, and SGBS preadipocytes[48] in three

(G) Oleic acid treatment in PHH affects lipid-related morphological features suggestive of increased lipid-droplet size and number. y axis shows LP units (normalized LP values across PHH data, see STAR Methods).
(H) Metformin treatment in PHH results in global changes affecting features across all channels.
(I) Metformin effect in hepatocytes is suggestive of increased mitochondrial activity, while lipid-droplet size and number are reduced. Metformin-treated hepatocytes are also smaller and show reduced cytoskeletal randomness. y axis shows LP units (normalized LP values across PHH data, see STAR Methods).

Please cite this article in press as: Laber et al., Discovering cellular programs of intrinsic and extrinsic drivers of metabolic traits using LipocyteProfiler, Cell Genomics (2023), https://doi.org/10.1016/j.xgen.2023.100346

Cell Genomics
Technology

CellPress
OPEN ACCESS

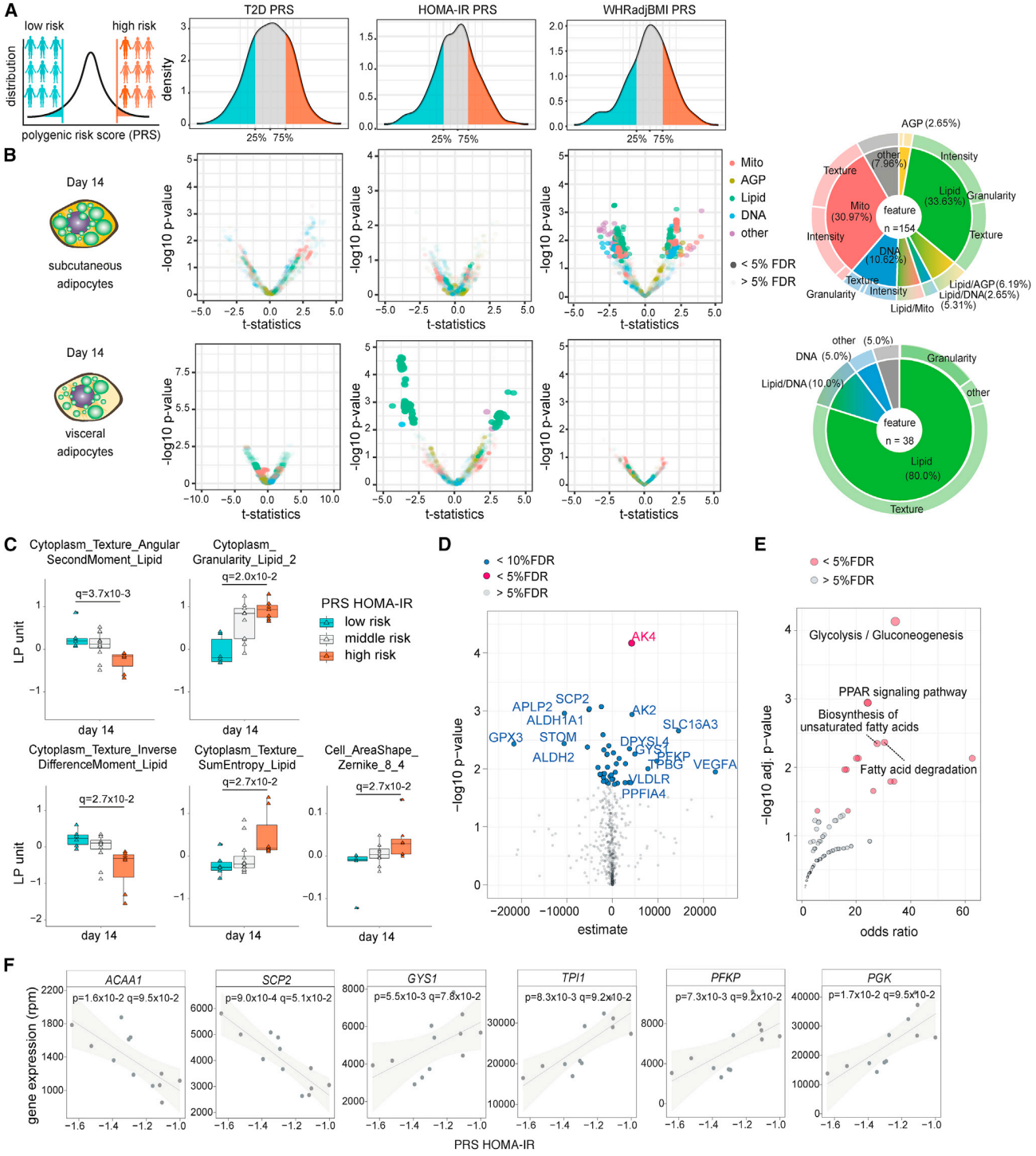


Figure 5. Polygenic risk effects for insulin resistance affect lipid degradation in differentiated visceral adipocytes

(A) Donors from the bottom and top 25 percentiles of genome-wide PRS for three T2D-related traits (HOMA-IR, T2D, WHRadjBMI) were selected to compare LipocyteProfiles across the time course of visceral and subcutaneous adipocyte differentiation.

(B) LipocyteProfiler applied to visceral and subcutaneous differentiating adipocytes reveals trait-specific polygenic effects on image-based cellular signatures for HOMA-IR in differentiated visceral AMSCs (day 14; largely *Lipid* features) and WHRadjBMI in subcutaneous adipocytes (day 14, largely *Mito* and *Lipid* features), but no effect for T2D. See also Figures S5A and S5D (days 0, 3, and 8).

(legend continued on next page)

Please cite this article in press as: Laber et al., Discovering cellular programs of intrinsic and extrinsic drivers of metabolic traits using LipocyteProfiler, Cell Genomics (2023), https://doi.org/10.1016/j.xgen.2023.100346

independent experiments and found no significant batch effect (BEscore 0.0047, 0.0001, and 0.0003; Figure S4A). We also used a machine-learning-based classification model[49] to predict variables such as batch and cell type, based on the morphological profile. We show that the accuracy of predicting cell type is substantially higher than predicting batch (Figure S4A), indicating that our LipocyteProfiler framework can detect intrinsic versus extrinsic variance in our dataset with low batch effect and high accuracy. Second, we performed a variance component analysis across 65 donor-derived differentiating AMSCs to assess the contribution of intrinsic genetic variation compared with the contribution of other possible confounding factors such as batch, T2D status, age, sex, body mass index (BMI), cell density, and passage number. In total, we found that across all samples and batches, the largest contributor to feature variance was donor ID, accounting for 17.03% (interquartile range 11.45%–21.95%) of variance (Figure S4B). Other factors appeared to contribute only marginally to overall variance of the data, including extrinsic factors such as batch effect (6.02%, 3.94%–8.84%), plating density (3.75%, 1.55%–5.61%) and intrinsic factors such as sex (0.86%, 0.26%–2.44%), age (0.55%, 0.15%–1.39%), BMI (0.41%, 0.07%–1.33%), and T2D (0.19%, 0.03%–0.59%). These data suggest that LipocyteProfiler allows us to detect and distinguish interindividual genetic feature variation to a similar degree as reported for human induced pluripotent stem cells (iPSCs), where quantitative assays of cell morphology demonstrated a donor contribution to interindividual variation in the range of 8%–23%.[50] To account for the variable feature-specific contributions of batch, sex, age, and BMI to overall feature variance, we corrected for those covariables in our analyses. Together, these data suggest that LipocyteProfiler features can be used to study the effect of genetic contributions to morphological and cellular programs.

To ascertain the effect of polygenic risk for metabolic disease on cellular programs, we used the latest genome-wide association study (GWAS) summary statistics for T2D. We constructed individual genome-wide polygenic risk scores (PRSs) for three T2D-related traits that have been linked to adipose tissue: T2D,[51] insulin resistance by homeostasis measure assessment (HOMA-IR[52,53]), and waist-to-hip ratio adjusted for BMI (WHRadjBMI[54]). To evaluate whether HOMA-IR PRS effects are confounded by BMI, we compared the distribution of BMI between groups of high, medium, and low HOMA-IR PRS carriers and observed that HOMA-IR PRS appears to be largely independent of BMI (Kolmogorov-Smirnov tests: top 25% and bottom 25% p = 0.797, Figure S4C). Next, we selected donors from the bottom and top 25th percentiles of these genome-wide PRS distributions (referred to as low and high polygenic risk) and compared LipocyteProfiler features across the time course of visceral and subcutaneous adipocyte differentiation in high and low polygenic risk groups (Figures 5A, 5B, S5A, and S5D; Tables S7 and S8).

We found significant effects on image-based cellular signatures for HOMA-IR and WHRadjBMI, but no polygenic effect for T2D (Figures 5B, S5A, and S5D; Tables S7 and S8). More specifically, we observed an effect of HOMA-IR polygenic risk on cellular profiles at day 14 in visceral adipocytes (38 features, FDR < 5%, Figure 5B and Table S7A), indicating a spatiotemporal and depot-specific effect of polygenic risk for insulin resistance. The features that differed between the high and low HOMA-IR PRS carriers were mostly *Lipid* features (Figure 5B). Visceral adipocytes from high polygenic risk individuals showed increased *Lipid Granularity* ($q = 2.0 \times 10^{-2}$), increased *Cytoplasm_Texture_SumEntropy_Lipid* ($q = 2.7 \times 10^{-2}$), increased *Cells_AreaShape_Zernike_8_4* ($q = 2.7 \times 10^{-2}$), decreased *Cytoplasm_Texture_InverseDifferenceMoment_Lipid* ($q = 2.7 \times 10^{-2}$), and reduced *Cytoplasm_Texture_AngularSecondMoment_Lipid* ($q = 3.7 \times 10^{-3}$) in the cytoplasm compared with low polygenic risk individuals (Figure 5C). The data further reveal that the pattern that contrasts between high and low HOMA-IR polygenic risk carriers is driven by lipid-informative LipocyteProfiler core features. Cellular signatures of high HOMA-IR polygenic risk carriers include core features that describe increased *Lipid Granularity*, increased *Lipid Radial Distribution* in the middle rings of the cell, and increased *Cell Area* (Tables S7B–S7D). These data indicate that visceral adipocytes from individuals with high compared with low polygenic risk for insulin resistance are characterized by a lipid-rich cellular profile, driven by key features informative for increased number of small to medium-sized lipid droplets, less homogeneous lipid-droplet distribution, and larger adipocytes, indicating excessive lipid accumulation in visceral adipocytes from individuals at high polygenic risk. Notably, the pattern that differentiates individuals at high and low polygenic risk recapitulates signatures that resemble an inhibition of lipolysis, as demonstrated by the inverse direction of effect in isoproterenol-stimulated visceral AMSCs shown in Figure 4. Furthermore, we observed that lower HOMA-IR PRS increases the number of small lipid droplets in visceral adipocytes, which are precisely the features affected in response to isoproterenol (Figures S5B and 4E). Together, these image-derived rich representations of cellular signatures describe a cellular program that is characterized by a metabolic switch toward lipid accumulation rather than lipolysis in visceral adipocytes derived from individuals at high polygenic risk for insulin resistance.

To further resolve the cellular program underlying HOMA-IR PRSs in visceral adipocytes and ascertain the effects of polygenic risk for HOMA-IR on gene expression, we integrated image-based information from LipocyteProfiler with RNA-seq data from the same donor-derived samples. Looking at mRNA levels for 512 genes known to be involved in adipocyte differentiation and function (gene set enrichment analysis hallmark gene sets for adipogenesis, fatty acid metabolism, and glycolysis[55,56]), we identified 51 genes under the polygenic control of HOMA-IR (FDR < 10%)

(C) HOMA-IR polygenic risk in visceral AMSCs manifested in altered lipid texture, lipid granularity, and cell shape features, resembling an inhibition of lipolysis. y axis shows LP units (normalized LP values across eight batches, see STAR Methods). See also Figure 4B (isoproterenol stimulation).
(D) Linear regression of gene expression levels of 512 genes known to be involved in adipocyte function with HOMA-IR PRS.
(E) Pathway enrichment analysis of genes that correlate with HOMA-IR PRSs (FDR < 10%) in visceral adipocytes highlight biological processes related to glucose metabolism, fatty acid transport, degradation, and lipolysis (KEGG pathways 2019).
(F) Representative genes that associate with HOMA-IR PRS in visceral adipocytes.

Please cite this article in press as: Laber et al., Discovering cellular programs of intrinsic and extrinsic drivers of metabolic traits using LipocyteProfiler, Cell Genomics (2023), https://doi.org/10.1016/j.xgen.2023.100346

in fully differentiated visceral adipocytes (Figure 5D and Table S9). Genes correlating with the HOMA-IR PRS were enriched for biological processes related to glucose metabolism, fatty acid transport, degradation, and lipolysis (Figure 5E and Table S10). Negatively correlated genes include *ACAA1* ($p = 1.6 \times 10^{-2}$, $q = 9.5 \times 10^{-2}$) and *SCP2* ($p = 9.0 \times 10^{-4}$, $q = 5.1 \times 10^{-2}$) (Figure 5F), consistent with an inhibition of lipolysis and lipid degradation in visceral adipocytes from individuals at high polygenic risk for HOMA-IR. Positively correlated genes include *GYS1*, which is a regulator of glycogen biosynthesis shown to causally link glycogen metabolism to lipid-droplet formation in brown adipocytes[57] ($p = 5.5 \times 10^{-3}$, $q = 7.8 \times 10^{-2}$, Figure 5F). Additionally, multiple critical enzymes of the glycolysis pathway (*TPI1* [$q = 9.2 \times 10^{-2}$], *PFKP* [$q = 9.2 \times 10^{-2}$], *PGK* [$q = 9.5 \times 10^{-2}$], Figure 5F), and marker genes of energy metabolism (*AK2* and *AK4*; Figure S5C) are positively correlated with HOMA-IR PRS, suggesting a metabolic switch from lipolytic degradation of triglycerides to glycolytic activity. Although a causal link between visceral adipose mass and insulin resistance has been widely observed,[58] the mechanism behind this observation is not understood. Together, orthogonal evidence from both high-content image- and RNA-based profiling experiments in subcutaneous and visceral AMSCs suggests that individuals with high polygenic risk for HOMA-IR are characterized by blocking lipid degradation in visceral adipocytes.

Polygenic risk for lipodystrophy-like phenotype manifests in cellular programs indicating reduced lipid accumulation capacity in subcutaneous adipocytes

To resolve polygenic effects on adipocyte cellular programs beyond heterogeneous T2D and insulin resistance traits, we used the clinically informed process-specific partitioned PRS of lipodystrophy,[59] and tested for association of the lipodystrophy-like PRS and LipocyteProfiler features throughout adipocyte differentiation (linear regression adjusted for BMI, age, sex, and principal component 1 [PC1], FDR < 5%; Table S11 and Figure 6A). The lipodystrophy PRS was constructed based on 20 T2D-associated loci that were grouped together as having similar associations with a lipodystrophy-like phenotype, signifying insulin resistance with a lower BMI[59] (Figure 6A). We found that polygenic risk of lipodystrophy associates with distinctive features in the *Mito*, *AGP*, and *Lipid* categories in subcutaneous AMSCs at day 8 and day 14 of differentiation, whereas increased lipodystrophy PRS associates primarily with *Lipid* features in visceral adipocytes at nominal significance (Figures 6B–6D and S6A; Table S11A). This highlights a depot- and spatiotemporal-dependent effect of polygenic risk on cellular profiles captured with LipocyteProfiler. Using the LipocyteProfiler core feature set, we identified *Mito Intensity*, *Texture*, and *Granularity* features, *AGP Granularity* features, and *Lipid Intensity* features to be most informative for driving the lipodystrophy PRS cellular process in subcutaneous adipocytes (Table S11B). More specifically, the profiles that associate with lipodystrophy polygenic risk include core features informative for increased mitochondrial membrane potential (e.g., *Cells_Intensity_Integrated_Intensity_Mito* $q = 3.4 \times 10^{-2}$; *Cells_Intensity_Mean Intensity_Mito* $q = 3.4 \times 10^{-2}$; Table S11B), changes to the actin cytoskeleton indicating decreased cortical actin at the plasma membrane (e.g., *Cells_RadialDistribution_FracAtD_AGP* ring 2 of 4 $q = 3.4 \times 10^{-2}$ and 3 of 4 $q = 3.4 \times 10^{-2}$; Figure S6B and Table S11B), and decreased lipid accumulation in subcutaneous adipocytes (e.g., *Cells_RadialDistribution_RadialCV_Lipid_4of4* $q = 3.4 \times 10^{-2}$, *Cells_Texture_DifferenceEntropy_Lipid_10_00* $q = 3.4 \times 10^{-2}$; Table S11B). Strikingly, representative images of subcutaneous adipocytes derived from individuals at the tail ends of lipodystrophy PRS (high risk [25th percentiles] compared with low risk [bottom 25th percentiles]) confirm that adipocytes from high PRS carriers have increased mitochondrial stain intensity—indicating higher mitochondrial membrane potential[60]—accompanied by smaller lipid droplets on average compared with adipocytes from individuals with low PRS (Figure 6D). We also note that CRISPR-Cas9-mediated knockout of the monogenic familial partial lipodystrophy gene *PLIN1* maps to features informative for decreased number of medium- and large-sized lipid droplets (Figure 1H), matching the polygenic risk effect. To assess whether the identified cellular changes underlying lipodystrophy polygenic risk resemble cellular drivers of monogenic forms of lipodystrophy, we next correlated expression of marker genes of monogenic familial partial lipodystrophy syndromes (*PPARG*, *LIPE*, *PLIN1*, *AKT2*, *CIDEC*, *LMNA*, and *ZMPSTE24*) with LipocyteProfiler features across subcutaneous adipocytes from 26 individuals. We found similar cellular signatures between profiles from monogenic lipodystrophy-associated genes and the polygenic lipodystrophy profile, with high effect sizes of *Mito* and *AGP* features (Figure S6C). These results suggest that polygenic and monogenic forms of lipodystrophy converge on similar cellular mechanisms involving increased mitochondrial activity and decreased lipid accumulation in subcutaneous adipocytes from high PRS donors. This finding is consistent with the fact that different monogenic forms of lipodystrophy showed similar consequences on mitochondrial OXPHOS in patient samples.[61]

To further resolve the cellular pathways of lipodystrophy polygenic risk that could underlie the morphological signature in subcutaneous adipocytes, we created a network of genes linked to features identified to be under the control of lipodystrophy polygenic risk. This analysis identified 23 genes that had ten or more connections to features derived from the lipodystrophy PRS LipocyteProfiler (FDR < 0.1%, Figure 6E). Sixteen of those genes are significantly (FDR < 10%) correlated with the lipodystrophy PRS (Figure S6D). For example, we found *EHHADH* (a marker gene of peroxisomal β-oxidation) and *NFATC3* (a gene involved in mitochondrial fragmentation and previously linked to a lipodystrophic phenotype in mice[62]) to be positively correlated with increased polygenic risk ($q < 0.1$ in both cases; Figure S6D), suggesting that gene networks identified through LipocyteProfiler signatures recapitulate mechanisms of polygenic risk and that LipocyteProfiler can be used to identify molecular mechanisms of disease risk.

Together, these data map aggregated polygenic risk for a lipodystrophy-like phenotype onto cellular programs characterized by increased mitochondrial activity and decreased lipid accumulation in subcutaneous adipocytes, which is consistent with the notion that limited peripheral storage capacity of adipose tissue underlies polygenic lipodystrophy.[63]

Please cite this article in press as: Laber et al., Discovering cellular programs of intrinsic and extrinsic drivers of metabolic traits using LipocyteProfiler, Cell Genomics (2023), https://doi.org/10.1016/j.xgen.2023.100346

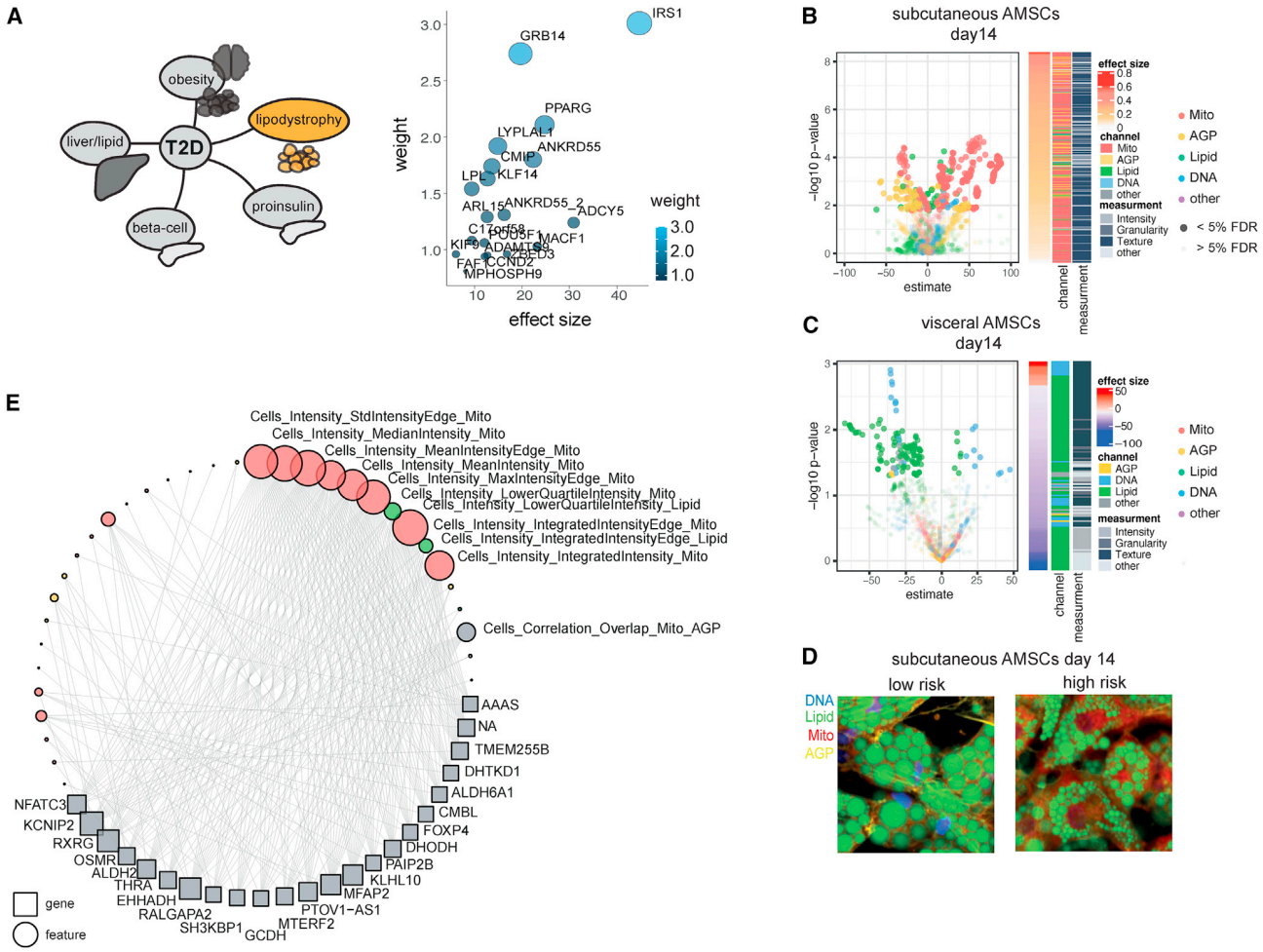


Figure 6. Polygenic risk for lipodystrophy-like phenotype manifests in cellular programs that indicate increased mitochondrial activity, reduced actin cytoskeleton remodeling, and reduced lipid accumulation capacity in subcutaneous adipocytes

(A) Schematic of T2D process-specific PRS (left panel). Lipodystrophy-specific PRS consists of 20 T2D-associated loci contributing to polygenic risk for a lipodystrophy-like phenotype.[59] y axis: weights of individual loci; x axis: effect size of individual loci contributing to polygenic risk for a lipodystrophy-like phenotype.

(B–D) Depot-specific effects on LipocyteProfiles in AMSCs at day 14 are under the polygenic control of the lipodystrophy cluster with a mitochondrial and AGP-driven profile in subcutaneous AMSCs (B), whereas in visceral AMSCs mostly *Lipid* features were associated with increased polygenic risk (C). See also Figure S6A (days 0, 3, and 8). Computationally averaged images of subcutaneous AMSCs from low- and high-risk allele carriers for lipodystrophy PRS show higher mitochondrial intensity, reduced cortical actin, and reduced lipid-droplet size in high-risk carriers (D).

(E) Gene-feature connections for lipodystrophy PRS-mediated differential features are enriched for Mitochondrial Intensity features informative for mitochondrial membrane potential in subcutaneous AMSCs at day 14 (FDR < 0.1%). See also Figure S6D.

Allele-specific effect of the *2p23.3* lipodystrophy-like locus on mitochondrial fragmentation and lipid accumulation in visceral adipocytes

To confirm that LipocyteProfiler can link an individual genetic risk locus to meaningful cellular profiles in visceral adipocytes, we investigated a locus on chromosome 2, spanning the *DNMT3A* gene at location *2p23.3*, which is one out of the 20 lipodystrophy process-specific risk loci included in the lipodystrophy PRS analyses. The *2p23.3* metabolic risk haplotype (minor allele frequency of 0.35 in 1000 Genomes Phase 3 combined populations) is associated with a higher risk for T2D and WHRadjBMI (Figure 7A). To map the *2p23.3* metabolic risk locus to cellular functions, we compared LipocyteProfiler features of subcutaneous and visceral AMSCs from risk and non-risk haplotype carriers at three time points during adipocyte differentiation: before (day 0), early (day3), and terminal (day 14) differentiation (Figure 7B). In visceral AMSCs, we identified 92 and 23 core features that are significantly different between haplotypes at day 3 and day 14 of differentiation, respectively (Figure 7C and Table S12). At day 3, 70% of significantly different image-based features are mitochondrial, and on day 14, 80% of differential features are lipid-related. These findings suggest that the *2p23.3* locus is associated with a mitochondrial function phenotype during early differentiation, which then progress to altered

Please cite this article in press as: Laber et al., Discovering cellular programs of intrinsic and extrinsic drivers of metabolic traits using LipocyteProfiler, Cell Genomics (2023), https://doi.org/10.1016/j.xgen.2023.100346

CellPress
OPEN ACCESS

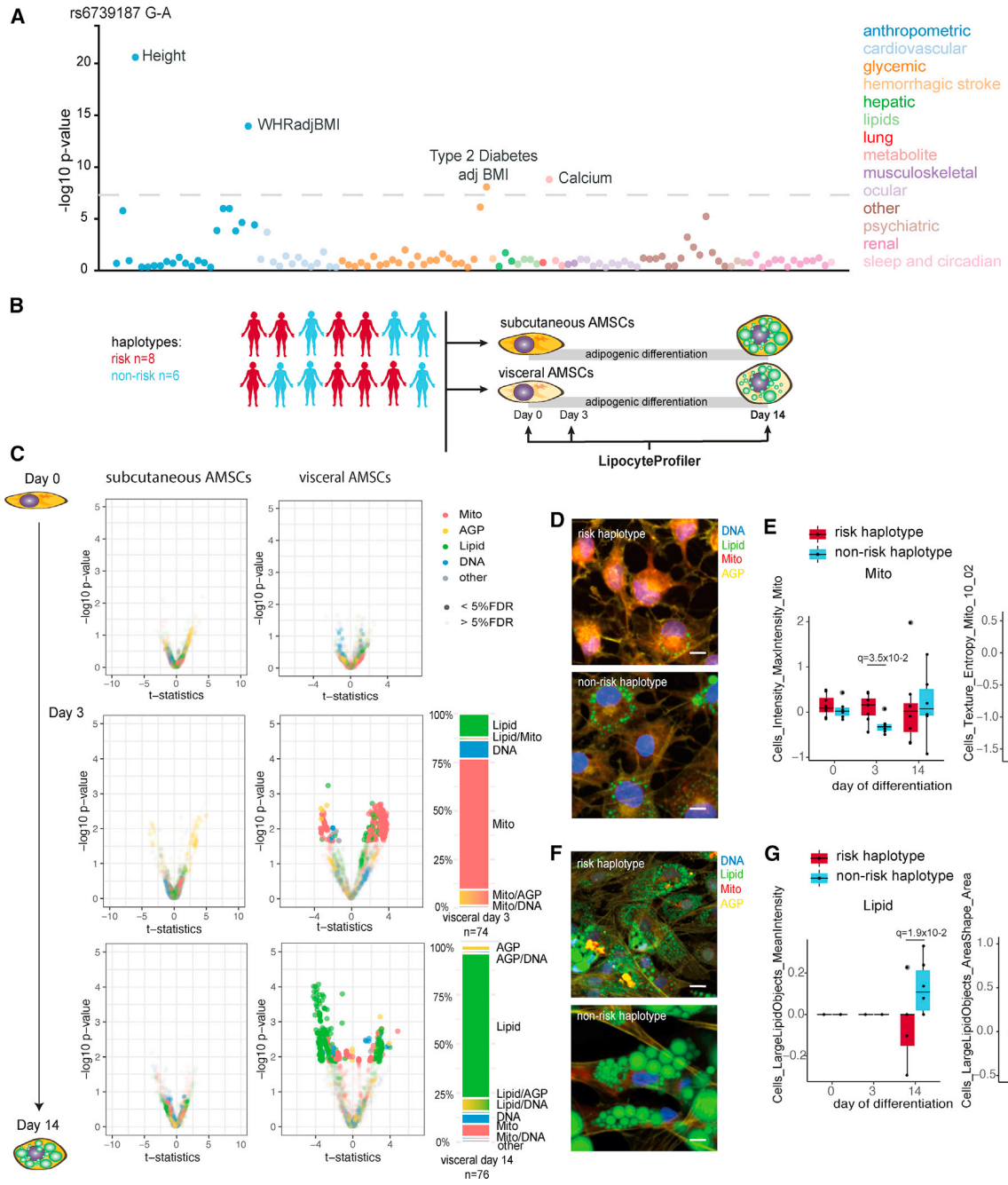


Figure 7. *2p23.3* lipodystrophy-like locus effect on mitochondrial fragmentation and lipid accumulation in visceral adipocytes

(A) PheWAS[64] at the *2q23.3* risk locus shows associations with height, WHRadjBMI, T2D, and Calcium.

(B) LipocyteProfiler was performed in subcutaneous and visceral AMSCs of eight risk and six non-risk haplotype carriers across adipocyte differentiation (days 0, 3, and 14).

(C) In visceral AMSCs, 74 and 76 features were different between haplotypes at day 3 and day 14 of differentiation, respectively, with 70% of differential features at day 3 being mitochondrial and 80% lipid-related at day 14.

(D) Representative images of visceral AMSCs from risk (top) and non-risk (bottom) haplotype at day 3 of differentiation stained using LipocytePainting. Scale bars, 10 μm.

(E) *Mito MaxIntensity* and *Mito Texture Entropy* were higher at day 3 of differentiation in visceral AMSCs from six risk haplotype carriers, suggesting more fragmented and higher mitochondrial membrane potential. y axis shows LP units (normalized LP values across eight batches, see STAR Methods).

(legend continued on next page)

Please cite this article in press as: Laber et al., Discovering cellular programs of intrinsic and extrinsic drivers of metabolic traits using LipocyteProfiler, Cell Genomics (2023), https://doi.org/10.1016/j.xgen.2023.100346

lipid-droplet formation in mature visceral adipocytes. Representative microscopic images from day 3 of differentiation show higher mitochondrial stain intensities in risk haplotype carriers (Figure 7D). The top-scoring, most differential mitochondrial features (*Cells_MaxIntensity_Mito* q = 3.5×10^{-2}, *Cells_Texture_Entropy_Mito* q = 3.5×10^{-2}, and *Cytoplasm_Granularity_7_Mito* q = 3.8×10^{-2}; Figures 7E and S7A) are increased in metabolic risk carriers, suggestive of less tubular mitochondria with increased mitochondrial membrane potential and altered function. At day 14 of differentiation, AMSCs from metabolic risk haplotype carriers show smaller lipid droplets in representative microscopic images (Figure 7F). More specifically, we observed that risk haplotype carriers have decreased *Lipid Intensity* (q = 1.9×10^{-2}; Figure 7G) in the cell and a smaller area of *large Lipid objects* (*LargeLipidobjects_AreaShape* q = 3.5×10^{-2}; Figure 7G), suggesting a lipid phenotype characterized by reduced lipid-droplet stabilization and/or formation. Distinct core features drive the genetic effect in visceral adipocytes at day 3, including *Mito Texture* (e.g., *DifferenceVariance* and *Entropy*) and at day 14 *Lipid Texture* features (Table S12A), highlighting cellular processes. This profile is associated with increased mature adipocyte diameter estimates (Figure 2) and suggests that risk haplotype carriers have a cellular profile that is consistent with visceral WAT hypertrophy. We further note that our findings in human adipocytes are corroborated by organismal perturbation of the candidate effector transcript *DNMT3A* in mice, where deletion of *Dnmt3a* results in changes of whole-body fat mass (Figure S7B)[65] and protects from high-fat-diet-induced insulin resistance, which is mainly attributed to actions in visceral adipose tissue.[66] Together, these data demonstrate that LipocyteProfiler captures complex cellular phenotypes associated with genetic risk for cardiometabolic diseases and traits and allows the effective resolution of spatial-temporal context of action. With LipocyteProfiler, we generated a resource that enables unbiased mechanistic interrogation of the hundreds of cardiometabolic disease loci with unknown functions. We have provided all data and software as open-access and open-source for the community.

DISCUSSION

We present a novel high-content image-based profiling framework, LipocyteProfiler, for enabling the identification of causal relationships between natural genetic variation, effect of drugs, and physiologically relevant stimulations, and the identification of effector genes with cellular programs in the context of cardiometabolic disease. We provide proof-of-principle results showcasing that we can link natural genetic variation to distinct morphological and cellular profiles using LipocyteProfiler-based deep phenotypic profiles generated in primary AMSCs with a given genetic background. This demonstrates that LipocyteProfiler is useful for unraveling disease-relevant complex cellular programs beyond hypothesis-driven cell-based readouts alone. We show that the information gained from LipocyteProfiler can report on both physiological and pathological states of the cell and identify cellular traits underlying cell-state transitions, providing a controlled toolkit to interrogate dynamic rather than static programs. Using LipocyteProfiler in defined cell states, we can robustly detect subtle phenotypic differences driven by drug treatment, genetic perturbation, and natural genetic variation. Our ability to detect these subtle changes might be a consequence of cell traits capturing the downstream manifestations of genomic, transcriptional, and proteomic effects. We show that polygenic risk for metabolic traits converges into discrete pathways and mechanisms. LipocyteProfiler elucidates morphological and cellular signatures underlying differential polygenic metabolic risk specific to distinct adipose depots, metabolic traits, and cell-developmental time points. For example, we observed polygenic effects on lipid degradation in visceral adipocytes in the context of insulin resistance, and mitochondrial activity and cytoskeleton remodeling in subcutaneous adipocytes under the control of lipodystrophy-specific polygenic contributors to T2D risk. We note that the mitochondrial and actin cytoskeleton informative cellular programs which associate with a lipodystrophy-like phenotype show similarities to morphological signatures associated with genome-wide polygenic risk for WHRadjBMI, a proxy of unfavorable fat distribution. Future work using other adiposity PRSs such as described in Martin et al.[67] will help to identify genetic drivers of cardiometabolic disease and further deconvolve the cellular programs underlying favorable und unfavorable adiposity.

LipocyteProfiler enables scalable, unbiased, mechanistic interrogation of metabolic disease loci whose functions remain unknown. By linking image-based profiles to transcriptional states, we provide a rich resource of gene-cellular trait connections that relate image-based features to biological processes. We envision that LipocyteProfiler-generated quantitative, high-dimensional representations of morphological and cellular features will complement the palette of omics-based profiling readouts. Combined with forward and reverse genetic screens, this can link genetic perturbations to cellular programs in lipid-accumulating cells. We also note that LipocyteProfiler may generate a suitably complex readout to contribute to ongoing endeavors in the community to improve differentiation protocols of iPSCs and discover cellular programs underlying genetic perturbations in high-throughput genetic screens. Moreover, we expect that the power to identify genetic drivers for metabolic diseases will be demonstrated using a plethora of univariate and multivariate genome-wide polygenic scores to resolve the molecular heterogeneities of T2D and other cardiometabolic traits. Finally, we expect that, with increased sample sizes, our approach will help to pave the way to map cellular quantitative trait loci in population-scale image-based profiling endeavors (GWAS-in-a-dish) to link common genetic risk variation to lipocyte phenotypes and accelerate therapeutic pathway discoveries. Our

(F) Representative images of visceral AMSCs from risk (top) and non-risk (bottom) haplotype at day 14 of differentiation stained using LipocytePainting. Scale bars, 10 μm.

(G) *LargeLipidObject MedianIntensity* was lower and *Lipid Texture AngularSecondMoment* was higher at day 14 of differentiation in visceral AMSCs from six risk haplotype carriers, suggesting a perturbed lipid phenotype characterized by reduced lipid-droplet stabilization and/or formation. y axis shows LP units (normalized LP values across eight batches, see STAR Methods).

Please cite this article in press as: Laber et al., Discovering cellular programs of intrinsic and extrinsic drivers of metabolic traits using LipocyteProfiler, Cell Genomics (2023), https://doi.org/10.1016/j.xgen.2023.100346

disease-oriented LipocyteProfiler image-based profiling tool can be modified by swapping or adding different disease-relevant dyes or antibodies of interest and could be applied to disease-relevant models for any disease of interest.

Limitations of the study

A primary limitation in the current study is the low sample size to link genetic variants to cellular and morphological processes, and as such we were not able to stratify by sex or other covariates. Findings presented here in our proof-of principle study need to be replicated in larger population-scale experiments in the future, which will help to evaluate reproducibility of our results and evaluate sex-specific polygenic risk effects in the context of cardiometabolic traits. We further acknowledge that the AMSCs are derived from patients undergoing abdominal laparoscopic surgery, and as such the BMI distribution of the patient cohort is skewed to high BMI (mean 49.34 ± SD 11.28).

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Human primary AMSC isolation/abdominal laparoscopy cohort–Munich obesity BioBank/MOBB
  - Differentiation of human AMSCs
  - Primary human hepatocyte culture
  - MGB Biobank cohort
- METHOD DETAILS
  - LipocytePainting
  - Staining and microscopy of actin-cytoskeleton in subcutaneous AMSCs
  - Isolation and adipocyte diameter determination of floating mature adipocytes
  - CRISPR-Cas9-mediated knockout of adipocyte marker genes
- QUANTIFICATION AND STATISTICAL DETAILS
  - LipocyteProfiling
  - LipocyteProfiler feature reduction using ARACNE
  - Generating of average cells
  - RNA-seq
  - Gene expression and LipocyteProfiler feature network
  - Quality control of genotyping data
  - Constructing PRSs

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.xgen.2023.100346.

ACKNOWLEDGMENTS

This work was supported by the Novo Nordisk Foundation (NNF21SA0072102); FNIH AMP-T2D RFB8b; NIDDK UM1 DK126185; NIDDK DK102173; NIDDK K24 DK110550; K23DK114551; NIDDK U01 DK105554; NIH OD-026839-01; a Novo Nordisk Postdoctoral Fellowship run in partnership with the University of Oxford; American Diabetes Association Innovative and Clinical Translational Award 1-19-ICTS-068; NHGRI FAIN# U01HG011723; the Clinical Cooperation Group "Nutrigenomics and Type 2 Diabetes" from the Helmholtz Center Munich and the German Center for Diabetes Research; the Else Kroner-Fresenius Foundation, Bad Homburg, Germany; the Li Ka Shing Foundation; NIHR Oxford Biomedical Research Centre, Oxford; NIH (1P50HD104224-01); Gates Foundation (INV-024200); Wellcome Trust Investigator award 221782/Z/20/Z; and the Doris Duke Charitable Foundation grant 2020096. M.C. and S.B.R.J. received a Next Generation Award at the Broad Institute. The hWAT and hBAT cell lines were kindly shared by Yu-Hua Tseng at the Joslin Diabetes Center, Harvard Medical School, and the SGBS line was shared by Martin Wabitsch, University of Ulm, Germany. Q14

AUTHOR CONTRIBUTIONS

Conceptualization, S.L., S.S., J.M.M., A.A., J.C.F., S.B.R.J., and M.C.; methodology, S.L., S.S., M.K.-A., D.R.S., G.P.W., B.A.C., A.E.C., and M.C.; formal analysis, S.L., S.S., J.M.M., K.F., G.D., D.R.S., G.P.W., B.A.C., M.S.U., F.R.C.d.S., D.R.S., and M.J.; investigation, S.L., S.S., A.A., J.H., G.G., E.M., A. Saadat, A.H., A. Stefek, S.H., V.C.-N., P.K., F.R.C.d.S., and D.R.S.; resources, H.H., M.C., and J.C.F.; data curation, H.D. and S.R.-P.; visualization, S.L. and S.S.; supervision, B.A.C., H.H., M.S.U., J.C.F., A.E.C., C.L., S.B.R.J., and M.C.; funding acquisition, H.H., J.C.F., A.E.C., C.L., S.B.R.J., and M.C.; writing – original draft, S.L., S.S., H.D., and M.C.; writing – review & editing, J.M.M., H.D., A.A., A.E.C., S.B.R.J., J.C.F., M.S.U., C.L., and H.H.

DECLARATION OF INTERESTS

J.C.F. has received consulting honoraria from Goldfinch Bio and Astra Zeneca and speaking honoraria from Novo Nordisk, Astra Zeneca, and Merck for research presentations over which he had full control of content. M.C. holds equity in Waypoint Bio, serves as a consultant for Pfizer, and is a member of the Nestle Scientific Advisory Board. The authors have filed a provisional patent application (63/218,656).

INCLUSION AND DIVERSITY Q8

We support inclusive, diverse, and equitable conduct of research.

Received: October 21, 2021
Revised: August 22, 2022
Accepted: May 26, 2023
Published: June 20, 2023

REFERENCES

1. Feng, Y., Mitchison, T.J., Bender, A., Young, D.W., and Tallarico, J.A. (2009). Multi-parameter phenotypic profiling: using cellular effects to characterize small-molecule compounds. Nat. Rev. Drug Discov. *8*, 567–578. https://doi.org/10.1038/nrd2876.
2. Green, R.A., Kao, H.-L., Audhya, A., Arur, S., Mayers, J.R., Fridolfsson, H.N., Schulman, M., Schloissnig, S., Niessen, S., Laband, K., et al. (2011). A high-resolution C. elegans essential gene network based on phenotypic profiling of a complex tissue. Cell *145*, 470–482. https://doi.org/10.1016/j.cell.2011.03.037.
3. Hughes, R.E., Elliott, R.J.R., Dawson, J.C., and Carragher, N.O. (2021). High-content phenotypic and pathway profiling to advance drug discovery in diseases of unmet need. Cell Chem. Biol. *28*, 338–355. https://doi.org/10.1016/j.chembiol.2021.02.015.

4. Neumann, B., Walter, T., Hériché, J.K., Bulkescher, J., Erfle, H., Conrad, C., Rogers, P., Poser, I., Held, M., Liebel, U., et al. (2010). Phenotypic profiling of the human genome by time-lapse microscopy reveals cell division genes. Nature *464*, 721–727. https://doi.org/10.1038/nature08869.

5. Bray, M.-A., Singh, S., Han, H., Davis, C.T., Borgeson, B., Hartland, C., Kost-Alimova, M., Gustafsdottir, S.M., Gibson, C.C., and Carpenter, A.E. (2016). Cell Painting, a high-content image-based assay for morphological profiling using multiplexed fluorescent dyes. Nat. Protoc. *11*, 1757–1774. https://doi.org/10.1038/nprot.2016.105.

6. Kepiro, M., Varkuti, B.H., and Davis, R.L. (2018). High content, phenotypic assays and screens for compounds modulating cellular processes in primary neurons. Methods Enzymol. *610*, 219–250. https://doi.org/10.1016/bs.mie.2018.09.021.

7. Scheeder, C., Heigwer, F., and Boutros, M. (2018). Machine learning and image-based profiling in drug discovery. Curr. Opin. Struct. Biol. *10*, 43–52. https://doi.org/10.1016/j.coisb.2018.05.004.

8. Simm, J., Klambauer, G., Arany, A., Steijaert, M., Wegner, J.K., Gustin, E., Chupakhin, V., Chong, Y.T., Vialard, J., Buijnsters, P., et al. (2018). Repurposing high-throughput image assays enables biological activity prediction for drug discovery. Cell Chem. Biol. *25*, 611–618.e3. https://doi.org/10.1016/j.chembiol.2018.01.015.

9. Wawer, M.J., Li, K., Gustafsdottir, S.M., Ljosa, V., Bodycombe, N.E., Marton, M.A., Sokolnicki, K.L., Bray, M.-A., Kemp, M.M., Winchester, E., et al. (2014). Toward performance-diverse small-molecule libraries for cell-based phenotypic screening using multiplexed high-dimensional profiling. Proc. Natl. Acad. Sci. USA *111*, 10911–10916. https://doi.org/10.1073/pnas.1410933111.

10. Rohban, M.H., Singh, S., Wu, X., Berthet, J.B., Bray, M.-A., Shrestha, Y., Varelas, X., Boehm, J.S., and Carpenter, A.E. (2017). Systematic morphological profiling of human gene and allele function via Cell Painting. Elife *6*, e24060. https://doi.org/10.7554/eLife.24060.

11. Olzmann, J.A., and Carvalho, P. (2019). Dynamics and functions of lipid droplets. Nat. Rev. Mol. Cell Biol. *20*, 137–155. https://doi.org/10.1038/s41580-018-0085-z.

12. Grandl, M., and Schmitz, G. (2010). Fluorescent high-content imaging allows the discrimination and quantitation of E-LDL-induced lipid droplets and Ox-LDL-generated phospholipidosis in human macrophages. Cytometry A. *77*, 231–242.

13. Liu, L., Zhang, K., Sandoval, H., Yamamoto, S., Jaiswal, M., Sanz, E., Li, Z., Hui, J., Graham, B.H., Quintana, A., and Bellen, H.J. (2015). Glial lipid droplets and ROS induced by mitochondrial defects promote neurodegeneration. Cell *160*, 177–190. https://doi.org/10.1016/j.cell.2014.12.019.

14. Robichaud, S., Fairman, G., Vijithakumar, V., Mak, E., Cook, D.P., Pelletier, A.R., Huard, S., Vanderhyden, B.C., Figeys, D., Lavallée-Adam, M., et al. (2021). Identification of novel lipid droplet factors that regulate lipophagy and cholesterol efflux in macrophage foam cells. Autophagy *17*, 3671–3689. https://doi.org/10.1080/15548627.2021.1886839.

15. Wang, H., Quiroga, A.D., and Lehner, R. (2013a). Analysis of lipid droplets in hepatocytes. Methods Cell Biol. *116*, 107–127. https://doi.org/10.1016/b978-0-12-408051-5.00007-3.

16. Cruz, A.L.S., Barreto, E.d.A., Fazolini, N.P.B., Viola, J.P.B., and Bozza, P.T. (2020). Lipid droplets: platforms with multiple functions in cancer hallmarks. Cell Death Dis. *11*, 105. https://doi.org/10.1038/s41419-020-2297-3.

17. Hershey, B.J., Vazzana, R., Joppi, D.L., and Havas, K.M. (2019). Lipid droplets define a sub-population of breast cancer stem cells. J. Clin. Med. *9*, 87. https://doi.org/10.3390/jcm9010087.

18. Prats, C., Donsmark, M., Qvortrup, K., Londos, C., Sztalryd, C., Holm, C., Galbo, H., and Ploug, T. (2006). Decrease in intramuscular lipid droplets and translocation of HSL in response to muscle contraction and epinephrine. J. Lipid Res. *47*, 2392–2399. https://doi.org/10.1194/jlr.M600247-JLR200.

19. Wang, Z., Jiang, T., Li, J., Proctor, G., McManaman, J.L., Lucia, S., Chua, S., and Levi, M. (2005). Regulation of renal lipid metabolism, lipid accumulation, and glomerulosclerosis in FVBdb/db mice with type 2 diabetes. Diabetes *54*, 2328–2335. https://doi.org/10.2337/diabetes.54.8.2328.

20. Weinert, S., Poitz, D.M., Auffermann-Gretzinger, S., Eger, L., Herold, J., Medunjanin, S., Schmeisser, A., Strasser, R.H., and Braun-Dullaeus, R.C. (2013). The lysosomal transfer of LDL/cholesterol from macrophages into vascular smooth muscle cells induces their phenotypic alteration. Cardiovasc. Res. *97*, 544–552. https://doi.org/10.1093/cvr/cvs367.

21. Xu, S., Zhang, X., and Liu, P. (2018). Lipid droplet proteins and metabolic diseases. Biochim. Biophys. Acta, Mol. Basis Dis. *1864*, 1968–1983. https://doi.org/10.1016/j.bbadis.2017.07.019.

22. Bray, M.-A., and Carpenter, A. (2017). Advanced assay development guidelines for image-based high content screening and analysis. In Assay Guidance Manual, S. Markossian, G.S. Sittampalam, A. Grossman, K. Brimacombe, M. Arkin, D. Auld, C.P. Austin, J. Baell, J.M.M. Caaveiro, and T.D.Y. Chung, et al., eds. (Eli Lilly & Company and the National Center for Advancing Translational Sciences).

23. Caicedo, J.C., Cooper, S., Heigwer, F., Warchal, S., Qiu, P., Molnar, C., Vasilevich, A.S., Barry, J.D., Bansal, H.S., Kraus, O., et al. (2017). Data-analysis strategies for image-based cell profiling. Nat. Methods *14*, 849–863. https://doi.org/10.1038/nmeth.4397.

24. Gustafsdottir, S.M., Ljosa, V., Sokolnicki, K.L., Anthony Wilson, J., Walpita, D., Kemp, M.M., Petri Seiler, K., Carrel, H.A., Golub, T.R., Schreiber, S.L., et al. (2013). Multiplex cytological profiling assay to measure diverse cellular states. PLoS One *8*, e80999. https://doi.org/10.1371/journal.pone.0080999.

25. Butte, A.J., and Kohane, I.S. (1999). Mutual information relevance networks: functional genomic clustering using pairwise entropy measurements. Pac. Symp. Biocomput., 418–429. https://doi.org/10.1142/9789814447331_0040.

26. Lachmann, A., Giorgi, F.M., Lopez, G., and Califano, A. (2016). ARACNe-AP: gene network reverse engineering through adaptive partitioning inference of mutual information. Bioinformatics *32*, 2233–2235. https://doi.org/10.1093/bioinformatics/btw216.

27. Li, Y., Liang, C., Wong, K.-C., Jin, K., and Zhang, Z. (2014). Inferring probabilistic miRNA-mRNA interaction signatures in cancers: a role-switch approach. Nucleic Acids Res. *42*, e76. https://doi.org/10.1093/nar/gku182.

28. Tachmazidou, I., Süveges, D., Min, J.L., Ritchie, G.R.S., Steinberg, J., Walter, K., Iotchkova, V., Schwartzentruber, J., Huang, J., Memari, Y., et al. (2017). Whole-genome sequencing coupled to imputation discovers genetic signals for anthropometric traits. Am. J. Hum. Genet. *100*, 865–884. https://doi.org/10.1016/j.ajhg.2017.04.014.

29. Xue, R., Lynes, M.D., Dreyfuss, J.M., Shamsi, F., Schulz, T.J., Zhang, H., Huang, T.L., Townsend, K.L., Li, Y., Takahashi, H., et al. (2015). Clonal analyses and gene profiling identify genetic biomarkers of the thermogenic potential of human brown and white preadipocytes. Nat. Med. *21*, 760–768. https://doi.org/10.1038/nm.3881.

30. Kanzaki, M. (2006). Insulin receptor signals regulating GLUT4 translocation and actin dynamics. Endocr. J. *53*, 267–293. https://doi.org/10.1507/endocrj.kr-65.

31. Cedikova, M., Kripnerová, M., Dvorakova, J., Pitule, P., Grundmanova, M., Babuska, V., Mullerova, D., and Kuncova, J. (2016). Mitochondria in white, Brown, and beige adipocytes. Stem Cell. Int. *2016*, 6067349. https://doi.org/10.1155/2016/6067349.

32. Fei, W., Du, X., and Yang, H. (2011). Seipin, adipogenesis and lipid droplets. Trends Endocrinol. Metabol. *22*, 204–210. https://doi.org/10.1016/j.tem.2011.02.004.

33. Gandotra, S., Le Dour, C., Bottomley, W., Cervera, P., Giral, P., Reznik, Y., Charpentier, G., Auclair, M., Delépine, M., Barroso, I., et al. (2011). Perilipin deficiency and autosomal dominant partial lipodystrophy. N. Engl. J. Med. *364*, 740–748. https://doi.org/10.1056/NEJMoa1007487.

34. Shijun, L., Khan, R., Raza, S.H.A., Jieyun, H., Chugang, M., Kaster, N., Gong, C., Chunping, Z., Schreurs, N.M., and Linsen, Z. (2020). Function and characterization of the promoter region of perilipin 1 (PLIN1): roles of E2F1, PLAG1, C/EBPβ, and SMAD3 in bovine adipocytes. Genomics *112*, 2400–2409. https://doi.org/10.1016/j.ygeno.2020.01.012.

35. Brasaemle, D.L., Barber, T., Wolins, N.E., Serrero, G., Blanchette-Mackie, E.J., and Londos, C. (1997). Adipose differentiation-related protein is an ubiquitously expressed lipid storage droplet-associated protein. J. Lipid Res. *38*, 2249–2263.

36. Tsai, T.-H., Chen, E., Li, L., Saha, P., Lee, H.-J., Huang, L.-S., Shelness, G.S., Chan, L., and Chang, B.H.-J. (2017). The constitutive lipid droplet protein PLIN2 regulates autophagy in liver. Autophagy *13*, 1130–1144. https://doi.org/10.1080/15548627.2017.1319544.

37. Gao, A.W., and Houtkooper, R.H. (2014). Mitochondrial fission: firing up mitochondria in brown adipose tissue. EMBO J. *33*, 401–402. https://doi.org/10.1002/embj.201487798.

Q15

38. Qiu, P., Gentles, A.J., and Plevritis, S.K. (2011). Discovering biological progression underlying microarray samples. PLoS Comput. Biol. *7*, e1001123. https://doi.org/10.1371/journal.pcbi.1001123.

39. Baglioni, S., Cantini, G., Poli, G., Francalanci, M., Squecco, R., Di Franco, A., Borgogni, E., Frontera, S., Nesi, G., Liotta, F., et al. (2012). Functional differences in visceral and subcutaneous fat pads originate from differences in the adipose stem cell. PLoS One *7*, e36569. https://doi.org/10.1371/journal.pone.0036569.

40. Wang, Q.A., Tao, C., Gupta, R.K., and Scherer, P.E. (2013b). Tracking adipogenesis during white adipose tissue development, expansion and regeneration. Nat. Med. *19*, 1338–1344. https://doi.org/10.1038/nm.3324.

41. Kusminski, C.M., and Scherer, P.E. (2012). Mitochondrial dysfunction in white adipose tissue. Trends Endocrinol. Metabol. *23*, 435–443. https://doi.org/10.1016/j.tem.2012.06.004.

42. Miller, C.N., Yang, J.-Y., England, E., Yin, A., Baile, C.A., and Rayalam, S. (2015). Isoproterenol increases uncoupling, glycolysis, and markers of beiging in mature 3T3-L1 adipocytes. PLoS One *10*, e0138344. https://doi.org/10.1371/journal.pone.0138344.

43. Hoffstedt, J., Arner, P., Hellers, G., and Lönnqvist, F. (1997). Variation in adrenergic regulation of lipolysis between omental and subcutaneous adipocytes from obese and non-obese men. J. Lipid Res. *38*, 795–804.

44. Morigny, P., Boucher, J., Arner, P., and Langin, D. (2021). Lipid and glucose metabolism in white adipocytes: pathways, dysfunction and therapeutics. Nat. Rev. Endocrinol. *17*, 276–295. https://doi.org/10.1038/s41574-021-00471-8.

45. Liu, F., Wang, C., Zhang, L., Xu, Y., Jang, L., Gu, Y., Cao, X., Zhao, X., Ye, J., and Li, Q. (2014). Metformin prevents hepatic steatosis by regulating the expression of adipose differentiation-related protein. Int. J. Mol. Med. *33*, 51–58. https://doi.org/10.3892/ijmm.2013.1560.

46. Demine, S., Renard, P., and Arnould, T. (2019). Mitochondrial uncoupling: a key controller of biological processes in physiology and diseases. Cells *8*. https://doi.org/10.3390/cells8080795.

47. Forkink, M., Manjeri, G.R., Liemburg-Apers, D.C., Nibbeling, E., Blanchard, M., Wojtala, A., Smeitink, J.A.M., Wieckowski, M.R., Willems, P.H.G.M., and Koopman, W.J.H. (2014). Mitochondrial hyperpolarization during chronic complex I inhibition is sustained by low activity of complex II, III, IV and V. Biochim. Biophys. Acta *1837*, 1247–1256. https://doi.org/10.1016/j.bbabio.2014.04.008.

48. Fischer-Posovszky, P., Newell, F.S., Wabitsch, M., and Tornqvist, H.E. (2008). Human SGBS cells - a unique tool for studies of human fat cell biology. Obes. Facts *1*, 184–189. https://doi.org/10.1159/000145784.

49. Venables, W.N., and Ripley, B.D. (2002). Random and mixed effects. Modern Applied Statistics with S, 271–300. Springer.

50. Kilpinen, H., Goncalves, A., Leha, A., Afzal, V., Alasoo, K., Ashford, S., Bala, S., Bensaddek, D., Casale, F.P., Culley, O.J., et al. (2017). Common genetic variation drives molecular heterogeneity in human iPSCs. Nature *546*, 370–375. https://doi.org/10.1038/nature22403.

51. Mahajan, A., Taliun, D., Thurner, M., Robertson, N.R., Torres, J.M., Rayner, N.W., Payne, A.J., Steinthorsdottir, V., Scott, R.A., Grarup, N., et al. (2018). Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. Nat. Genet. *50*, 1505–1513. https://doi.org/10.1038/s41588-018-0241-6.

52. Dupuis, J., Langenberg, C., Prokopenko, I., Saxena, R., Soranzo, N., Jackson, A.U., Wheeler, E., Glazer, N.L., Bouatia-Naji, N., Gloyn, A.L., et al. (2010). New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. Nat. Genet. *42*, 105–116. https://doi.org/10.1038/ng.520.

53. Matthews, D.R., Hosker, J.P., Rudenski, A.S., Naylor, B.A., Treacher, D.F., and Turner, R.C. (1985). Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. Diabetologia *28*, 412–419. https://doi.org/10.1007/BF00280883.

54. Pulit, S.L., Stoneman, C., Morris, A.P., Wood, A.R., Glastonbury, C.A., Tyrrell, J., Yengo, L., Ferreira, T., Marouli, E., Ji, Y., et al. (2019). Meta-analysis of genome-wide association studies for body fat distribution in 694 649 individuals of European ancestry. Hum. Mol. Genet. *28*, 166–174. https://doi.org/10.1093/hmg/ddy327.

55. Mootha, V.K., Lindgren, C.M., Eriksson, K.-F., Subramanian, A., Sihag, S., Lehar, J., Puigserver, P., Carlsson, E., Ridderstråle, M., Laurila, E., et al. (2003). PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat. Genet. *34*, 267–273. https://doi.org/10.1038/ng1180.

56. Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. USA *102*, 15545–15550. https://doi.org/10.1073/pnas.0506580102.

57. Mayeuf-Louchart, A., Lancel, S., Sebti, Y., Pourcet, B., Loyens, A., Delhaye, S., Duhem, C., Beauchamp, J., Ferri, L., Thorel, Q., et al. (2019). Glycogen dynamics drives lipid droplet biogenesis during Brown adipocyte differentiation. Cell Rep. *29*, 1410–1418.e6. https://doi.org/10.1016/j.celrep.2019.09.073.

58. Lebovitz, H.E., and Banerji, M.A. (2005). Point: visceral adiposity is causally related to insulin resistance. Diabetes Care *28*, 2322–2325. https://doi.org/10.2337/diacare.28.9.2322.

59. Udler, M.S., Kim, J., von Grotthuss, M., Bonàs-Guarch, S., Cole, J.B., Chiou, J., Anderson, C.D. on behalf of METASTROKE and the ISGC; Boehnke, M., Laakso, M., Atzmon, G., et al. (2018). Type 2 diabetes genetic loci informed by multi-trait associations point to disease mechanisms and subtypes: a soft clustering analysis. PLoS Med. *15*, e1002654. https://doi.org/10.1371/journal.pmed.1002654.

60. Xiao, B., Deng, X., Zhou, W., and Tan, E.-K. (2016). Flow cytometry-based assessment of mitophagy using MitoTracker. Front. Cell. Neurosci. *10*, 76. https://doi.org/10.3389/fncel.2016.00076.

61. Sleigh, A., Stears, A., Thackray, K., Watson, L., Gambineri, A., Nag, S., Campi, V.I., Schoenmakers, N., Brage, S., Carpenter, T.A., et al. (2012). Mitochondrial oxidative phosphorylation is impaired in patients with congenital lipodystrophy. J. Clin. Endocrinol. Metab. *97*, E438–E442. https://doi.org/10.1210/jc.2011-2587.

62. Hu, L., He, F., Huang, M., Peng, M., Zhou, Z., Liu, F., and Dai, Y.-S. (2018). NFATc3 deficiency reduces the classical activation of adipose tissue macrophages. J. Mol. Endocrinol. *61*, 79–89. https://doi.org/10.1530/JME-18-0070.

63. Lotta, L.A., Gulati, P., Day, F.R., Payne, F., Ongen, H., van de Bunt, M., Gaulton, K.J., Eicher, J.D., Sharp, S.J., Luan, J., et al. (2017). Integrative genomic analysis implicates limited peripheral adipose storage capacity in the pathogenesis of human insulin resistance. Nat. Genet. *49*, 17–26. https://doi.org/10.1038/ng.3714.

64. Gagliano Taliun, S.A., VandeHaar, P., Boughton, A.P., Welch, R.P., Taliun, D., Schmidt, E.M., Zhou, W., Nielsen, J.B., Willer, C.J., Lee, S., et al.

(2020). Exploring and visualizing large-scale genetic associations by using PheWeb. Nat. Genet. *52*, 550–552.

65. Dickinson, M.E., Flenniken, A.M., Ji, X., Teboul, L., Wong, M.D., White, J.K., Meehan, T.F., Weninger, W.J., Westerberg, H., Adissu, H., et al. (2016). High-throughput discovery of novel developmental phenotypes. Nature *537*, 508–514. https://doi.org/10.1038/nature19356.
66. You, D., Nilsson, E., Tenen, D.E., Lyubetskaya, A., Lo, J.C., Jiang, R., Deng, J., Dawes, B.A., Vaag, A., Ling, C., et al. (2017). Dnmt3a is an epigenetic mediator of adipose insulin resistance. Elife *6*, e30766. https://doi.org/10.7554/eLife.30766.
67. Martin, S., Cule, M., Basty, N., Tyrrell, J., Beaumont, R.N., Wood, A.R., Frayling, T.M., Sorokin, E., Whitcher, B., Liu, Y., et al. (2021). Genetic evidence for different adiposity phenotypes and their opposing influences on ectopic fat and risk of cardiometabolic disease. Diabetes *70*, 1843–1856. https://doi.org/10.2337/db21-0129.
68. Pinello, L., Canver, M.C., Hoban, M.D., Orkin, S.H., Kohn, D.B., Bauer, D.E., and Yuan, G.-C. (2016). Analyzing CRISPR genome-editing experiments with CRISPResso. Nat. Biotechnol. *34*, 695–697. https://doi.org/10.1038/nbt.3583.
69. Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M., and Lee, J.J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience *4*, 7. https://doi.org/10.1186/s13742-015-0047-8.
70. Delaneau, O., Zagury, J.-F., and Marchini, J. (2013). Improved whole-chromosome phasing for disease and population genetic studies. Nat. Methods *10*, 5–6. https://doi.org/10.1038/nmeth.2307.
71. Ge, T., Chen, C.-Y., Ni, Y., Feng, Y.-C.A., and Smoller, J.W. (2019). Polygenic prediction via Bayesian regression and continuous shrinkage priors. Nat. Commun. *10*, 1776. https://doi.org/10.1038/s41467-019-09718-5.
72. Vilhjálmsson, B.J., Yang, J., Finucane, H.K., Gusev, A., Lindström, S., Ripke, S., Genovese, G., Loh, P.-R., Bhatia, G., Do, R., et al. (2015). Modeling linkage disequilibrium increases accuracy of polygenic risk scores. Am. J. Hum. Genet. *97*, 576–592. https://doi.org/10.1016/j.ajhg.2015.09.001.
73. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21. https://doi.org/10.1093/bioinformatics/bts635.
74. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550. https://doi.org/10.1186/s13059-014-0550-8.
75. Margolin, A.A., Nemenman, I., Basso, K., Wiggins, C., Stolovitzky, G., Dalla Favera, R., and Califano, A. (2006). ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. BMC Bioinf. *7*, S7. https://doi.org/10.1186/1471-2105-7-S1-S7.
76. McInnes, L., Healy, J., Saul, N., and Großberger, L. (2018). UMAP: uniform manifold approximation and projection. J. Open Source Softw. *3*, 861. https://doi.org/10.21105/joss.00861.
77. Gu, Z., Eils, R., and Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics *32*, 2847–2849. https://doi.org/10.1093/bioinformatics/btw313.
78. Akulenko, R., Merl, M., and Helms, V. (2016). BEclear: batch effect detection and adjustment in DNA methylation data. PLoS One *11*, e0159921. https://doi.org/10.1371/journal.pone.0159921.
79. Skurk, T., and Hauner, H. (2012). Primary culture of human adipocyte precursor cells: expansion and differentiation. Methods Mol. Biol. *806*, 215–226. https://doi.org/10.1007/978-1-61779-367-7_15.
80. Raajendiran, A., Ooi, G., Bayliss, J., O'Brien, P.E., Schittenhelm, R.B., Clark, A.K., Taylor, R.A., Rodeheffer, M.S., Burton, P.R., and Watt, M.J. (2019). Identification of metabolically distinct adipocyte progenitor cells in human adipose tissues. Cell Rep. *27*, 1528–1540.e7. https://doi.org/10.1016/j.celrep.2019.04.010.
81. Karlson, E.W., Boutin, N.T., Hoffnagle, A.G., and Allen, N.L. (2016). Building the Partners HealthCare Biobank at Partners personalized medicine: informed consent, return of research results, recruitment lessons and operational considerations. J. Personalized Med. *6*, 2. https://doi.org/10.3390/jpm6010002.
82. Yu, S., Liao, K.P., Shaw, S.Y., Gainer, V.S., Churchill, S.E., Szolovits, P., Murphy, S.N., Kohane, I.S., and Cai, T. (2015). Toward high-throughput phenotyping: unbiased automated feature extraction and selection from knowledge sources. J. Am. Med. Inf. Assoc. *22*, 993–1000. https://doi.org/10.1093/jamia/ocv034.
83. Fischer, B., Schöttl, T., Schempp, C., Fromme, T., Hauner, H., Klingenspor, M., and Skurk, T. (2015). Inverse relationship between body mass index and mitochondrial oxidative phosphorylation capacity in human subcutaneous adipocytes. Am. J. Physiol. Endocrinol. Metab. *309*, E380–E387. https://doi.org/10.1152/ajpendo.00524.2014.
84. Shalem, O., Sanjana, N.E., Hartenian, E., Shi, X., Scott, D.A., Mikkelson, T., Heckl, D., Ebert, B.L., Root, D.E., Doench, J.G., and Zhang, F. (2014). Genome-scale CRISPR-Cas9 knockout screening in human cells. Science *343*, 84–87. https://doi.org/10.1126/science.1247005.
85. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al. (2011). Scikit-learn: machine learning in Python. J. Mach. Learn. Res. *12*, 2825–2830.
86. Jang, I.S., Margolin, A., and Califano, A. (2013). hARACNe: improving the accuracy of regulatory model reverse engineering via higher-order data processing inequality tests. Interface Focus *3*, 20130011. https://doi.org/10.1098/rsfs.2013.0011.
87. Wang, X., Hripcsak, G., and Friedman, C. (2009). Characterizing environmental and phenotypic associations using information theory and electronic health records. BMC Bioinf. *10* (*Suppl 9*), S13. https://doi.org/10.1186/1471-2105-10-s9-s13.
88. Krueger, F. (2015). A Wrapper Tool Around Cutadapt and FastQC to Consistently Apply Quality and Adapter Trimming to FastQ Files (Babraham Institute).
89. Liao, Y., Smyth, G.K., and Shi, W. (2019). The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. Nucleic Acids Res. *47*, e47. https://doi.org/10.1093/nar/gkz114.
90. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., and Sham, P.C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. *81*, 559–575. https://doi.org/10.1086/519795.
91. McCarthy, S., Das, S., Kretzschmar, W., Delaneau, O., Wood, A.R., Teumer, A., Kang, H.M., Fuchsberger, C., Danecek, P., Sharp, K., et al. (2016). A reference panel of 64,976 haplotypes for genotype imputation. Nat. Genet. *48*, 1279–1283. https://doi.org/10.1038/ng.3643.

Please cite this article in press as: Laber et al., Discovering cellular programs of intrinsic and extrinsic drivers of metabolic traits using LipocyteProfiler, Cell Genomics (2023), https://doi.org/10.1016/j.xgen.2023.100346

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
Human adipose-derived mesenchymal stem cells	Munich Obesity BioBank (MOBB)	NA
Primary human hepatocytes	BioIVT	YNZ
Chemicals, peptides, and recombinant proteins		
MitoTracker™ Deep Red FM	Molecular Probes, Inc.	M22426
BODIPY™ 505/515	Molecular Probes, Inc.	D3921
Alexa Fluor™ 568 Phalloidin	Life Technologies Corp.	A12380
Hoechst 33342	Molecular Probes, Inc.	H3570
Wheat Germ Agglutinin, Alexa Fluor™ 555 Conjugate	Molecular Probes, Inc.	W32464
SYTO™ 14 Green Fluorescent Nucleic Acid Stain	Molecular Probes, Inc.	S7576
16% Paraformaldehyde, methanol-free	Electron Microscopy Sciences	15710-S
Hank's Balanced Salt Solution (1x), HBSS	Life Technologies Corp.	14025076
Triton X-100	Merck KGaA	X100
Phalloidin-Atto-565	Merck KGaA	94072
Critical commercial assays		
Infinium HTS assay + GSA Bead Chips	Illumina, Inc.	NA
Deposited data		
Raw data and code	GitHub	https://github.com/ClaussnitzerLab/Lipocyte-Profiler
Experimental models: Cell lines		
hWAT	Xue et al.[29]	https://doi.org/10.1038/nm.3881
hBAT	Xue et al.[29]	https://doi.org/10.1038/nm.3881
SGBS	Wabitsch et al. 2001	https://doi.org/10.1038/sj.ijo.0801520
Cas9 expressing hWAT	This paper	NA
Oligonucleotides		
guide sequences targeting: *PPARG*: ATACACAGGTGCAATCAAAG and CAACTTTGGGATCAGCTCCG; *PPARGC1A* TATTGAACGCACCTTAAGTG and AGTCCTCACTGGTGGACACG; *MFN1*: CACCAGGTCATCTCTCAAGA and TTATATGGCCAATCCCACTA; *PLIN1*: TCACGGCAGATACTTACCAG and TCTGCACGGTGTATCGAGAG; *INSR*: TTATCGGCGATATGGTGATG and AGTGAGTATGAGGATTCGGC; *IRS1* CCCAGGACCCGCATTCAAAG and CCGAAGCACTAGATCGCCGT	This paper	NA
non-targeted controls (control guide sequences): ATCAGGCCTTGTCCGTGATT; TACGTCATTAAGAGTTCAAC; GACAGTGAAATTAGCTCCCA; GATTCATACTAAACACTCTAx; CCTAGTTCATAAGCTACGCC	This paper	NA
Software and algorithms		
LipocyteProfiler	This paper	https://github.com/ClaussnitzerLab/Lipocyte-Profiler

(*Continued on next page*)

Please cite this article in press as: Laber et al., Discovering cellular programs of intrinsic and extrinsic drivers of metabolic traits using LipocyteProfiler, Cell Genomics (2023), https://doi.org/10.1016/j.xgen.2023.100346

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
CRISPResso	Pinello et al.[68]	https://doi.org/10.1038/nbt.3583
PLINK	Purcell et al.[69] and Chang et al. 2015	https://doi.org/10.1086/519795 and https://doi.org/10.1186/s13742-015-0047-8
SHAPEIT2	Delaneau et al.[70]	https://doi.org/10.1038/nmeth.2307
PRS-CS	Ge et al.[71]	https://doi.org/10.1038/s41467-019-09718-5
LDpred	Vilhjálmsson et al.[72]	https://doi.org/10.1016/j.ajhg.2015.09.001
igraph	Csardi et al. 2006	https://igraph.org
Enrichr	Chen et al. 2013	https://maayanlab.cloud/Enrichr/
FastQC	Babraham Bioinformatics	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/
STAR	Dobin et al.[73]	https://doi.org/10.1093/bioinformatics/bts635
DESeq2	Love et al.[74]	https://doi.org/10.1186/s13059-014-0550-8
CellProfiler 3.1.9	Carpenter et al. 2006	https://doi.org/10.1186/gb-2006-7-10-r100
Harmony 4.9	PerkinElmer Inc.	HH17000010
ARACNE	Margolin et al.[75]	https://doi.org/10.1186/1471-2105-7-S1-S7
UMAP R package 0.2.7.0	McInnes et al.[76]	https://doi.org/10.21105/joss.00861
ComplexHeatmap Bioconductor package 2.7.7	Gu et al.[77]	https://doi.org/10.1093/bioinformatics/btw313
Sample Progression Discovery (SPD)	Qui et al. 2011	https://doi.org/10.1371/journal.pcbi.1001123.g001
BEclear	Akulenko et al.[78]	https://doi.org/10.1371/journal.pone.0159921
R 3.6.1	The R Foundation for Statistical Computing	NA
Other		
CellCarrier Ultra 96 well plate, black (now: PhenoPlate™ 96-well)	PerkinElmer Inc.	#6005550
CellCarrier Ultra 96 well plate, black (now: PhenoPlate™ 96-well); collagen-coated	PerkinElmer Inc.	#6055700
Opera Phenix® High-Content Screening System	PerkinElmer Inc.	NA
Leica DMi8 microscope with HC PL APO ×63/1.40 oil objective	Leica Microsystems GmbH	NA
MGB Biobank data	Partners HealthCare hospitals	NA

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Melina Claussnitzer (melina@broadinstitute.com).

Materials availability

This study did not generate new unique reagents.

Data and code availability

The codes are publicly available on GitHub https://github.com/ClaussnitzerLab/Lipocyte-Profiler and https://zenodo.org/record/7341916#.ZCtCjuzMKJ9. The high content imaging data are available at the Cell Painting Gallery on the Registry of Open Data on AWS (https://registry.opendata.aws/cellpainting-gallery/) under accession number cpg0011. The transcriptomics data are available on the GEO (Accession number GSE184089).

Please cite this article in press as: Laber et al., Discovering cellular programs of intrinsic and extrinsic drivers of metabolic traits using LipocyteProfiler, Cell Genomics (2023), https://doi.org/10.1016/j.xgen.2023.100346

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Human primary AMSC isolation/abdominal laparoscopy cohort–Munich obesity BioBank/MOBB

We obtained AMSCs from subcutaneous and visceral adipose tissue from patients undergoing a range of abdominal laparoscopic surgeries (sleeve gastrectomy, fundoplication or appendectomy). The visceral adipose tissue is derived from the proximity of the angle of His and subcutaneous adipose tissue obtained from beneath the skin at the site of surgical incision. Additionally, human liposuction material was obtained. Each participant gave written informed consent before inclusion and the study protocol was approved by the ethics committee of the Technical University of Munich (Study No 5716/13). Isolation of AMSCs was performed as previously described in.[79] For a subset of donors, purity of AMSCs was assessed as previously described in.[80] Briefly, cells were stained with 0.05μg CD34, 0.125μg CD29, 0.375μg CD31, 0.125μg CD45 per 250K cells and analyzed on CytoFlex together with negative control samples of corresponding AMSCs.

Differentiation of human AMSCs

For imaging, cells were seeded at 10K cells/well in 96-well plates (Cell Carrier, Perkin Elmer #6005550) and induced 4 days after seeding. For RNAseq, cells were seeded at 40K cells/well in 12-well dishes (Corning). Before Induction cells were cultured in proliferation medium (Basic medium consisting of DMEM-F12 1% Penicillin - Streptomycin, 33μM Biotin and 17μM Pantothenate supplemented with 0.13μM Insulin, 0.01μg/ml EGF, 0.001μg/ml FGF, 2.5%FCS). Adipogenic differentiation was induced by changing culture medium to induction medium. (Basic medium supplemented with 0.861μM Insulin, 1nM T3, 0.1μM Cortisol, 0.01 mg/ml Transferrin, 1μM Rosiglitazone, 25nM Dexamethasone, 2.5nM IBMX). On day 3 of adipogenic differentiation culture medium was changed to differentiation medium (Basic medium supplemented with 0.861μM Insulin, 1nM T3, 0.1μM Cortisol, 0.01 mg/ml Transferrin). Medium was changed every 3 days. Visceral-derived AMSCs were differentiated by further adding 2% FBS as well as 0.1mM oleic and linoleic acid to the induction and differentiation media. For isoproterenol stimulation experiments, 1uM isoproterenol was added to the differentiation media and cells treated overnight.

Primary human hepatocyte culture

Primary human hepatocytes (PHH) were purchased from BioIVT. Donor lot YNZ was used in this study. PHH were thawed and immediately resuspended in CP media (BioIVT) supplemented with torpedo antibiotic (BioIVT). Cell count and viability were assessed by trypan blue exclusion test prior to plating. Hepatocytes were plated onto collagen-coated Cellcarrier-96 Ultra Microplates (Perkin Elmer) at a density of 50,000 cells per well in CP media supplemented. Four hours after plating, media was replaced with fresh CP media. After 24 h, media was replaced with fresh CP media or CP media containing oleic acid (0.3mM) or metformin (5mM). Hepatocytes were incubated for an additional 24 h prior to processing.

MGB Biobank cohort

The MGB Biobank[81] maintains blood and DNA samples from more than 60,000 consented patients seen at Partners HealthCare hospitals, including Massachusetts General Hospital, Brigham and Women's Hospital, McLean Hospital, and Spaulding Rehabilitation Hospital, all in the USA. Patients are recruited in the context of clinical care appointments at more than 40 sites, clinics, and electronically through the patient portal at Partners HealthCare. Biobank subjects provide consent for the use of their samples and data in broad-based research. The Partners Biobank works closely with the Partners Research Patient Data Registry (RPDR), the Partners' enterprise scale data repository designed to foster investigator access to a wide variety of phenotypic data on more than 4 million Partners HealthCare patients. Approval for analysis of Biobank data was obtained by Partners IRB, study 2016P001018.

Type 2 diabetes status was defined based on ''curated phenotypes'' developed by the Biobank Portal team using both structured and unstructured electronic medical record (EMR) data and clinical, computational and statistical methods. Natural Language Processing (NLP) was used to extract data from narrative text. Chart reviews by disease experts helped identify features and variables associated with particular phenotypes and were also used to validate results of the algorithms. The process produced robust phenotype algorithms that were evaluated using metrics such as sensitivity, the proportion of true positives correctly identified as such, and positive predictive value (PPV), the proportion of individuals classified as cases by the algorithm.[82]

a. Control selection criteria.
   1. Individuals determined by the ''curated disease'' algorithm employed above to have no history of type 2 diabetes with NPV of 99%.
   2. Individuals at least age 55.
   3. Individuals with HbA1c less than 5.7
b. Case selection criteria.
   1. Individuals determined by the ''curated disease'' algorithm employed above to have type 2 diabetes with PPV of 99%
   2. Individuals at least age 30 given the higher rate of false positive diagnoses in younger individuals.

Please cite this article in press as: Laber et al., Discovering cellular programs of intrinsic and extrinsic drivers of metabolic traits using LipocyteProfiler, Cell Genomics (2023), https://doi.org/10.1016/j.xgen.2023.100346

Genomic data for 30,240 participants was generated with the Illumina Multi-Ethnic Genotyping Array, which covers more than 1.7 million markers, including content from over 36,000 individuals, and is enriched for exome content with >400,000 markers missense, nonsense, indels, and synonymous variants.

METHOD DETAILS

LipocytePainting

Human primary AMSCs and PHH were plated in 96-well CellCarrier Black plates (PerkinElmer #6005550). AMSCs were differentiated for 14 days, and high content imaging was performed at day 0, day 3, day 8 and day 14 of adipogenic differentiation in replicates of 4 per donor/time point/depot (inter-replicate variance of 0.075). Primary human hepatocytes were stained after 48 h in culture, and 24h following treatment with oleic acid or metformin. On the respective day of the assay, cell culture media was removed and replaced by 0.5uM Mitotracker staining solution (1mM MitoTracker Deep Red stock (Invitrogen #M22426) diluted in culture media) to each well followed by 30 min incubation at 37°C protected from light. After 30min Mitotracker staining solution was removed and cells were washed twice with Dulbecco's Phosphate-Buffered Saline (1X), DPBS (Corning #21-030-CV) and 2.9uM BODIPY staining solution (3.8mM BODIPY 505/515 stock (Thermofisher #D3921) diluted in DPBS) was added followed by 15 min incubation at 37°C protected from light. Subsequently, cells were fixed by adding 16% Methanol-free Paraformaldehyde, PFA (Electron Microscopy Sciences #15710-S) directly to the BODIPY staining solution to a final concentration of 3.2% and incubated for 20 min at RT protected from light. PFA was removed and cells were washed once with Hank's Balanced Salt Solution (1x), HBSS (Gibco #14025076). To permeabilize cells 0.1% Triton X-100 (Sigma Aldrich #X100) was added and incubated at RT for 10 min protected from light. After Permeabilization multi-stain solution (10 units of Alexa Fluor 568 Phalloidin (ThermoFisher #A12380), 0.01 mg/ml Hoechst 33342 (Invitrogen #H3570), 0.0015 mg/ml Wheat Germ Agglutinin, Alexa Fluor 555 Conjugate (ThermoFisher #W32464), 3uM SYTO 14 Green Fluorescent Nucleic Acid Stain (Invitrogen #S7576) diluted in HBSS) was added and cells were incubated at RT for 10 min protected from light. Finally, staining solution was removed and cells were washed three times with HBSS. Cells were imaged using a Opera Phenix High content screening system using confocal, 20× objective. Per well we imaged 25 fields.

Staining and microscopy of actin-cytoskeleton in subcutaneous AMSCs

To stain the actin cytoskeleton, and nuclei, cells were washed twice with ice-cold PBS and fixed with paraformaldehyde Roti-Histofix 4% (Roth, Karlsruhe, Germany) for 15 min. Cells were washed twice with ice-cold PBS for 5 min and incubated with ice-cold 0.1% Triton X-/PBS (Roth, Karlsruhe, Germany) for 5 min. Cells were washed twice with PBS and stained with 0.46% Bisbenzimide H 33258 (Sigma-Aldrich, Steinheim, Germany), and 1% Phalloidin-Atto-565 (Sigma-Aldrich, Steinheim, Germany). Cells were incubated for 1 h at RT in the dark. Afterward, cells were washed twice with PBS for 5 min and kept in PBS at 4°C until imaging. Images were acquired on a Leica DMi8 microscope using the HC PL APO ×63/1.40 oil objective. Images were processed using the Leica LasX software.

Isolation and adipocyte diameter determination of floating mature adipocytes

Mature adipocyte isolation was carried out as described earlier.[83] Immediately after isolation, approximately 50 μL of the adipocyte suspension was pipetted onto a glass slide and the diameter of 100 cells was manually determined under a light microscope.

CRISPR-Cas9-mediated knockout of adipocyte marker genes

We generated a hWAT cell-line stably expressing Cas9 as previously described in Shalem et al.[84] We validated the generated line by assessing Cas9 activity (90%) and adipocyte differentiation capacity using adipocyte marker gene expression and morphological profiling. CRISPR-Cas9 mediated knockdown of *PPARG*, *PPARGC1AA*, *MFN1*, *PLIN1*, *INSR*, and *IRS1* was performed in pre-adipocytes (5 days before differentiation) using three replicates per guide and two guides per gene (guide sequences targeting *PPARG*: ATACACAGGTGCAATCAAAG and CAACTTTGGGATCAGCTCCG; *PPARGC1A* TATTGAACGCACCTTAAGTG and AGTCCTCACT GGTGGACACG; *MFN1*: CACCAGGTCATCTCTCAAGA and TTATATGGCCAATCCCACTA; *PLIN1*: TCACGGCAGATACTTACC AG and TCTGCACGGTGTATCGAGAG; *INSR*: TTATCGGCGATATGGTGATG and AGTGAGTATGAGGATTCGGC; *IRS1* CCCAGGA CCCGCATTCAAAG and CCGAAGCACTAGATCGCCGT) as well as five non-targeted controls (control guide sequences: ATCAG GCCTTGTCCGTGATT, TACGTCATTAAGAGTTCAAC, GACAGTGAAATTAGCTCCCA, GATTCATACTAAACACTCTA, CCTAGTTATA AGCTACGCC) in an 96-well arrayed format. Guide on-target efficiency was assessed using Next-generation sequencing followed by CRISPResso analysis.[68] AMSCs were stained using LipocytePainting (see above) on day 14 of differentiation. After feature extraction and QC steps (see also LipocyteProfiling), we removed samples where guide cutting efficiency was <10% or where discrepancy between the two guides was equal or above 10%. For visualizations we used one non-targeted control that showed lowest standard deviation of replicates and was closest to the median of all five non-targeted controls across all LipocyteProfiler features.

QUANTIFICATION AND STATISTICAL DETAILS

LipocyteProfiling

Quantitation was performed using CellProfiler 3.1.9. Prior to processing, flat field illumination correction was performed using functions generated from the median intensity across each plate. Nuclei were identified using the DAPI stain and then expanded to identify

Please cite this article in press as: Laber et al., Discovering cellular programs of intrinsic and extrinsic drivers of metabolic traits using LipocyteProfiler, Cell Genomics (2023), https://doi.org/10.1016/j.xgen.2023.100346

whole cells using the Phalloidin/WGA and BODIPY stains. Regions of cytoplasm were then determined by removing the Nuclei from the Cell segmentations. Speckles of BODIPY staining were enhanced to assist in detection of small and large individual Lipid objects. For each object set measurements were collected representing size, shape, intensity, granularity, texture, colocalization and distance to neighboring objects. After LipocyteProfiler (LP) feature extraction data was filtered by applying automated and manual quality control steps. First, fields with a total cell count less than 50 cells were removed. Second, every field was assessed visually and fields that were corrupted by experimental induced technical artifacts were removed. Furthermore, blocklisted features(Way, 2020), LP-features measurement category *Manders*, *RWC and Costes*, that are known to be noisy and generally unreliable were removed. Additionally, LP-features named *SmallLipidObjetcs*, that measure small objects stained by SYTO14 rather than lipid informative objects, were also removed. After filtering data were normalized per plate using a robust scaling approach[85] that subtracts the median from each variable and divides it by the interquartile range. Individual wells were aggregated for downstream analysis by cell depot and day of differentiation. Subsequent data analyses were performed in R3.6.1 and MATLAB using base packages unless noted.

To assess batch effects we visualized the data using a Principal component analysis and quantified it using a Kolmogorov-Smirnov test implemented in the "BEclear" R package.[78] Additionally we applied a k-nearest neighbor (knn) supervised machine learning based classification algorithm implemented in the "class" R package[49] to investigate the accuracy of predicting biological and technical variation. For this analysis the dataset, consisting of 3 different cell types (hWAT, hBAT, SGBS) distributed on the 96-well plate, imaged at 4 days of differentiation, was split into equally balanced testing (n = 18) and training (n = 56) sets. Accuracy of this classification model was predicted based on three different categories, i.e. cell type, batch and column of the 96-well plate. (https://github.com/ClaussnitzerLab/Lipocyte-Profiler)

For dimensionality reduction visualization Uniform manifold approximation and projection maps (UMAP) were created using the UMAP R package version 0.2.7.0[76] (https://github.com/ClaussnitzerLab/Lipocyte-Profiler). To visualize LipocyteProfiler features and their effect size ComplexHeatmap Bioconductor package version 2.7.7[77] was used (https://github.com/ClaussnitzerLab/Lipocyte-Profiler)

To identify patterns of adipocyte differentiation underlying the morphological profiles a sample progression discovery analysis (SPD) was performed using the algorithm previously described in Qiu et al.[38] Briefly, the two adipose depots were analyzed separately, and features were clustered into modules based on correlation (correlation coefficient 0.6). Minimal spanning trees (MST) were constructed for each module and MSTs of each module are correlated with each other. Modules that support common MST were selected and an overall MST based on features of all selected modules were reconstructed.

Variance component analysis was performed by fitting multivariable linear regression models - yi $\sim$ xi + zi + . - where y denotes an LipocyteProfiler feature of individual i and x, z, etc. independent variables that could confound identification of biological sources of variability of the dataset. Independent variables are experimental batch, adipose depot, passaging before freezing, season and year of AMSCs isolation, sex, age, BMI, T2D status of individual, LipocyteProfiler feature Cells_Neighbors_PercentTouching_Adjacent corresponding to density of cell seeding and identification numbers of individuals. (https://github.com/ClaussnitzerLab/Lipocyte-Profiler)

To test whether there is a difference of morphological profiles on the tail ends of polygenic risk scores (PRS) for T2D, HOMA-IR and WHRadjBMI a multi-way analysis of variance (ANOVA) was performed. Individuals belonging to top 25% and bottom 25% of PRS score distribution are categorized into a categorical variable with 2 levels, top 25% or 25% bottom, according to their PRS percentile. Differences of morphological profiles are predicted using the categorized PRS variable adjusted for sex, age, BMI and batch. Additionally linear regression models were fitted adjusted for sex, age, BMI, batch and PC1 to predict differences of morphological and cellular profiles based polygenic risk for metabolic traits. To overcome multiple testing burden p values were corrected using false positive rate (FDR) described in R package "qvalue" (qvalue, no date). Features with FDR <5% were classified to be significantly impacted by the PRS variable. (https://github.com/ClaussnitzerLab/Lipocyte-Profiler) To decrease complexity we first removed features based on effect size and measurement type/class, and second removed features that correlate r > 0.85 with at least 10% of features of remaining features (https://github.com/ClaussnitzerLab/Lipocyte-Profiler).

LipocyteProfiler feature reduction using ARACNE

Algorithm for the Reconstruction of Accurate Cellular Networks, ARACNE[26,75,86] is a software package designed to capture regulatory networks from gene expression data. The method makes use of mutual information (MI) ranking to prioritize first order relationships (also known as direct regulatory relations) among genes, and to generate accurate maps of the regulatory network. The use of MI for the prioritization of biomolecular and clinical characteristics has been demonstrated previously calculated by the algorithm can be used to prioritize interactions between other cellular and molecular characteristics.[25,27,87] In this study, we use the ARACNE algorithm to construct a graph where its nodes represent the LipocyteProfiler's features and its weighted edges represent MI between prioritized first order relations between the nodes. We then build on this network to reduce the dimensionality of the LipocyteProfiler total feature space. Since the graph is constructed based on the calculated MI, the number of significant interactions of a node can be used to: (1) identify information hubs in the graph based on the number of connectivities of the nodes. After ranking the nodes based on their number of edges, a 75% upper quantile cutoff was applied to define the information hubs, (2) identify nodes with low connectivity that share minimal information. We calculated the average of weights (MI) of outgoing edges from every node and considered this average as the MI represented by the node. By ranking the nodes based on their calculated MI and applying a 25% lower quantile cutoff, we could identify the non-hub nodes that share minimal information (Figures S1B and S1C). The characterization of these two types of nodes play an important role in reducing the dimensionality of the LipocyteProfiler feature space while mitigating the overall information loss.

Please cite this article in press as: Laber et al., Discovering cellular programs of intrinsic and extrinsic drivers of metabolic traits using LipocyteProfiler, Cell Genomics (2023), https://doi.org/10.1016/j.xgen.2023.100346

We classify those LipocyteProfiler features identified as information hubs or that passed the MI cutoff as "LipocyteProfiler core features". We applied the two criteria on the total 3,005 LipocyteProfiler input features of visceral and subcutaneous derived AMSCs resulting in two sets of adipose depot-dependent LipocyteProfiler core feature sets. As a result, 986 features were labeled as LipocyteProfiler core features for visceral adipocytes and 1,002 for subcutaneous adipocytes. Between the two sets, there are 770 shared core features. The scripts for executing the ARACNE algorithm on the LipocyteProfiler features and the post processing steps for identifying the LipocyteProfiler core features are available through the GitHub page (https://github.com/ClaussnitzerLab/Lipocyte-Profiler).

Generating of average cells

For each group of interest, cells were pooled and divided into 100 clusters via K-Means clustering (scikit-learn). Individual cells were then sampled from the cluster closest to a theoretical point representing the mean of all object measurements, as determined by a Euclidean distance matrix.

RNA-seq

RNA-seq data were processed using FastQC[88] and spliced reads were aligned to human genome assembly (hg19) using STAR[73] followed by counting gene levels using Rsubread R package.[89] Next, raw read counts were normalized using the DESseq2 R package.[74] For differential expression analysis on the tail ends of polygenic risk scores (PRS) for HOMA-IR a multi-way analysis of variance (ANOVA) was performed on subset of 512 genes (GSEA hallmark gene sets for adipogenesis, fatty acid metabolism and glycolysis). Individuals belonging to top 25% and bottom 25% of PRS score distribution are categorized into a categorical variable with 2 levels, top 25% or 25% bottom, according to their PRS percentile. Differences in transcriptional profiles are predicted using categorized PRS variable adjusted for sex, age, BMI and batch. To overcome multiple testing burden p values were corrected using false positive rate (FDR) described in R package "qvalue" (qvalue). Genes with FDR <10% were classified to be significantly impacted by PRS and were uploaded to Enrichr to analyze them as a gene list against the WikiPathways. (https://github.com/ClaussnitzerLab/Lipocyte-Profiler)

Gene expression and LipocyteProfiler feature network

A linear regression model was fitted of 2,760 LP-features and global transcriptome RNA-seq data adjusted for sex, age, BMI and batch in subcutaneous AMSCs at day 14 of differentiation. Gene LP features associations were declared to be significant when passing the FDR cut-off of FDR< 0.01% (FDR<0.1%). LP features belonging to Cells category were used for further analysis. Associations between genes and LP features were visualized using "igraph" R package (Csardi, Nepusz and Others, 2006 - https://igraph.org/) (https://github.com/ClaussnitzerLab/Lipocyte-Profiler). Genes that are connected to top scoring LP features were uploaded to Enrichr to analyze them as a gene list against WikiPathways or BioPlanet. Adipocyte marker genes, *SCD*, *PLIN2*, *LIPE*, *INSR*, *GLUT4 and TIMM22*, were chosen to demonstrate morphological profiles matching their known pathways, by identifying LP features that associate with those genes with a global significant level of FDR<5%. (https://github.com/ClaussnitzerLab/Lipocyte-Profiler)

Quality control of genotyping data

Genotyping of all samples was performed in two separate batches using the Infinium HTS assay on Global Screening Array beadchips. Since the two sets of samples were genotyped with different versions of the beadchips and in different batches, we Qced, imputed, and generated the genome-wide polygenic scores separately and combined the results afterward.

A 3-step quality control protocol was applied using PLINK,[69,90] and included 2 stages of SNP removal and an intermediate stage of sample exclusion. The exclusion criteria for genetic markers consisted of: proportion of missingness ≥ 0.05, HWE p $\leq$ 1 x 10–20 for all the cohort, and MAF <0.001. This protocol for genetic markers was performed twice, before and after sample exclusion. For the individuals, we considered the following exclusion criteria: gender discordance, subject relatedness (pairs with PI-HAT ≥ 0.125 from which we removed the individual with the highest proportion of missingness), sample call rates ≥ 0.02 and population structure showing more than 4 standard deviations within the distribution of the study population according to the first seven principal components. After QC, 35 subjects remained for the analysis for which we had matched LipocyteProfiler imaging data.

Genotypes were phased with SHAPEIT2,[70] and then performed genotype imputation with the Michigan Imputation server, using Haplotype Reference Consortium (HRC)[91] as reference panel. We excluded variants with an info imputation r-squared <0.5 and a MAF <0.005.

Constructing PRSs

Genome-wide polygenic scores were computed using PRS-CS[71] and using the "auto" parameter to specify the phi shrinkage parameter. We computed the PRS-CS polygenic scores for the following traits: T2D,[51] BMI, waist-to-hip ratio adjusted and unadjusted by BMI, and stratified by sex and combined.[54] Genome-wide PRS for HOMA-IR were computed with LdPred[72] using summary statistics from Dupuis et al. (Dupuis et al., 2010). Process-specific PRSs were constructed based on five clusters defined in Udler et al.[59] by selecting the SNPs that had weight larger than 0.75 for each of a given cluster.

All PRSs were tested for association with T2D and with BMI using the 30,240 MGB Biobank samples from European Ancestry defined based on self-reported and principal components.

nature metabolism

Article

https://doi.org/10.1038/s42255-023-00807-w

A non-coding variant linked to metabolic obesity with normal weight affects actin remodelling in subcutaneous adipocytes

Received: 20 August 2021

Accepted: 12 April 2023

Check for updates

A list of authors and their affiliations appears at the end of the paper

Recent large-scale genomic association studies found evidence for a genetic link between increased risk of type 2 diabetes and decreased risk for adiposity-related traits, reminiscent of metabolically obese normal weight (MONW) association signatures. However, the target genes and cellular mechanisms driving such MONW associations remain to be identified. Here, we systematically identify the cellular programmes of one of the top-scoring MONW risk loci, the *2q24.3* risk locus, in subcutaneous adipocytes. We identify a causal genetic variant, rs6712203, an intronic single-nucleotide polymorphism in the *COBLL1* gene, which changes the conserved transcription factor motif of POU domain, class 2, transcription factor 2, and leads to differential *COBLL1* gene expression by altering the enhancer activity at the locus in subcutaneous adipocytes. We then establish the cellular programme under the genetic control of the *2q24.3* MONW risk locus and the effector gene *COBLL1*, which is characterized by impaired actin cytoskeleton remodelling in differentiating subcutaneous adipocytes and subsequent failure of these cells to accumulate lipids and develop into metabolically active and insulin-sensitive adipocytes. Finally, we show that perturbations of the effector gene *Cobll1* in a mouse model result in organismal phenotypes matching the MONW association signature, including decreased subcutaneous body fat mass and body weight along with impaired glucose tolerance. Taken together, our results provide a mechanistic link between the genetic risk for insulin resistance and low adiposity, providing a potential therapeutic hypothesis and a framework for future identification of causal relationships between genome associations and cellular programmes in other disorders.

Obesity and type 2 diabetes (T2D)-related traits are intimately linked by both environmental and genetic factors. The global prevalence of obesity and T2D has risen dramatically over the past century and both diseases constitute a serious increasing public health concern worldwide, with T2D predicted to rise in prevalence from 451 to 693 million people between 2017 and 2045 (ref. 1). Most epidemiological and genetic studies have linked obesity to the pathogenesis of T2D through positive phenotypic correlations between adiposity and T2D. However, a small number of loci have been identified that do not follow this pattern or even correlate in the opposite phenotypic direction[2]. Also, up to 45% of individuals with obesity do not present with poor glycaemic or lipid profiles, commonly called the metabolically healthy

e-mail: melina@broadinstitute.org

obese (MHO). Concurrently, up to 30% of normal-weight individuals present with significant cardiometabolic risk factors, known as the metabolically obese normal weight (MONW)[3–10].

Genome-wide association studies (GWAS) identified more than 700 genomic loci associated with glycaemic traits and T2D and more than 1,000 loci associated with adiposity-related traits, including body mass index (BMI) as a proxy of overall obesity, waist:hip ratio (WHR) as a proxy of body fat distribution, body fat percentage and direct measures of subcutaneous adipose tissue (SAT) and visceral adipose tissue (VAT) mass[2,11–20]. A small proportion of T2D genetic risk loci have been associated with decreased body fat percentage and a decreased SAT:VAT ratio[2,21–23]. Notably, these MONW/MHO genetic risk loci are characterized by distinct association signatures implying locus-specific cellular programmes[2,21–23].

The metabolic risk haplotype at *2q24.3* displays cross-phenotype association signatures that are reminiscent of the MONW phenotype and are associated with an increased risk of T2D, increased homeostatic model assessment for insulin resistance, increased WHR adjusted for BMI and decreased body fat percentage, decreased estimated SAT mass and cardiometabolic trait risk[2,15,24–26]. Consistent with these associations, the *2q24.3* locus falls into the lipodystrophy cluster of T2D loci[27], suggesting adipocytes as the mediating cell type at this locus. Notably, among the 20 loci identified in the T2D lipodystrophy process-specific cluster[27], the *2q24.3* locus is the top-scoring one, suggesting the strongest contribution to a lipodystrophic-like phenotype among the T2D GWAS loci.

However, like the vast majority of genetic risk loci identified through GWAS, the mechanism of the *2q24.3* metabolic risk locus is currently unknown. There are multiple factors that still confound the goal of converting genetic associations into specific knowledge, that is, pinpointing the causal variant(s), target gene(s) and mechanism(s). First, the resolution of association mapping is inherently limited by the haplotype structure of the human genome because a common variant in the population is usually strongly genetically linked to many neighbouring variants. Second, the genetic architecture of common diseases is fundamentally different to that of rare diseases, as over 80% of associated regions do not contain any protein-altering common variants[28]. Importantly, genetics and descriptive catalogues of molecular activity alone are, for non-coding signals at least, insufficient to build compelling proof for a chain of causation that stretches from variant to phenotype. In this study, we set out to combine statistical and experimental methods to identify the mechanism by which *2q24.3* confers decreased adiposity and increased risk of T2D (Extended Data Fig. 1a), possibly enabling further development of therapeutic strategies to alter glycaemic control.

Results

The *2q24.3* risk locus maps to adipocyte enhancer signatures

To identify diseases and traits associated with the *2q24.3* locus tagged by the variant rs3923113, we visualized large-scale phenome-wide associations from the UK Biobank (UKB) and from a meta-analysis of a series of metabolic traits (Fig. 1a and Extended Data Fig. 1b). We observed that the *2q24.3* locus was associated with increased T2D risk, increased fasting insulin levels and a series of body fat-related traits, including increased WHR adjusted for BMI, but decreased trunk fat percentage, arm fat percentage, hip circumference and whole-body fat mass, suggesting a complex pleiotropic risk locus consistent with a MONW association signature, that is, a lean, metabolically unhealthy phenotype.

The *2q24.3* locus tagged by rs3923113 encompasses 55 kb, spanning from the *COBLL1* intronic regions to the intergenic region between *GRB14* and *COBLL1* (Extended Data Fig. 1c). The MONW locus harbours 20 non-coding single0nucleotide polymorphisms (SNPs) in high linkage disequilibrium (LD) ($r^2 > 0.8$, 1000 Genomes Project Phase 1 EUR). These single-nucleotide variants (referred to as the candidate regulatory variants) define two alternative haplotypes: the ancestral haplotype 1 (frequency 38% in European individuals), associated with a decreased risk for the MONW association pattern; and haplotype 2 (frequency 62%), associated with an increased MONW risk. To connect genetic variants at the *2q24.3* locus to relevant cell types and cell states, we examined chromatin state maps across 127 reference epigenomes from the Roadmap Epigenomics and the ENCODE consortium (Fig. 1b and Extended Data Fig. 1d). We found that the locus maps to multiple enhancer signatures, including active enhancers in mesenchymal stem cells, adipocyte progenitors and adipocytes (Fig. 1b). Several of the 20 non-coding variants map within or in the vicinity of regions with active enhancer chromatin states, suggesting that the *2q24.3* locus acts in adipocytes through gene regulatory mechanisms.

Next, we examined whether the two haplotypes (risk compared to non-risk) showed differences in chromatin structure during adipocyte differentiation. Specifically, we performed assays for enhancer activity (H3K27 acetylation (H3K27ac) chromatin immunoprecipitation followed by sequencing (ChIP–seq)) and chromatin accessibility (assay for transposase-accessible chromatin with sequencing (ATAC–seq)) on adipose-derived mesenchymal stem cells (AMSCs) from heterozygous individuals across a time course of differentiation (before induction (day 0), early differentiation (day 2), intermediate differentiation (day 6) and terminal differentiation (day 14)), and compared the numbers of reads from the two haplotypes (Extended Data Fig. 1e). The MONW risk haplotype (haplotype 2) was associated with a decrease in H3K27ac, a proxy of enhancer activity, and chromatin accessibility, with the MONW risk haplotype enriched by roughly 1.5-fold. The allele-specific difference in chromatin accessibility was reproducible across three heterozygous lines (Supplementary Tables 1 and 2), was most pronounced at day 0 of differentiation and declined after induction of differentiation (Extended Data Fig. 1f). These results indicate that haplotype 1 is associated with an active enhancer state, whereas haplotype 2 is associated with a weak enhancer state primarily in adipocyte progenitors.

The rs6712203 variant affects adipocyte *COBLL1* expression

To identify which of the 20 candidate regulatory variants is likely to mediate the differential enhancer activity in adipocyte progenitors, we used two orthogonal computational approaches to prioritize variants, phylogenetic module complexity analysis (PMCA)[28–30] and Bassett[31] (Fig. 2a–c and Supplementary Tables 3 and 4). PMCA assesses evolutionary conservation of sequence, order and distance (in human and at least one other vertebrate species) of groups of at least three transcription factor binding motifs within a 120-bp region. Basset uses a sequence-based deep convolutional neural network (CNN) approach to predict the effects of non-coding variants on regulatory activity by training on the sequence content of a given epigenomic mark in a tissue or cell type of interest. After training on genome-wide chromatin accessibility (ATAC–seq) data in numerous cell types, including AMSC progenitors, before induction (day 0), one variant, rs6712203, stood out as consistently showing the highest score for PMCA and Bassett (Fig. 2a). Bassett predicted that the T allele on the protective haplotype increases chromatin accessibility relative to the C allele on the risk haplotype in adipocyte progenitors. These sequence-based estimates of rs6712203 C>T single-nucleotide change importance are consistent with the variant overlapping an active enhancer associated with H3K27ac and H3K4 mono-methylation in adipocyte progenitors. In line with the variant importance at rs6712203, conditional analyses of anthropometric and glycaemic traits defining MONW in the UKB confirmed an rs6712203 C/T association consistent with a primary effect in female participants for fat mass ($\beta = 0.022$, $P = 4.1 \times 10^{-08}$ in females and $\beta = 0.0084$, $P = 0.052$ in males, difference $P = 0.019$), hip circumference ($\beta = 0.037$, $P = 4.3 \times 10^{-15}$ in females and $\beta = 0.0097$, $P = 0.008$ in males, difference $P = 4.3 \times 10^{-06}$) and T2D in both females and males (females $\beta = -0.11$, $P = 1.4 \times 10^{-07}$ and males $\beta = -0.068$, $P = 2.4 \times 10^{-05}$,

https://doi.org/10.1038/s42255-023-00807-w

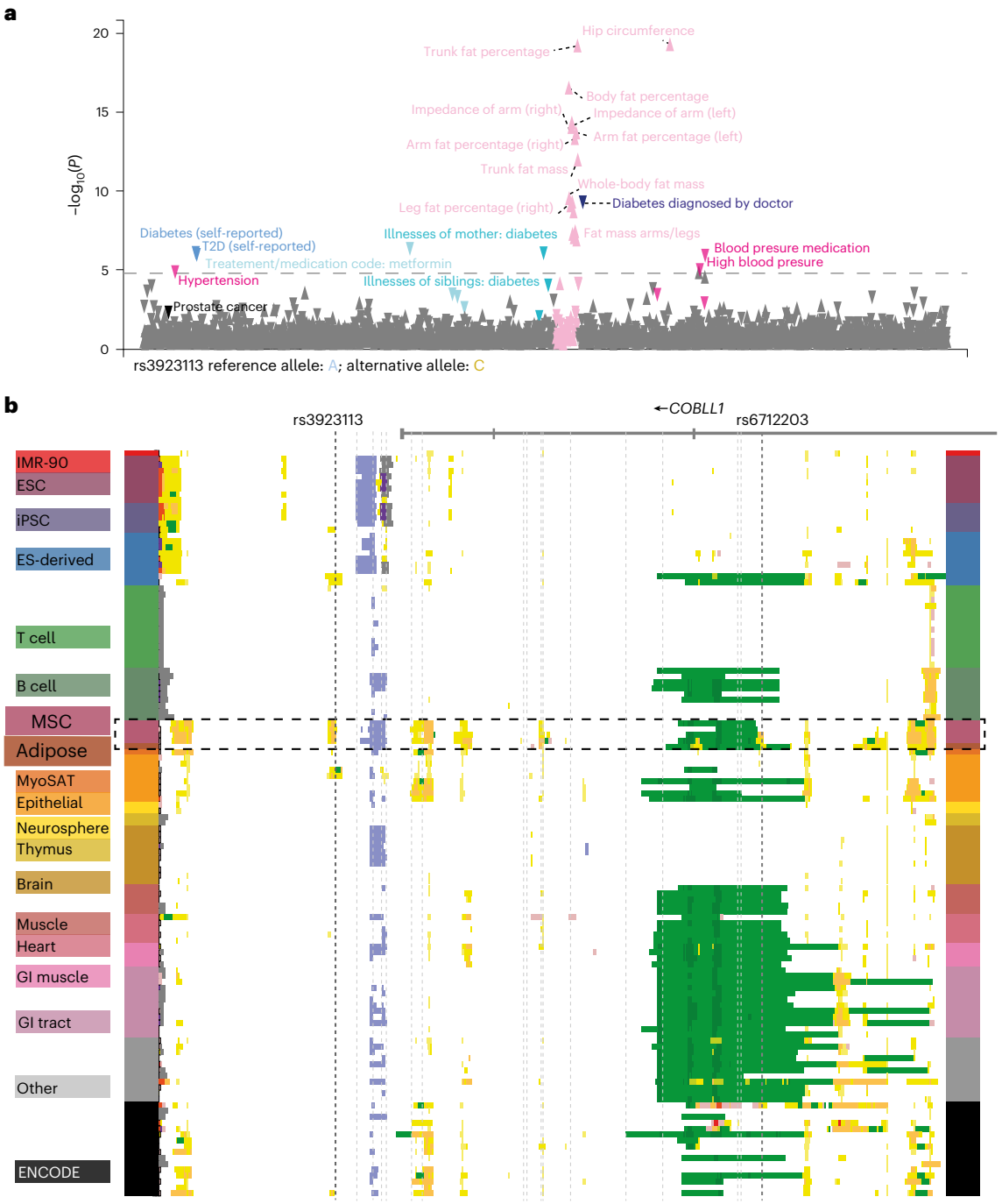


Fig. 1 | The pleiotropic *2q24.3* MONW locus is associated with increased risk for T2D and decreased adiposity-related traits, and maps to sparse enhancer signatures in adipocytes. a, Phenome-wide association studies of trait associations at the rs3923113-tagged haplotype in the UKB[87]. The colours represent trait classes while individual rs3923113 variant association *P* values are shown on the *y* axis. The direction of effect is indicated by the orientation of the triangles: upward, increase; downward, decrease. **b**, Chromatin state annotations for the 55-kb-long MONW risk locus. Genomic intervals are shown across 127 human cell types and tissues reference epigenomes profiled by the Roadmap Epigenomics project, based on a 25-state chromatin state model (for the colours, see Fig. S1) learned from 12 epigenomic marks using imputed signal tracks at 25-nucleotide resolution[88]. Chromatin states considered included Polycomb repressed states (grey, H3K27me3), weak enhancers (yellow, H3K4me1 only), strong enhancers (orange, also H3K27ac) and transcribed enhancers (green, also H3K36me3) (https://www.freepatentsonline.com/y2022/0243178.html). ES, embryonic stem; ESC, embryonic stem cell; GI, gastrointestinal; iPSC, induced pluripotent stem cell; MSC, mesenchymal stem cell.

Q9

difference $P = 0.098$, Supplementary Table 5 and Extended Data Figs. 1g and 2). Furthermore, we observed that the rs6712203 association with T2D was dependent on BMI ($\beta = -0.0028$, $P = 0.70$ in normal weight and underweight and $\beta = -0.093$, $P = 2.92 \times 10^{-06}$ in obese, difference $P = 2.0 \times 10^{-05}$, Supplementary Table 6).

We next used three-dimensional genome conformation data from Hi-C assays in normal human epidermal keratinocytes and human fibroblasts[32] to define the physical boundaries of potential proximal and long-distant target genes. We found that the locus lies in a well-defined contact domain containing only two genes, cordon-bleu

https://doi.org/10.1038/s42255-023-00807-w

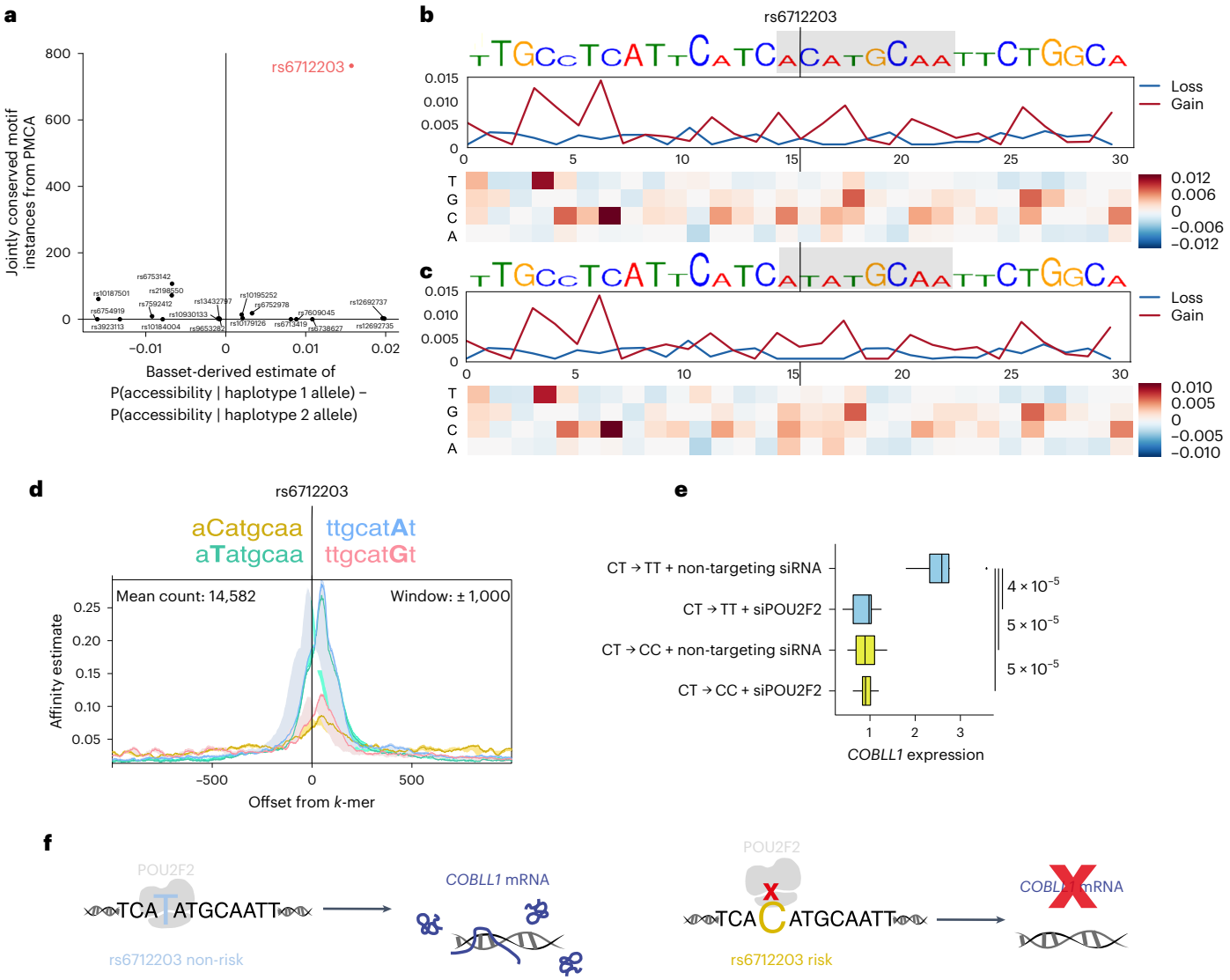


Fig. 2 | rs6712203 is a functional variant at the *2q24.3* MONW locus. **a**, Phylogenetic conservation analysis and CNN-based prediction of chromatin accessibility for 20 highly linked (LD = $r^2 > 0.8$) variants at the *2q24.3* locus (tag SNP rs3923113). Phylogenetic conservation scores of jointly conserved motifs using PMCA are shown on the *x* axis[29]. PMCA was used to identify orthologous regions in 20 vertebrates and to scan the 120-bp sequence context for groups of transcription factor binding site motifs whose sequence, order and distance range is conserved across species. Scores indicate the count of non-overlapping jointly conserved transcription factor binding site motifs whose relative positions within the window are conserved. Predicted relative change in chromatin accessibility (SNP accessibility difference SAD scores) in preadipocytes (day 0 of differentiation) for each SNP comparing alleles on haplotype 1 and haplotype 2 is shown on the *y* axis. A deep CNN Basset[31] was trained on genome-wide ATAC–seq data assayed in preadipocytes. Alleles were assigned to each SNP in the haplotype and evaluated for predicted accessibility. **b**, Both PMCA and Basset predicted rs6712203 as a functional variant. For the C allele, there are no nucleotide variants that reduce binding in the rs6712203 region. Each position on the *x* axis represents a nucleotide and the four values in the heatmap correspond to all possible substitutions. **c**, For the T allele, in silico saturation mutagenesis suggests that binding loss of the *POU2F* motif that overlaps rs6712203, including the C allele itself, results in reduced predicted chromatin accessibility. **d**, Intragenomic replicates[35] predict a higher binding affinity of POU2 family transcription factors for the T than C allele to both strands. Offsets from instances of the given *k*-mer sequence are shown on the *x* axis. Estimated affinity of binding (https://www.freepatentsonline.com/y2022/0243178.html) is shown on the *y* axis. A model with 8-mers is shown; alternatives with 6-mers through to 9-mers are shown in Extended Data Fig. 3b. **e**, Generation of rs6712203 CRISPR–Cas9 engineered lines starting from SGBS preadipocytes (heterozygous for rs6712203) edited to the homozygous risk (CC, yellow) and non-risk (TT, blue) alleles and qPCR-based gene expression measurement of *COBLL1* conditional on the regulator POU2F2. $n = 3$ biologically independent experiments. **f**, Schematic model of the regulatory circuitry under the genetic control of rs6712203.

WH2 repeat protein like 1 (*COBLL1*) and growth factor receptor bound protein 14 (*GRB14*) (Extended Data Fig. 3a), without any evidence for long-range chromatin interactions. To dissect which of these two genes are targeted by rs6712203, we undertook a CRISPR interference (CRISPRi) approach by engineering immortalized human preadipocytes (in human white adipose tissue (hWAT))[33] to stably express a catalytically dead Cas9 fused to a Krüppel-associated box (KRAB) domain[34] (Extended Data Fig. 3b). Transduction of Cas9 Endonuclease Dead (dCas9)-hWAT preadipocytes with increasing amounts of lentiviral particles carrying single-guide RNAs (sgRNAs) targeting the transcriptional start sites (TSS) of *COBLL1* and *GRB14* genes led to dose-dependent downregulation of mRNA expression by 72% and

96% compared to non-targeting negative controls (Extended Data Fig. 3c). At the highest dose, COBLL1 protein was reduced by 56% without effects of sgRNAs targeting the GRB14 TSS on *COBLL1* expression (Extended Data Fig. 3d). After confirming the knockdown efficiency and specificity of our system, we next designed a panel of six sgRNAs surrounding the rs6712203 variant (Extended Data Fig. 3e) and additional non-targeting and COBLL1/GRB14 TSS controls to assess its effects on *COBLL1* and *GRB14* mRNA expression. All TSS sgRNAs led to a reduction in *COBLL1* and *GRB14* mRNA content by 64% and 94% without any interference on each other (Extended Data Fig. 3f,g). More importantly, all tested sgRNAs, with exception of sgRNA 1, led to significant downregulation of *COBLL1* over a range of 20–40% compared to negative controls (Extended Data Fig. 3f), while no effects were observed on *GRB14* (Extended Data Fig. 3g), thus pointing to *COBLL1* as the primary target of rs6712203.

We next performed in silico saturation mutagenesis to evaluate the predicted change in chromatin accessibility from mutation at every position into each alternative nucleotide within a 30-bp region surrounding rs6712203 using ATAC–seq data during AMSC differentiation. We found that the rs6712203 T allele is critical for a POU2F2 motif (Fig. 2d). The C allele of this SNP converts the chromatin in this site into a less accessible chromatin, supporting a model in which a transcription factor, possibly POU2F2, differentially binds to these allelic variants of rs6712203. To estimate the preferential binding affinity of POU2F2 to the C risk compared to the T non-risk allele, we used the intragenomic replicate (IGR) method[35] on publicly available POU2F2 ChIP–seq data from the ENCODE project. By comparing the frequency of *k*-mers matching the rs6712203 T allele versus the C allele, we confirmed that POU2F2 preferentially binds to the T allele (9-mer change in affinity −0.38, two-tailed permutation $P < 0.0034$) (Fig. 2d and Extended Data Fig. 3h). These data suggest an increased POU2F2 binding to the rs6712203 T non-risk allele and suggest POU2F2 as the upstream regulator of variant action at this locus.

We next sought to establish rs6712203 causality by directly confirming that the haplotype-specific effects on enhancer activity and POU2F2 binding are mediated by rs6712203 using CRISPR-based genome editing at this SNP. We edited the human Simpson–Golabi–Behmel syndrome (SGBS) preadipocytes ($n = 5$) that are heterozygous at rs6712203 to create isogenic lines for the homozygous TT (non-risk genotype) and homozygous CC (risk genotype) alleles. In line with the CRISPRi rs6712203 enhancer repression experiments outlined above, we observed that cells harbouring the CC homozygous risk showed 2.4-fold lower *COBLL1* expression levels compared to the TT non-risk genotype (Fig. 2e), pointing towards *COBLL1* as a target gene of the rs6712203 regulatory circuitry. To test a *cis*/*trans*-conditional effect of the rs6712203 variant and the upstream regulator POU2F2 on target gene expression, we performed targeted small interfering RNA (siRNA)-mediated ablation of POU2F2 in SGBS cells and found that silencing of POU2F2 in TT allele carriers reduces *COBLL1* gene expression to the level of CC allele carriers in preadipocytes (Fig. 2e), confirming POU2F2 as a functional regulator at the locus.

Together, our computational and experimental approaches support a model where the causal MONW variant rs6712203 regulates *COBLL1* gene expression in a POU2F2-dependent manner (Fig. 2f).

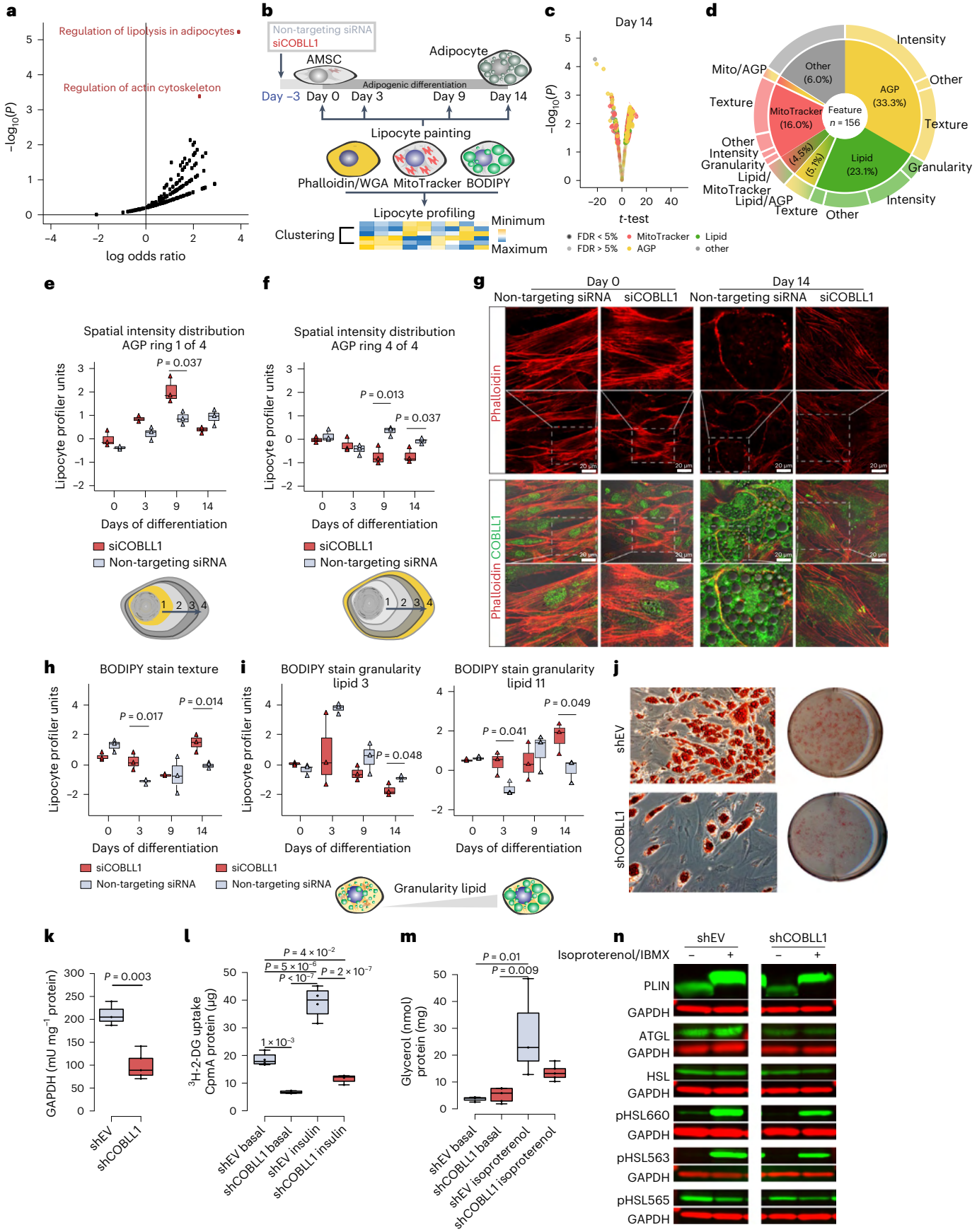
COBLL1 affects actin remodelling in subcutaneous adipocytes

To understand the role of COBLL1 in adipocyte cellular programmes, we examined the gene expression and cellular localization of *COBLL1* in differentiating adipocytes. We observed that *COBLL1* is expressed at all stages of adipocyte differentiation with an increase in mRNA over the course of differentiation (Extended Data Fig. 4a). We observed consistently higher *COBLL1* mRNA levels in subcutaneous compared to visceral adipocytes across the adipocyte differentiation process (Extended Data Fig. 4a). Overall, we found an enrichment of *COBLL1* gene expression in adipose tissue compared to 146 other tissues and cell types[36] (Extended Data Fig. 4b).

To connect the *2q24.3* locus to cellular functions in adipose tissue, we used genome-wide coexpression matrices in adipocytes matched with a series of cellular assays. We identified *COBLL1* coregulated genes in genome-wide expression data from primary human AMSCs in a cohort of 12 healthy, non-obese individuals. *COBLL1* coexpressed genes were highly enriched in biological processes related to regulation of actin cytoskeleton and regulation of lipolysis in adipocytes, including integrin subunit alpha M (*ITGAM*), phosphatidylinositol -4,5-bisphosphate 3-kinase catalytic subunit alpha (*PIK3CA*), rho-associated protein kinase 2 (*ROCK2*), integrin alpha-1 (*ITGA1*), and rho guanine nucleotide exchange factor 7 (*ARHGEF7*), *CRK*, fibroblast growth factor receptor 2 (*FGFR2*) and rho guanine nucleotide exchange factor 6 (*ARHGEF6*) (Fig. 3a, Extended Data Fig. 4c,d and Supplementary Tables 7–10), which are implicated in actin remodelling processes and insulin responsiveness[37–40]. This is consistent with recent studies showing that COBLL1 possesses a Wiskott–Aldrich syndrome protein homology 2 actin monomer-binding domain, and promotes filamentous actin (F-actin) formation in Cos-7, neuronal and prostate cancer cells[41,42]. Finally, we found a significantly decreased relative *COBLL1* gene expression after a weight loss of 6.9 ± 1.9 kg in 18 women with obesity (BMI = 34.9 ± 3.8 kg m^{-2}) (Extended Data Fig. 4e).

Fig. 3 | The *2q24.3* effector gene *COBLL1* affects actin remodelling processes in differentiating adipocytes. a, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment of genes correlated with *COBLL1* in adipocytes using Enrichr[89,90]. $n = 30$ donors, *P* values were derived from a hypergeometric test. **b**, Schematic of siCOBLL1 experiments in AMSCs across differentiation. AMSCs from a normal-weight female donor were silenced 3 d before induction and LipocyteProfiling (days 0, 3, 8 and 14 of differentiation), $n = 3$. **c**, Image-based profiles of si*COBLL1*-treated compared to non-targeting siRNA-treated AMSCs on day 14; two-sided *t*-test, significance level FDR < 5%, $n = 3$. **d**, Pie chart illustrating differential features per channel and measurement class comparing si*COBLL1* and the non-targeting siRNA control on day 14. **e**,**f**, Spatial intensity distribution of AGP informative for the actin cytoskeleton in the centre of the cytoplasm: Cytoplasm_RadialDistribution_FracAtD_AGP_1of4 (**e**) and juxtaposed to the plasma membrane (Cytoplasm_RadialDistribution_RadialCV_AGP_4of4) (**f**). Two-sided *t*-test, $n = 3$ biologically independent experiments, median ± 95% confidence interval (CI). **g**, Representative images of *COBLL1* knockdown and non-targeting siRNA control at days 0 and 14 of differentiation. COBLL1 staining: anti-COBLL1 primary and donkey anti-rabbit IgG H&L secondary antibodies. Actin staining: phalloidin–Atto 565. Olympus FLUOVIEW FV1000 CLSM Inverse microscope (40× magnification). Images were processed with Image J. Dotted square: image zoom-in. **h**,**i**, Texture of BODIPY stain (Cells_Texture_Correlation_Lipid_10_01) (**h**) and granularity measures of the BODIPY stain (Cells_Granularity_3_BODIPY) (**i**) in *siCOBLL1* knockdown and non-targeting siRNA; two-sided *t*-test, $n = 3$ biologically independent experiments, median ± 95% CI. **j**, Oil Red O staining in SGBS adipocytes after stable lentiviral *COBLL1* knockdown (shCOBLL1) versus empty vector control (shEV); scale bar, 15 mm. **k**, GAPDH metabolic activity in differentiated shCOBLL1 compared to non-targeting siRNA adipocytes. Paired Student's *t*-test, median ± 95% CI, $n = 3$. **l**, Basal and insulin-stimulated ^{3}H-2-DG uptake in differentiated shCOBLL1 compared to shEV adipocytes. One-way analysis of variance (ANOVA) with Tukey's honestly significant difference (HSD) test, median ± 95% CI, $n = 3$ biologically independent experiments. **m**, Basal and isoproterenol-stimulated lipolysis rate measured using glycerol release in differentiated shCOBLL1 compared to shEV adipocytes. One-way ANOVA with Tukey's HSD test, median ± 95% CI, $n = 3$ biologically independent experiments. **n**, Western blots for lipolysis-relevant proteins assayed in basal or isoproterenol/IBMX-stimulated shCOBLL1 versus shEV adipocytes ($n = 2$).

https://doi.org/10.1038/s42255-023-00807-w



To identify morphological and cellular traits associated with altered *COBLL1* expression, we used siRNA-mediated knockdown of *COBLL1* in AMSCs coupled with a high-content imaging assay that we recently developed, LipocyteProfiler[43]. LipocyteProfiler reports on general and adipocyte-specific cellular traits for preadipocytes before differentiation (day 0), and at three time points of adipocyte differentiation (3 d (day 3), 9 d (day 9) and 14 d (day 14) after adipogenic induction) (Fig. 3b). We examined 1,175 quantitative features, spread across two cellular compartments (whole-cell and cytoplasm only) and five dyes informative for morphological and adipocyte cellular traits (BODIPY, phalloidin, wheat germ agglutinin (WGA), SYTO14 and MitoTracker (Methods) imaged in four fluorescence channels (Fig. 3b). We observed that *COBLL1* knockdown in proliferating preadipocytes with 80% knockdown efficiency (Extended Data Fig. 5a) results in changes of diverse morphological and cellular features across adipocyte differentiation with a peak at the later stages of differentiation (Fig. 3c and Extended Data Fig. 5b–d). On day 14 of differentiation, 156 features differed significantly (false discovery rate (FDR) < 5%) between *COBLL1* knockdown and non-targeting control, spread across lipid (23.1%), actin, Golgi, plasma membrane (AGP)-related (33.3%) and mitochondrial (16.0%) channels (Fig. 3d and Supplementary Table 11). For AGP-related cellular processes, we observed that *COBLL1* knockdown alters the spatial intensity distribution of AGP across the cytoplasm. We previously observed that these AGP-related features can be indicative of actin cytoskeleton remodelling during differentiation[43]. After *COBLL1* silencing, we observed increased actin-associated intensity in the centre of the cell (day 9 $P = 0.037$; Fig. 3e) and decreased actin-associated intensity at the cell cortex (day 9 $P = 0.013$, day 14 $P = 0.026$; Fig. 3f) in differentiated subcutaneous adipocytes. This indicates that *COBLL1* has a role in remodelling the actin cytoskeleton during adipocyte maturation, as reduced levels of *COBLL1* disturb the disassembly of F-actin stress fibres across the cytoplasm and the reassembling to cortical F-actin (F-actin juxtaposed to the plasma membrane).

The disturbed actin remodelling process was accompanied by lowered differentiation capacity as shown by decreased lipid droplet formation. More specifically, we confirmed that *COBLL1* knockdown was associated with a decreased disassembly of stiff F-actin stress fibres reaching in the middle of the cell body, at the expense of F-actin structure assembly at the cell cortex in differentiated cells (Fig. 3g and Extended Data Fig. 5e). Consistent with the notion that remodelling of F-actin stress fibres into cortical actin is linked to adipocyte differentiation, *COBLL1*-ablated adipocytes showed significant changes in both texture of lipid-related pixels (Cells_Texture_Correlation_Lipid; pixel intensities of BODIPY stain within the cell are more similar, day 3 $P = 0.017$ and day 14 $P = 0.039$; Fig. 3h) and lipid-related granularity measures (a class of metrics that captures the typical sizes of bright spots for the BODIPY stain) within the cell compared to adipocytes expressing *COBLL1*. More specifically, we observed that silencing of *COBLL1* results in adipocytes with fewer smaller-sized lipid droplets (Cells_Granularity_Lipid 3, day 14 $P = 0.0483$; Fig. 3i) and more medium-to-large-sized lipid droplets in mature adipocytes (Cells_Granularity_Lipid 11, day 3 $P = 0.0412$ and day 14 $P = 0.0486$; Fig. 3i), and an overall decrease in lipid droplet numbers (Cellls_Children_LargeLipidObjects_Count, $P = 0.006$; Extended Data Fig. 5f). These data indicate a disturbed lipid droplet formation and adipogenic differentiation in *COBLL1*-ablated cells, driven by an altered cytoskeleton remodelling.

To investigate if the COBLL1 effect on actin remodelling in adipocytes impacts adipocyte cellular programmes related to metabolic disease, we performed stable ablation of *COBLL1* using lentivirus (shCOBLL1) in differentiating adipocytes. We observed that ablation of *COBLL1* resulted in a decreased capacity to differentiate into metabolically active, round-shaped, lipid-filled mature adipocytes, as shown by decreased Oil Red O staining of accumulated triglycerides (Fig. 3j), adipocyte differentiation marker gene expression (Extended Data Fig. 5g) and glyceraldehyde-3-phosphate dehydrogenase (GAPDH) activity measurements (2.1-fold, $P = 3 \times 10^{-3}$; Fig. 3k). We further found a correlation between the mRNA levels of *COBLL1* and leptin, an adipokine produced in proportion to the size of fat depots[44] in SAT ($r = 0.74$, $P = 5 \times 10^{-5}$) (Extended Data Fig. 5h,i). This effect on leptin is consistent with GWAS of serum leptin levels (rs6712203 C allele $\beta = 0.0308$, $P = 9 \times 10^{-6}$ and $\beta = 0.0236$, $P = 1 \times 10^{-5}$ (BMI-adjusted)) in Kilpeläinen et al.[45]; and $\beta = 0.0285$, $P = 0.005889$ in Folkersen et al.[46]. We further found a 3.4-fold ($P = 2 \times 10^{-7}$) decrease of insulin-responsive glucose uptake in shCOBLL1 adipocytes compared to non-targeting control, as measured by radiolabelled ^{3}H-2-deoxyglucose (^{3}H-2-DG) uptake assays (Fig. 3l). The failure of shCOBLL1 adipocytes to respond to insulin may have resulted from both lowered differentiation efficiency and failure of the cortical actin remodelling that mediates GLUT4 vesicle trafficking. Finally, we observed a failure of shCOBLL1 adipocytes to break down triglycerides to free fatty acids and glycerol through lipolysis after β-adrenergic stimulation using isoproterenol and the phosphodiesterase inhibitor IBMX compared to their control cells (Fig. 3m). This was accompanied with decreased protein levels of the lipolytic enzymes adipocyte triglyceride lipase, hormone-sensitive lipase (HSL), protein kinase A-dependent serine phosphorylated HSL (pHSL660, pHSL563) and on the lipid droplet-associated protein perilipin-1 (PLIN) (Fig. 3n). Notably, we did not observe an effect on cellular and morphological features when silencing *COBLL1* after induction of differentiation (Extended Data Fig. 5j,k), suggesting that *COBLL1* acts early in differentiation, with phenotypic effects primarily manifesting in mature adipocytes. We also did not observe an effect when *COBLL1* was ablated in visceral AMSCs (Extended Data Fig. 5l), indicating that *COBLL1* is critically involved in actin remodelling processes in subcutaneous adipocytes.

We additionally examined the effect of *GRB14* stable knockdown in AMSCs and observed that *GRB14* ablation (knockdown efficiency of 61%) did not significantly decrease in adipocyte differentiation capacity as measured by Oil Red O staining, GAPDH activity (Extended Data Fig. 5m,n) or insulin-responsive glucose uptake and *GLUT4* gene expression (Extended Data Fig. 5o,p), supporting *COBLL1* as the effector gene at this locus.

The MONW haplotype affects the adipocyte actin cytoskeleton

To confirm that the changes on the actin cytoskeleton and subsequent effects on adipocyte functions are under the genetic control of the rs6712203 MONW risk haplotype, we used LipocyteProfiler[43] to phenotypically profile primary human adipocytes differentiation throughout from individuals carrying the risk haplotype ($n = 6$) compared the non-risk haplotype ($n = 7$) (Fig. 4a). The data revealed that AGP and lipid features informative for the actin cytoskeleton and lipid accumulation differed in subcutaneous adipocytes from rs6712203 metabolic risk versus non-risk haplotype carriers (Fig. 4b,c and Extended Data Fig. 6a–c). More specifically, we found that 77 morphological features, spread across lipid (16.9%), actin-associated AGP (45.5%) and mitochondrial (26.0%) channels, significantly differed between the haplotypes on day 14 of differentiation (FDR = 5%; Supplementary Table 12). We did not observe any significant difference in visceral adipocytes (Fig. 4d and Extended Data Fig. 6d–f), which is consistent with the depot-specific effect of *COBLL1* knockdown (Extended Data Fig. 5k). Notably, we found that the risk haplotype associates with increased actin-associated intensity in the centre of the cell (day 0 $P = 0.018$, day 3 $P = 0.042$, day 9 $P = 0.011$ day 14 $P = 0.009$; Fig. 4e) and decreased actin-associated intensity at the cell cortex (day 9 $P = 0.024$, day 14 $P = 0.009$; Fig. 4f). These data recapitulate our findings after *COBLL1* knockdown and confirms that adipocytes from risk allele carriers are characterized by less cortical actin, which is required for insulin-stimulated glucose uptake in those cells and therefore directly relevant to fasting insulin levels and T2D. Furthermore, we observed a risk haplotype association with higher lipid object counts (day 8 $P = 0.043$, day 14 $P = 0.034$;

https://doi.org/10.1038/s42255-023-00807-w

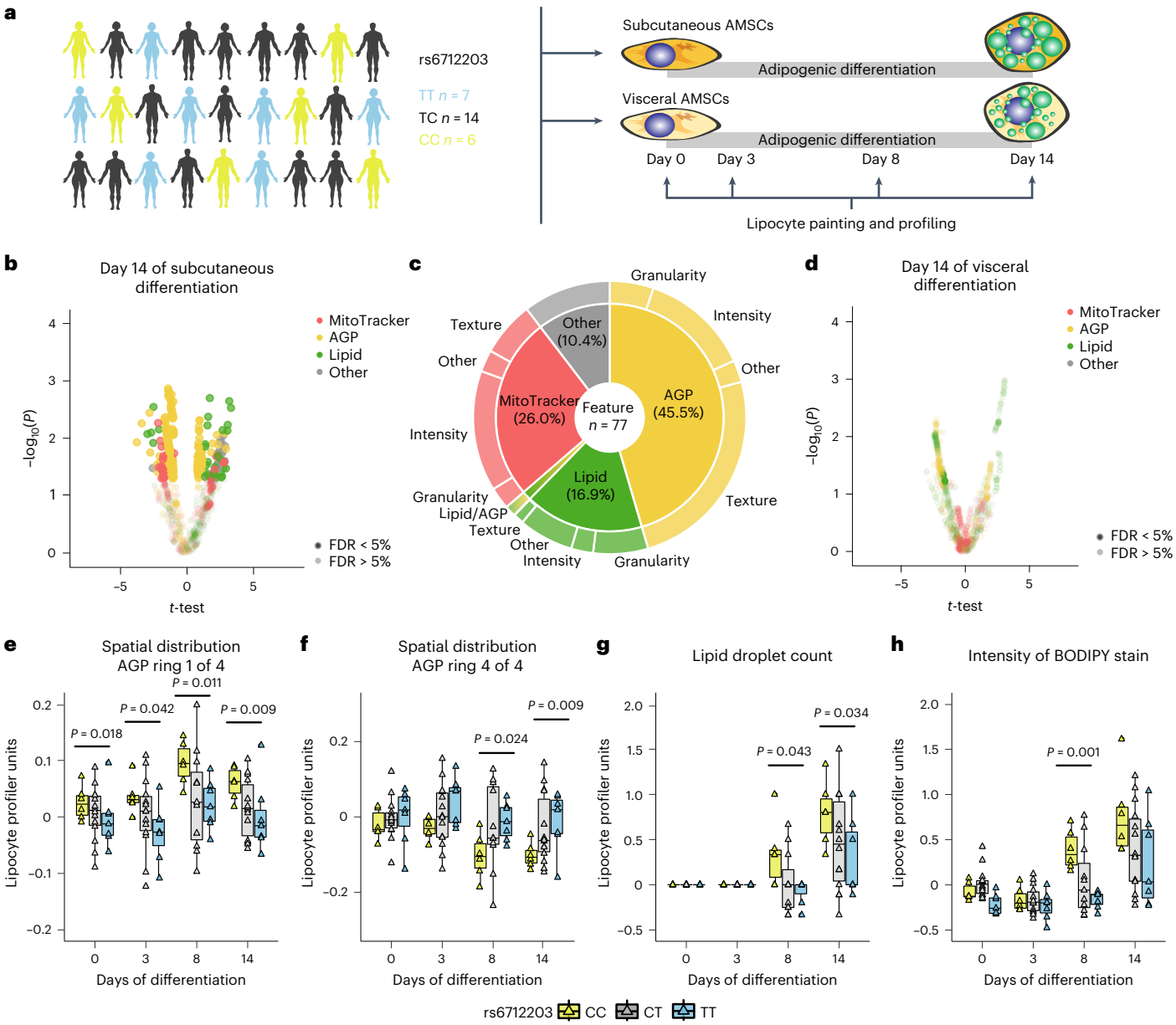


Fig. 4 | The rs6712203 MONW risk haplotype affects actin remodelling in adipocytes and adipocyte lipid storage capacity. a, Schematic of adipocyte differentiation and LipocyteProfiling of subcutaneous AMSCs derived from TT ($n = 7$) and CC ($n = 6$) allele carriers of rs6712203 using LipocyteProfiler[43]. **b**,**d**, Differences in morphological profiles between TT ($n = 7$) and CC ($n = 6$) allele carriers at day 14 in subcutaneous **(b)** and visceral **(d)** AMSCs (multi-way ANOVA, significance level FDR < 5%). **c**, Pie chart illustrating non-redundant differential features per channel and class of measurement at day 14 of subcutaneous adipocyte differentiation in rs6712003 homozygous risk carriers compared to non-risk carriers. **e**,f, Spatial intensity distribution of AGP in the centre of the cytoplasm near the nucleus in subcutaneous adipocytes derived from TT ($n = 7$), TC ($n = 14$) and CC ($n = 6$) carriers of rs6712203 (Cytoplasm_RadialDistribution_FracAtD_AGP_1of4) (**e**) and juxtaposed to the plasma membrane (Cytoplasm_RadialDistribution_RadialCV_AGP_4of4) (**f**) throughout differentiation. Multi-way ANOVA; data represent the median ± 95% CI. **g**,**h**, Lipid droplet count (Cells_Children_LargeLipidObjects_Count) (**g**) and intensity of BODIPY stain (Cells_Intensity_IntegratedIntensity_Lipid) (**h**) throughout differentiation. Multi-way ANOVA; data represent the median ± 95% CI.

Fig. 4g), which is representative of several lipid droplets, and higher lipid-related intensity (day 8 $P = 0.001$; Fig. 4h), which is indicative of dysfunctional lipid droplet formation. These genetic effects on the actin cytoskeleton dynamics and lipid accumulation in AMSCs are coherent with the effects we observed after the *COBLL1* knockdown experiments (Fig. 3b–f,h,i), indicating that altered *COBLL1* expression in the risk haplotype underlies the observed phenotypic effects in adipocytes. Summarized, these data show that the rs6712203 MONW risk locus, by altering *COBLL1* mRNA expression levels, impacts actin remodelling in differentiating adipocytes, thereby strongly affecting fat mass-relevant and T2D-relevant cellular programmes including adipocyte differentiation, lipid droplet formation, stimulated lipolysis rate and insulin-stimulated glucose uptake.

Cobll1-deficient mice are metabolically obese normal weight

We generated a CRISPR-engineered *Cobll1* knockout (*Cobll1*$^{-/-}$) mouse model to determine a potential role for *Cobll1* in the regulation of metabolic function in vivo (Extended Data Fig. 7a,b). First, we sought to assess the effect of *Cobll1* knockout on morphological and cellular profiles in differentiating murine SAT AMSCs by LipocyteProfiler (days 0, 2 and 10 of differentiation; Fig. 5a). We found that mostly lipid features significantly (FDR < 5%) differed between knockout and control at

https://doi.org/10.1038/s42255-023-00807-w

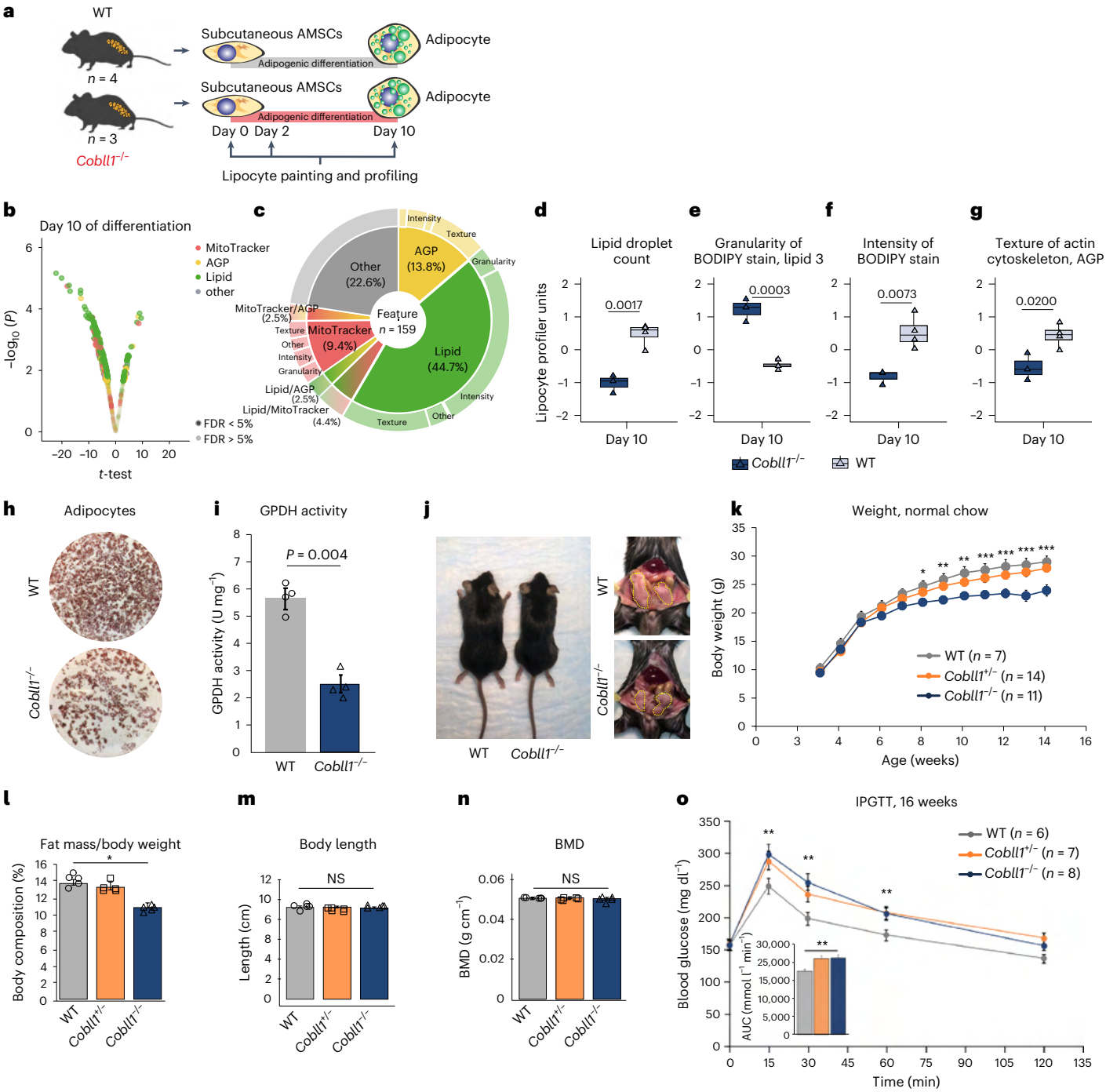


Fig. 5 | *Cobll1*-deficient mice are leaner and display metabolically dysfunctional phenotypes. a, Schematic of differentiation and LipocyteProfiling at three time points (days 0, 2 and 10) of AMSCs derived from *Cobll1*$^{-/-}$ mice ($n = 3$) and WT mice ($n = 4$). **b**, Morphological profiles of the AMSCs of *Cobll1*$^{-/-}$ mice compared to the AMSCs of WT mice at day 10; two-sided *t*-test, significance level FDR ≤ 5%. **c**, Pie chart illustrating non-redundant differential features per channel and class of measurement comparing the AMSCs of *Cobll1*$^{-/-}$ and WT mice at day 10 of differentiation. **d**–**g**, Lipid droplet count (Cells_Children_LargeLipidObjects_Count) (**d**), granularity of BODIPY stain (Cells_Granularity_3_BODIPY) (**e**), intensity of BODIPY stain (Cells_Intensity_IntegratedIntensity_Lipid) (**f**) and texture of actin cytoskeleton (Cytoplasm_Texture_Entropy_AGP) (**g**) at day 10 of differentiation. Two-sided *t*-test; data represent the median ± 95% CI. **h**, Oil Red O staining of differentiated murine AMSCs. **i**, GAPDH activity of differentiated murine AMSCs was assessed by measuring the decrease in NADH at 340 nm. Data represent the mean ± s.e.m. $^{*}P < 0.05$ compared to the WT group. **j**, Representative photograph of 14-week-old WT and *Cobll1*$^{-/-}$ mice fed a normal chow. The yellow dotted lines delineate perigonadal WAT. **k**, Mouse body weight across time. Data are expressed as the mean ± s.e.m. $^{*}P < 0.05$ and $^{**}P < 0.01$ compared to the WT group. **l**, Body composition (fat mass/body weight). Data are expressed as the mean ± s.e.m. $^{*}P < 0.05$ compared to the WT group. **m**, Body length measurements of WT and *Cobll1*$^{-/-}$ mice ($n = 6$). Data are expressed as the mean ± s.e.m. $^{*}P < 0.05$ compared to the WT group. NS, not significant. **n**, BMD analysed by dual energy X-ray absorptiometry (DEXA). Data are expressed as the mean ± s.e.m. $^{*}P < 0.05$ compared to the WT group. **o**, IPGTT in WT and *Cobll1*$^{-/-}$ and *Cobll1*$^{+/-}$ mice. The inset graph shows the area under the curve (AUC) of the blood glucose concentration levels measured during IPGTT. Data represent the mean ± s.e.m. $^{**}P < 0.01$ compared to the WT group. In **i**,**k**–**o** statistical significance was determined by Student's *t*-test. In **l**–**o** the experiment was repeated independently three times with similar results.

day 10 of differentiation (Fig. 5b,c and Extended Data Fig. 7c–e). More specifically, the data revealed that the AMSCs of *Cobll1* knockout mice showed fewer and smaller lipid droplets (Lipid object_count $P = 0.0017$, Lipid_Granularity size measure 3 $P = 0.0003$, Fig. 5d,e), lower lipid intensity ($P = 0.0073$; Fig. 5f) and, in line with the lipid-related observation, decreased actin cytoskeleton-related heterogeneity across the cytoplasm (Cytoplasm_Texture_Entropy_AGP, $P = 0.0200$; Fig. 5g). These findings indicate that *Cobll1* knockout in mice affects actin cytoskeleton remodelling and lipid accumulation during in vitro adipocyte differentiation, mimicking our observations in human adipocytes. Indeed, when examining the effect of *Cobll1* knockout adipocytes on lipid accumulation using Oil Red O, we observed fewer differentiated adipocytes in *Cobll1*$^{-/-}$ compared to wild-type (WT) cells (Fig. 5h). We also observed significantly lower GAPDH activity, an indicator of adipocyte differentiation, in *Cobll1*$^{-/-}$ mice compared to their WT littermates ($P = 0.004$) (Fig. 5i), suggesting that the ablation of *Cobll1*$^{-/-}$ leads to impaired adipogenesis, supporting our finding in human adipocytes.

To assess the impact of the *2q24.3* MONW locus effector COBLL1 on organismal processes, we assayed for growth and body composition phenotypes in *Cobll1*$^{-/-}$ mice. At 10 weeks of age, *Cobll1*$^{-/-}$ homozygous animals displayed 20–25% less weight gain compared to the WT control and *Cobll1* heterozygous (*Cobll1*$^{+/-}$) littermates (Fig. 5j,k), reflecting a significant reduction in total fat mass percentage (3–5%), but with no difference in body length or in bone mineral density (BMD), suggesting that the phenotype of *Cobll1*$^{-/-}$ is due to reduced fat mass (Fig. 5l–n). Next, we examined glucose homeostasis by performing intraperitoneal glucose tolerance tests (IPGTTs). *Cobll1*$^{-/-}$ mice displayed impaired glucose tolerance compared to WT and heterozygous littermates (Fig. 5o). In conclusion, the phenotypic characteristics of the *Cobll1* knockout mouse model recapitulate the MONW association patterns observed in humans and demonstrate how abrogation of *Cobll1* links molecular and cellular phenotypes to organismal-level metabolic phenotypes associated with genetic variation in the *2q24*.3 locus in humans.

Discussion

The *2q24.3* locus is pleiotropic in nature and, intriguingly, is associated with increased risk of T2D and simultaneously with decreased body fat percentage, reminiscent of a MONW phenotype association signature. In this study, we applied a series of experimental and computational approaches to systematically dissect the *2q24.3* metabolic risk locus and link it to a causal variant (rs6712203), its effector gene (*COBLL1*), its causal cell type, cell context (developmental time point, adipose depot) and cellular mechanisms (actin remodelling). These altered cellular functions, that is, adipocyte differentiation into metabolically active subcutaneous adipocytes, lipid metabolism and insulin-responsive glucose uptake, are relevant for T2D, and body fat percentage and distribution. When ablating *Cobll1* in mice, we showed a lipodystrophy-like phenotype, recapitulating the pleiotropic association with T2D and decreased body fat mass in humans. Thus, we provide genetic and mechanistic evidence that a common genetic variant limits peripheral energy storage capacity and simultaneously impairs insulin responsiveness.

The results of this study lend support to the common hypothesis that the individual risk of T2D and fasting insulin is modified by changes to the mass, distribution and function of adipose tissue[23,47], and that a metabolically healthy state is largely dependent on SAT expandability. Inherited and acquired lipodystrophies, as characterized by the selective or global perturbation of adipose tissue function, mass and distribution, result in severe forms of insulin resistance and diabetes; shared molecular mechanisms between rare familial partial lipodystrophy type 1 and common forms of insulin resistance at the genetic level have been suggested previously[23]. Several common metabolic risk loci are characterized by a MONW and MHO association; distinct association signatures suggest that multiple mechanisms are involved, most of which are to be identified[21,22,48]. Previous work has convincingly implicated variants at the *FAM13A* locus to affect metabolic disease risk by affecting subcutaneous adipocyte differentiation[48]. In this work, we implicate for the first time actin cytoskeleton remodelling as a critical factor for subcutaneous adipocyte function and as causally involved in metabolic disease progression in humans, stressing the notion that MONW and MHO predisposing loci control distinct cellular programmes.

We observed evidence of sex-dimorphic effects when conditioning MONW traits on rs6712203, which is in line with a reported sexual dimorphism for WHR consistently conveying stronger effects in women[15,26,49,50] and a sex-independent effect on T2D[11–14], and with a sex-dimorphic effect on gene expression for *COBLL1* but not for *GRB14* (ref. 19). While we did not observe a phenotypic effect of Cobll1 knockout in female mice, we cannot exclude the possibility that this is due to well-known sexual dimorphism in mouse metabolic phenotypes[51]. Q15

The COBLL1 protein has been introduced as a biomarker of high prognostic value for different types of cancer, a modulator of cell morphology in prostate cancer[42], and lipid metabolism and insulin signalling in adipocytes[52]. In this study, we established chain of causation linking the *2q24.3* locus to its functional variant, its adipocyte cell type-specific and context-specific effects, its regulatory element, its effector gene *COBLL1* and finally its causal cellular function, that is, actin remodelling in differentiating adipocytes, which is under the genetic control of both locus and target gene. These findings are in line with recent reports linking actin dynamics, regulated by the F-actin:G-actin ratio, and insulin-stimulated trafficking and fusion of GLUT4 vesicles[53–55]. Actin remodelling also occurs in hypertrophically differentiated adipocytes[55,56], an essential adaptive mechanism of post-mitotic cells, such as mature adipocytes, and was recently linked to cellular senescence in states of obesity and insulin resistance[57]. An unresolved question to be addressed in future work is whether COBLL1-mediated actin dynamics contributes to metabolic functions of certain adipocyte subpopulations with different degrees of insulin sensitivity as recently identified by single-cell and spatial transcriptomics[58,59]. Furthermore, while we did not observe a haplotype-dependent effect on actin cytoskeleton remodelling in visceral adipocytes, further studies are warranted to elucidate how actin depolymerization and repolymerization during the differentiation process may differ in subcutaneous compared to visceral adipocytes.

Besides *COBLL1*, the insulin receptor adaptor protein GRB14 has an intuitive effector target gene at the *2q24.3* locus. GRB14 is a negative regulator of insulin signalling[60,61]. In line with our findings, physiological studies in *GRB14* knockout mice show increased glucose tolerance and insulin sensitivity linked to higher IRS1 and Akt phosphorylation in liver and skeletal muscle but not adipose tissue[62], which may rely more on regulation of proximal insulin signalling by GRB10 (refs. 62–65). Somewhat contradictory to these findings and to the results reported in this present work, a recent study reported that CRISPR–Cas9-mediated ablation of GRB14 in human SGBS preadipocytes resulted in decreased preadipocyte proliferation, adipocyte differentiation and adipocyte insulin responsiveness[52]. This contradiction might be explained by the fact that different perturbation schemes were applied, that is, small hairpin RNA (shRNA)-mediated silencing of GRB14 as described in this study compared to CRISPR–Cas9 knockout of GRB14, which may be confounded by the number of cells in a population that do not produce knockout phenotypes[66] (González et al.[67]); also partial loss-of-function or gain-of-function phenotypes can be generated by Cas9-induced in-frame insertion and deletions, and hypomorphic alleles[67], which can obscure the readout. *COBLL1* as the effector gene that underlies the T2D association is further corroborated by recent exome chip and exome sequencing studies linking an rs7607980 coding variant in the *COBLL1* gene to T2D and glycaemic traits (minor allele frequency (MAF) = 0.12, $P = 4.7 \times 10^{-11}$)[68–70]. Furthermore, recent rare variant aggregation analyses at *COBLL1* revealed nominal association with WHR[71] but not with GRB14, which agrees with our findings that *COBLL1* drives at Q16 Q17

least part of the *2q24.3* genetic risk. Our sequence-based predictive models scored rs6712203 highest across all *2q24.3* haplotype variants; our experimental data provide evidence that the SNP rs6712203 acts as a causal regulatory variant which, at least in part, drives effects at the *2q24.3* MONW locus tagged by rs3923113. Notably, our sequence-based models predict additional variants at the locus to affect regulatory activity; at least two additional variants, rs12692737 and rs10179126, are associated with allelic chromatin accessibility bias. Therefore, while beyond the scope of our study, we note that our work does not preclude multiple variants acting in concert at this locus in a cell state-dependent manner, potentially implicating *GRB14* along with *COBLL1* as effector genes. Furthermore, besides rs3923113, recent large-scale genetic association studies identified multiple additional signals tagged by other lead SNPs in reasonably high LD that associate with metabolic disease[13,25]; also, the effect of indels should be evaluated in future studies.

In conclusion, the *2q24.3* locus is a prime example of a common genetic locus that predisposes to limited peripheral adipose storage capacity and insulin resistance, driven by an impairment of dynamic actin cytoskeleton remodelling of the differentiating subcutaneous adipocyte. Our study epitomizes the potential of translating pleiotropic genetic loci into mechanistic understanding towards therapeutic hypotheses.

Methods

Human primary AMSCs isolation and needle biopsies

We obtained AMSCs from the adipose tissue of patients undergoing a range of abdominal laparoscopic surgeries (sleeve gastrectomy, fundoplication or appendectomy). VAT was derived from the proximity of the angle of His and SAT was obtained from beneath the skin at the site of surgical incision. Human liposuction material was obtained from a private plastic surgery clinic. Isolation of AMSCs was performed as described previously[72,73]. Abdominal specimens from SAT were obtained by needle aspiration before and 4 weeks after a very-low-calorie diet (800 kcal). Study design and the abdominal needle aspiration procedure were previously described in Ott et al.[74].

Each participant gave written informed consent before inclusion and the study protocol was approved by the ethics committee of the Technical University of Munich (147 study no. 5716/13).

Differentiation of human AMSCs

For imaging, cells were seeded at 10,000 cells per well in 96-well plates (high-content imaging; CellCarrier plates, catalogue no. 6005550; PerkinElmer) or seeded at 18,000 cells per well in collagen IV-coated 8-well µ-slides (higher-resolution imaging; catalogue no. 80822, ibidi) and induced 4 d after seeding. For RNA sequencing (RNA-seq), cells were seeded at 40,000 cells per well in 12-well dishes (Corning). Before induction, cells were cultured in proliferation medium (DMEM/F-12, 1% penicillin-streptomycin, 33 µM biotin, 17 µM pantothenate supplemented with 0.13 µM insulin, 0.01 µg ml^{-1} epidermal growth factor, 0.001 µg ml^{-1} fibroblast growth factor, 2.5% FCS). Subcutaneous and visceral adipogenic differentiation was induced by changing culture medium to induction medium (basic medium supplemented with 0.861 µM insulin, 1 nM T3, 0.1 µM cortisol, 0.01 mg ml^{-1} transferrin, 1 µM rosiglitazone, 25 nM dexamethasone and 2.5 nM IBMX; for visceral AMSCs, 0.1 mM oleic and linoleic acid and 2% FCS were also used). On day 3 of adipogenic differentiation, culture medium was changed to differentiation medium (basic medium supplemented with 0.861 µM insulin, 1 nM T3, 0.1 µM cortisol, 0.01 mg ml^{-1} transferrin; for visceral AMSCs, 0.1 mM oleic and linoleic acid and 2% FCS were also used.

Genotyping

Genotyping was performed using the Illumina Global Screening beadchip array. DNA was extracted using QIAGEN's DNeasy Blood and Tissue Kit (catalogue no. 69504) and sent to the Oxford Genotyping Centre for genotyping on the Infinium HTS assay on Global Screening Array bead-chips (Illumina). Genotype quality control (QC) was done using GenomeStudio and genotypes were converted into PLINK format. We checked sample missingness but found no sample with missingness greater than 5%. For the remaining sample QC steps, we reduced the genotyping data down to a set of high-quality SNPs. These SNPs were: (1) common (MAF > 0%); (2) missingness < 0.1%; (3) independent, pruned at an LD (R^2) threshold of 0.2; (4) autosomal only; (5) outside the lactase locus (chromosome 2), the major histocompatibility complex (chromosome 6), and outside the inversions on chromosomes 8 and 17; and (6) in Hardy–Weinberg equilibrium ($P > 1 \times 10^{-3}$). Using the remaining approximately 65,000 SNPs, we checked samples for inbreeding (--het in PLINK), but found no samples with excess homozygosity or heterozygosity (no sample more than 6 s.d. from the mean). We also checked for relatedness (--genome in PLINK) and found one pair of samples to be identical; we kept the sample with the higher overall genotyping rate. Finally, we performed principal component analysis using EIGENSTRAT and projected the samples onto data from HapMap 3, which includes samples from 11 global populations. Six samples had some non-European ancestral background, while most samples were of European descent. We did not remove any samples at this step, selecting to adjust for principal components (PCs) in genome-wide testing. Adjustment for PCs failed to eliminate population stratification; therefore, we restricted the process to samples of European descent only, defined as samples falling within ±10 s.d. of the first and second PCs of the CEU (Utah residents with Northern and Western European ancestry) and TSI (Tuscans in Italy) samples included in the HapMap 3 dataset.[42,43] Finally, sex information was received after the initial sample QC was complete. One sample with potentially mismatching sex information (comparing genotype and phenotype information) was discovered after analyses were complete and therefore remained in the analysis.

SNP QC. We removed monomorphic SNPs and all SNPs with missingness greater than 5% and out of Hardy–Weinberg equilibrium ($P < 1 \times 10^{-6}$). We set heterozygous haploid sites to missing to enable downstream imputation. The final cleaned dataset included 190 samples and approximately 700,000 SNPs.

Genotype imputation. We performed imputation via the Michigan Imputation Server as described in Glastonbury et al.[75]. Briefly, we aligned SNPs to the positive strand and then uploaded the data (in VCF format) to the server. We imputed the data using the Haplotype Reference Consortium panel. We selected EAGLE as the phasing tool to phase the data. To impute chromosome X, we followed the server protocol for imputing this chromosome (including using SHAPEIT to perform the phasing step).

ATAC–seq in immortalized AMSCs

ATAC–seq was performed by adapting the protocol from Buenrostro et al.[76] by adding a nucleus preparation step. Differentiating cells were lysed directly in the culture plate before adipogenesis was induced (days 0, 3, 6 and 14). Ice-cold lysis buffer was added directly onto cells grown in a 12-well plate. Plates were incubated on ice for 10 min until cells were permeabilized and nuclei released. Cells in lysis buffer were gently scraped off the well and transferred into a chilled 1.5-ml tube to create crude nuclei. Nuclei were spun down at 600*g* for 10 min at 4 °C. Nuclei pellets were then resuspended in 40 µl Tagmentation DNA Buffer (catalogue no. FC-121-1031, Nextera) and the quality of the nuclei was assessed using trypan blue. The volume of 50,000 nuclei was determined using a haemocytometer. The transposition reaction was performed as described previously[76]. All tagmented DNA was PCR-amplified for eight cycles using the following PCR conditions: 72 °C for 5 min, 98 °C for 30 s, followed by thermocycling at 98 °C for 10 s, 63 °C for 30 s and 72 °C for 1 min. The quality of the ATAC–seq libraries was assessed using a Bioanalyzer High Sensitivity Chip (Applied Biosystems). All libraries had a mean fragment size of

approximately 200 bp and characteristic nucleosome patterning, indicating good quality. Libraries were pooled and sequenced on a HiSeq 4000 system (Illumina), generating 50 million 75-bp paired-end reads per sample. To reduce bias due to PCR amplification of libraries, duplicate reads were removed. Sequencing reads were aligned to the hs37d5 reference genome; the Burrows–Wheeler Aligner-MEM was used for mapping. All experiments were performed in technical duplicates.

LipocyteProfiling in human and mouse AMSCs

The full LipocyteProfiler protocol is described in Laber et al.[43]. Briefly, human and mouse AMSCs were plated and differentiated in 96-well CellCarrier plates for high-content imaging at days 0, 3, 8 and 14 of differentiation. On the respective day of the assay, cell culture medium was removed and replaced with 0.5 µM MitoTracker staining solution (1 mM MitoTracker Deep Red stock (catalogue no. M22426, Invitrogen) diluted in culture medium) to each well followed by 30-min incubation at 37 °C protected from light. After 30 min, the MitoTracker staining solution was removed and cells were washed twice with Dulbecco's PBS (1×) (catalogue no. 21-030-CV, Corning) and 2.9 µM BODIPY staining solution (3.8 mM BODIPY 505/515 stock (catalogue no. D3921, Thermo Fisher Scientific) diluted in Dulbecco's PBS) was added followed by 15-min incubation at 37 °C protected from light. Subsequently, cells were fixed by adding 16% methanol-free paraformaldehyde (PFA) (catalogue no. 15710-S, Electron Microscopy Sciences) directly to the BODIPY staining solution to a final concentration of 3.2% and incubated for 20 min at room temperature protected from light. PFA was removed and cells were washed once with Hank's Balanced Salt Solution (HBSS) (1×) (catalogue no. 14025076, Gibco). To permeabilize cells, 0.1% Triton X-100 (catalogue no. X100, Sigma-Aldrich) was added and incubated at room temperature for 10 min protected from light. After permeabilization, multi-stain solution (10 U Alexa Fluor 568 phalloidin (catalogue no. A12380, Thermo Fisher Scientific), 0.01 mg ml^{-1} Hoechst 33342 (catalogue no. H3570, Invitrogen), 0.0015 mg ml^{-1} WGA, Alexa Fluor 555 conjugate (catalogue no. W32464, Thermo Fisher Scientific), and 3 µM SYTO 14 Green Fluorescent Nucleic Acid Stain (catalogue no. S7576, Invitrogen) diluted in HBSS) was added and cells were incubated at room temperature for 10 min protected from light. The staining solution was removed and cells were washed three times with HBSS. Cells (25 fields per well) were imaged with an Opera Phenix High-Content Screening System.

Staining and microscopy

To stain the actin cytoskeleton, COBLL1 and nuclei, cells were washed twice with ice-cold PBS and fixed with PFA Roti-Histofix 4% (Roth) for 15 min. Cells were washed twice with ice-cold PBS for 5 min and incubated with ice-cold 0.1% Triton X-100/PBS (Roth). Cells were washed again twice with PBS and incubated for 1 h at room temperature with 4% BSA, then incubated with 1:100 primary COBLL1 antibody (catalogue no. HPA053344, Atlas Antibodies, lot no. 000010746) overnight at 4 °C, followed by 1 h at room temperature (final concentration = 2 µg ml^{-1}). Cells were washed twice with PBS and stained with 0.46% bisBenzimide H 33258 (Sigma-Aldrich), 1% phalloidin–Atto 565 (Sigma-Aldrich) and a secondary antibody against COBLL1 (1:200 Alexa Fluor 488, Abcam) at a final concentration of 0.01 mg ml^{-1}). Cells were incubated for 1 h at room temperature in the dark. Afterwards, cells were washed twice with PBS for 5 min and kept in PBS at 4 °C until imaging. Images were acquired on a Leica DMi8 microscope (Leica Microsystems) using the HC PL APO 63×/1.40 oil objective. Images were processed with the Leica LASX software.

LipocyteProfiler

The full LipocyteProfiler protocol is described in Laber et al.[43]. Quantitation was performed using CellProfiler v.3.1.9. Before processing, flat-field illumination correction was performed using functions generated from the mean intensity across each plate. Nuclei were identified using with 4,6-diamidino-2-phenylindole stain and then expanded to identify whole cells using the AGP and BODIPY stains. Regions of the cytoplasm were determined by removing the nuclei from the cell segmentations. Speckles of BODIPY staining were enhanced to assist in the detection of small and large individual BODIPY objects. For each object set, measurements were collected representing size, shape, intensity, granularity, texture, colocalization and distance to neighbouring objects. After the feature extraction, data were filtered by applying automated and manual QC steps. First, fields with a total cell count less than 50 cells were removed. Second, fields that were corrupted by experimentally induced technical artefacts were removed by applying a manually defined QC mask. Furthermore, blocklisted features known to be noisy and generally unreliable were removed. After filtering, data were normalized per plate using a robust scaling approach that subtracts the median from each variable and divides it by the interquartile range. For each individual, wells were aggregated for downstream analysis by cell depot and day of differentiation. Subsequent data analyses were performed in R v.3.6.1 using base packages unless noted. For dimensionality reduction visualization, uniform manifold approximation and projection (UMAP) maps were created using the UMAP R package (https://github.com/lmcinnes/umap) with default settings.

To test for a difference in morphological profiles at any day of differentiation due to *COBLL1* knockdown, individuals were analysed separately using a *t*-test. To test for a difference in morphological profiles at any day of differentiation between risk and non-risk haplotypes, a multi-way ANOVA was performed. Differences in morphological profiles between TT ($n = 7$) and CC ($n = 6$) allele carriers were adjusted for sex, age, BMI and batch. To overcome multiple-testing burden, *P* values were corrected using the FDR described in the R package qvalue (https://github.com/StoreyLab/qvalue). Features with an FDR < 5% were classified as significant and filtered based on redundancy and effect size.

COBLL1 silencing using siRNA

All silencing experiments were performed on four technical replicates. One day before silencing, AMSCs were plated into 96-well plates with 10,000 cells per well or collagen IV-coated 8-well glass µ-slides with 18,000 cells per well using growth medium. RNA-based silencing of *COBLL1* was performed using the RNAiMAX Reagent (catalogue no. 13778075, Thermo Fisher Scientific) according to the manufacturer's protocol. Briefly, Lipofectamine RNAiMAX Transfection Reagent was diluted in Opti-MEM medium (catalogue no. 11058021, Gibco). At the same time, siRNA was diluted in Opti-MEM medium. Then, diluted siRNA was added to the diluted Lipofectamine RNAiMAX reagent at a 1:1 ratio and incubated for 5 min. Coated 8-well glass µ-slides were incubated for 20 min at room temperature. The concentration of reagents per well in a 96-well plate were 0.5 µl (10 µM) of silencing oligonucleotides (catalogue no. 4392420, Ambion) or negative control duplex (catalogue no. 4390846, Ambion) and 1.5 µl Lipofectamine RNAiMAX Reagent. The plate was gently swirled and placed in a 37 °C incubator at 5% CO_2 for 3 d. Cells were then induced to differentiate using a standard differentiation cocktail or collected for gene expression analysis to assess knockdown efficiency.

RNA preparation and quantitative PCR

Total RNA was extracted with TRIzol (catalogue no. 15596026, Ambion) and the Direct-zol RNA MiniPrep Kit (catalogue no. R2052, Zymo Research) according to the manufacturer's instructions. Complementary DNA (cDNA) was synthesized with the High-Capacity cDNA Reverse Transcription Kit (catalogue no. 4368814, Applied Biosystems) according to the manufacturer's instructions. Quantitative PCR (qPCR) was performed using the PCR Master Mix (catalogue no. K0172, Thermo Fisher Scientific) and TaqMan probes for the target gene *COBLL1* (catalogue no. 4448892, ID: Hs01117513_m1, Thermo Fisher Scientific) and

the housekeeping gene *CANX* (catalogue no. 4448892, ID Hs01558409_m1, Thermo Fisher Scientific). Relative gene expression was calculated by the $2-^{\Delta\Delta_{Ct}}$ method. Target gene expression was normalized to the expression of *CANX*.

RNA-seq and splicing analysis

RNA-seq reads were trimmed using SeqPurge with the following command: SeqPurge -a1 CTGTCTCTTATACACATCTCCGAGCCCACGAGAC -a2 CTGTCTCTTATACACATCTGACGCTGCCGACGA.

For the transcript-level quantification, trimmed reads were analysed with Kallisto (with 25 bootstraps) and the transcripts per million (TPM) estimates were log-transformed; the top 10 PCs were computed. Next, reads were summed across all transcripts of a given gene to obtain gene-level estimates of the expression in each sample.

For the splicing analysis with Leafcutter, reads were mapped with STAR using the following parameters: STAR--twopassMode Basic--outSAMstrandField intronMotif--readFilesCommand zcat--outSAMtype BAM Unsorted.

They were then processed using Samtools and RegTools to convert to a JUNC file with the following command: regtools junctions extract -s 1 -a 8 -m 50 -M 500000.

Finally, reads were clustered into splicing events with the following command from the Leafcutter project: leafcutter_cluster_regtools.py -j <files> -m 50 -l 500000. These clusters were then converted to TPM and modelled as a function of the rs6712203 genotype.

RNA pathway enrichment analysis

Transcript-level (log) RNA expression was compared between *COBLL1* and all other quantified genes using linear regression. The effect of *COBLL1* on other genes was compared and adjusted for the expression PCs, sample depot source, cell line and day of differentiation. This resulted in effect sizes of individual genes in terms of how similar they were to *COBLL1*; those with estimates that had a Bonferroni-adjusted $P > 1 \times 10^{-13}$ and an absolute effect size less than 0.1 or greater than 10 were excluded. Similarly expressed genes with strong association with *COBLL1* were uploaded to Enrichr and analysed as a gene list against the KEGG, and the WikiPathways and HCI pathways. The full set of tests is available at https://maayanlab.cloud/Enrichr/enrich?dataset=1a9a07019bfd8bbddc6eb6c26641bfef; the sensitivity evaluation in which very-lowly-expressed and highly-expressed genes were not excluded are available at https://maayanlab.cloud/Enrichr/enrich?dataset=231b12708d04818007d93364e489fab7.

PMCA variant conservation analysis

The PMCA results were replicated from Claussnitzer et al.[29]. Briefly, transcription factor binding sites and their cooccurrence across species were tallied and classified into complex and non-complex regions. Complex regions were counted on the basis of motifs aligned across species and were then plotted against Basset scores to discover putative causal variants.

Basset variant effect prediction analysis

Basset models were trained and evaluated as in Sobreira et al.[77]. Models were trained to capture chromatin regulation relevant to adipocyte differentiation and a number of other metabolic tissues; these effects were estimated by determining the difference in effect between alleles at each variant. The variants with the largest effect on accessibility were considered the most important and most likely to be causal.

Allele-specific accessibility analysis

Allele-specific analyses were performed as in Sinnott-Armstrong et al.[78]. Briefly, reads were aligned from a heterozygous individual on the basis of the variant; the number of reads supporting each allele were tallied at each time point and across variants on the haplotype.

Conditional and BMI-dependent variant association analysis

Variants (n = 6,167) within 100 kb of rs6712203 were included in the analysis. White British individuals in the UKB were analysed with phenotype type 2 diabetes (as described in Eastwood et al.[79]), log WHR adjusted for BMI, hip circumference and whole-body fat mass. Individuals were stratified on the basis of reported sex and filtered to White British unrelated individuals as described in Sinnott-Armstrong et al.[80]. Conditional analyses and all associations were performed using PLINK 2.0.

POU2F2 affinity modelling using the IGR method

The IGR method was used for POU2F2 affinity modelling using the POU2F2 ChIP–seq data[35]. To correct for systematic bias in the sequencing depth around particular *k*-mers, all scores were offset by a baseline value, defined as the average signal between the forward and reverse complement instances of the *k*-mer between −200 and −195 and between 195 and 199 bases away from the *k*-mer centre. To include only large-effect binding differences, the prominence was defined as the maximum score across any point in the context of either the forward or reverse complement version of the *k*-mer for both alleles and the maximum difference as the maximum absolute difference in scores between the two alleles at any point in the window. The baseline ratio was defined as the ratio of the maximum difference to the prominence, which varied between 0 (if the two alleles were equal at all points) and 2 (if they were perfectly complementary at their highest absolute point).

To map high-quality putative disrupted binding sites, the *k*-mer sequence that gave the highest affinity under the germline was recorded as reference and the *k*-mer sequence that gave the highest affinity under the somatic variant as alternate. The quality of a given *k*-mer was defined as the correlation between the average context plot forward and the reverse of the average context plot of the reverse complement; the symmetry of a given *k*-mer was the correlation between the average context plot forward and the average context plot reverse. Quality was high when the antiparallel binding was preserved, and symmetry was high when the peak signal was centred with respect to the variant. The results were included as passed when the Bonferroni-corrected *P* value for the comparison was less than 0.05, the baseline ratio was greater than 0.5, quality and symmetry were both greater than 0.85 for one of the alleles, and quality and symmetry were both greater than 0.5 for the other allele.

Microarray expression data

A global gene expression measurement was performed, using Illumina HumanRef-8 v.3 BeadChip microarrays from whole abdominal SAT. Signal intensities were quantile-normalized before the correlation analysis.

SGBS genome editing

To edit the rs6712203 heterozygous allele in SGBS preadipocytes to the homozygous risk (CC) and non-risk (TT) alleles, we applied the CRISPR–Cas9 homology-directed repair genome editing approach. The hCas9 vector was purchased from Addgene (plasmid no. 41815). The guide sequence was selected using the design tool (Zhang Lab, MIT) with a predicted number of 228 potential off-target sites, located 211 bp upstream of rs6712203. It was cloned in front of the U6 promoter into the BbsI cloning site of the sgRNA expression vector (R. Kühn, Helmholtz Zentrum München), using the double-stranded oligonucleotides 5′-CACCGACTCTCCACTACCATTGCCA-3′ and 5′-AAACTGGCAATGGTAGTGGAGAGTC-3′. To amplify the 2,009-bp homology region with the risk or non-risk allele of rs6712203 at mid-position, the genomic DNA (gDNA) of SGBS cells was amplified with the primers 5′-GGTGGTCCCATTAAAAAGAAAGAAGCTTGG-3′ and 5′-CTTCTCTTTTACCCTGCTGGCTACTGGTTG-3′ using High-Fidelity Q5 DNA polymerase (New England Biolabs). The gel-purified PCR

product was cloned into the blunt end pJet1.2 vector using the CloneJET PCR Cloning Kit (Fermentas). A clone with the rs6712203 C allele was selected and the corresponding T allele vector was generated using the Q5 Site-Directed Mutagenesis Kit (New England Biolabs) with the primers 5′-TCATTCATCATATGCAATTCTGG-3′ and 5′-GGCAAATTAATATTTAGGATTATATC-3′. To avoid Cas9 reactivity after genome editing, the NGG guide target sequence was mutated to NCG in both homology vectors with the primers 5′-CCATTGCCAACGGCTGAGTCAG-3′ and 5′-TAGTGGAGAGTTCTCACAAAAC-3′. SGBS cells were cotransfected with green fluorescent protein (Lonza), hCas9, the respective sgRNA and pMACS v.4.1 (Miltenyi Biotec) plasmids using the Amaxa Nucleofector device (programme U-033) (Lonza). Cells were sorted using the MACSelect Transfected Cell Selection Kit (Miltenyi Biotec). The integrity of each edited vector construct and the SGBS cell nucleotide exchange was confirmed by DNA sequencing (Eurofins).

Lentiviral SGBS cell transduction

The MISSON Lentiviral Packaging Mix (Sigma-Aldrich) was used according to the manufacturer's instructions. Packaging HEK 293T cells were grown in a low-antibiotic growth medium (DMEM, 10% FCS, 0.1% penicillin-streptomycin). When cells were approximately 70% confluent, they were cotransfected using X-treme GENE HP (Roche), with the packaging plasmid pCMVdeltaR8.91, the envelope plasmid pMD2.G and the pLKO-based plasmid containing shRNA against the human target gene *COBLL1* (NM_014900.2-3071s1c1), *COBLL1* (NM_014900.2-4440s1c1), *GRB14* (NM_004490.1-1243s1c1) or empty vector MISSION TRC2 pLKO.5-puro plasmid (Sigma-Aldrich). Cells were incubated for 24 h, and the medium was discarded and replaced with a serum-rich medium (30% FCS). The supernatant containing the viable virus particles was collected 48 and 72 h after transfection, centrifuged to remove cellular debris and stored at −80 °C.

SGBS cells were seeded at a concentration of 2.6×10^4 cells per 6-well plate and grown in normal growth medium. After 24 h the medium was replaced and supplemented with 8 μg ml^{-1} Polybrene (Sigma-Aldrich) and virus supernatant with a multiplicity of infection of 2. On consecutive days, cells were washed with PBS and medium was replaced to remove the virus. The medium was supplemented with 0.5 μg ml^{-1} puromycin 96 h after infection to select stable clones. When cells were grown confluent, puromycin was removed from the medium and cells were differentiated until day 16. Target gene silencing was confirmed by qPCR with reverse transcription (RT–qPCR).

Generation of dCas-hWAT preadipocyte cell line

We generated an hWAT cell line, an established white adipocyte line[33], stably expressing dCas9-KRAB. Lentivirus was delivered to preadipocytes using spinfection. For single-copy integration, the virus titre was optimized to reach an integration efficiency of 30-40%. Four days after infection, we positively selected for cells expressing the integrated transgene using blasticidin. We validated the generated line by assessing dCas9-KRAB activity delivering a sgRNA targeting CD81, stably expressed in adipocytes, and measuring CD81 gene expression.

CRISPRi

sgRNAs targeting TSS were designed using CRISPick (https://portals.broadinstitute.org/gppx/crispick/public) and sgRNAs targeting non-coding regions were designed with assistance from the Genetic Perturbation Platform at the Broad Institute. Guide sequences are as follows: non-targeting control 1: AAAAAGCTTCCGCCTGATGG; non-targeting control 2: CGTCCCTTCGTCTCTGCTTA; non-targeting control 3: TCACCTCCGAACGAACACCT; non-targeting control 4: GGACGCACCATTCCGGGTGA; COBLL1_1: AGCTCGCCTGTTCTCCCTCG; COBLL1_2: GAGTAGGAGAGGAAGCCGCG; COBLL1_3: AGAGCCGAGCGGCAAGAGCG; COBLL1_4: GAGCCGAGCGGCAAGAGCGC; GRB14_1: ACCACACCTGCAGAGCGCTC; GRB14_2: GCCCGAGCGCTCTGCAGGTG; GRB14_3: CGAGCGCTCTGCAGGTGTGG; GRB14_4: CACCACACCTGCAGAGCGCT; RS4: AGAGTTCTCACAAAACTCCA; RS9: CTCTCCACTACCATTGCCAA; RS20: CCATTGCCAAGGGCTGAGTC; RS1: CATCACATGCAATTCTGGCA; RS8: AATACAAATAGAGCAAGTAT; RS11: TTGTAAGTGGAAAAAAAGCT. sgRNAs were cloned into the BsmBI site of a lentiviral pXPR_050 vector (plasmid no. 96925, Addgene)[81] and verified using Sanger sequencing.

Lentiviral particles were generated by transfecting HEK 293T cells (8×10^5 per well of a 6-well plate seeded the day before transfection) with 1,250 ng sgRNA, 1,250 ng psPAX2 and 250 ng vesicular stomatitis virus G vectors using a TransIT-LT1 transfection reagent at 8.5 μl per well (Mirus Bio). Viral particles were collected 72 h after transfection, passed through a 0.45-μm filter, aliquoted and stored at −80 °C until use. For lentiviral transduction by spinfection, cells were seeded at 1×10^5 per well of 12-well plates in the presence of 50, 250 or 500 μl viral supernatant and Polybrene (4 μg ml^{-1}) and spun at 930 *g* for 2 h. Cells were selected with 1 μg ml^{-1} puromycin for 3 d, split once and collected 6 d after transduction for RNA extraction and gene expression by qPCR. Cas9 expression and *COBLL1* knockdown were verified by western blot using specific antibodies (Cas9: catalogue no. C15310258, Diagenode; COBLL1: catalogue no. HPA053344, Sigma-Aldrich). Ponceau-S staining or β-actin (catalogue no. 4970, Cell Signaling Technology).

Measurement of GAPDH activity

Cells were grown to confluence and differentiated until day 16 in 6-well plates. Cells were collected in a GPADH buffer with 0.05 M Tris-HCl (pH 7.4), 1 mM EDTA and 1 mM mercaptoethanol before they were stored at −80 °C until further use. Samples were gently defrosted at 4 °C and sonified for 7 s at 29%, and centrifuged for 10 min at 10,000*g* at 4 °C. GAPDH activity was measured as described previously[82]. Briefly, GAPDH activity was assessed, measuring the conversion of dihydroxyacetone phosphate (Sigma-Aldrich) in the presence of the coenzyme nicotinamide adenine dinucleotide (NADH) (Omnilab) at a wavelength of 340 nm, using the Infinite 200 (Tecan). Protein concentrations were assessed using the BCA-reducing agent compatible protein assay kit (Thermo Fisher Scientific) with BSA standard samples in GAPDH buffer for quantification. The value for each condition was calculated using the ratio between GAPDH activity and protein concentration.

Glucose uptake, lipolysis and western blot analysis

For glucose, glycerol and western blot analysis, shRNA COBLL1, shRNA empty vector and shRNA GRB14 SGBS cells were differentiated until day 16 in 6-well plates.

The insulin-stimulated 2-DG uptake experiment was performed as described previously[83]. Briefly, cells were incubated in glucose-free DMEM and F-12 (1:1) containing 1% penicillin-streptomycin, 16 μM biotin, 36 μM pantothenic acid, 14.3 mM $NaHCO_3$ and 0.5 mM sodium pyruvate(Sigma-Aldrich) for 12 h. The medium was replaced with 118 mM NaCl, 1.2 mM KH_2PO_4, 4.8 mM KCl, 1.2 mM $MgSO_4$, 2.5 mM $CaCl_2$, 10 mM HEPES, 2.5 mM sodium pyruvate (Sigma-Aldrich) and 0.5% BSA (Sigma-Aldrich), pH 7,35. After 1.5 h the same buffer was added fresh either without supplement or with 1 μM insulin for 30 min. Radioactive uptake was started by adding KRH ^{3}H-2-DG at an activity of 1 μCi ml^{-1} and 50 μM 2-DG. Cells were incubated for 30 min and then washed with PBS. Cells were scraped off after adding 200 μl IGEPAL and 150 μM phloretin. Radioactivity was measured using liquid scintillation counting with an external standard. Protein concentrations to normalize 2-DG uptake were measured.

To measure glycerol release, cells were washed with PBS and incubated for 3 h in phenol red-free DMEM containing 2% free fatty acid-free BSA (Roth). The medium was changed and cells were incubated for 1 h without supplement for basal lipolysis or the addition of 10 μM isoproterenol (Sigma-Aldrich) and 0.5 mM IBMX for stimulated lipolysis. The supernatant was collected for spectrophotometric glycerol measurement in a Sirius tube luminometer (Berthold Technologies)

using glycerol kinase (Sigma-Aldrich) and the ATP Kit SL (BioThema). Remaining cells were collected for protein quantification and western blot analysis in radioimmunoprecipitation assay buffer containing 50 mM Tris-HCl (pH 8), 150 mM NaCl, 0.2% SDS, 1% NP-40, 0.5% deoxycholate, 1 mM phenylmethylsulfonyl fluoride, phosphatase and protease inhibitors. Western blot analysis was performed using a mouse anti-human GAPDH IgG (Ambion) and the Lipolysis Activation Antibody Sampler Kit (catalogue no. 8334, Cell Signaling Technology) according to the manufacturer's protocol. Secondary IRDye IgG (LI-COR) was used to generate the fluorescence, detected by the Odyssey scanner (LI-COR).

Relative gene expression RT–qPCR

Primer pairs were designed using published nucleotide sequences from the human genome (GenBank NCBI/UCSC) and Ensembl, primer3input[84], net primer (PREMIER Biosoft) and primer blast (NCBI GenBank)[85]. Primers against the human target genes *COBLL1* (forward TGATCCTGACTCAGCCCTT, reverse CAACTACCTTCTTTCCATTGC) *LEPTIN* (forward TGGGAAGGAAAATGCATTGGG, reverse ATAAGGTCAGGATGGGGTGG) and *GLUT4* (forward CTGTGCCATCCTGATGACTG, reverse CCAGGGCCAATCTCAAAA) and the reference genes *HPRT1* (forward TGAAAAGGACCCCACGAAG, reverse AAGCAGATGGCCACAGAACTAG), *PPIA* (forward TGGTTCCCAGTTTTTCATC, reverse CGAGTTGTCCACAGTCAGC) and *IPO8* (forward CGGATTATAGTCTCTGACCATGTG, reverse TGTGTCACCATGTTCTTCAGG) were synthesized by Eurofins.

Total RNA was extracted using the RNeasy Mini Kit (QIAGEN) and 0.5 µg was reverse-transcribed using the High-Capacity cDNA Reverse Transcription Kit. RT–qPCR was performed using 96-well plates (black frame, white wells) with heat-sealing films, fixed by the 4s2 Automated Heat Sealer (all from 4titude). The Maxima SYBR Green Mix (Thermo Fisher Scientific) was used for amplification in a RT–qPCR Mastercycler ep realplex (Eppendorf), with a denaturation step of 95 °C for 10 min and 40 cycles of 95 °C for 15 s and 60 °C for 40 s, followed by a melting curve. Relative gene expression was calculated by the $2^{-\Delta\Delta_{Ct}}$ method[86] with a reference gene index of *HPRT*, *PPIA* and *IPO8* or the reference gene *HPRT* (weight loss study).

Mice

All mice (C57BL/6J) were originally obtained from the Charles River Laboratories. To genetically engineer a *Cobll1* whole-body knockout (*Cobll1*$^{-/-}$) model, we used CRISPR–Cas9 genome editing system. Male mice were weaned at 4 weeks of age and body weight was measured every week from 4 to 14 weeks of age. All mice were housed in a temperature-controlled room (22 °C with approximately 50–65% humidity) on a 12-h light–dark cycle with ad libitum access to food (normal diet: 14% fat, 64.8% carbohydrate and 21.2% protein, catalogue no. 2920X, Harlan Teklad) and under pathogen-free conditions. To analyse body fat mass (%), body length (cm) and BMD (g cm^{-2}), we used the DEXA scan. Before scanning, animals were anaesthetized with ketamine. All procedures were conducted with approval from the Institutional Animal Care and Use Committee of the University of Chicago (no. ACUP-71656; IBC0934).

CRISPR–Cas9-mediated generation of a *Cobll1* knockout mouse model. To confirm directly that ablation of *Cobll1* affects T2D-related phenotypes in vivo, we applied the CRISPR–Cas9 system to genetically engineer a *Cobll1* whole-body knockout (Cobll1$^{-/-}$) model. Using specific sgRNAs, we targeted the *Cobll1* gene in the C57BL/6 genetic background. We used guides with the following sequences: gRNA (exon 2) 5′-TTGCTCACTAGTGGGGTCGCAGG-3′ and gRNA (exon 6) 5′-CTTCCTCCGGCCGAGACGAAGGG-3′.

Genotyping. The genotypes of *Cobll1* mutant mice were determined by PCR amplification of gDNA extracted from tails. PCR was performed for 30 cycles at 95 °C for 30 s, 60 °C for 15 s and 72 °C for 30 s, with a final extension at 72 °C for 5 min. PCR amplification was performed using the following primer sets: forward 5′-AAAAGTTTCCTGATGTGAAAGTCA-3′ and reverse 5′-AAAAACAGATGCTCCCCAGA-3′. The PCR products were size-separated by electrophoresis on a 4% agarose gel for 1 h.

In vivo glucose tolerance test. At 16 weeks old, mice were tested for glucose sensitivity by IPGTT. Before IPGTT, mice were fasted for 4 h and an initial blood glucose reading was taken. This fast was followed by intraperitoneal injection of 2 mg kg^{-1} dextrose (catalogue no. CAS 50-99-7, Merck Millipore) and subsequent blood glucose checks using an Accu-Chek Aviva glucometer (Roche). Blood glucose readings were taken at 15, 30, 60 and 120 min after dextrose injection. After IPGTT, mice resumed a high-fat diet. An unpaired, two-sided Student's *t*-test was used to test for significance.

Real-time qPCR. After establishment of stable *Cobll1* knockout mice, ablation of *Cobll1* expression was confirmed by RT–qPCR in relevant tissues, which showed significant decrease in the mRNA fold change of *Cobll1* knockout mice compared to WT and heterozygous littermates. Total RNA was isolated from the inguinal white fat pad, kidney and liver using the RNA extraction reagent RNeasy Mini Kit (catalogue no. 74104, QIAGEN). cDNA synthesis was performed using the SuperScript III First-Strand Synthesis System (catalogue no. 18080-044, Thermo Fisher Scientific). Real-time qPCR reactions were performed using the SsoAdvanced Universal SYBR Green Supermix (catalogue no. 1725270, Bio-Rad Laboratories). Real-time qPCR amplification was performed using the primer sets: qPcrF 5′-CGTCACAGAGCAACAAGACA-3′ and qPcrR 5′-ACTGAGCACAGAGGAACACG-3′.

Isolation, culture and differentiation of mouse preadipocytes. Primary adipocytes were isolated from dissected SAT of 6-week-old mice and digested in 1 g ml^{-1} type I collagenase solution (catalogue no. LS004174, Worthington Biochemical Corporation) solution (containing 3.5% BSA v/v (catalogue no. A9418, Merck Millipore)) in a 37 °C water bath with shaking at 120 rpm for 45 min. The suspension was centrifuged at 250*g* for 5 min; then, the cell pellet was resuspended in culture medium (high-glucose DMEM, catalogue no. D5796, Merck Millipore), 20% FCS (catalogue no. A31605-01, Thermo Fisher Scientific), 100 U ml^{-1} penicillin and 0.1 mg ml^{-1} streptomycin (catalogue no. 15140-122, Thermo Fisher Scientific), was filtered through a 45-µm strainer and was seeded in 25-cm^2 flasks. Confluent preadipocytes were induced for 2 d with an adipogenic medium (high-glucose DMEM, 10% FCS, penicillin-streptomycin (10,000 U ml^{-1}, 10,000 µg ml^{-1}), 850 nM insulin (catalogue no. 12585014, Thermo Fisher Scientific), 1 nM T3 (catalogue no. T6397, Merck Millipore), 500 µM IBMX (catalogue no. I5879, Merck Millipore), 1 µM dexamethasone (catalogue no. D4902, Merck Millipore), 125 µM indomethacin (catalogue no. 70270, Cayman Chemical) and 1 µM rosiglitazone (catalogue no. 71740, Cayman Chemical), and then switched to differentiation medium (adipogenic medium without IBMX, dexamethasone and indomethacin). Cells were collected on the eight day of differentiation and used for further analysis.

Oil Red O and GAPDH assay of mouse preadipocytes. Oil Red O staining was used to assess the presence of lipids in mature adipocytes. For Oil Red O staining, cells were washed with PBS (catalogue no. P38135, Merck Millipore) and fixed with 4% PFA (catalogue no. P6148, Merck Millipore). Fixed cells were then covered with 3 mg ml^{-1} Oil Red O (catalogue no. O0625, Merck Millipore) dissolved in 60% isopropanol (v/v) for 20 min; then the dye was washed away with H_2O. To determine GAPDH activity, we used a commercially available kit from Takara Bio (catalogue no. MK426), by monitoring the dihydroxyacetone phosphate-dependent oxidation of NADH at 340 nm. Enzyme activity was calculated by the formula described in the manufacturer's protocol; GAPDH activity was expressed as U mg^{-1} protein.

Statistics and reproducibility

Statistical analyses were performed using a two-tailed Student's *t*-test or ANOVA to compare the means of two or multiple groups, respectively. No data were excluded from the analyses.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The ATAC–seq data for the immortalized human adipocytes are deposited in the database of Genotypes and Phenotypes under accession no. PRJNA539992. The high-content imaging data are available at the Cell Painting Gallery on the Registry of Open Data on Amazon Web Services (https://registry.opendata.aws/cellpainting-gallery/) under accession no. cpg0011. Source data are provided with this paper.

Code availability

The code used is publicly available on GitHub (https://github.com/ClaussnitzerLab/). The GitHub LP-2q24.3-metabolic-risk-locus repository containing the code and files used to generate LipocyteProfiles can be accessed at https://github.com/sophiestrobel/LP-2q24.3-metabolic-risk-locus.git.

References

1. Cho, N. H. et al. IDF Diabetes Atlas: global estimates of diabetes prevalence for 2017 and projections for 2045. *Diabetes Res. Clin. Pract.* **138**, 271–281 (2018).
2. Lu, Y. et al. New loci for body fat percentage reveal link between adiposity and cardiometabolic disease risk. *Nat. Commun.* **7**, 10495 (2016).
3. Hosseinpanah, F., Barzin, M., Sheikholeslami, F. & Azizi, F. Effect of different obesity phenotypes on cardiovascular events in Tehran Lipid and Glucose Study (TLGS). *Am. J. Cardiol.* **107**, 412–416 (2011).
4. Caleyachetty, R. et al. Metabolically healthy obese and incident cardiovascular disease events among 3.5 million men and women. *J. Am. Coll. Cardiol.* **70**, 1429–1437 (2017).
5. Arnlöv, J., Ingelsson, E., Sundström, J. & Lind, L. Impact of body mass index and the metabolic syndrome on the risk of cardiovascular disease and death in middle-aged men. *Circulation* **121**, 230–236 (2010).
6. Aung, K., Lorenzo, C., Hinojosa, M. A. & Haffner, S. M. Risk of developing diabetes and cardiovascular disease in metabolically unhealthy normal-weight and metabolically healthy obese individuals. *J. Clin. Endocrinol. Metab.* **99**, 462–468 (2014).
7. Calori, G. et al. Prevalence, metabolic features, and prognosis of metabolically healthy obese Italian individuals: the Cremona Study. *Diabetes Care* **34**, 210–215 (2011).
8. Wildman, R. P. et al. The obese without cardiometabolic risk factor clustering and the normal weight with cardiometabolic risk factor clustering: prevalence and correlates of 2 phenotypes among the US population (NHANES 1999–2004). *Arch. Intern. Med.* **168**, 1617–1624 (2008).
9. Primeau, V. et al. Characterizing the profile of obese patients who are metabolically healthy. *Int. J. Obes.* **35**, 971–981 (2011).
10. Yaghootkar, H. et al. Genetic evidence for a normal-weight 'metabolically obese' phenotype linking insulin resistance, hypertension, coronary artery disease, and type 2 diabetes. *Diabetes* **63**, 4369–4377 (2014).
11. Vujkovic, M. et al. Discovery of 318 new risk loci for type 2 diabetes and related vascular outcomes among 1.4 million participants in a multi-ancestry meta-analysis. *Nat. Genet.* **52**, 680–691 (2020).
12. Spracklen, C. N. et al. Identification of type 2 diabetes loci in 433,540 East Asian individuals. *Nature* **582**, 240–245 (2020).
13. Shungin, D. et al. New genetic loci link adipose and insulin biology to body fat distribution. *Nature* **518**, 187–196 (2015).
14. Fox, C. S. et al. Genome-wide association for abdominal subcutaneous and visceral adipose reveals a novel locus for visceral fat in women. *PLoS Genet.* **8**, e1002695 (2012).
15. Morris, A. P. et al. Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat. Genet.* **44**, 981–990 (2012).
16. Chu, A. Y. et al. Multiethnic genome-wide meta-analysis of ectopic fat depots identifies loci associated with adipocyte development and differentiation. *Nat. Genet.* **49**, 125–130 (2017).
17. Wang, H. et al. Genotype-by-environment interactions inferred from genetic effects on phenotypic variability in the UK Biobank. *Sci. Adv.* **5**, eaaw3538 (2019).
18. Gao, C. et al. Genome-wide study of subcutaneous and visceral adipose tissue reveals novel sex-specific adiposity loci in Mexican Americans. *Obesity* **26**, 202–212 (2018).
19. Lagou, V. et al. Sex-dimorphic genetic effects and novel loci for fasting glucose and insulin variability. *Nat. Commun.* **12**, 24 (2021).
20. Chen, J. et al. The trans-ancestral genomic architecture of glycemic traits. *Nat. Genet.* **53**, 840–860 (2021).
21. Loos, R. J. F. & Kilpeläinen, T. O. Genes that make you fat, but keep you healthy. *J. Intern. Med.* **284**, 450–463 (2018).
22. Kilpeläinen, T. O. et al. Genetic variation near *IRS1* associates with reduced adiposity and an impaired metabolic profile. *Nat. Genet.* **43**, 753–760 (2011).
23. Lotta, L. A. et al. Integrative genomic analysis implicates limited peripheral adipose storage capacity in the pathogenesis of human insulin resistance. *Nat. Genet.* **49**, 17–26 (2017).
24. Kooner, J. S. et al. Genome-wide association study in individuals of South Asian ancestry identifies six new type 2 diabetes susceptibility loci. *Nat. Genet.* **43**, 984–989 (2011).
25. DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium, Asian Genetic Epidemiology Network Type 2 Diabetes (AGEN-T2D) Consortium, South Asian Type 2 Diabetes (SAT2D) Consortium et al. Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. *Nat. Genet.* **46**, 234–244 (2014).
26. Heid, I. M. et al. Meta-analysis identifies 13 new loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. *Nat. Genet.* **42**, 949–960 (2010).
27. Udler, M. S. et al. Type 2 diabetes genetic loci informed by multi-trait associations point to disease mechanisms and subtypes: a soft clustering analysis. *PLoS Med.* **15**, e1002654 (2018).
28. Hindorff, L. A. et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc. Natl Acad. Sci. USA* **106**, 9362–9367 (2009).
29. Claussnitzer, M. et al. Leveraging cross-species transcription factor binding site patterns: from diabetes risk loci to disease mechanisms. *Cell* **156**, 343–358 (2014).
30. Claussnitzer, M. et al. FTO obesity variant circuitry and adipocyte browning in humans. *N. Engl. J. Med.* **373**, 895–907 (2015).
31. Kelley, D. R., Snoek, J. & Rinn, J. L. Basset: learning the regulatory code of the accessible genome with deep convolutional neural networks. *Genome Res.* **26**, 990–999 (2016).
32. Dixon, J. R. et al. Chromatin architecture reorganization during stem cell differentiation. *Nature* **518**, 331–336 (2015).
33. Xue, R. et al. Clonal analyses and gene profiling identify genetic biomarkers of the thermogenic potential of human brown and white preadipocytes. *Nat. Med.* **21**, 760–768 (2015).
34. Gilbert, L. A. et al. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **154**, 442–451 (2013).

35. Cowper-Sal lari, R. et al. Breast cancer risk-associated SNPs modulate the affinity of chromatin for FOXA1 and alter gene expression. *Nat. Genet.* **44**, 1191–1198 (2012).
36. Benita, Y. et al. Gene enrichment profiles reveal T-cell development, differentiation, and lineage-specific transcription factors including ZBTB25 as a novel NF-AT repressor. *Blood* **115**, 5376–5384 (2010).
37. Kawaguchi, N. et al. ADAM12 induces actin cytoskeleton and extracellular matrix reorganization during early adipocyte differentiation by regulating β1 integrin function. *J. Cell Sci.* **116**, 3893–3904 (2003).
38. Qian, Y. et al. PI3K induced actin filament remodeling through Akt and p70S6K1: implication of essential role in cell migration. *Am. J. Physiol. Cell Physiol.* **286**, C153–C163 (2004).
39. Morandi, E. M. et al. ITGAV and ITGA5 diversely regulate proliferation and adipogenic differentiation of human adipose derived stem cells. *Sci. Rep.* **6**, 28889 (2016).
40. Truebestein, L., Elsner, D. J., Fuchs, E. & Leonard, T. A. A molecular ruler regulates cytoskeletal remodelling by the Rho kinases. *Nat. Commun.* **6**, 10029 (2015).
41. Izadi, M. et al. Cobl-like promotes actin filament formation and dendritic branching using only a single WH2 domain. *J. Cell Biol.* **217**, 211–230 (2018).
42. Takayama, K.-I., Suzuki, T., Fujimura, T., Takahashi, S. & Inoue, S. COBLL1 modulates cell morphology and facilitates androgen receptor genomic binding in advanced prostate cancer. *Proc. Natl Acad. Sci. USA* **115**, 4975–4980 (2018).
43. Laber, S. et al. Discovering cellular programs of intrinsic and extrinsic drivers of metabolic traits using LipocyteProfiler. Preprint at *bioRxiv* https://doi.org/10.1101/2021.07.17.452050 (2021).
44. Harris, R. B. S. Direct and indirect effects of leptin on adipocyte metabolism. *Biochim. Biophys. Acta* **1842**, 414–423 (2014).
45. Kilpeläinen, T. O. et al. Genome-wide meta-analysis uncovers novel loci influencing circulating leptin levels. *Nat. Commun.* **7**, 10494 (2016).
46. Folkersen, L. et al. Genomic and drug target evaluation of 90 cardiovascular proteins in 30,931 individuals. *Nat. Metab.* **2**, 1135–1148 (2020).
47. Small, K. S. et al. Regulatory variants at *KLF14* influence type 2 diabetes risk via a female-specific effect on adipocyte size and body composition. *Nat. Genet.* **50**, 572–580 (2018).
48. Fathzadeh, M. et al. *FAM13A* affects body fat distribution and adipocyte function. *Nat. Commun.* **11**, 1465 (2020).
49. Randall, J. C. et al. Sex-stratified genome-wide association studies including 270,000 individuals show sexual dimorphism in genetic loci for anthropometric traits. *PLoS Genet.* **9**, e1003500 (2013).
50. Sung, Y. J. et al. Genome-wide association studies suggest sex-specific loci associated with abdominal and visceral fat. *Int. J. Obes.* **40**, 662–674 (2016).
51. Palmer, B. F. & Clegg, D. J. The sexual dimorphism of obesity. *Mol. Cell. Endocrinol.* **402**, 113–119 (2015).
52. Chen, Z. et al. Functional screening of candidate causal genes for insulin resistance in human preadipocytes and adipocytes. *Circ. Res.* **126**, 330–346 (2020).
53. Chen, L., Hu, H., Qiu, W., Shi, K. & Kassem, M. Actin depolymerization enhances adipogenic differentiation in human stromal stem cells. *Stem Cell Res.* **29**, 76–83 (2018).
54. Kanzaki, M. & Pessin, J. E. Insulin-stimulated GLUT4 translocation in adipocytes is dependent upon cortical actin remodeling. *J. Biol. Chem.* **276**, 42436–42444 (2001).
55. Kim, J. I. et al. During adipocyte remodeling, lipid droplet configurations regulate insulin sensitivity through F-actin and G-actin reorganization. *Mol. Cell. Biol.* **39**, e00210–e00219 (2019).
56. Hansson, B. et al. Adipose cell size changes are associated with a drastic actin remodeling. *Sci. Rep.* **9**, 12941 (2019).
57. Li, Q. et al. Obesity and hyperinsulinemia drive adipocytes to activate a cell cycle program and senesce. *Nat. Med.* **27**, 1941–1953 (2021).
58. Bäckdahl, J. et al. Spatial mapping reveals human adipocyte subpopulations with distinct sensitivities to insulin. *Cell Metab.* **33**, 1869–1882 (2021).
59. Emont, M. P. et al. A single-cell atlas of human and mouse white adipose tissue. *Nature* **603**, 926–933 (2022).
60. Bérézíat, V. et al. Inhibition of insulin receptor catalytic activity by the molecular adapter Grb14. *J. Biol. Chem.* **277**, 4845–4852 (2002).
61. Depetris, R. S. et al. Structural basis for inhibition of the insulin receptor by the adaptor protein Grb14. *Mol. Cell* **20**, 325–333 (2005).
62. Cooney, G. J. et al. Improved glucose homeostasis and enhanced insulin signalling in Grb14-deficient mice. *EMBO J.* **23**, 582–593 (2004).
63. Wang, L. et al. Peripheral disruption of the *Grb10* gene enhances insulin signaling and sensitivity in vivo. *Mol. Cell. Biol.* **27**, 6497–6505 (2007).
64. Holt, L. J. et al. Dual ablation of Grb10 and Grb14 in mice reveals their combined role in regulation of insulin signaling and glucose homeostasis. *Mol. Endocrinol.* **23**, 1406–1414 (2009).
65. Cariou, B. et al. Increased adipose tissue expression of Grb14 in several models of insulin resistance. *FASEB J.* **18**, 965–967 (2004).
66. Shi, J. et al. Discovery of cancer drug targets by CRISPR–Cas9 screening of protein domains. *Nat. Biotechnol.* **33**, 661–667 (2015).
67. González, F. et al. An iCRISPR platform for rapid, multiplexable, and inducible genome editing in human pluripotent stem cells. *Cell Stem Cell* **15**, 215–226 (2014).
68. Fuchsberger, C. et al. The genetic architecture of type 2 diabetes. *Nature* **536**, 41–47 (2016).
69. Albrechtsen, A. et al. Exome sequencing-driven discovery of coding polymorphisms associated with common metabolic phenotypes. *Diabetologia* **56**, 298–310 (2013).
70. Wessel, J. et al. Low-frequency and rare exome chip variants associate with fasting glucose and type 2 diabetes susceptibility. *Nat. Commun.* **6**, 5897 (2015).
71. Kan, M. et al. Rare variant associations with waist-to-hip ratio in European-American and African-American women from the NHLBI-Exome Sequencing Project. *Eur. J. Hum. Genet.* **24**, 1181–1187 (2016).
72. Claussnitzer, M. C. et al. Leveraging transcription factor binding site patterns: from diabetes risk loci to disease mechanisms. *Cell* **156**, 343–358 (2014).
73. Hauner, H., Skurk, T. & Wabitsch, M. Cultures of human adipose precursor cells. *Methods Mol. Biol.* **155**, 239–247 (2001).
74. Ott, B. et al. Effect of caloric restriction on gut permeability, inflammation markers, and fecal microbiota in obese women. *Sci. Rep.* **7**, 11955 (2017).
75. Glastonbury, C. A. et al. Machine learning based histology phenotyping to investigate the epidemiologic and genetic basis of adipocyte morphology and cardiometabolic traits. *PLoS Comput. Biol.* **16**, e1008044 (2020).
76. Buenrostro, J. D., Wu, B., Chang, H. Y. & Greenleaf, W. J. ATAC-seq: a method for assaying chromatin accessibility genome-wide. *Curr. Protoc. Mol. Biol.* **109**, 21.29.1–21.29.9 (2015).
77. Sobreira, D. R. et al. Extensive pleiotropism and allelic heterogeneity mediate metabolic effects of *IRX3* and *IRX5*. *Science* **372**, 1085–1091 (2021).

78. Sinnott-Armstrong, N. et al. A regulatory variant at *3q21.1* confers an increased pleiotropic risk for hyperglycemia and altered bone mineral density. *Cell Metab.* **33**, 615–628 (2021).
79. Eastwood, S. V. et al. Algorithms for the capture and adjudication of prevalent and incident diabetes in UK Biobank. *PLoS ONE* **11**, e0162388 (2016).
80. Sinnott-Armstrong, N. Genetics of 35 blood and urine biomarkers in the UK Biobank. *Nat. Genet.* **53**, 185–194 (2021).
81. Sanson, K. R. et al. Optimized libraries for CRISPR–Cas9 genetic screens with multiple modalities. *Nat. Commun.* **9**, 5416 (2018).
82. Pairault, J. & Green, H. A study of the adipose conversion of suspended 3T3 cells by using glycerophosphate dehydrogenase as differentiation marker. *Proc. Natl Acad. Sci. USA* **76**, 5138–5142 (1979).
83. Claussnitzer, M., Skurk, T., Hauner, H., Daniel, H. & Rist, M. J. Effect of flavonoids on basal and insulin-stimulated 2-deoxyglucose uptake in adipocytes. *Mol. Nutr. Food Res.* **55**, S26–S34 (2011).
84. Untergasser, A. et al. Primer3—new capabilities and interfaces. *Nucleic Acids Res.* **40**, e115 (2012).
85. Ye, J. et al. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics* **13**, 134 (2012).
86. Pfaffl, M. W. A new mathematical model for relative quantification in real-time RT–PCR. *Nucleic Acids Res.* **29**, e45 (2001).
87. Elliott, L. et al. The genetic basis of human brain structure and function: 1,262 genome-wide associations found from 3,144 GWAS of multimodal brain imaging phenotypes from 9,707 UK Biobank participants. Preprint at *bioRxiv* https://doi.org/10.1101/178806 (2017).
88. Kundaje, A. et al. Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317–330 (2015).
89. Chen, E. Y. et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* **14**, 128 (2013).
90. Kuleshov, M. V. et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **44**, W90–W97 (2016).

Acknowledgements

Work in the Claussnitzer lab is supported by the Foundation for the National Institutes of Health (NIH) (AMP-T2D RFB8b), the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) (UM1 DK126185 and NIDDK DK102173), the Novo Nordisk Foundation (NNF21SA0072102) and by a Next Generation Award from the Broad Institute of MIT and Harvard. We acknowledge support from the Center for Advanced Light Microscopy of the Technical University of Munich School of Life Sciences, T. Kufer at the University of Hohenheim for access to microscopes, the Clinical Cooperation Group 'Nutrigenomics and Type 2 Diabetes' from the Helmholtz Center Munich and the German Center for Diabetes Research, the Else Kröner-Fresenius-Foundation, a Stanford Graduate Fellowship, a Center for Computational, Evolutionary and Human Genomics Graduate Fellowship, the Novo Nordisk Foundation Challenge Grant (no. NNF18OC0033754), NIH grant nos. R01HL128075 and P30DK020595, the Federal Ministry of Education and Research (award no. 0315674), the German Research Foundation (award no. 338582098) and the Federal Ministry of Education and Research (grant no. 0315674). We thank A. Flaccus, A. Kozza, D. Mvondo, M. Hubersberger, E. Hofmair, J. Röttgen, C. Deuschle, A. Saadat and S. Harken for technical support.

Author contributions

V.G. and M.C. conceptualized the study. V.G., S.L., N.S.-A., D.R.S., S.M.S., T.M.B., P.K., B.N.M., H.E., Y.H., B.B., G.G., J.H., D.R.S., N.A., V.C.-N., T.S., S.O., K.S., B.A.C., A.E.C., S.N.D., C.M.L., H.H., M.A.N. and M.C. devised the methodology. V.G., S.L., N.S.-A., D.R.S., S.M.S., T.M.B., P.K., B.N.M., H.E., Y.H., B.B., D.S. and B.C. carried out the formal analysis. V.G., S.L., N.S.-A., D.R.S., S.M.S., T.M.B., P.K., B.N.M., H.E., Y.H., B.B., D.S. and B.C. carried out the investigation. H.H., C.M.L., M.A.N. and M.C. obtained the resources. A.E.C., H.H., C.M.L., M.A.N. and M.C. supervised the study. H.H., C.M.L., M.A.N. and M.C. acquired the funding. V.G., N.S.-A., D.R.S., S.M.S. and M.C. wrote the original draft. V.G., S.L., N.S.-A., D.R.S., S.M.S., T.M.B., P.K., B.N.M., H.E., Y.H., B.B., G.G., J.H., D.S., N.A., V.C.-N., T.S., S.O., K.S., B.C., A.C., S.N.D., C.M.L., H.H., M.A.N. and M.C. reviewed and edited the manuscript.

Competing interests

M.C. holds equity in Waypoint Bio and is a member of the Nestle scientific advisory board. The authors have filed a provisional patent application (63/218,656).

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s42255-023-00807-w.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42255-023-00807-w.

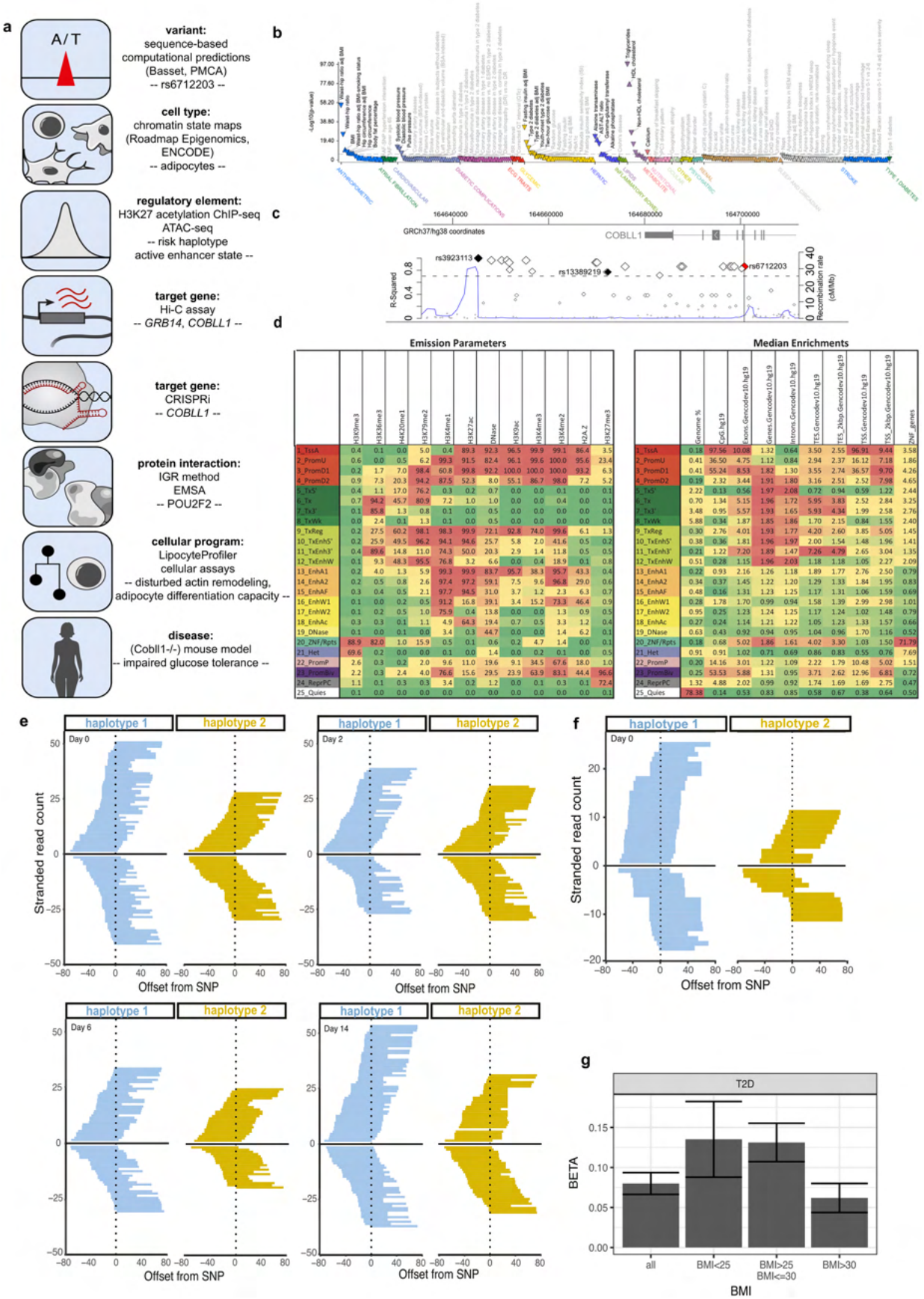
Correspondence and requests for materials should be addressed to Melina Claussnitzer.

Peer review information *Nature Metabolism* thanks Inês Cebola, Hannah Maude, Camilla Scheele and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Ashley Castellanos-Jankiewicz, in collaboration with the *Nature Metabolism* team.

Viktoria Glunk[1,2,18], Samantha Laber[3,18], Nasa Sinnott-Armstrong[3,4,5,18], Debora R. Sobreira[6,18], Sophie M. Strobel[1,2,3,18], Thiago M. Batista[3], Phil Kubitz[1,2,3], Bahareh Nemati Moud[1,2], Hannah Ebert[7], Yi Huang[3], Beate Brandl[1,2], Garrett Garbo[3], Julius Honecker[1,2], David R. Stirling[8], Nezar Abdenur[3], Virtu Calabuig-Navarro[3,7], Thomas Skurk[1,2], Sören Ocvirk[9,10], Kerstin Stemmer[11,12,13], Beth A. Cimini[8], Anne E. Carpenter[8], Simon N. Dankel[14], Cecilia M. Lindgren[3,15], Hans Hauner[1,2], Marcelo A. Nobrega[6,19] & Melina Claussnitzer[3,16,17,19] ✉

[1]Institute of Nutritional Medicine, School of Medicine, Technical University of Munich, Else Kröner-Fresenius-Center for Nutritional Medicine, School of Life Sciences, Freising-Weihenstephan, Freising, Germany. [2]ZIEL Institute for Food & Health, Technical University of Munich, Freising, Germany. [3]Broad Institute of MIT and Harvard, Medical and Population Genetics Program & Type 2 Diabetes Systems Genomics Initiative, Cambridge, MA, USA. [4]Department of Genetics, Stanford University, Stanford, CA, USA. [5]Herbold Computational Biology Program, Publich Health Sciences Division, Fred Hutchinson Cancer Center, Seattle, WA, USA. [6]Department of Human Genetics, University of Chicago, Chicago, IL, USA. [7]Institute of Nutritional Sciences, University of Hohenheim, Stuttgart, Germany. [8]Imaging Platform, Broad Institute of MIT and Harvard, Cambridge, MA, USA. [9]Division of Gastroenterology, Hepatology and Nutrition, Department of Medicine, University of Pittsburgh, Pittsburgh, PA, USA. [10]Intestinal Microbiology Research Group, Department of Molecular Toxicology, German Institute of Human Nutrition Potsdam-Rehbruecke, Nuthetal, Germany. [11]Molecular Cell Biology, Institute for Theoretical Medicine, University of Augsburg, Augsburg, Germany. [12]Institute for Diabetes and Obesity, Helmholtz Zentrum München, Neuherberg, Germany. [13]German Center for Diabetes Research, Neuherberg, Germany. [14]Department of Clinical Science, University of Bergen, Bergen, Norway. [15]Big Data Institute at the Li Ka Shing Centre for Health Information and Discovery, University of Oxford, Oxford, UK. [16]Diabetes Unit and Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA, USA. [17]Department of Medicine, Harvard Medical School, Boston, MA, USA. [18]These authors contributed equally: Viktoria Glunk, Samantha Laber, Nasa Sinnott-Armstrong, Debora R. Sobreira, Sophie M. Strobel. [19]These authors jointly supervised this work: Marcelo A. Nobrega, Melina Claussnitzer. ✉e-mail: melina@broadinstitute.org

https://doi.org/10.1038/s42255-023-00807-w

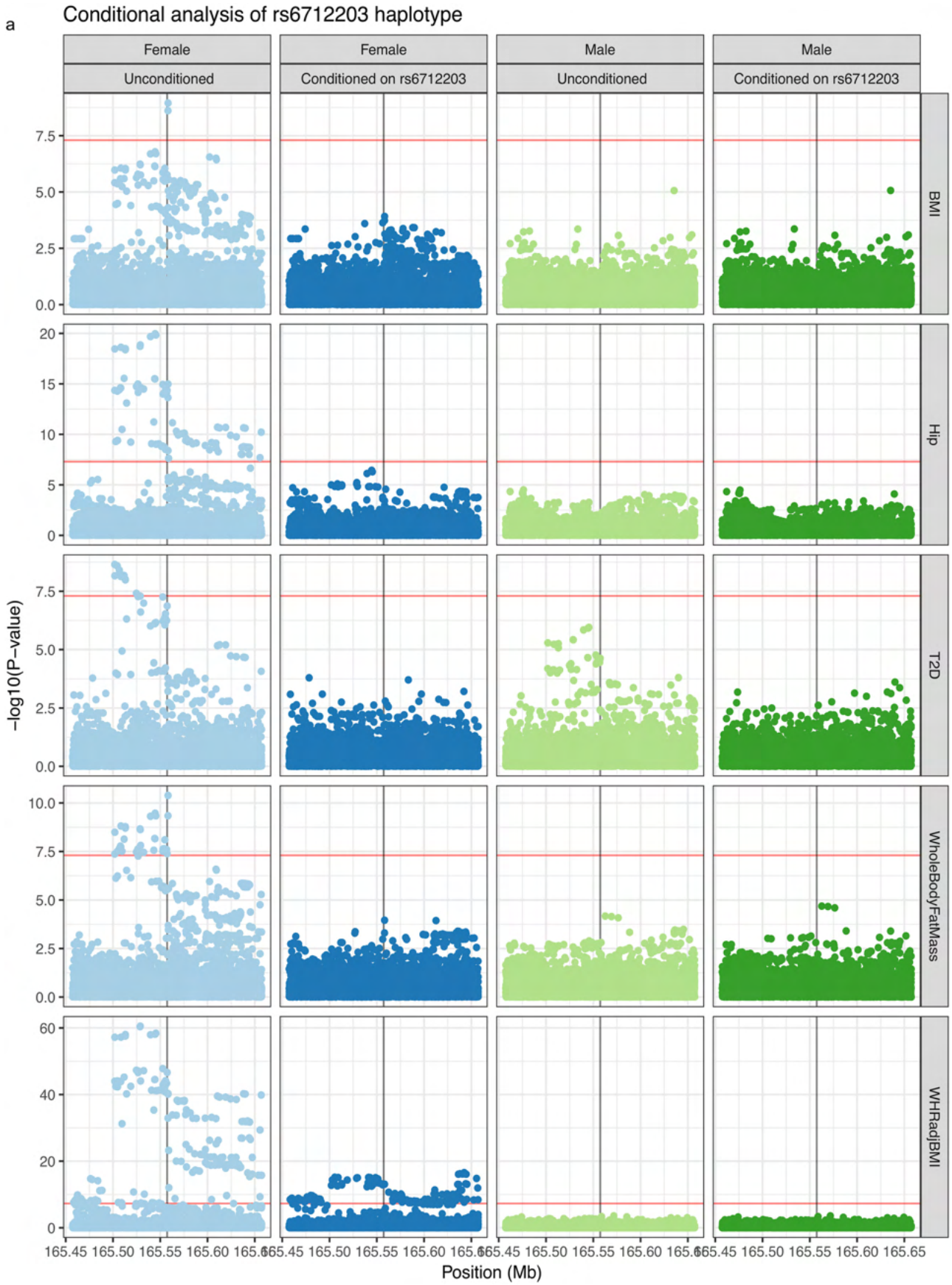


Extended data Fig. 1 | See next page for caption.

https://doi.org/10.1038/s42255-023-00807-w

Extended data Fig. 1 | The pleiotropic *2q24.3* MONW locus is associated with increased risk for type 2 diabetes and decreased adiposity related traits and maps to sparse enhancer signatures in adipocytes. (**a**) Schematic overview for the 2q24.3 metabolic risk locus dissection. Aim of step (top, bold); methods/experiments used (middle); key finding/result of each step (bottom). (**b**) PheWAS of trait associations at the rs3923113-tagged haplotype of a meta-analysis https://t2d.hugeamp.org/. Colors represent trait classes while individual rs3923113 variant association p-values are shown on the Y axis. Direction of effect is indicated by orientation of triangles, upward: increase, downward: decrease (**c**) The *2q24.3* MONW locus spans 19 non-coding SNPs in high linkage disequilibrium with rs3923113 (LD r2 > 0.8). The region of association localizes to a >55 kb interval in an intergenic region between *COBLL1* and *GRB14*. (**d**) Annotation panel and color key for the twenty-five state chromatin model70. Rows represent chromatin states abbreviations, columns are emission parameters, corresponding to the frequency with which each mark is expected in each state (left table) and genome coverage and median enrichments of relevant genomic annotations (right panel). TssA: Active TSS, TssAFlnk: Flanking Active TSS, TxFlnk: Transcription at gene 5′ and 3′, Tx: Strong Transcription, TxWk: Weak Transcription, EnhG: Genic enhancers: Enh: Enhancers, ZNF/Rpts: ZNF genes & repeats, Het: Heterochromatin, TssBiv: Bivalent/Poised TSS, BivFlnk: Flanking Bivalent TSS/Enhancer, EnhBiv: Bivalent Enhancer, ReprPC: Repressed Polycomb, ReprPCWk: Weak Repressed Polycomb, Quies: Quiescent/Low. (**e**) Stranded allele-specific chromatin accessibility measures at the haplotype using ATAC-seq data in differentiating adipocytes from a heterozygous individual. For each day of differentiation of an individual heterozygous, the number of reads overlapping with 20 non-coding SNPs in the haplotype, ordered by their start position and strand relative to the position of the variant, are shown. More reads indicate higher activity in haplotype 1 (non-risk, blue) compared to haplotype 2 (risk yellow). x-axis: offset from SNP position (bp), y-axis: stranded read count. (**f**) Replication of the effect at time 0 (mesenchymal stem cells) with ATAC-seq. (**g**) BMI-dependent variant association analysis. Bar plots represent the beta of the rs6712203 association with type 2 diabetes following BMI stratification. The cohort analysed is the UK Biobank self-identified white British individuals (total N = 327,960; N = 109198 with BMI < 25, N = 140539 with BMI between 25 and 30, and N = 78223 with BMI > = 30), and overlay of data points is not practical. Betas and 95% confidence intervals are shown, derived from a two-sided generalized linear model on outcome adjusted for demographic covariates (age, sex, genotyping array, 40 PCs).

https://doi.org/10.1038/s42255-023-00807-w

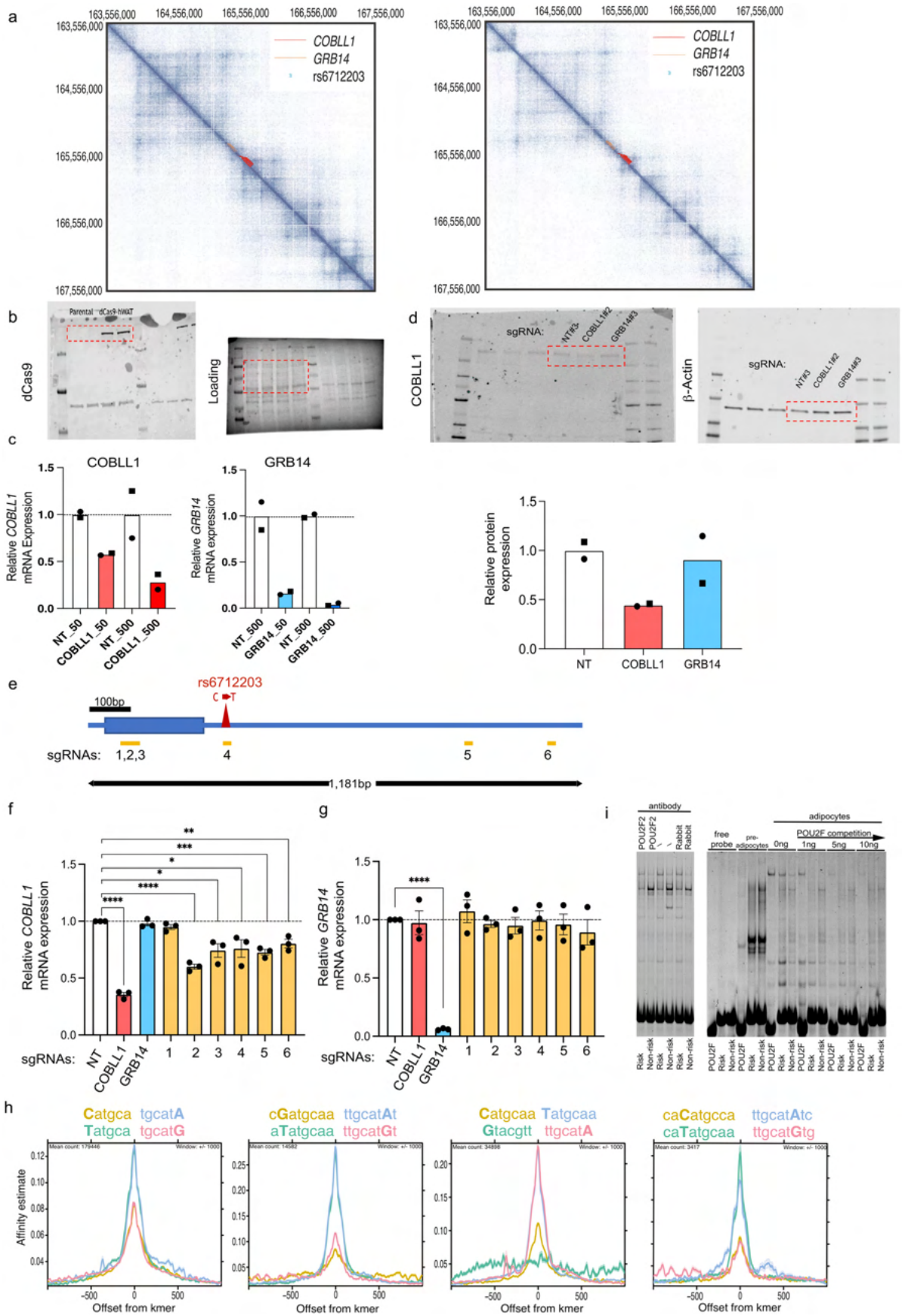


Extended data Fig. 2 | See next page for caption.

https://doi.org/10.1038/s42255-023-00807-w

Extended data Fig. 2 | Conditional ana- lyses implicating rs6712203 in the genetic control of anthropometric traits and type 2 diabetes. a, Conditional analyses implicating rs6712203 in the genetic control of anthropometric traits and type 2 diabetes. Each panel represents a different trait/sex/conditional analysis window, and all panels have an X axis corresponding to 100 kb on either side of the rs6712203 variant. The Y axis shows, for each variant in the window, the association strength for the given trait conditioned on the variants noted in White British participants in UK Biobank with the sex shown, and red lines indicate the significance threshold 5 × 10-8). −log10 p-values are shown, derived from a two-sided generalized linear model on outcome adjusted for demographic covariates (age, sex, genotyping array, 40 PCs).

https://doi.org/10.1038/s42255-023-00807-w

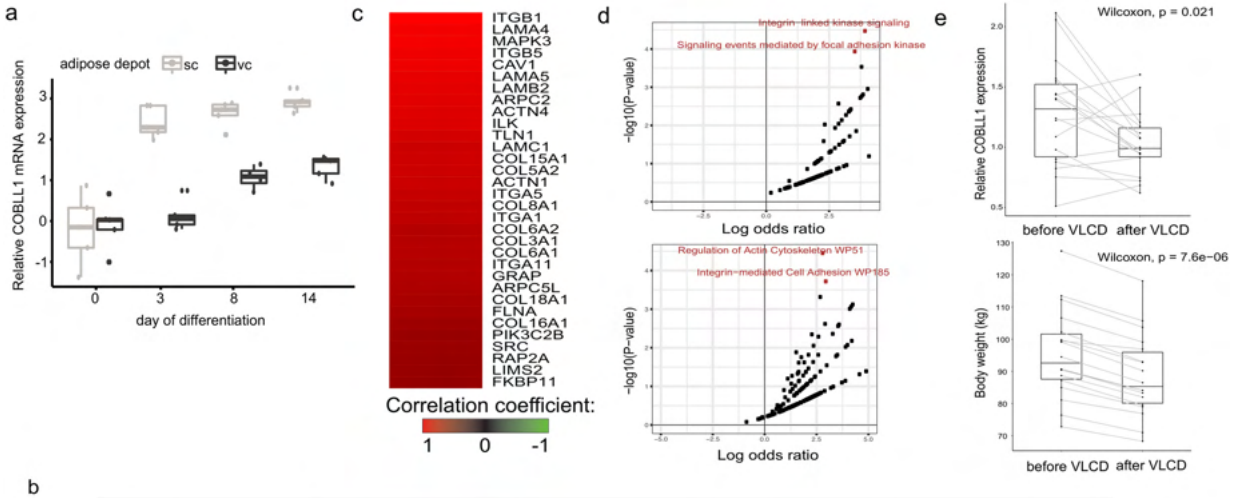


Extended data Fig. 3 | See next page for caption.

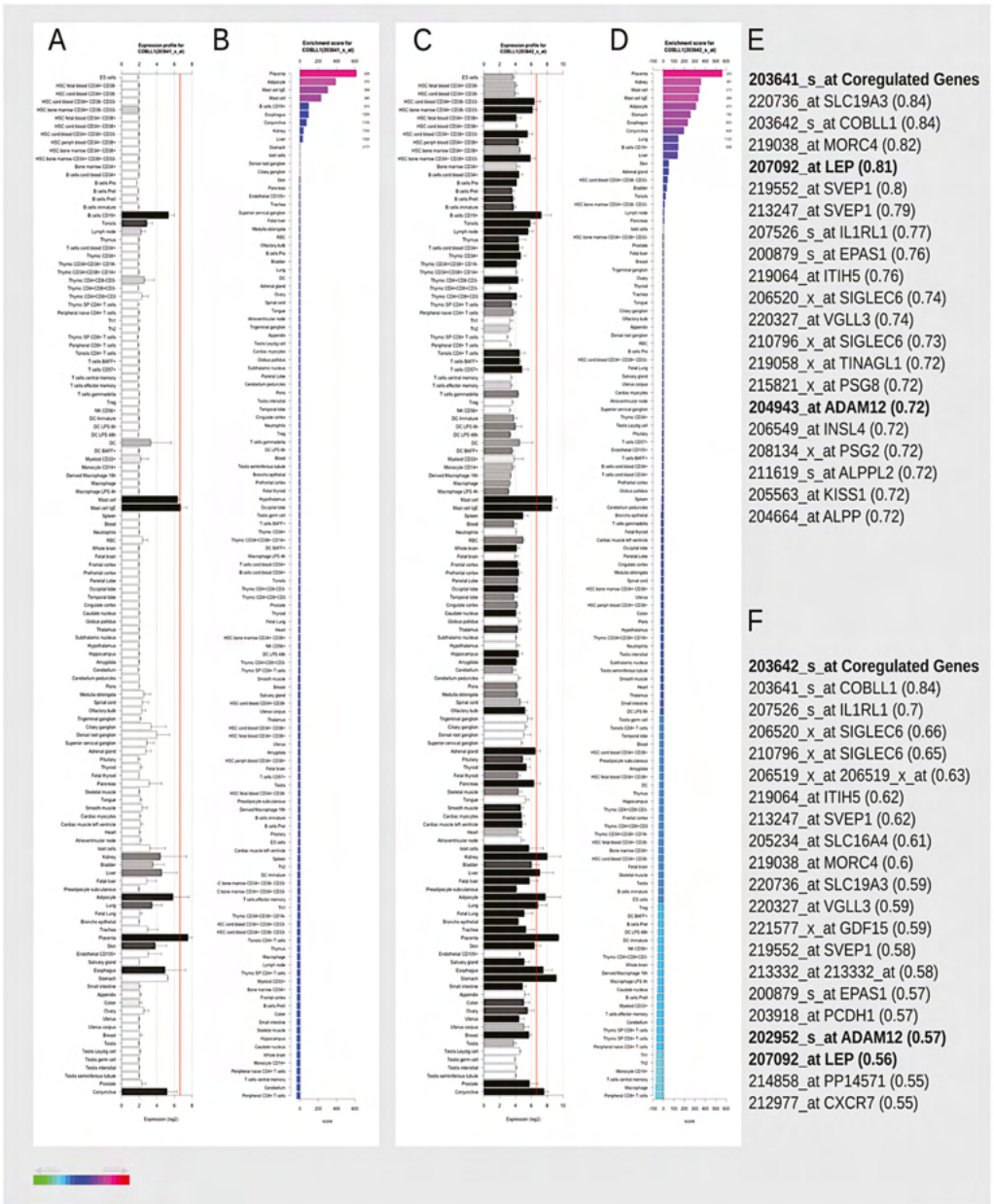
https://doi.org/10.1038/s42255-023-00807-w

Extended data Fig. 3 | Chromatin inter- actions and CRISPRi of 2q24.3 locus identify *COBLL1* as target genes. (**a**) Cross-cell type conserved genome-wide higher order chromatin interactions for the 2q24.3 locus analyzed by Hi-C assays in human fibroblasts (left) and NHEK primary normal human epidermal keratinocytes (right), chr2: 163,556,000 - 167,558,000 (hg19), binned at 2 kb resolution. (**b**) Cas9 protein expression in dCas9 hWAT compared to the parental hWAT cell line. (**c**) mRNA expression of COBLL1 and GRB14 in response to increasing amounts of lentiviral sgRNA vectors (2 sgRNAs, virus volume 50 µl and 500 µl) targeting TSS regions of each gene compared to non-targeting controls (NT, 2 sgRNAs). Columns are means of individual sgRNAs indicated by different symbols. (**d**) COBLL1 protein expression normalized to b-actin in dCas9 hWATs transduced with sgRNAs targeting COBLL1 or GRB14 compared to controls. Top panel: Image of gel of representative sgRNA targeting NT, COBLL1 or GRB14. Bottom panel: plot of protein expression; 2 sgRNA for each target in 2 replicates. (**e**) Representation of 1,181 bp region flanking the COBLL1 intronic variant rs6712203 at the 2q24.3 MONW locus showing individual sgRNAs (n = 6) targeting the rs6712203 flanking regulatory region used in the CRISPRi experiments. (**f**, **g**) mRNA expression of (**f**) COBLL1 and (**g**) GRB14 in undifferentiated dCas9-hWAT preadipocytes at 6 days post lentiviral transduction with sgRNAs targeting TSS regions (red: COBLL1 TSS; blue GRB14 TSS) and the rs6712203-flanking regulatory element at position 1 to 6 as depicted in (**e**). Data are mean +/− SEM of 3 independent experiments. **** P < 0.0001, *** P = 0.0004, ** P = 0.006, * P = 0.013 – 0.036, two-tailed Student's *t* test. (**h**) Predicted binding of POU2F2 between the two alleles using the Intragenomic Replicate Method (Cowper-Sal lari et al. 2012). As in Fig. 2d with different kmer counts.

https://doi.org/10.1038/s42255-023-00807-w



b

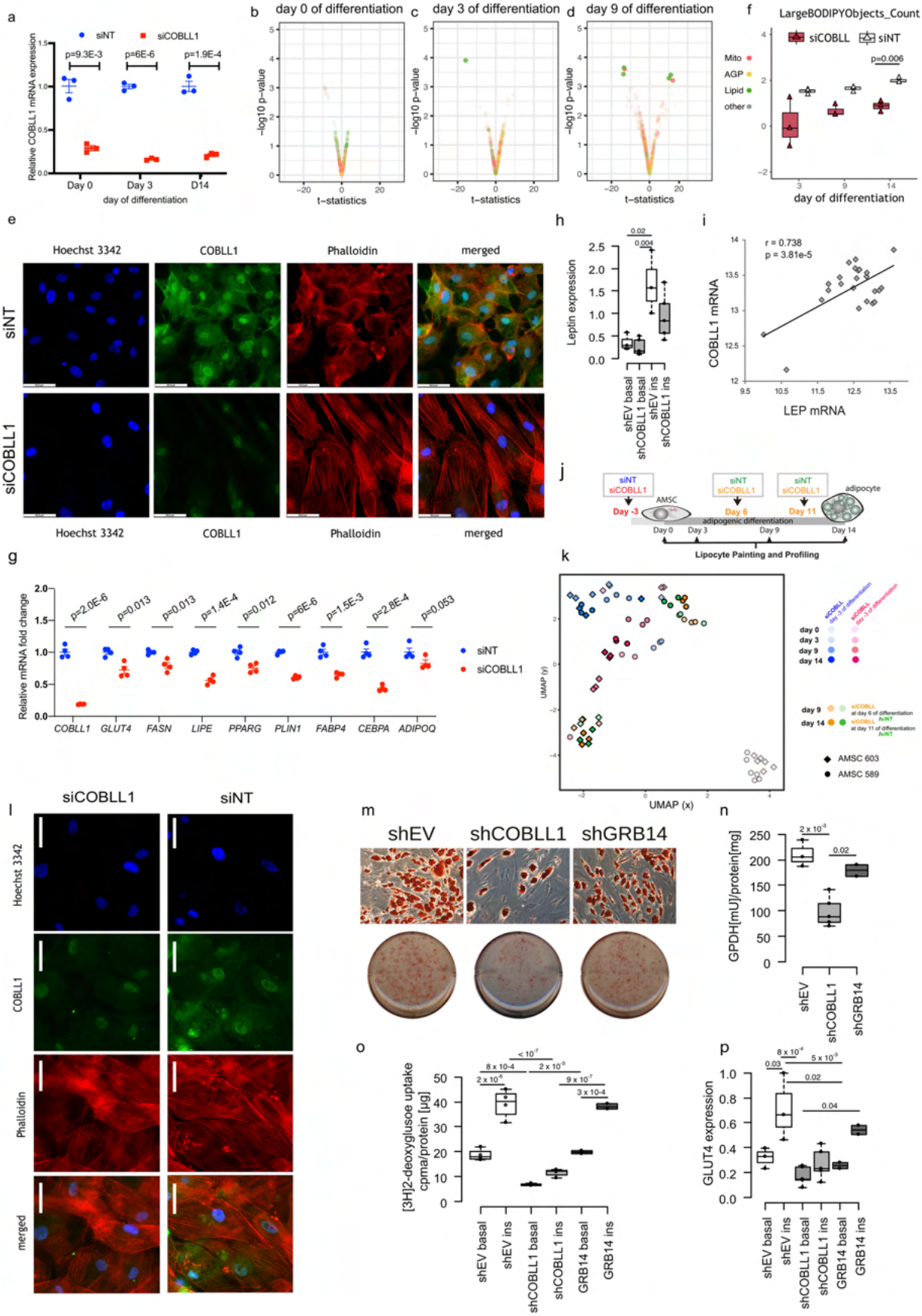


Extended data Fig. 4 | See next page for caption.

https://doi.org/10.1038/s42255-023-00807-w

Extended data Fig. 4 | *COBLL1* regulates actin cytoskeleton remodeling. (**a**) COBLL1 expression in subcutaneous and visceral AMSCs throughout adipogenic differentiation, N = 4 biologically independent experiments, t-test two-sided, data represent median + 95% CI. (**b**) COBLL1 gene expression enrichment across 142 tissues (A-D) from enrichment profiler36. COBLL1 probes 203641_s_at and 203642_s_at were used for coregulation analysis (E-F). (**c**) Correlation with COBLL1 probe ILMN_1761260 using microarray data from lean and individuals with obesity. (**d**) Enrichment of pathways in the HCI (upper panel) and WikiPathways (lower panel) gene set lists from Enrichr, plotted as in Fig. 3A (KEGG), with p-value thresholds corresponding to the FDR cutoffs in those data. p-values are derived from a hypergeometric test. (**e**) COBLL1 expression in subcutaneous adipose tissue before and after a very low caloric diet (VLCD, upper panel, n = 18), corresponding body weight (lower panel), Wilcoxon signed-rank test.

https://doi.org/10.1038/s42255-023-00807-w

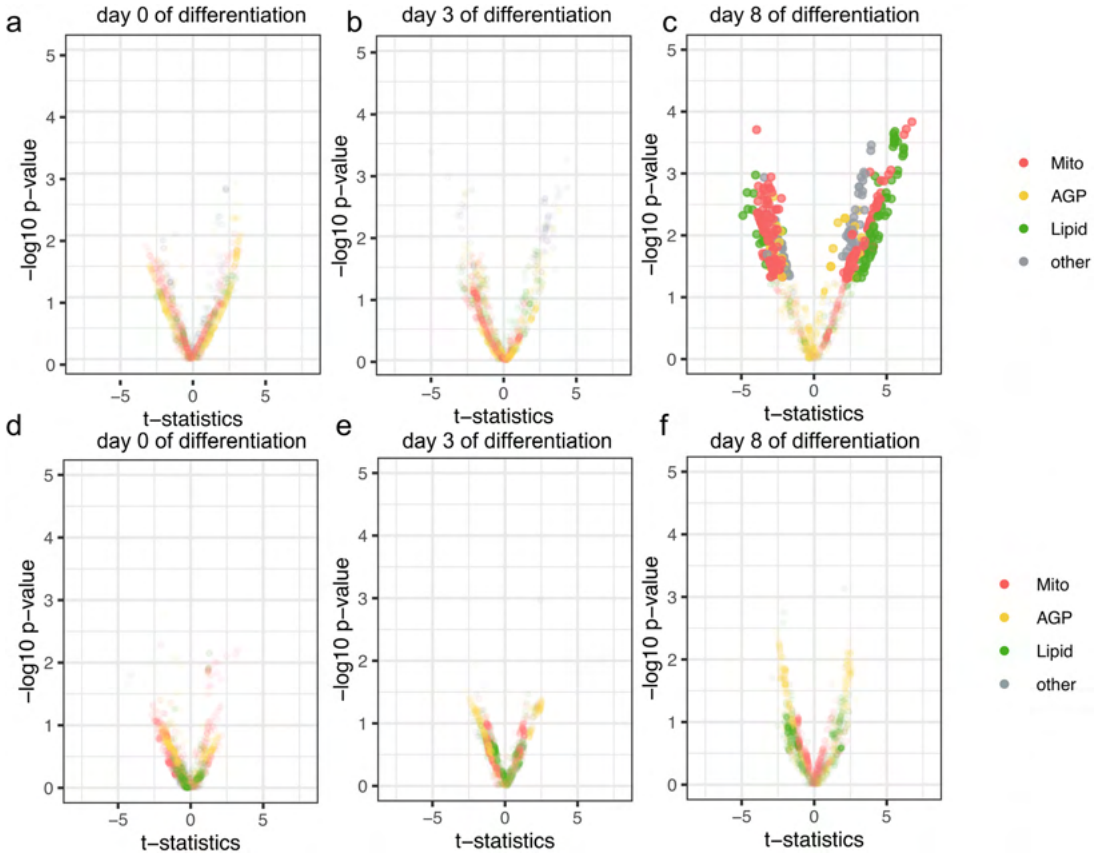


Extended data Fig. 5 | See next page for caption.

https://doi.org/10.1038/s42255-023-00807-w

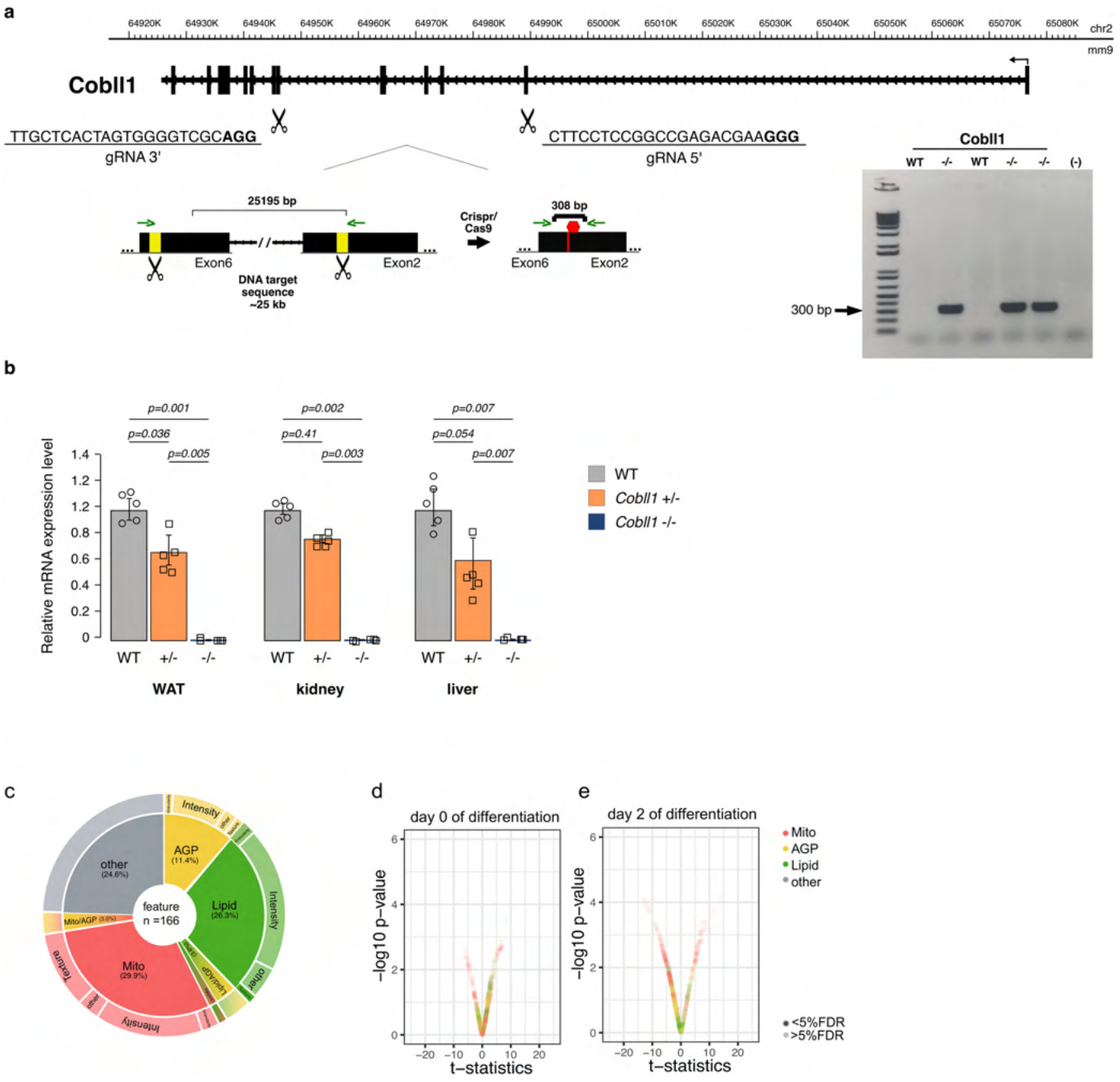
Extended data Fig. 5 | Knockdown of *COBLL1* affects actin remodeling processes in differentiating adipocytes along with adipocyte differentiation, insulin sensitivity and lipolysis rate. (**a**) COBLL1 expression in siCOBLL1 and siNT at day 0, 3 and 14 of differentiation, N = 3 biologically independent experiments, t-test two-sided. knock-down efficiency 80%, mean values + SEM. (**b**–**d**) Morphological profiles of siCOBLL compared to siNT AMSCs at day 0 (**b**) day 3 (**c**) and day 9 (**d**) of differentiation, t-test two-sided, significance level < 5%FDR. (**e**)ActinandCOBLL1stainginginsiCOBLL1comparedtosiNT subcutaneous adipocytes at day 9 using phalloidin and COBLL1 antibody staining (HPA053344, Alexa-Fluor 488), magnification x63/oil. Scale bar = 52.8 um. Representative results from N = 3 independent experiments. (**f**) Cells_Children_LargeBODIPY_objects_count in siCOBLL1- and siNT AMSCs at day 3, 9, 14, N = 3 biologically independent experiments, t-test two-sided, significance level < 5% FDR. (**G**) qPCR-based gene expression of COBLL1 and adipocyte marker genes GLUT4, FASN, LIPE, PPARG, PLIN1, FABP4, CEBPA, ADIPOQ in siCOBLL1 and siNT AMSCs at day 14 of differentiation, t-test two-sided, N = 4 biologically independent experiments, mean values +/− SEM. (**h**) qPCR-based leptin gene expression in shCOBLL1 compared to shEV adipocytes. Data are represented as median + 95% CI, one-way ANOVA with Tukey's HSD test, N = 4 biologically independent experiments (**i**) Correlation of COBLL1 mRNA with LEP mRNA in subcutaneous adipose tissue from 24 lean individuals measured by Illumina microarrays. The pearson's correlation coefficient r and p-value are depicted (**j**) Schematic of siCOBLL1 KD and AMSCs differentiation. (**k**) UMAP-based dimensionality reduction of LipocyteProfiler features in siCOBLL1 and siNT AMSCs. (**l**) Actin and COBLL1 staining in siCOBLL1 and siNT visceral adipocytes at day 14 using phalloidin and COBLL1 antibody staining (HPA053344, Alexa-Fluor 488), magnification x63/oil. Representative result from N = 2 independent experiments, scale bar = 52,8um (**m**) Representative Oil-Red-O lipid staining in SGBS adipocytes following lentiviral *COBLL1* knock-down (shCOBLL1, knock-down efficiency 69%) and GRB14 (shGRB14, knock-down efficiency 61%) compared to empty vector control (shEV), scale bar = 15 mm. (**n**) GPDH metabolic activity in shCOBLL1, shGRB14 and shEV SGBS adipocytes, one-way ANOVA with Tukey's HSD test, mean + 95% CI, N = 4 biologically independent experiments (**o**) Basal and insulin-stimulated 3H-2-deoxyglucose uptake in shCOBLL1, shGRB14 and shEV SGBS adipocytes, one-way ANOVA with Tukey's HSD test, mean + 95% CI, N = 4 biologically independent experiments, 1st and 3rd quartiles (box) and median (middle line) are indicated, p = 4.3 × 10-8. (**p**) qPCR-based *GLUT4* gene expression in shCOBLL1, shGRB14 and shEV adipocytes, one-way ANOVA with Tukey's HSD test, mean + 95% CI, N = 4 biologically independent experiments.

https://doi.org/10.1038/s42255-023-00807-w



Extended data Fig. 6 | Lipocyte profiles of risk versus non-risk haplotype carriers. (**a**–**c**) Differences in morphological profiles between TT (n = 7) and CC (n = 6) allele carriers at day 0 (**a**), day 3 (**b**) and day 8 (**c**) in subcutaneous AMSCs (multi-way ANOVA, significance level < 5% FDR). (**d**–**f**) Differences in morphological profiles between TT (n = 7) and CC (n = 6) allele carriers at (**d**) day 0, (**e**) day 3 and (**f**) day 8 in visceral AMSCs (multi-way ANOVA, significance level < 5% FDR).

https://doi.org/10.1038/s42255-023-00807-w



Extended data Fig. 7 | Generation of COBLL1 mutant mice using CRISPR/ Cas9 editing. (**a**) Overview of the CRISPR/- Cas9 strategy to delete ~20 kb of the Cobll1 gene. The gRNA-targeting sequences (gRNAs) are underlined, and the PAM sequences are indicated in bold. Exons are represented as thick black boxes, introns are indicated as black lines with arrows, and the yellow boxes indicate the DNA-targeting region. Red hexagon indicates a stop codon gen- erating a Cobll1 truncated protein. Agarose gel showing the PCR products generated from DNA containing success- fully targeted Cobll1 from F0 mouse tail genomic DNA. The 308 bp band corresponds to the genomic deletion. (**b**) A real-time quantitative PCR of levels of Cobll1 mRNA in white adipose tissue (WAT), liver and kidney of Cobll1 WT, Cobll1 heterozygous (+/−) and null knockout Cobll1 (−/−) animals to confirm the Cobll1 ablation in knockout animals. Each group was analyzed using 5 different mice and the values were expressed as the mean ± s.e.m and P values by Student's t-test. the experiment was repeated independently two times with similar results. (**c**) Pie chart illustrating non-redundant differential features per channel and class of measurement at day 8 of subcutaneous adipocyte differentiation in rs6712003 homozygous risk compared to non-risk carriers. (**d**, **e**) Differences in morphological profiles between AMSCs from Cobll1−/− mice (n = 3) and WT (n = 4) at (**h**) day 0 (**i**) day 2 (t-test two-sided, significance level < 5%FDR).

Manuscript in preparation

A non-coding metabolic risk variant, rs12454712, associates with context-dependent adipocyte molecular and cellular programs in humans

Samantha Laber*[1], Sophie Strobel*[1,2], Hesam Dashti[1,3], Josep M Mercader[1,4], Saaket Agrawal[1], Alina Ainbinder[1], Julius Honecker[2], Garrett Garborcauskas[1], David R. Stirling[5], Alham Saadat[1], Maria Kost-Alimova[5], Giacomo Deodato[1], Alycen Harney[1], Gregory P. Way[5], Sierra Harken[1], Nasa Sinott-Armstrong[1], Angela Ma[6], Johan LM Björkegren[6], Cecilia Lindgren[1,7,8], Jose C. Florez[1,4,9], Suzanne B. R. Jacobs[1,4], Beth A. Cimini[5], Jesse M. Engreitz[1,10,11], Amit Khera[1,9,12], Miriam S. Udler[1,4,9], Anne E. Carpenter[5], Hans Hauner[2], Melina Claussnitzer[#1,4,9]

[1]Programs in Metabolism and Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge.
[2]Institute of Nutritional Medicine, School of Medicine, Technical University of Munich, Germany; Else Kröner-Fresenius-Centre for Nutritional Medicine, School of Life Sciences, Technical University of Munich, Munich, Germany.
[3]Division of Preventive Medicine, Brigham and Women's Hospital, Boston, MA 02215
[4]Diabetes Unit and Center for Genomic Medicine, Massachusetts General Hospital, Boston, Massachusetts, 02114, USA.
[5]Broad Institute of MIT and Harvard, Cambridge, 02142, MA.
[6]Department of Genetics & Genomic Sciences, Institute of Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York, NY, 10029-6574, USA and Department of Medicine, Karolinska Institutet, Karolinska Universitetssjukhuset, Huddinge, Sweden
[7]Big Data Institute, at the Li Ka Shing Centre for Health Information and Discovery, University of Oxford, Oxford, OX3 7FZ, UK.
[8]Wellcome Centre for Human Genetics, University of Oxford, Oxford, OX3 7BN, UK.
[9]Department of Medicine, Harvard Medical School, Boston, MA.
[10]Department of Genetics, Stanford University School of Medicine, Stanford, CA, 94305, USA.
[11]BASE Initiative, Betty Irene Moore Children's Heart Center, Lucile Packard Children's Hospital, Stanford University School of Medicine, Stanford, CA, 94305, USA.

*shared lead authors, alphabetically ordered

[#]corresponding author email address: melina@broadinstitute.org

Abstract:

Hypothesis: Most disease-associated genetic loci map to more than one disease or trait, suggesting they may give rise to complex disease phenotypes through effects on multiple genes, cell types, cell states, and/or cellular programs. In this context, a single genetic variant

may unfold its effect on molecular, cellular and organismal functions in a highly context-dependent manner, reminiscent of molecular and cellular pleiotropy. Here, we test this paradigm by honing in on a genetic risk locus on chromosome 18 that is associated with multiple metabolic traits reminiscent of a lipodystrophy association signature and has been fine-mapped to the same variant, rs12454712, for each trait.
Methods: We analyze a suite of molecular and cellular data types and perform deep phenotypic profiling of cells derived from metabolic risk and non-risk carriers using integrated high-dimensional image-based morphological and transcriptomics profiling data to mechanistically interrogate a genetic risk loci in an unbiased manner.
Results: We show that the risk genotype of rs12454712 modifies expression levels of at least three effector transcripts at specific temporal stages in different metabolic cell types, and results in distinct cellular and morphological representations that might converge to modulate disease susceptibility. Most notably, we found that rs12454712 decreases expression of *BCL2* and *KDSR*, and that both rs12454712 and *BCL2* knockdown lead to increased mitochondrial fragmentation, decreased lipid accumulation, and increased apoptosis in subcutaneous adipocytes. In visceral adipocytes we observed that rs12454712 increases expression of *VPS4B*, and is associated with decreased mitochondrial activity and thermogenesis rate.
Conclusions: Together, our data reveal highly context-dependent effects of a single metabolic risk variant rs12454712 on molecular and cellular functions in metabolic cells, highlighting the complexities underlying disease-associated loci in humans. We showcase a framework of unbiased dissection of mediating molecular and cellular mechanisms.

Main Text:

The overwhelming majority of trait-associated loci identified through genome-wide association studies (GWAS) overlap with loci from multiple traits (Watanabe et al. 2019; Franzén et al. 2016). Such pleiotropy at individual genetic trait loci might arise from a single variant in a given trait-associated haplotype affecting multiple cell types and cell states of action, multiple effector genes, and cellular programs. In fact, genetic variants have been shown to modulate gene expression and splicing processes dependent on cellular context and environmental factors (Findley et al. 2021). Cell context can differ depending on environmental factors, differentiation dynamics, and anatomical location of a given tissue. For instance, adipose tissue present in subcutaneous areas and visceral areas in the abdominal cavity are characterized by substantial

differences at the anatomical, molecular and cellular level, and are known to have physiological, clinical and prognostic implications for cardiometabolic traits (Karlsson et al. 2019).

Here, we showcase a scenario in which a single non-coding genetic variant links to multiple effector genes, and cellular programs depending on its cellular context. More specifically, we chose the *18q21.33* locus which harbors a single variant, rs12454712, in the trait-associated credible set that robustly associates with increased risk of type 2 diabetes (T2D), and falls into a subtype of insulin resistance with a “lipodystrophy-like” fat distribution (low body mass index (BMI), low adiponectin levels, and high-density lipoprotein cholesterol levels, and high triglyceride levels) (Udler et al. 2018; Yaghootkar et al. 2014).

By combining novel experimental and statistical methods, we mechanistically dissect this locus to identify mediating cell types and target genes, as well as developmental time-points of action and cellular functions that could account for the associated phenotypes in humans.

rs12454712 affects three target genes in a context-dependent manner

First, we used PheWAS to jointly analyze hundreds of traits and disease states in several biobanks, and we found rs12454712 to be associated with a number of related metabolic traits (Figure 1a, PheWAS from AMP T2D Knowledge Portal; Figure S1a, PheWAS in UKBB (Taliun et al. 2020)), including increased risk for T2D adjusted for BMI, increased waist-to-hip ratio adjusted for BMI (WHRadjBMI), increased triglyceride levels, decreased BMI, decreased weight, and decreased hip circumference (Walford et al. 2016); (Mahajan et al. 2018); (Pulit et al. 2019). This association signature indicates that the rs12454712 T allele associates with a lean but metabolically unhealthy phenotype compared to the C allele, which is consistent with a clinical presentation of lipodystrophy in TT carriers. While the GWAS association with WHRadjBMI has previously been reported to be female-specific (Shungin et al. 2015), the GWAS association with BMI is not sex specific (Pulit et al. 2019). To identify the potential tissues of action, we next overlapped rs12454712 with chromatin state maps across reference epigenomes from 833 cell types and tissues (Boix et al. 2021) and found that this locus maps to an active regulatory element (Figure 1b, Figure S1c), particularly in adipose and skeletal muscle, suggesting that this locus mediates T2D risk through actions in these two T2D-relevant tissues (Figure 1b). These findings are further supported by predictions from the computational ‘Tissue of Action’ model (Torres et al. 2020) scoring both skeletal muscle and adipose as the tissues that most likely mediate T2D risk at this genetic signal.

The variant rs12454712 maps to the first intron of the *BCL2* (B-cell lymphoma 2) gene, and the *BCL2* promoter is the closest promoter to the locus. Fine-mapping of this locus showed that the 99% credible set for association with T2D consists of only this variant (Mahajan et al. 2018), and there are no variants in linkage disequilibrium with rs12454712, suggesting that rs12454712 is the single causal variant at the locus. Based on these fine-mapping data, we focused our characterization of this locus on this single variant. To identify possible effector transcript(s) mediating risk at the *18q21.33* locus, we used orthogonal genetics and epigenetics based approaches to assess the regulatory architecture surrounding rs12454712 in skeletal muscle and adipose tissue. More specifically, we examined (1) Hi-C in human mesenchymal stem cells (MSCs) which give rise to adipocytes and skeletal myocytes (Dixon et al. 2015), (2) promoter capture Hi-C in adipocytes, (3) Activity-by-contact predictions linking enhancers to genes in adipocytes, adipose tissue and psoas muscle (Fulco et al. 2019), and (4) eQTL data in skeletal muscle from the STARNET consortium (Franzén et al. 2016) and in adipose from in-house differentiating subcutaneous and visceral adipocytes.

Three-dimensional chromosomal conformation HiC data in MSCs revealed that the locus has a concise topologically associated domain (TAD) structure, encompassing three protein-coding genes: *BCL2, KDSR, and VPS4B* (Figure S1b).

In skeletal muscle, activity-by-contact (ABC) (Fulco et al. 2019) target gene predictions implicated *BCL2* as the target gene of the rs12454712 locus. Indeed, we confirm an eQTL in skeletal muscle data from the STARNET consortium, with the TT risk genotype being associated with lower *BCL2* expression compared to the CC non-risk haplotype (Figure 1c, Figure S1e). This is in line with recent studies reporting that BCL2 is a crucial regulator of stimulus-induced autophagy *in vivo* and is required for muscle glucose homeostasis (Fernández et al. 2018; He et al. 2012).

In adipocytes, promoter Capture Hi-C showed that the regulatory region encompassing rs12454712 forms functional connections to the *BCL2* promoter and the ABC model predicts regulation of *BCL2, KDSR* and *VPS4B* (Figure 1c; Figure S1d). We and others have previously found that T2D and obesity risk variants can show context-specific effects during adipogenesis (Claussnitzer et al. 2015; Sobreira et al. 2021; Sinnott-Armstrong et al. 2021). Accordingly, we explored the potential genotype-driven gene expression differences at this locus across four

time points during differentiation of both subcutaneous and visceral adipocytes (Figure 1d-e). Because the rs12454712 genotype has been associated with sex-specific effects we balanced distribution of sexes as much as possible given limited sample size (see Methods) and we adjusted for sex for all subsequent analyses. We observed a dose-dependent effect of the rs12454712 genotype on *BCL2* and *KDSR* expression in undifferentiated subcutaneous AMSCs (day 0) with significantly reduced expression associated with the TT risk genotype compared to the CC genotype (Figure 1e). These genotype-dependent gene expression changes were blunted following induction of adipogenic differentiation (Figure S1f, upper panel). In visceral AMSCs, we found a dose-dependent effect where the TT risk genotype is associated with increased *VPS4B* expression at day 0 (Figure 1e), but had no effect after differentiation was induced (Figure S1f, lower panel). Thus, the effect of rs12454712 on *BCL2* and *KDSR* in subcutaneous and *VPS4B* expression in visceral AMSCs is specific to preadipocytes, which is consistent with the absence of eQTLs in mature adipocytes. In line with these findings there are no published eQTLs in adipose tissue from both depots in GTEx and STARNET (Consortium and GTEx Consortium 2017; Franzén et al. 2016).

Together, these data point towards a regulatory network in which rs12454712 affects at least three target genes (*BCL2, VPS4B, KDSR*) in at least three tissue types (subcutaneous preadipocytes, visceral preadipocytes, skeletal muscle) at specific developmental windows.

rs12454712 alters apoptotic and lipid degradation processes in subcutaneous adipocytes

Given that rs12454712 is most strongly associated with body fat distribution, which has been shown to be primarily driven by effects in adipose tissue (Shungin et al. 2015; Pulit et al. 2019), we next set out to dissect the mechanistic consequences of rs12454712 variation in adipocytes from both subcutaneous and visceral adipose depots.

To identify possible functional consequences of rs12454712 in adipocytes, we compared cell morphological profiles from homozygous TT, homozygous CC, and heterozygous CT allele carriers in primary human subcutaneous and visceral AMSCs throughout adipocyte differentiation (Figure 2a) using LipocyteProfiler, an unbiased high-content image-based profiling assay for metabolic cell types (Laber et al., n.d.). In brief, LipocyteProfiler generates high-dimensional representations of morphological and cellular profiles consisting of ~3,000 features describing the structure, function, and relationship between cellular organelles as imaged in four channels, namely AGP (actin cytoskeleton, Golgi and plasma membrane), Lipid

(lipid droplets and cytoplasmic RNA), Mito (mitochondria) and DNA (nucleic-acid) (Laber et al., n.d.).

We profiled subcutaneous and visceral adipocytes at four stages of differentiation, including the adipocyte progenitor state (Day 0), early differentiating adipocytes (Day 3), intermediate differentiation (Day 8) and fully differentiated adipocytes (Day 14). In light of limited sample sizes for deep phenotypic profiling of differentiating adipocytes in our experiments (TT: n=11, females n=7, males n=4: CC: n=5, females n=4, males n=1: CT: n=9, females n=8, males n=1), we were not able to stratify by sex. Instead we adjusted for sex in all analyses to account for sex-specific effects as observed for the association of rs12454712 with WHRadjBMI. We found that the morphological profiles of subcutaneous AMSCs with rs12454712 TT and CC genotypes differed significantly at the later stages of adipocyte differentiation (day 8 and day 14) rather than the preadipocyte cell state (Figure 2b). On day 8 of differentiation, we found 172 non-redundant features to be significantly different between the risk and the non-risk genotypes (FDR<0.05), most of which map to mitochondria-related features (Figure 2b). We visually confirmed a genotype-driven effect on mitochondrial features in subcutaneous adipocytes at day 8 in representative images of cells from each genotype. In detail, we observed higher mitochondrial stain intensity (Figure 2c, left) and an overall enrichment of mitochondrial features among upregulated features (Figure 2c, middle, Table S1) in the TT risk compared to the CC non-risk subcutaneous adipocyte. Three of the most significant feature differences between genotypes on day 8 were features informative for the structural appearance of mitochondria indicating a cellular signature of increased mitochondrial fragmentation. Other top-scoring features are mitochondrial intensity features (Figure 2c, right) indicating an increased mitochondrial membrane potential in TT risk cells (Laber et al., n.d.).

Intriguingly, although target gene expression changes are restricted to undifferentiated preadipocytes (*BCL2* and *KDSR* in subcutaneous adipocyte progenitors), the described genotype-driven cellular consequences on mitochondria manifest in maturing adipocytes. To further assess the effect of target gene expression changes in adipocyte progenitors on the function of mature adipocytes, we next correlated *BCL2*, *KDSR* and *VPS4B* gene expression across 26 donor-derived subcutaneous preadipocyte lines at day 0 (the cell stage in which we saw a genotype-driven effect on *BCL2* and *KDSR* gene expression) with their morphological profile at day 8 (the cell stage where we observed genotype-driven effects on mitochondrial morphology and function). Features that associate with *BCL2* gene expression at day 0 mostly

map to the mitochondrial and less to the AGP channel. Features that associate with *KDSR* expression levels map to mitochondria, concordant with genotype-driven effects on the morphological profile when comparing TT with CC carriers at this time-point (Figure S2a). We further observed that effect sizes of these *BCL2* and *KDSR* associated mitochondrial features correlated with rs12454712 genotype-dependent features on day 8 of differentiation (*BCL2* $p < 2.2\times10^{-16}$, cor = 0.676; *KDSR* $p < 2.2\times10^{-16}$, cor = 0.694, pearson correlation analysis) (Figure S2b-c). These data support the notion that *BCL2* and *KDSR* expression in preadipocytes is associated with the cellular genotype-dependent phenotype in maturing adipocytes, despite the absence of rs12454712-driven *BCL2* or *KDSR* gene expression changes at this developmental time point.

In terminally differentiated subcutaneous AMSCs (day 14), the TT risk genotype manifested in a cellular profile that differed in 171 features from the CC non-risk genotype. These genotype-dependent features spread across all organelles and feature classes (Figure 2b, d). Similar to day 8, day 14 adipocytes from TT risk genotype carriers showed a difference in mitochondrial stain patterns (Cells_Texture_AngularSecondMoment_Mito_20_01, Cells_Texture_Correlation_Mito_5_00, Cells_Texture_Entropy_Mito_20_01) and higher number of small mitochondrial fragments (Cytoplasm_Granularity_4_Mito) (Figure 2d, right, Figure S3h, Table S2), indicating that mitochondrial structure continues to be altered in adipocytes from TT risk genotype carriers in a manner consistent with increased mitochondrial fragmentation. Further, adipocytes from TT risk genotype carriers show more and larger lipid droplets compared to adipocytes from non-risk genotype carriers (Cells_Mean_LargeLipidObjects_AreaShape_Area, Cytoplasm_Granularity_4_Lipid) (Figure 2d right, Table S2), which is also visible when comparing representative cells from both genotypes (Figure 2d, left). Additionally, adipocytes from TT genotype carriers had smaller nuclei (Nuclei_AreaShape_MedianRadius), fewer cell neighbors (Cells_Neighbors_NumberOfNeighbors_Adjacent), and a more condensed cytoplasm (Cytoplasm_AreaShape_Compactness) (Table S2). Increased mitochondrial membrane potential and mitochondrial fission/fragmentation are early and fundamental hallmarks of apoptosis, a process that progresses into a distinct set of physical changes involving the cytoplasm, nucleus, and plasma membrane (Ly, Grubb, and Lawen 2003). Consistent with an apoptotic phenotype, we observed genotype-dependent morphological changes in different cellular organelles of adipocytes that have been previously described for apoptotic processes, including accumulation of cytoplasmic lipid droplets composed largely of neutral lipids (Boren

and Brindle 2012), smaller nuclei potentially due to condensed chromatin following endonuclease fragmentation, and round-shaped cells which lose cell-cell contact and shrink potentially due to disintegration of cell junctions (Ly, Grubb, and Lawen 2003) (Table S2).

To further resolve rs12454712-driven molecular pathways, we next integrated RNAseq data with LipocyteProfiler image-based genotype-dependent features at day 8 of subcutaneous adipocyte differentiation. Using a linear regression model, we identified 2,539 genes significantly associated with the morphological and cellular profile of rs12454712 in subcutaneous adipocytes (FDR 0.1%); these genes are significantly enriched for pathways characterizing fatty acid catabolic process (GO:0009062) and associated with apoptosis (GO:0043065) (Table S3). Together, both morphological profiling and gene expression results indicate rs12454712 alters apoptotic and lipid degradation processes in primary human subcutaneous adipocytes.

BCL2 silencing mimics morphological and cellular consequences of rs12454712 in subcutaneous adipocytes

We find that *BCL2* is the target gene most strongly altered by rs12454712 in subcutaneous adipocytes (Figure 1), which is a well-known anti-apoptotic gene (Nagel et al. 2014). To test if decreased *BCL2* induces a cellular signature that resembles the morphological profile of rs12454712 (i.e. a state of increased apoptosis), we used siRNA-mediated perturbation of *BCL2* in primary subcutaneous AMSCs from five normal-weight female individuals (~60% knockdown efficiency; Figure 2e, Figure S3a) and assessed cell number, cell morphology, and mitochondrial respiration. By day 14, *BCL2* knockdown (KD) reduced cell numbers by ~50% (Figure S3b, Figure 2f). Interestingly, the pro-apoptotic consequences of *BCL2* loss are restricted to mature adipocytes (days 8 and 14), as there was no difference in cell numbers in AMSCs before induction of adipogenesis and in early differentiation (days 0 and 3) (Figure 2f). This is in line with previous studies reporting anti-apoptotic functions of BCL2 family members are restricted to differentiating cells, and not detected in mesenchymal stem cells (Oliver et al. 2011). When comparing LipocyteProfiles between BCL2-KD and non-targeting control AMSCs (siNT), we found that BCL2-KD alters cellular profiles throughout differentiation, with the strongest effect in day 14 adipocytes where mitochondrial and lipid-related features predominate (Figure 2g, Figure S3c, Table S4). We found that BCL2-KD increases Mito texture and intensity (Figure S3d), which was also visible when comparing representative cells from *BCL2* silenced and siNT (Figure S3e). We also observed that smaller Mito granularity features are increased in BCL2-KD adipocytes, while the larger granularity measurements are decreased (Figure S3f), suggesting

more fragmented mitochondria in *BCL2* knockdown adipocytes (BCL2-KD) compared to siNT. This phenotype suggests that cells following BCL2-KD are characterized by a mitochondrial fission process, which is supported by a negative correlation of gene expression levels of *hFis*, a mitochondrial fission gene, with larger granularity measures across adipocytes from 26 individuals (Figure S3g). Inversely, gene expression of *MFN2*, a mitochondrial fusion gene, correlated negatively with smaller and positively with larger granularity measures, suggesting that the mitochondrial fragmentation phenotype observed in adipocytes from TT risk genotype carriers and BCL2-KD adipocytes is indicative of increased mitochondrial fission (Figure S3g-h). Using RNA-seq based transcriptome-wide gene expression measurements of *BCL2* knockdown versus siNT adipocytes showed that pro-apoptotic genes (e.g. *TNFSF10*, *DCN, CALU, TAGLN*) were significantly upregulated in subcutaneous adipocytes at day 14, whereas genes related to lipid metabolism (e.g. *LIPE, PLIN4, FASN* and *APOE*) were downregulated (Figure 2h, Table S5), similar to what we had observed for the rs12454712 risk genotype.

We next investigated whether these BCL2-KD-induced morphological changes translate into altered mitochondrial respiration in a mitochondrial stress test using the Seahorse Bioflux Analyser. BCL2-KD increases the oxygen consumption rate (OCR) and extracellular acidification rate (ECR) (Figure S3i), revealing a more energetic profile in BCL2-KD adipocytes compared to siNT cells (Figure S3i). This is consistent with a recent study reporting an increase in maximal respiration in subcutaneous adipocytes of metabollically unhealthy obese subjects compared to metabolically healthy obese (Böhm et al. 2020), suggesting that mitochondrial impairment and a possible increased mitochondrial permeability could be an underlying mechanism in subcutaneous adipocytes.

Together, these results suggest that BCL2-KD induces mitochondrial fission, reduces cell number, and increases OCR and extracellular acidification rate in subcutaneous adipocytes. These cellular consequences are consistent with effects of the rs12454712 TT risk genotype on mitochondrial fragmentation processes and are consistent with initiation of the apoptotic program in primary human subcutaneous adipocytes.

rs12454712 associates with reduced mitochondrial thermogenic capacity and increased lipid accumulation in mature visceral adipocytes

Next, we sought to investigate which cellular programs underlie the association of rs12454712 with a lipodystrophy-like phenotype in primary human visceral AMSCs. We observed an

rs12454712 genotype-driven effect predominantly on mitochondrial-associated morphological features at day 14 of differentiation (Figure 3a, Table S6). Specifically, we observed higher cell number, decreased mitochondrial stain intensity (indicative of lower mitochondrial membrane potential), and features indicating differences in mitochondrial structure in TT risk genotype carriers (Figure 3b). To assess whether these genotype-associated cellular changes in visceral adipocytes could be driven by *VPS4B* expression changes, we correlated *VPS4B* expression at day 0 (where we observed genotype-mediated effect on target gene expression) with both transcriptome-wide gene expression and LipocyteProfiler-derived features on day 14 of differentiation. We observed that *VPS4B* expression in preadipocytes correlates with genes enriched for OXPHOS pathways (Figure 3c) in differentiated adipocytes. In line with this effect on oxidative phosphorylation processes we found an enrichment of image-based features mapping to mitochondria among the correlated features (Figure S4a, upper panel). Indeed, OXPHOS genes that negatively correlate with *VPS4B* expression were also negatively correlated with the TT risk genotype (Figure 3d). Higher expression of *VPS4B* (as seen in preadipocytes of the TT risk genotype) correlated negatively with Mito intensity and texture, and positively with lipid texture features and a feature describing overlap between mitochondrial and lipid stain (Cells_Mean_LargeLipidObjects_Correlation_K_Lipid_Mito; Figure S4a, bottom panel), suggesting a profile of reduced mitochondrial membrane potential and higher colocalization of mitochondria and lipid droplets. This cellular profile may indicate altered thermogenesis, as mitochondria are anchored to lipid droplets during ATP production and lipid droplet expansion, but dissociate from lipid droplets during browning-induced fatty acid oxidation (Benador et al. 2018).

To assess whether visceral AMSCs from rs12454712 TT risk genotype carriers resemble the morphological profiles that we would expect in a state of reduced thermogenic capacity, we used two thermogenic stimulatory agents, i.e. isoproterenol and free fatty acids (FFA). We first compared rs12454712-associate image-based signatures with that of isoproterenol-treated visceral AMSCs at day 14 (see (Laber et al., n.d.). Isoproterenol is an adrenergic receptor agonist known to induce adipocyte browning and increase thermogenic capacity (Miller et al. 2015)). We found 313 features that are significantly different following isoproterenol treatment and 159 features that differ between the rs12454712 genotypes (FDR<5%) (Figure 3e). We found that the LipocyteProfile associated with the protective C allele resembled the response to isoproterenol, particularly pointing to features informative for lipid and mitochondria (Figure 3e). As an orthogonal approach, we differentiated visceral AMSCs with FFA (oleic and linoleic acid;

see methods), which are known stimulators of thermogenesis (Barquissau et al. 2016). We characterized rs12454712-driven morphological consequences at three time points of differentiation (day 3, 8, 14) and confirmed a predominately mitochondrial and lipid-related morphological profile which is similar to the response to isoproterenol (Figure S4b-c). To further support that the rs12454712 risk genotype affects thermogenesis rate in visceral AMSC derived adipocytes following FFA treatment, we compared the percentage of predicted thermogenesis-active adipocytes between genotypes using the BATLAS webtool (Perdikari et al. 2018) on RNA-Seq transcripts of cell cultures. We confirmed that FFA treatment increases the percentage of thermogenically active adipocytes by 36% across all individuals (Figure 3f, right). Intriguingly, TT risk homozygotes show a lower percentage of thermogenically active adipocytes compared to CC homozygotes following FFA treatment, despite similar basal levels (Figure 3f, left), suggesting that visceral AMSCs from TT risk homozygotes have a blunted response to thermogenic stimuli.

Together, these data support a model in which elevated *VSP4B* expression in TT risk genotype carriers causes reduced mitochondrial thermogenic capacity and increased lipid accumulation in mature visceral adipocytes. We note that we did not observe a similar association with mitochondrial processes for the other potential target genes, *BCL2* and *KDSR*, in visceral adipocytes, furthering the evidence that *VPS4B* and not *BCL2* or *KDSR* is the effector gene in visceral adipocytes. Our work establishes an adipose depot-dependent effect, in which the T risk allele associates with distinct target genes and cellular signatures in adipocytes from subcutaneous and visceral adipose tissue.

Polygenic risk of adverse body fat distribution converges on mitochondrial impairment in subcutaneous adipocytes

Finally, we sought to decipher whether the mechanisms identified for rs12454712 would align with global cellular drivers of polygenic risk for increased WHRadjBMI. We compared morphological and cellular profiles of high and low polygenic risk female individuals for WHRadjBMI (Figure 4a) ((Laber et al., n.d.), see Methods for details) and observed 113 significant differences between subcutaneous adipocytes of low and high polygenic risk groups (Figure 4b). This effect was specific for subcutaneous adipocytes as we did not observe an effect on morphological or cellular programs in visceral adipocytes when ascertaining the effects of polygenic risk scores for WHRadjBMI on image-based feature representations. Subcutaneous adipocytes from high polygenic risk carriers are characterized by higher

mitochondrial intensity and mitochondrial granularity features and higher count of large lipid-related objects (Figure 4c-d). To identify possible mediating molecular pathways for the morphological and cellular changes, we used a linear regression model of LipocyteProfiler features and transcriptome-wide gene expression data from matched AMSCs at day 14 of differentiation and identified 7,146 genes that correlated with WHRadjBMI-mediated features. More specifically, the identified WHRadjBMI morphological profile was enriched for genes involved in deficiency of tricarboxylic acid cycle (TCA) pathway (WP2453; WP78), fatty acid oxidation (WP143, WP368), and apoptosis modulation and signaling (WP1772) (Figure 4E) (Table S7), similar to what we observed for the rs12454712 genotype in subcutaneous AMSCs (see Figure 2). Our data suggest that polygenic risk of adverse body fat distribution at least in part involves apoptotic pathways characterized by mitochondrial impairment in subcutaneous adipocytes.

Discussion

In our work we decipher distinct granular mechanisms underlying the single metabolic risk variant, rs12454712, which associates with a trait signature reminiscent of polygenic lipodystrophy. We show that the rs12454712 T risk allele affects (1) *BCL2* gene expression in skeletal muscle, (2) *BCL2* and *KDSR* gene expression in primary human subcutaneous adipocyte progenitors, and (3) *VPS4B* gene expression in primary human visceral adipocyte progenitors in a cell-autonomous manner. We further associated variant effects to depot-specific cellular programs, namely mitochondrial fragmentation and increased apoptosis rate in subcutaneous adipocytes, and decreased thermogenesis rate in visceral adipocytes (Figure 4f). Consistent with our results, rs12454712 has recently been implicated in affecting *BCL2* expression in Simpson-Golabi-Behmel syndrome adipose cells, and BCL2 has been implicated by the same group in apoptosis rate in murine 3T3-L1 adipocytes (Dong et al., n.d.).

In light of the fact that the *18q21.33* locus harbors only a single variant in the fine-mapped credible set, our work highlights the complexities in dissecting disease-associated loci in humans, which typically harbor dozens of tightly linked genetic variants. This complexity is further observed in analyses of adipose depots in humans. When we investigated sex-specific associations of rs12454712 with MRI-quantified visceral (VAT), abdominal subcutaneous (ASAT), and gluteofemoral (GFAT) adipose tissue volumes adjusted for BMI and height in 37,641 UK Biobank participants (Agrawal, Klarqvist, et al., n.d.) 2021; (Agrawal, Wang, et al.,

n.d.) rs12454712 is associated with a nominal reduction in GFATadjBMI volume in females (beta female = -0.032, P = 0.002; beta male = -0.018, P = 0.09); a nominal increase in ASATadjBMI volume (beta male = 0.015, P = 0.15; beta female = 0.024, P = 0.02), and has no effect on VATadjBMI volume. These data indicate that–similar to depot-specific cellular effects observed in our study–rs12454712 has depot-specific effects on complex adipose volume traits. As we were not able to test cellular effects in gluteofemoral AMSCs, future phenotypic profiling experiments in AMSCs derived from the gluteofemoral region will further our understanding of rs12454712 depot-specific effects on adipocyte metabolism and how these translate to complex physiological phenotypes.

Here, we showcase a framework based on integration of high-content image based profiling coupled with transcriptomics in a relatively small set of primary human AMSCs that enables systematic and unbiased mechanistic interrogation of genetic risk loci. Specifically, our framework allowed us to i) unravel the spatio-temporal complexities of a risk locus that modulated target gene expression at a specific developmental window and manifested in cellular phenotypes at another, and to ii) identify cellular mechanism by comparing genotype-driven morphological profiles with signatures of cellular traits (e.g. ROS, apoptosis and thermogenesis).

Due to their apoptotic properties, BCL2 inhibitors are currently used in the clinic for chronic lymphocytic leukemia and small lymphocytic lymphoma. Importantly, pharmacological inhibition of *BCL2* using venetoclax has been reported to cause hyperglycemia in 16% of patients and severe hyperglycemia in 5% in a 1 year follow up clinical trial (Roberts et al. 2016), as well as loss of body weight (de Vos et al. 2018; Konopleva et al. 2014; Leung, Thomé, and Dispenzieri 2018), indicating that BCL2 inhibition can lead to systemic metabolic adverse effects including reduced insulin sensitivity, even during weight loss.

Our study has several limitations. While this study focused on molecular and cellular effects of the rs12454712 target gene *BCL2,* future studies need to validate the role of VPS4B and KDSR on cellular functions using directed perturbations. Further, it will be necessary to address the role of rs124554712 in other tissues, especially in skeletal muscle, which will likely be driving at least part of the insulin resistance association at this locus. We exploited genotype-driven morphological profiles in primary cells from a relatively small cohort of 26 donors and our study was potentially underpowered to detect effects of polygenic risk for WHRadjBMI in visceral cells.

The rs12454712 association with WHRadjBMI has been shown to be sex-specific, while the association with BMI and T2D does not appear to be sex-specific (Shungin et al. 2015; Pulit et al. 2019) Due to the small sample size we were not able to stratify the data by sex and chose to adjust for sex instead. Future studies comparing effects in females and males will be imprtant to evaluate sex specificity of the effects we observed. Together, future population-scale profiling of primary human cells, enabled by larger and standardized biobanks of human primary cells, will be critical and transformative to further validate the findings and dissect the mechanisms presented in this study.

Figure Legends

Fig.1 rs12454712 is associated with a lipodystrophy-like phenotype and is predicted to regulate target genes in adipose tissue.

a. Phenome-wide association study (PheWAS) (Taliun et al. 2020) for rs12454712 shows associations with a number of metabolic traits, including insulin sensitivity, BMI, BMI-adjusted T2D, T2D, BMI-adjusted waist-to-hip ratio (WHRadjBMI) and WHR. See also Figure S1a
b. The *18q21.33* locus contains no other variants in high linkage disequilibrium with the lead variant rs12454712 and overlaps active regulatory marks in adipose tissue and skeletal muscle. Image visualized using the T2D knowledge portal. See also Figure S1c
c. Promoter Capture Hi-C and ABC. See also Figure S1b+d
d. Schematic description of RNA-seq profiling in patient-derived subcutaneous and visceral adipose-derived mesenchymal stem cells (AMSCs) throughout adipogenesis (day 0, 3, 8, and 14).
e. Gene expression of target genes *BCL2*, *KDSR* and *VPS4P* in visceral and subcutaneous pre-adipocytes at day 0 of differentiation (subcutaneous: n=3 CC non-risk, n=10 CT heterozygotes; n=11 TT risk; visceral: n=4 CC non-risk, n=9 CT heterozygotes; n=14 TT risk). See also Figure S1e

Fig.2 Genotype-specific effect of rs12454712 on ROS and apoptosis in subcutaneous adipocytes

a. LipocyteProfiler (LipocytePainting staining of cells and LipocyteProfiling feature extraction) was performed in subcutaneous and visceral AMSCs of eleven risk and five

non-risk genotype carriers for rs12454712 at four time points during adipocyte differentiation (day 0, 3, 8, and 14).

b. LipocyteProfiler reveals rs12454712-mediated effect on morphological phenotype manifest at later stages of differentiation (day 8 and day 14).
c. Representative images of subcutaneous AMSCs from TT risk (top) and CC non-risk (bottom) genotypeat day 8 of differentiation stained using LipocytePainting. Scale bar = 10um. At day 8 of differentiation, 172 non-redundant significant features (< 5%FDR) differ between the genotypes, mostly mapping to mitochondria-related features. Representative significant features different between genotypes are Cytoplasm_Texture_InfoMeas2_Mito_5_01, Cells_Intensity_MaxIntensity_Mito, Cytoplasm_Texture_Entropy_Mito_20_00. See also Table S1 + S8. Y-axis shows LP units (normalised LP values across eight batches, see methods).
d. Representative images of subcutaneous AMSCs from TT risk (top) and CC non-risk (bottom) genotypes at day 14 of differentiation stained using LipocyteProfiler. Scale bar = 10um. At day 14 of differentiation, 174 non-redundant significant features (< 5%FDR) differ between the genotypes. Representative significant features different between genotypes are Cells_Texture_Correlation_Mito_5_02, Cytoplasm_Granularity_4_Mito, Cells_Mean_LargeLipidObjects_AreaShape_MedianRadius. See also Table S2 + S8 Y-axis shows LP units (normalised LP values across eight batches, see methods)
e. *BCL2* was silenced in subcutaneous AMSCs using siRNA from five normal-weight female individuals for assessment of cell number, cell morphology throughout differentiation using LipocyteProfiler and mitochondrial respiration using the Seahorse Bioflux Analyser at day14 of differentiation. AMSCs were treated with siBCL2 for 3 days before induction, at which point knockdown efficiency was ~60% and maintained until terminal differentiation (measured in n=3). See also Figure S3.
f. BCL2-KD reduced cell numbers measured by Cells_Number_Object_Number in subcutaneous AMSCs at day 8 and day 14 of differentiation (n=5). See also Figure S3b
g. LipocyteProfiles of BCL2-KD and non-targeting control AMSCs (n=5) show differences with highest effect size in mitochondrial and lipid-related features at day 14 of adipocyte differentiation. See also Figure S3c for other time-points and individual features. See also Table S4.
h. BCL2-KD mediated gene expression changes show upregulation of pro-apoptotic genes and down-regulation of genes involved in lipid metabolism.

Fig. 3 Genotype-dependent effect of rs12454712 on thermogenesis and lipid accumulation in visceral adipocytes

a. LipocyteProfiler in visceral AMSCs of eleven risk and five non-risk genotype carriers for rs12454712 at four time points during adipocyte differentiation (day 0, 3, 8 and 14) revealed a genotype-driven change of mitochondrial features specifically on day 14. See also Figure 4b with free fatty acid treatment
b. Genotype-dependent differential features are Cells_Texture_InverseDifference_Moment_Mito_10_02, Cells_Intensity_Mean_Intensity_Mito, and Cells_Neighbors_NumberOfNeighbors_10 at day 14 of differentiation. See also Table S6 + S8. Y-axis shows LP units (normalised LP values across eight batches, see methods). Representative images of visceral AMSCs from TT risk (right) and CC non-risk (left) genotype at day 14 of differentiation stained using LipocytePainting. Scale bar = 10um.
c. *VSPB4* expression in preadipocytes is linked to gene expression changes enriched in OXPHOS pathways.
d. Identified OXPHOS genes associated with *VPS4B* expression at day 0 showed the same direction of effect between genotypes at day 14 of differentiation.
e. Significant features of morphological profile of isoproterenol treatment and/or between the rs12454712 risk and non-risk genotype (<FDR 5%) mapped predominantly to the lipid and mitochondria channels and overlapped in their direction of effect. See Figure S4c with FFA treatment
f. BATLAS-based estimates of thermogenesis-active adipocytes from RNA-seq data in differentiated visceral adipocytes in response to FFA (Perdikari et al. 2018). FFA treatment increased the percentage of thermogenic active visceral adipocytes. TT risk genotype carriers were characterized by a lower percentage of thermogenesis active adipocytes compared to CC genotype carriers following FFA treatment, despite similar basal levels.

Fig. 4 Polygenic risk for WHRadjBMI manifests in an apoptotic cellular profile in subcutaneous adipocytes

a. Schematic of distribution of polygenic risk for WHRadjBMI. Low risk (cyan) refers to the bottom 25%, high risk (orange) refers to the top 25%, percent within the assessed cohort.

b. Lipocyte profile of WHRadjBMI in subcutaneous AMSCs at day 14 mapped to predominantly lipid- and mitochondrial-related features.
c. Genotype-dependent differential representative features are Cells_Intensity_Mean_Intensity_Mito, Cytoplasm_Granularity_3_Mito, and Cells_LargeLipidobjects_Count. Y-axis shows LP units (normalised LP values across eight batches, see methods)
d. Representative images of subcutaneous AMSCs at day 14 from low (left) and high (right) polygenic risk female carriers for WHRadjBMI stained using LipocytePainting. Scale bar = 10um.
e. Pathway enrichment analysis of genes connected to morphological profile of WHRadjBMI reveal enrichment of lipid degradation and apoptotic processes.
f. Schematic summarizing effect of rs12454712 and high polygenic risk for WHRadjBMI on mitochondrial function, lipid metabolism and cell number in subcutaneous AMSCs, and thermogenesis capacity in visceral AMSCs, resulting in a lipodystrophy-like phenotype. See also Table S9.

Supplementary Figure Legends

Fig. S1 Regulatory landscape around rs12454712.

a. Phenome-wide association study (PheWAS) using UKBB data visualised in the browser http://big.stats.ox.ac.uk/ (Elliott et al. 2018) for rs12454712 shows associations with a number of obesity- and fat-related (e.g. Hip circumference) as well as muscle-related (e.g. Basal metabolic rate) traits.
b. Hi-C in MSCs (Dixon et al. 2015) visualised using the 3D Genome browser (Wang et al. 2018)
c. rs12454712 lies within an active regulatory element assessed by overlapping the locus with chromatin state maps across 833 reference epigenomes (Boix et al. 2021).
d. Activity-by-contact (ABC) target gene prediction (Fulco et al. 2019) in adipocyte nuclei (ENCODE Project Consortium 2004; Zhou et al. 2015) and adipocytes differentiated from adipose-derived mesenchymal stem cells (Schmidt et al. 2015).
e. eQTL starnet browser. (mssm.STARNET.edu)
f. Rs12454712 genotype-dependent expression of target genes *BCL2*, *KDSR* and *VPS4B* at day 3, 8, and 14 in subcutaneous (top panel) and subcutaneous (bottom panel) AMSCs.

Fig. S2 rs12454712 affects target gene expression in preadipocytes manifesting on morphology in differentiated adipocytes

a. Gene expression of *BCL2* and *KDSR* at day 0 are associated with morphological consequences at day 8 of differentiation of predominantly mitochondrial features in subcutaneous AMSCs. *VPS4B* expression at day 0 shows no morphological consequences at day 8 of differentiation in subcutaneous adipocytes. Barplots show percentages of significantly affected features (FDR <5%) from all features of each channel separately.
b. Correlation between effect sizes of mitochondrial features calculated based on gene expression of *BCL2*, *KDSR*, and *VPS4B* at day 0 and genotype-driven effect at day 8 demonstrate mitochondrial profiles at day 8 of *BCL2* and *KDSR* gene expression at day 0 and rs12454712-driven effect overlap in terms of effect size and significance. Red = feature is significantly affected by gene expression at day 0 and rs12454712; blue = feature is significantly affected by either gene expression at day 0 or rs12454712; grey = feature is significantly affected by neither gene expression at day 0 nor rs12454712 .
c. Correlations of effect size of lipid, AGP (actin, golgi, plasma membrane), and DNA features calculated based on gene expression of *BCL2*, *KDSR*, and *VPS4B* at day 0 and haplotype-driven effect at day 8.

Fig. S3 BCL2-KD affects mitochondrial structure and function

a. BCL2-KD efficiencies at day 0 and day 14 of differentiation as evaluated by qPCR-based BCL2 mRNA levels, n=3.
b. Cell numbers as assessed using Hoechst staining intensity in BCL2-KD and control subcutaneous AMSCs of three individuals at day 14 shows reduced cell numbers after BCL2 knockdown, n=3.
c. LipocyteProfiles of BCL2-KD and non-targeting control AMSCs at day 0, 3, 8 of differentiation, n=5. See also Figure S2h for day 14.
d. BCL2-KD increases mitochondrial texture (Cytoplams_Texture_InfoMeas2_Mito_5_01; Cytoplasm_Texture_Entropy_Mito_20_00) and intensity (Cells_Intensity_MaxIntensity_Mito), resembling the TT risk genotype.
e. Representative images of BCL2-KD (right) and control (left) at day 14 of differentiation.
f. Mitochondrial granularity features (1-16 dot size) were increased in BCL2-KD adipocytes specifically for the smaller measurements (*Cytoplams_Granularity_Mito_2*), and

decreased for the larger measurements, indicating more fragmented mitochondria in BCL2-KD adipocytes compared to siNT-treated cells.

g. Gene expression of *hFIS1*, a mitochondrial fission gene, correlated negatively larger granularity measures across adipocytes derived from 26 individuals. Gene expression of *MFN2*, a mitochondrial fusion gene, correlated negatively for smaller and positively with larger granularity measures.
h. Adipocytes from TT risk genotype carriers showed increased mitochondrial fission compared to CC non-risk genotype carriers.
i. Mitochondrial stress test using the Seahorse Bioflux Analyser in BCL2-KD and control AMSCs (n=3) at day 14 of differentiation shows increased maximal OCR (top) in BCL2-KO AMSCs. A combined measure of OCR and ECR (a measure of extracellular acidification), revealed that BCL2-KD resulted in a more energetic profile compared to control.

Fig. S4 rs12454712-driven consequences on visceral AMSCs under free fatty acid treatment

a. Gene expression of *VPS4B* at day 0 shows morphological consequences at day 14 of differentiation of predominantly mitochondrial and lipid features in visceral AMSCs treated with free fatty acid during adipogenesis. Barplots show percentages of significantly affected features (FDR <5%) from all features of each channel separately. Heatmap shows t-statistics of significantly affected features at day 14 from *VPS4B* expression at day 0.
b. LipocyteProfiler in visceral AMSCs of eleven risk and five non-risk genotype carriers for rs12454712 at three time points during adipocyte differentiation (day 3, 8 and 14) under free fatty acid treatment revealed a genotype-driven effect specifically on day 14 on lipid and mitochondrial features. Cells_Median_Intensity_Lipid, Cells_Mean_Intensity_Mito are reduced in TT genotype carriers. See also Figure 3a without free fatty acid treatment
c. Schematic of LipocyteProfiling on visceral AMSCs differentiated under FFA treatment comparing morphological profiles haploytes and isoproterenol stimulation at day 14. Significantly affected features (5%FDR) of both comparisons are predominantly lipid and mitochondrial features and show the same direction of effect.

Supplementary Tables

Methods

Human primary AMSCs isolation, Munich Obesity BioBank/MOBB

We obtained AMSCs from subcutaneous and visceral adipose tissue from patients undergoing a range of abdominal laparoscopic surgeries (sleeve gastrectomy, fundoplication or appendectomy). The visceral adipose tissue is derived from the proximity of the angle of His and subcutaneous adipose tissue obtained from beneath the skin at the site of surgical incision (anterior abdominal region). Additionally, human liposuction material was obtained. Each participant gave written informed consent before inclusion and the study protocol was approved by the ethics committee of the Technical University of Munich (Study № 5716/13). Isolation of AMSCs was performed as previously described (Skurk and Hauner 2012). Briefly, Krebs-ringer phosphate buffer (KRP) containing 4 % bovine serum albumin (BSA) and 200 u/ml collagenase were added to adipose tissue material and incubated for one hour at 37 °C under strong agitation in a water bath. After digestion, the supernatant containing mostly lipids from ruptured mature adipocytes was discarded, while the remaining volume was filtered through a nylon mesh with 2 mm pore size. The stromal vascular fraction was pelleted by centrifugation at 200 g for 10 minutes. The supernatant was removed and the pellet was resuspended in an ammonium chloride-containing buffer for the lysis of erythrocytes. The suspension then was filtered through 250 µm and 150 µm nylon meshes followed by centrifugation at 200 g for 10 minutes. The supernatant was discarded and the pellet was resuspended in an isolation medium. Subsequently, cells were seeded into T75 cell culture flasks. On the next day, cells were washed three times with phosphate-buffered saline (PBS) and the medium was changed to proliferation medium (PM), which was shown to enrich the stromal vascular fraction for adipose-derived mesenchymal stem cells. Following, the medium was changed every three days, and cells underwent proliferation until confluency and cryopreserved afterwards. For a subset of donors, purity of AMCSs was assessed as previously described (Raajendiran et al. 2019). Briefly, cells were stained with 0.05ug CD34, 0.125ug CD29, 0.375ug CD31, 0.125ug CD45 per 250K cells and analyzed on CytoFlex together with negative control samples of corresponding AMSCs.

Differentiation of human AMSCs

For imaging, cells were seeded at 10K cells/well in 96-well plates (Cell Carrier, Perkin Elmer #6005550) and induced 4 days after seeding. For RNAseq, cells were seeded at 40K cells/well

in 12-well dishes (Corning) (subcutaneous CC female n=3, male n=0; CT female n=9, male n=1; TT female n=9, male n=2; visceral CC female n=3, male n=1; CT female n=8, male n=1; TT female n=10, male n=4). Before Induction cells were cultured in proliferation medium (Basic medium consisting of DMEM-F12 1% Penicillin - Streptomycin, 33µM Biotin and 17µM Pantothenate supplemented with 0.13µM Insulin, 0.01ug/ml EGF, 0.001ug/ml FGF, 2.5%FCS). Adipogenic differentiation was induced by changing culture medium to induction medium. (Basic medium supplemented with 0.861µM Insulin, 1nM T3, 0.1µM Cortisol, 0.01mg/ml Transferrin, 1µM Rosiglitazone, 25nM Dexamethasone, 2.5nM IBMX). On day 3 of adipogenic differentiation culture medium was changed to differentiation medium (Basic medium supplemented with 0.861µM Insulin, 1nM T3, 0.1µM Cortisol, 0.01mg/ml Transferrin). Medium was changed every 3 days. Visceral-derived AMSCs were differentiated by further adding 2% FBS as well as 0.1mM oleic and linoleic acid to the induction and differentiation media. For isoproterenol stimulation experiments, 1uM isoproterenol was added to the differentiation media and cells treated overnight.

Genotyping and quality control

Genotyping of all samples was performed in two separate batches using the Infinium HTS assay on Global Screening Array bead-chips. Since the two sets of samples were genotyped with different versions of the beadchips and in different batches, we QCed, imputed, and generated the genome-wide polygenic scores separately and combined the results afterwards.

A 3-step quality control protocol was applied using PLINK(Purcell et al. 2007; Chang et al. 2015), and included 2 stages of SNP removal and an intermediate stage of sample exclusion. The exclusion criteria for genetic markers consisted of: proportion of missingness ≥ 0.05, HWE $p \leq 1 \times 10^{-20}$ for all the cohort, and MAF < 0.001. This protocol for genetic markers was performed twice, before and after sample exclusion. For the individuals, we considered the following exclusion criteria: gender discordance, subject relatedness (pairs with PI-HAT ≥ 0.125 from which we removed the individual with the highest proportion of missingness), sample call rates ≥ 0.02 and population structure showing more than 4 standard deviations within the distribution of the study population according to the first seven principal components. After QC, 35 subjects remained for the analysis for which we had matched LipocyteProfiler imaging data.

BCL2 silencing using siRNA

All silencing experiments were performed on 4 technical replicates. One day before silencing, AMSCs were plated into 96-well plates with 10K cells/well using growth medium. RNA-based silencing of *BCL2* was performed using RNAiMAX Reagent (ThermoFisher #13778075) and following the manufacturer's protocol. Briefly, Lipofectamine® RNAiMAX Reagent was diluted in Opti-MEM medium (Gibco, Cat# 11058021). At the same time, siRNA was diluted in Opti-MEM medium. Then, diluted siRNA was added to the diluted Lipofectamine® RNAiMAX reagent at a ratio 1:1 and incubated for 5min. The concentration of reagents per well in a 96-well plate were 0.5µl (10µM) of silencing oligo (Ambion Cat# 4392421, ID s1915) or negative control duplex (Ambion Cat#4390844), and 1.5µl of lipofectamine RNAiMAX Reagent. The plate was gently swirled and placed in a 37°C incubator at 5% CO2 for three days. Cells were then induced to differentiate following the standard differentiation cocktail or harvested for gene expression analysis to assess knockdown efficiency.

RNA preparation and qPCR

Total RNA was extracted with Trizol (Ambion 15596026) and the Direct-zol RNA MiniPrep Kit (Zymo R2052) following the manufacturer's instructions. cDNA was synthesized with High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems 4368814) following the manufacturer's instructions. qPCR was performed using Thermo Scientific PCR Master Mix (Thermo Scientific K0172) and taqman probes for target gene *BCL2* (Thermo Scientific, Cat#4448892, ID Hs04986394_s1) and housekeeping gene *CANX* (Thermo Scientific, Cat#4448892, ID Hs01558409_m1). Relative gene expression was calculated by the delta delta Ct method. Target gene expression was normalized to expression levels of *CANX*.

Lipocyte Painting and Lipocyte Profiling in human AMSCs

Human primary AMSCs were plated and differentiated in 96-well CellCarrier plates (Perkinelmer #6005550) for 14 days for high content imaging at day 0, day 3, day 8 and day 14 of adipogenic differentiation. On the respective day of the assay cells were stained following the LipocytePainting protocol described in (Laber et al., n.d.).

Image processing, quantification, quality control steps, and normalization of LipocyteProfiler features were performed as previously described in (Laber et al., n.d.).

Subsequent data analyses were performed in R3.6.1using base packages unless noted.
To test whether there is a difference of morphological profiles between the risk versus non-risk genotype a multi-way analysis of variance (ANOVA) was performed adjusted for sex, age, BMI

and batch in visceral and subcutaneous AMSCs at four timepoints of differentiation (github). For *BCL2* knockdown analysis we performed a multi-way analysis of variance (ANOVA) comparing *BCL2* silenced cells versus controls (github). To overcome multiple testing burden p-values were corrected using false positive rate (FDR) described in R package “qvalue”(*Qvalue* n.d.). Features with FDR < 5% were classified to be significantly impacted by the PRS variable (github). To visualize LipocyteProfiles and their effect size ComplexHeatmap Bioconductor package version 2.7.7(Gu, Eils, and Schlesner 2016) was used (github)

RNA-seq

RNA-seq data were processed using FastQC (Krueger and Others 2015) and spliced reads were aligned to human genome assembly (hg19) using STAR (Dobin et al. 2013) followed by counting gene levels using Rsubread R package (Liao, Smyth, and Shi 2019). Next, raw read counts were normalized using DESseq2 R package (Love, Huber, and Anders 2014). For genome-wide differential expression analysis of *BCL2* silenced versus control subcutaneous adipocytes at day 14 of differentiation DESseq2 R package (Love, Huber, and Anders 2014) was used. Genes with an adjusted p-value less/equal than 10e-6 and a fold change of log2 FC > |0.75| were considered as significant (github).

Selecting representative cell images

For each group of interest, cells were pooled and divided into 100 clusters via K-Means clustering (scikit-learn). Individual cells were then sampled from the cluster closest to a theoretical point representing the mean of all object measurements, as determined by a euclidean distance matrix.

Gene expression and LipocyteProfiler feature network

A linear regression model was fitted of 2,760 LP-features and global transcriptome RNA-seq data adjusted for sex, age, BMI and batch in subcutaneous AMSCs at day 14 of differentiation. Gene LP features associations were declared to be significant when passing FDR cut-off of 0.1% FDR. LP features belonging to *Cells* category were used for further analysis. Associations between genes and LP features were visualised using “igraph” R package (Csardi, Nepusz, and Others 2006). Genes that were connected to the morphological and cellular profile between the genotypeat day8 were uploaded to Enrichr to analyse them as a gene list against Gene ontology (GO) pathways (github). Genes that were connected to the morphological and cellular

profile of polygenic risk of WHRadjBMI (generated in (Laber et al., n.d.) were uploaded to Enrichr to analyze against Wiki pathways (github)

Seahorse

The protocol for a standard bioenergetics profile is composed of basal mitochondrial respiration, ATP turnover, proton leak and mitochondrial respiratory capacity. First, oxygen consumption rate (OCR) in basal conditions was determined and used to calculate the basal mitochondrial respiration. After this, 2 μM oligomycin was injected from the first port to inhibit ATP synthase, resulting in an accumulation of protons in the mitochondrial intermembrane space and a reduced activity of the electron transport chain. The resulting decrease in OCR reveals the respiration driving ATP synthesis in the cells, indicating ATP turnover. Residual oxygen consumption capacity can be attributed to the proton leak maintaining a minimal ETC and non-mitochondrial respiration. Next, 2 μM of the mitochondrial uncoupler FCCP was injected which results in an increase in OCR as the proton gradient across the inner mitochondrial membrane is dissipated and ETC resumed. This measurement reflects the maximal mitochondrial respiratory capacity. Finally, 2 μM Rotenone / Antimycin A are injected to completely stop ETC activity and the OCR reading at this phase reflects non-mitochondrial respiration. We normalized all data to the relative number of live cells in each well of the 96-well Seahorse plate.

Oxygen Consumption and Bioenergetics Profile was measured using the XF24 extracellular flux analyzer from Seahorse Bioscience. The protocol used in this assay was adapted from Gesta et al., 2011. For this assay, preadipocytes were counted and 10K cells per well were seeded onto seahorse 96 well plate in 50μl of growth media and left to adhere overnight. The next day, silencing was performed as seen in the previous section. Three days later, cells were induced to differentiate within the seahorse plate following the adipogenic differentiation protocol as described previously. Each cell type was run in 8 replicates. When the cells were terminally differentiated at day 14 post adipogenic induction, the assay was performed. The evening before the assay, the seahorse XF-24 instrument cartridge was loaded with seahorse calibrant and placed in a CO2-free incubator at 37°C overnight.

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

37°C in a CO2-free incubator to allow CO2 to diffuse out of solution. According to the manufacturer's protocol, the ports of the seahorse XF-24 analyser cartridge were then loaded with the following compounds:

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

Before running the assay, the XF-24 instrument cartridge was calibrated. For total oxygen consumption rate (OCR) measurements, the minimum OCR reading after Rotenone / Antimycin A treatment was subtracted from the initial untreated level, following the manufacturer's protocol. To directly measure mitochondrial thermogenesis, uncoupled respiration (proton leak) was measured by subtracting the minimum OCR level after Rotenone / Antimycin from the minimum level after oligomycin treatment. Oxygen concentrations were measured over time periods of 4 min with 2 min waiting and 2 min mixing. Normalisation was performed based on total cell count assessed using Hoechst assay.

ABC predictions

We used the Activity-by-Contact (ABC) model (Fulco et al. 2019) to link noncoding variants in predicted enhancers to target genes. Briefly, the ABC model involves identifying chromatin accessible elements that are likely to regulate genes in a given cell type by multiplying an estimate of enhancer activity (estimated from ATAC-seq or DNase-seq and H3K27ac ChIP-seq signals) by an estimate of 3D enhancer-promoter contact frequency (estimated from Hi-C datasets). We used ABC enhancer-gene predictions for skeletal muscle from Nasser *et al.* Nature 2021, and further generated ABC predictions in differentiating adipocytes by collecting ATAC-seq and H3K27ac ChIP-seq datasets as described in (Sinnott-Armstrong et al. 2021). We then looked for ABC-predicted enhancers that overlap with noncoding variants.

MRI adipose volume associations

Cohort and Trait Definitions

The UK Biobank is an observational study that enrolled over 500,000 individuals between the ages of 40 and 69 years between 2006 and 2010, of whom 43,521 underwent MRI imaging between 2014 and 2020 (Agrawal, Klarqvist, et al., n.d.; Littlejohns et al. 2020; Sudlow et al.

2015) We previously quantified VAT, ASAT, and GFAT volumes in 40,032 individuals of the imaged cohort after imaging QC as described elsewhere (Agrawal, Klarqvist, et al., n.d.).

Body composition was also assessed using DEXA in 40,798 individuals in the UK Biobank. Android fat mass (field 23245) and gynoid fat mass (field 23262) were available for download in 5,170 individuals, and could be computed near perfectly in the remaining 35,628 individuals by multiplying either android total mass (field 23248) with android tissue fat percentage (field 23247) or gynoid total mass (field 23265) with gynoid tissue fat percentage (field 23264) (R^2 > 99.7% in the 5,170 individuals with explicit labels).

Study Design

We determined associations between variants rs12454712 and VATadjBMI, ASATadjBMI, GFATadjBMI, AndroidadjBMI, and GynoidadjBMI. Participants were excluded from analysis if they met any of the following criteria: (1) mismatch between self-reported sex and sex chromosome count, (2) sex chromosome aneuploidy, (3) genotyping call rate < 0.95, or (4) were outliers for heterozygosity. Up to 37,641 participants were available for association with MRI-derived VATadjBMI, ASATadjBMI, and GFATadjBMI, while up to 38,703 participants were available for association with DEXA-derived AndroidadjBMI and GynoidadjBMI.

Statistical analysis

MRI-derived fat depot volumes (VAT, ASAT, and GFAT) and DEXA-derived body composition masses (Android and Gynoid) were adjusted for BMI and height ("adjBMI" traits) by taking residuals of sex-specific linear regressions against age at the time of MRI, age squared, BMI, and height.[2] Each trait was scaled to mean 0 and variance 1 in sex-specific groups before being combined for analysis. Hence, effect sizes should be interpreted as relative to sex-specific standard deviations. Linear regressions between a given trait-variant pair were adjusted for age at the time of imaging, age squared, sex, the first 10 principal components of genetic ancestry, genotyping array, and – in the case of VATadjBMI, ASATadjBMI, and GFATadjBMI – MRI imaging center. Analyses were performed using R 3.6.0 (R Project for Statistical Computing).

Acknowledgements

Work in the Claussnitzer lab is supported by the FNIH AMP-T2D RFB8b, NIDDK UM1 556 DK126185, NIDDK DK102173. M.C. and S.B.R.J. are supported by a Next Generation Award at the Broad Institute of MIT and Harvard. H.H. and M.C. received funding within the Clinical

Cooperation Group “Nutrigenomics and Type 2 Diabetes” from the Helmholtz Center Munich and the German Center for Diabetes Research. H.H. was funded by the Else Kröner-Fresenius-Foundation, Bad Homburg, Germany. S.L. was supported by a Novo Nordisk Postdoctoral 558 Fellowship run in partnership with the University of Oxford. C.M.L is supported by the Li Ka Shing Foundation, NIHR Oxford Biomedical Research Centre, Oxford, NIH (1P50HD104224-01), Gates Foundation (INV-024200), and a Wellcome Trust Investigator Award (221782/Z/20/Z). M.U. is supported by 569 K23DK114551. S.B.R.J. was supported by NIDDK UM1 DK126185, and NIDDK U01 570 DK105554. This work was further supported by the Sarnoff Cardiovascular Research Foundation Fellowship (to S.A.), grants 1K08HG010155 and 1U01HG011719 (to A.V.K.) from the National Human Genome Research Institute, a Hassenfeld Scholar Award from Massachusetts General Hospital (to A.V.K.), a Merkin Institute Fellowship from the Broad Institute of MIT and Harvard (to A.V.K.), and a sponsored research agreement from IBM Research to the Broad Institute of MIT and Harvard (P.T.E., A.P., P.B., A.V.K.). M.K., S.A., P.B., and A.V.K. are listed as co-inventors on a patent application for the use of imaging data in assessing body fat distribution and associated cardiometabolic risk. J.M.M. is supported by American Diabetes Association Innovative and Clinical Translational Award 1-19-ICTS-068 and by NHGRI, grant FAIN# U01HG011723. The authors gratefully acknowledge the use of the Opera Phenix High-Content/High-Throughput imaging system at the Broad Institute, funded by the S10 Grant NIH OD-026839-01.

Disclosure Addendum

M.C. is a member of the scientific advisory board of Nestle and Waypoint Bio and holds equity in Waypoint Bio. S.A. has served as a scientific consultant to Third Rock Ventures. A.V.K. has served as a scientific advisor to Sanofi, Amgen, Maze Therapeutics, Navitor Pharmaceuticals, Sarepta Therapeutics, Novartis, Verve Therapeutics, Silence Therapeutics, Veritas International, Color Health, Third Rock Ventures, and Columbia University (NIH); received speaking fees from Illumina, MedGenome, Amgen, and the Novartis Institute for Biomedical Research; and received a sponsored research agreement from the Novartis Institute for Biomedical Research.

References

Agrawal, Saaket, Marcus D. R. Klarqvist, Nathaniel Diamant, Patrick T. Ellinor, Nehal N. Mehta, Anthony Philippakis, Kenney Ng, Puneet Batra, and Amit V. Khera. n.d. “Association of

Machine Learning-Derived Measures of Body Fat Distribution in >40,000 Individuals with Cardiometabolic Diseases." https://doi.org/10.1101/2021.05.07.21256854.
———. n.d. "Association of Machine Learning-Derived Measures of Body Fat Distribution in >40,000 Individuals with Cardiometabolic Diseases." https://doi.org/10.1101/2021.05.07.21256854.
Agrawal, Saaket, Minxian Wang, Marcus D. R. Klarqvist, Joseph Shin, Hesam Dashti, Nathaniel Diamant, Seung Hoan Choi, et al. n.d. "Inherited Basis of Visceral, Abdominal Subcutaneous and Gluteofemoral Fat Depots." https://doi.org/10.1101/2021.08.24.21262564.
Barquissau, V., D. Beuzelin, D. F. Pisani, G. E. Beranger, A. Mairal, A. Montagner, B. Roussel, et al. 2016. "White-to-Brite Conversion in Human Adipocytes Promotes Metabolic Reprogramming towards Fatty Acid Anabolic and Catabolic Pathways." *Molecular Metabolism* 5 (5): 352–65.
Benador, Ilan Y., Michaela Veliova, Kiana Mahdaviani, Anton Petcherski, Jakob D. Wikstrom, Essam A. Assali, Rebeca Acín-Pérez, et al. 2018. "Mitochondria Bound to Lipid Droplets Have Unique Bioenergetics, Composition, and Dynamics That Support Lipid Droplet Expansion." *Cell Metabolism* 27 (4): 869–85.e6.
Böhm, Anja, Michaela Keuper, Tobias Meile, Marty Zdichavsky, Andreas Fritsche, Hans-Ulrich Häring, Martin Hrabě de Angelis, Harald Staiger, and Andras Franko. 2020. "Increased Mitochondrial Respiration of Adipocytes from Metabolically Unhealthy Obese Compared to Healthy Obese Individuals." *Scientific Reports* 10 (1): 12407.
Boix, Carles A., Benjamin T. James, Yongjin P. Park, Wouter Meuleman, and Manolis Kellis. 2021. "Regulatory Genomic Circuitry of Human Disease Loci by Integrative Epigenomics." *Nature* 590 (7845): 300–307.
Boren, J., and K. M. Brindle. 2012. "Apoptosis-Induced Mitochondrial Dysfunction Causes Cytoplasmic Lipid Droplet Formation." *Cell Death and Differentiation* 19 (9): 1561–70.
Chang, Christopher C., Carson C. Chow, Laurent Cam Tellier, Shashaank Vattikuti, Shaun M. Purcell, and James J. Lee. 2015. "Second-Generation PLINK: Rising to the Challenge of Larger and Richer Datasets." *GigaScience* 4 (February): 7.
Claussnitzer, Melina, Simon N. Dankel, Kyoung-Han Kim, Gerald Quon, Wouter Meuleman, Christine Haugen, Viktoria Glunk, et al. 2015. "FTO Obesity Variant Circuitry and Adipocyte Browning in Humans." *The New England Journal of Medicine* 373 (10): 895–907.
Consortium, Gtex, and GTEx Consortium. 2017. "Genetic Effects on Gene Expression across Human Tissues." *Nature*. https://doi.org/10.1038/nature24277.
Csardi, Gabor, Tamas Nepusz, and Others. 2006. "The Igraph Software Package for Complex Network Research." *InterJournal, Complex Systems* 1695 (5): 1–9.
Dixon, Jesse R., Inkyung Jung, Siddarth Selvaraj, Yin Shen, Jessica E. Antosiewicz-Bourget, Ah Young Lee, Zhen Ye, et al. 2015. "Chromatin Architecture Reorganization during Stem Cell Differentiation." *Nature* 518 (7539): 331–36.
Dobin, Alexander, Carrie A. Davis, Felix Schlesinger, Jorg Drenkow, Chris Zaleski, Sonali Jha, Philippe Batut, Mark Chaisson, and Thomas R. Gingeras. 2013. "STAR: Ultrafast Universal RNA-Seq Aligner." *Bioinformatics* 29 (1): 15–21.
Dong, Shan-Shan, Dong-Li Zhu, Xiao-Rong Zhou, Yu Rong, Mengqi Zeng, Jia-Bin Chen, Feng Jiang, et al. n.d. "An Intronic Risk SNP rs12454712 for Central Obesity Acts as an Allele-Specific Enhancer to Regulate BCL2 Expression." https://doi.org/10.2337/figshare.14579283.v1.
Elliott, Lloyd T., Kevin Sharp, Fidel Alfaro-Almagro, Sinan Shi, Karla L. Miller, Gwenaëlle Douaud, Jonathan Marchini, and Stephen M. Smith. 2018. "Genome-Wide Association Studies of Brain Imaging Phenotypes in UK Biobank." *Nature* 562 (7726): 210–16.
ENCODE Project Consortium. 2004. "The ENCODE (ENCyclopedia Of DNA Elements) Project." *Science* 306 (5696): 636–40.

Fernández, Álvaro F., Salwa Sebti, Yongjie Wei, Zhongju Zou, Mingjun Shi, Kathryn L. McMillan, Congcong He, et al. 2018. “Disruption of the Beclin 1–BCL2 Autophagy Regulatory Complex Promotes Longevity in Mice.” *Nature* 558 (7708): 136–40.
Findley, Anthony S., Alan Monziani, Allison L. Richards, Katherine Rhodes, Michelle C. Ward, Cynthia A. Kalita, Adnan Alazizi, et al. 2021. “Functional Dynamic Genetic Effects on Gene Regulation Are Specific to Particular Cell Types and Environmental Conditions.” *eLife* 10 (May). https://doi.org/10.7554/eLife.67077.
Franzén, Oscar, Raili Ermel, Ariella Cohain, Nicholas K. Akers, Antonio Di Narzo, Husain A. Talukdar, Hassan Foroughi-Asl, et al. 2016. “Cardiometabolic Risk Loci Share Downstream Cis- and Trans-Gene Regulation across Tissues and Diseases.” *Science* 353 (6301): 827–30.
Fulco, Charles P., Joseph Nasser, Thouis R. Jones, Glen Munson, Drew T. Bergman, Vidya Subramanian, Sharon R. Grossman, et al. 2019. “Activity-by-Contact Model of Enhancer-Promoter Regulation from Thousands of CRISPR Perturbations.” *Nature Genetics* 51 (12): 1664–69.
Gu, Zuguang, Roland Eils, and Matthias Schlesner. 2016. “Complex Heatmaps Reveal Patterns and Correlations in Multidimensional Genomic Data.” *Bioinformatics* 32 (18): 2847–49.
He, Congcong, Michael C. Bassik, Viviana Moresi, Kai Sun, Yongjie Wei, Zhongju Zou, Zhenyi An, et al. 2012. “Exercise-Induced BCL2-Regulated Autophagy Is Required for Muscle Glucose Homeostasis.” *Nature* 481 (7382): 511–15.
Karlsson, Torgny, Mathias Rask-Andersen, Gang Pan, Julia Höglund, Claes Wadelius, Weronica E. Ek, and Åsa Johansson. 2019. “Contribution of Genetics to Visceral Adiposity and Its Relation to Cardiovascular and Metabolic Disease.” *Nature Medicine* 25 (9): 1390–95.
Konopleva, Marina, Daniel A. Pollyea, Jalaja Potluri, Brenda J. Chyla, Todd Busman, Evelyn McKeegan, Ahmed Salem, et al. 2014. “A Phase 2 Study of ABT-199 (GDC-0199) in Patients with Acute Myelogenous Leukemia (AML).” *Blood* 124 (21): 118–118.
Krueger, F., and Others. 2015. “A Wrapper Tool around Cutadapt and FastQC to Consistently Apply Quality and Adapter Trimming to FastQ Files.” *Cambridge, UK: Babraham Institute*.
Laber, Samantha, Sophie Strobel, Josep-Maria Mercader, Hesam Dashti, Alina Ainbinder, Julius Honecker, Garrett Garborcauskas, et al. n.d. “Discovering Cellular Programs of Intrinsic and Extrinsic Drivers of Metabolic Traits Using LipocyteProfiler.” https://doi.org/10.1101/2021.07.17.452050.
Leung, Nelson, Stephan D. Thomé, and Angela Dispenzieri. 2018. “Venetoclax Induced a Complete Response in a Patient with Immunoglobulin Light Chain Amyloidosis Plateaued on Cyclophosphamide, Bortezomib and Dexamethasone.” *Haematologica*. https://doi.org/10.3324/haematol.2017.183749.
Liao, Yang, Gordon K. Smyth, and Wei Shi. 2019. “The R Package Rsubread Is Easier, Faster, Cheaper and Better for Alignment and Quantification of RNA Sequencing Reads.” *Nucleic Acids Research* 47 (8): e47.
Littlejohns, Thomas J., Jo Holliday, Lorna M. Gibson, Steve Garratt, Niels Oesingmann, Fidel Alfaro-Almagro, Jimmy D. Bell, et al. 2020. “The UK Biobank Imaging Enhancement of 100,000 Participants: Rationale, Data Collection, Management and Future Directions.” *Nature Communications* 11 (1): 2624.
Love, Michael I., Wolfgang Huber, and Simon Anders. 2014. “Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2.” *Genome Biology* 15 (12): 550.
Ly, J. D., D. R. Grubb, and A. Lawen. 2003. “The Mitochondrial Membrane Potential (deltapsi(m)) in Apoptosis; an Update.” *Apoptosis: An International Journal on Programmed Cell Death* 8 (2): 115–28.
Mahajan, A., D. Taliun, M. Thurner, N. R. Robertson, J. M. Torres, N. W. Rayner, A. J. Payne, et al. 2018. “Fine-Mapping Type 2 Diabetes Loci to Single-Variant Resolution Using High-Density Imputation and Islet-Specific Epigenome Maps.” *Nature Genetics* 50 (11).

https://doi.org/10.1038/s41588-018-0241-6.
Miller, Colette, Jeong-yeh Yang, Emily England, Amelia Yin, Viral Gajjar, and Srujana Rayalam. 2015. “Isoproterenol Induces Beiging and Thermogenesis in Mature 3T3-L1 Adipocytes.” *The FASEB Journal*. https://doi.org/10.1096/fasebj.29.1_supplement.lb274.
Nagel, Stella A., Michaela Keuper, Ivana Zagotta, Eveliina Enlund, Azahara Iris Ruperez, Klaus-Michael Debatin, Martin Wabitsch, and Pamela Fischer-Posovszky. 2014. “Up-Regulation of Bcl-2 during Adipogenesis Mediates Apoptosis Resistance in Human Adipocytes.” *Molecular and Cellular Endocrinology* 382 (1): 368–76.
Oliver, Lisa, Erika Hue, Julien Rossignol, Gwenola Bougras, Philippe Hulin, Philippe Naveilhan, Dominique Heymann, Laurent Lescaudron, and François M. Vallette. 2011. “Distinct Roles of Bcl-2 and Bcl-Xl in the Apoptosis of Human Bone Marrow Mesenchymal Stem Cells during Differentiation.” *PloS One* 6 (5): e19820.
Perdikari, Aliki, Germán Gastón Leparc, Miroslav Balaz, Nuno D. Pires, Martin E. Lidell, Wenfei Sun, Francesc Fernandez-Albert, et al. 2018. “BATLAS: Deconvoluting Brown Adipose Tissue.” *Cell Reports* 25 (3): 784–97.e4.
Pulit, Sara L., Charli Stoneman, Andrew P. Morris, Andrew R. Wood, Craig A. Glastonbury, Jessica Tyrrell, Loïc Yengo, et al. 2019. “Meta-Analysis of Genome-Wide Association Studies for Body Fat Distribution in 694 649 Individuals of European Ancestry.” *Human Molecular Genetics* 28 (1): 166–74.
Purcell, Shaun, Benjamin Neale, Kathe Todd-Brown, Lori Thomas, Manuel A. R. Ferreira, David Bender, Julian Maller, et al. 2007. “PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses.” *American Journal of Human Genetics* 81 (3): 559–75.
Qvalue. n.d. Github. Accessed April 17, 2021. https://github.com/StoreyLab/qvalue.
Raajendiran, Arthe, Geraldine Ooi, Jackie Bayliss, Paul E. O’Brien, Ralf B. Schittenhelm, Ashlee K. Clark, Renea A. Taylor, Matthew S. Rodeheffer, Paul R. Burton, and Matthew J. Watt. 2019. “Identification of Metabolically Distinct Adipocyte Progenitor Cells in Human Adipose Tissues.” *Cell Reports* 27 (5): 1528–40.e7.
Roberts, Andrew W., Matthew S. Davids, John M. Pagel, Brad S. Kahl, Soham D. Puvvada, John F. Gerecitano, Thomas J. Kipps, et al. 2016. “Targeting BCL2 with Venetoclax in Relapsed Chronic Lymphocytic Leukemia.” *The New England Journal of Medicine* 374 (4): 311–22.
Schmidt, Søren Fisker, Bjørk Ditlev Larsen, Anne Loft, Ronni Nielsen, Jesper Grud Skat Madsen, and Susanne Mandrup. 2015. “Acute TNF-Induced Repression of Cell Identity Genes Is Mediated by NFκB-Directed Redistribution of Cofactors from Super-Enhancers.” *Genome Research* 25 (9): 1281–94.
Shungin, Dmitry, Thomas W. Winkler, Damien C. Croteau-Chonka, Teresa Ferreira, Adam E. Locke, Reedik Mägi, Rona J. Strawbridge, et al. 2015. “New Genetic Loci Link Adipose and Insulin Biology to Body Fat Distribution.” *Nature* 518 (7538): 187–96.
Sinnott-Armstrong, Nasa, Isabel S. Sousa, Samantha Laber, Elizabeth Rendina-Ruedy, Simon E. Nitter Dankel, Teresa Ferreira, Gunnar Mellgren, et al. 2021. “A Regulatory Variant at 3q21.1 Confers an Increased Pleiotropic Risk for Hyperglycemia and Altered Bone Mineral Density.” *Cell Metabolism* 33 (3): 615–28.e13.
Skurk, Thomas, and Hans Hauner. 2012. “Primary Culture of Human Adipocyte Precursor Cells: Expansion and Differentiation.” *Methods in Molecular Biology* 806: 215–26.
Sobreira, Débora R., Amelia C. Joslin, Qi Zhang, Iain Williamson, Grace T. Hansen, Kathryn M. Farris, Noboru J. Sakabe, et al. 2021. “Extensive Pleiotropism and Allelic Heterogeneity Mediate Metabolic Effects of and.” *Science* 372 (6546): 1085–91.
Sudlow, Cathie, John Gallacher, Naomi Allen, Valerie Beral, Paul Burton, John Danesh, Paul Downey, et al. 2015. “UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age.” *PLoS Medicine* 12 (3): e1001779.

Taliun, Sarah A. Gagliano, Peter VandeHaar, Andrew P. Boughton, Ryan P. Welch, Daniel Taliun, Ellen M. Schmidt, Wei Zhou, et al. 2020. “Exploring and Visualizing Large-Scale Genetic Associations by Using PheWeb.” *Nature Genetics* 52 (6): 550–52.

Torres, Jason M., Moustafa Abdalla, Anthony Payne, Juan Fernandez-Tajes, Matthias Thurner, Vibe Nylander, Anna L. Gloyn, Anubha Mahajan, and Mark I. McCarthy. 2020. “A Multi-Omic Integrative Scheme Characterizes Tissues of Action at Loci Associated with Type 2 Diabetes.” *American Journal of Human Genetics* 107 (6): 1011–28.

Udler, Miriam S., Jaegil Kim, Marcin von Grotthuss, Sílvia Bonàs-Guarch, Joanne B. Cole, Joshua Chiou, Christopher D. Anderson on behalf of METASTROKE and the ISGC, et al. 2018. “Type 2 Diabetes Genetic Loci Informed by Multi-Trait Associations Point to Disease Mechanisms and Subtypes: A Soft Clustering Analysis.” *PLoS Medicine* 15 (9): e1002654.

Vos, S. de, L. J. Swinnen, D. Wang, E. Reid, N. Fowler, J. Cordero, M. Dunbar, et al. 2018. “Venetoclax, Bendamustine, and Rituximab in Patients with Relapsed or Refractory NHL: A Phase Ib Dose-Finding Study.” *Annals of Oncology: Official Journal of the European Society for Medical Oncology / ESMO* 29 (9): 1932–38.

Walford, Geoffrey A., Stefan Gustafsson, Denis Rybin, Alena Stančáková, Han Chen, Ching-Ti Liu, Jaeyoung Hong, et al. 2016. “Genome-Wide Association Study of the Modified Stumvoll Insulin Sensitivity Index Identifies BCL2 and FAM19A2 as Novel Insulin Sensitivity Loci.” *Diabetes* 65 (10): 3200–3211.

Wang, Yanli, Fan Song, Bo Zhang, Lijun Zhang, Jie Xu, Da Kuang, Daofeng Li, et al. 2018. “The 3D Genome Browser: A Web-Based Browser for Visualizing 3D Genome Organization and Long-Range Chromatin Interactions.” *Genome Biology* 19 (1): 151.

Watanabe, Kyoko, Sven Stringer, Oleksandr Frei, Maša Umićević Mirkov, Christiaan de Leeuw, Tinca J. C. Polderman, Sophie van der Sluis, Ole A. Andreassen, Benjamin M. Neale, and Danielle Posthuma. 2019. “A Global Overview of Pleiotropy and Genetic Architecture in Complex Traits.” *Nature Genetics* 51 (9): 1339–48.

Yaghootkar, Hanieh, Robert A. Scott, Charles C. White, Weihua Zhang, Elizabeth Speliotes, Patricia B. Munroe, Georg B. Ehret, et al. 2014. “Genetic Evidence for a Normal-Weight ‘Metabolically Obese’ Phenotype Linking Insulin Resistance, Hypertension, Coronary Artery Disease, and Type 2 Diabetes.” *Diabetes* 63 (12): 4369–77.

Zhou, Xin, Daofeng Li, Bo Zhang, Rebecca F. Lowdon, Nicole B. Rockweiler, Renee L. Sears, Pamela A. F. Madden, Ivan Smirnov, Joseph F. Costello, and Ting Wang. 2015. “Epigenomic Annotation of Genetic Variants Using the Roadmap Epigenome Browser.” *Nature Biotechnology* 33 (4): 345–46.

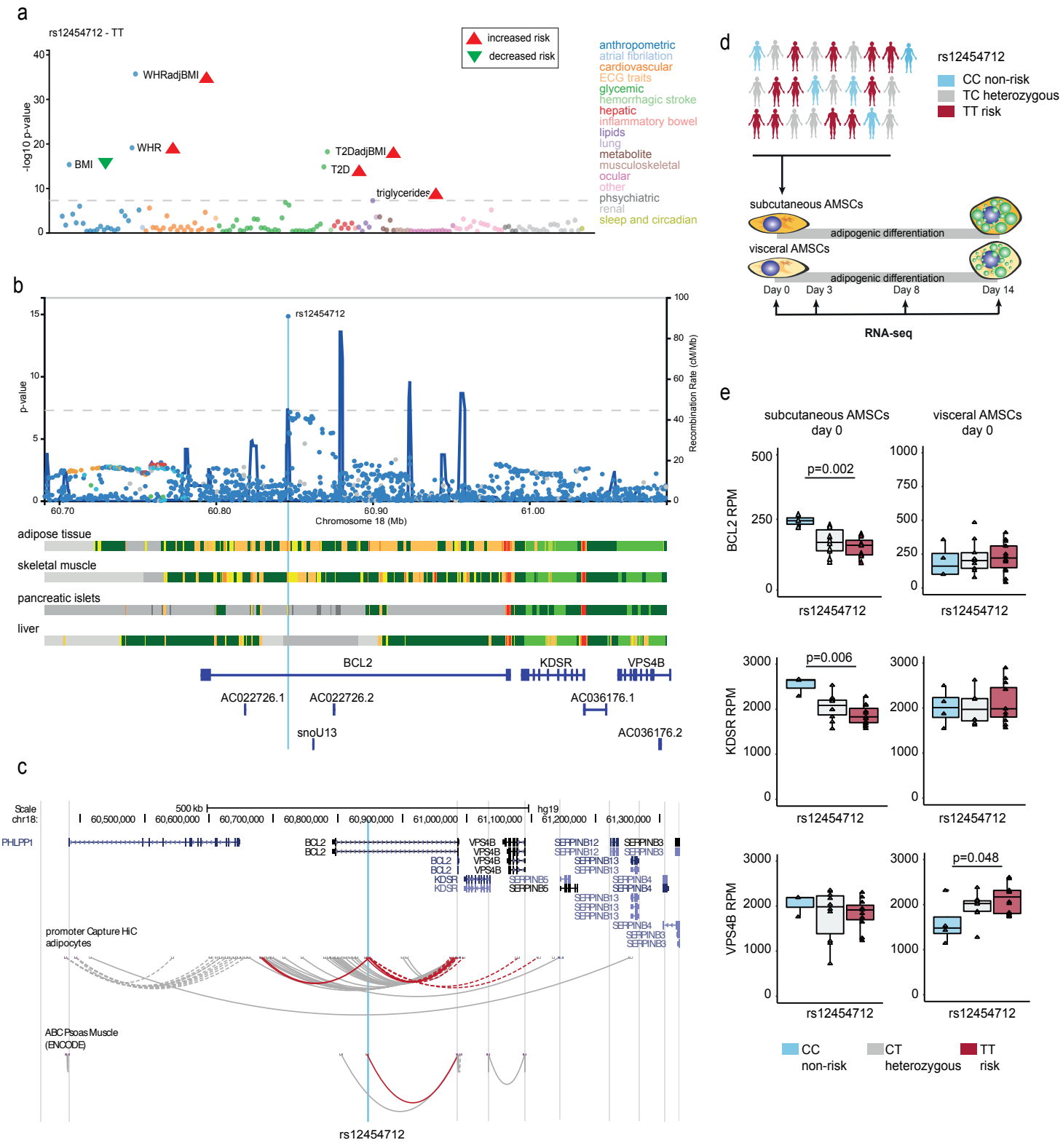

a
rs12454712 - TT
increased risk
decreased risk
-log10 p-value
WHRadjBMI
WHR
BMI
T2DadjBMI
T2D
triglycerides
anthropometric
atrial fibrilation
cardiovascular
ECG traits
glycemic
hemorrhagic stroke
hepatic
inflammatory bowel
lipids
lung
metabolite
musculoskeletal
ocular
other
physchiatric
renal
sleep and circadian
b
rs12454712
p-value
Recombination Rate (cM/Mb)
Chromosome 18 (Mb)
adipose tissue
skeletal muscle
pancreatic islets
liver
BCL2
KDSR
VPS4B
AC022726.1
AC022726.2
AC036176.1
snoU13
AC036176.2
c
Scale
chr18:
500 kb
hg19
PHLPP1
promoter Capture HiC
adipocytes
ABC Psoas Muscle
(ENCODE)
rs12454712
d
rs12454712
CC non-risk
TC heterozygous
TT risk
subcutaneous AMSCs
adipogenic differentiation
visceral AMSCs
adipogenic differentiation
Day 0
Day 3
Day 8
Day 14
RNA-seq
e
subcutaneous AMSCs
day 0
visceral AMSCs
day 0
BCL2 RPM
p=0.002
KDSR RPM
p=0.006
VPS4B RPM
p=0.048
rs12454712
CC
non-risk
CT
heterozygous
TT
risk

Figure 1

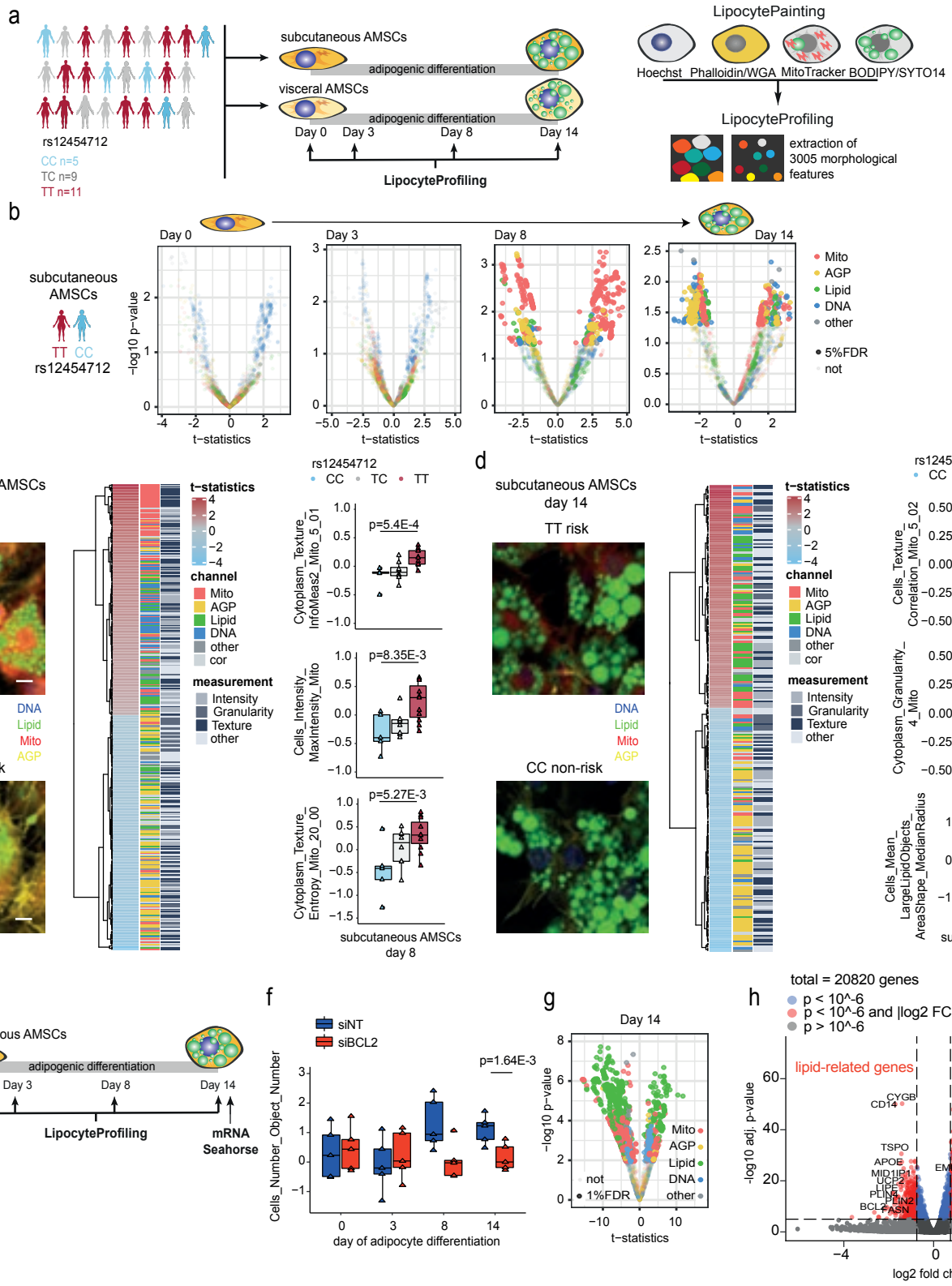

a
rs12454712
CC n=5
TC n=9
TT n=11
subcutaneous AMSCs
adipogenic differentiation
visceral AMSCs
adipogenic differentiation
Day 0
Day 3
Day 8
Day 14
LipocyteProfiling
LipocytePainting
Hoechst
Phalloidin/WGA
MitoTracker
BODIPY/SYTO14
LipocyteProfiling
extraction of 3005 morphological features
b
subcutaneous AMSCs
TT CC
rs12454712
Day 0
Day 3
Day 8
Day 14
−log10 p−value
t−statistics
Mito
AGP
Lipid
DNA
other
5%FDR
not
c
subcutaneous AMSCs day 8
TT risk
DNA
Lipid
Mito
AGP
CC non-risk
t−statistics
channel
Mito
AGP
Lipid
DNA
other
cor
measurement
Intensity
Granularity
Texture
other
rs12454712
CC TC TT
Cytoplasm_Texture_InfoMeas2_Mito_5_01
p=5.4E-4
Cells_Intensity_MaxIntensity_Mito
p=8.35E-3
Cytoplasm_Texture_Entropy_Mito_20_00
p=5.27E-3
subcutaneous AMSCs day 8
d
subcutaneous AMSCs day 14
TT risk
DNA
Lipid
Mito
AGP
CC non-risk
t−statistics
channel
Mito
AGP
Lipid
DNA
other
cor
measurement
Intensity
Granularity
Texture
other
rs12454712
CC TC TT
Cells_Texture_Correlation_Mito_5_02
p=2.68E-2
Cytoplasm_Granularity_4_Mito
p=4.73E-2
Cells_Mean_LargeLipidObjects_AreaShape_MedianRadius
p=2.47E-2
subcutaneous AMSCs day 14
e
subcutaneous AMSCs
adipogenic differentiation
Day -3
Day 0
Day 3
Day 8
Day 14
siNT
siBCL2
LipocyteProfiling
mRNA Seahorse
f
siNT
siBCL2
p=1.64E-3
Cells_Number_Object_Number
day of adipocyte differentiation
g
Day 14
−log10 p−value
t−statistics
Mito
AGP
Lipid
DNA
other
not
1%FDR
h
total = 20820 genes
p < 10^-6
p < 10^-6 and |log2 FC| > 0.75
p > 10^-6
pro-apoptotic genes
lipid-related genes
PDGFD
CYGB
CD14
CALU
TAGLN
COL11A1
WEE1
LOX
TMEM176B
CTSC
TSPO
APOE
MID1IP1
UCP2
IRF7
PDK4
PLIN2
FASN
BCL2
EMP1
DCN
TNFSF10
TMEM176A
-log10 adj. p-value
log2 fold change

Figure 2

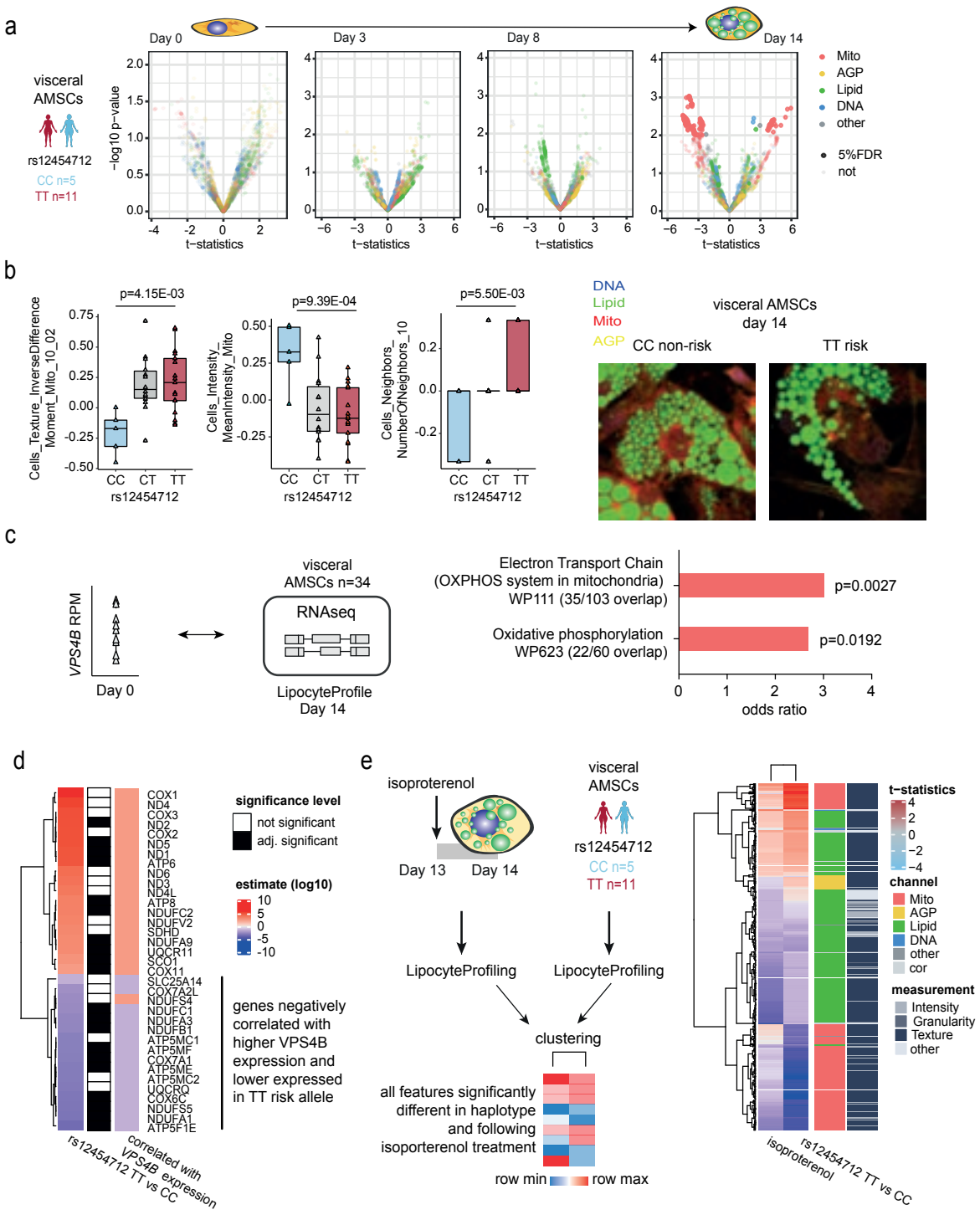

a
Day 0
Day 3
Day 8
Day 14
visceral AMSCs
rs12454712
CC n=5
TT n=11
−log10 p−value
t−statistics
Mito
AGP
Lipid
DNA
other
5%FDR
not
b
p=4.15E-03
p=9.39E-04
p=5.50E-03
Cells_Texture_InverseDifference Moment_Mito_10_02
Cells_Intensity_ MeanIntensity_Mito
Cells_Neighbors_ NumberOfNeighbors_10
CC CT TT
rs12454712
DNA
Lipid
Mito
AGP
visceral AMSCs day 14
CC non-risk
TT risk
c
VPS4B RPM
Day 0
visceral AMSCs n=34
RNAseq
LipocyteProfile Day 14
Electron Transport Chain (OXPHOS system in mitochondria) WP111 (35/103 overlap)
p=0.0027
Oxidative phosphorylation WP623 (22/60 overlap)
p=0.0192
odds ratio
d
COX1
ND4
COX3
ND2
COX2
ND5
ND1
ATP6
ND6
ND3
ND4L
ATP8
NDUFC2
NDUFV2
SDHD
NDUFA9
UQCR11
SCO1
COX11
SLC25A14
COX7A2L
NDUFS4
NDUFC1
NDUFA3
NDUFB1
ATP5MC1
ATP5MF
COX7A1
ATP5ME
ATP5MC2
UQCRQ
COX6C
NDUFS5
NDUFA1
ATP5F1E
significance level
not significant
adj. significant
estimate (log10)
genes negatively correlated with higher VPS4B expression and lower expressed in TT risk allele
correlated with VPS4B expression
rs12454712 TT vs CC
e
isoproterenol
Day 13
Day 14
visceral AMSCs
rs12454712
CC n=5
TT n=11
LipocyteProfiling
LipocyteProfiling
clustering
all features significantly different in haplotype and following isoporterenol treatment
row min
row max
t−statistics
channel
Mito
AGP
Lipid
DNA
other
cor
measurement
Intensity
Granularity
Texture
other
isoproterenol
rs12454712 TT vs CC
f
FFA
Day 0
Day 14
RNAseq
visceral AMSCs day 14
p<0.0001
% of brown adipocytes
FFA (-)
FFA (+)
Estimation Plot - CC
Estimation Plot - TT
Mean of differences
FFA (+) - FFA (-)
p=0.030
% increase in brown adipocytes following FFA treatment
CC
TT
FFA (+) - FFA (-)

Figure 3

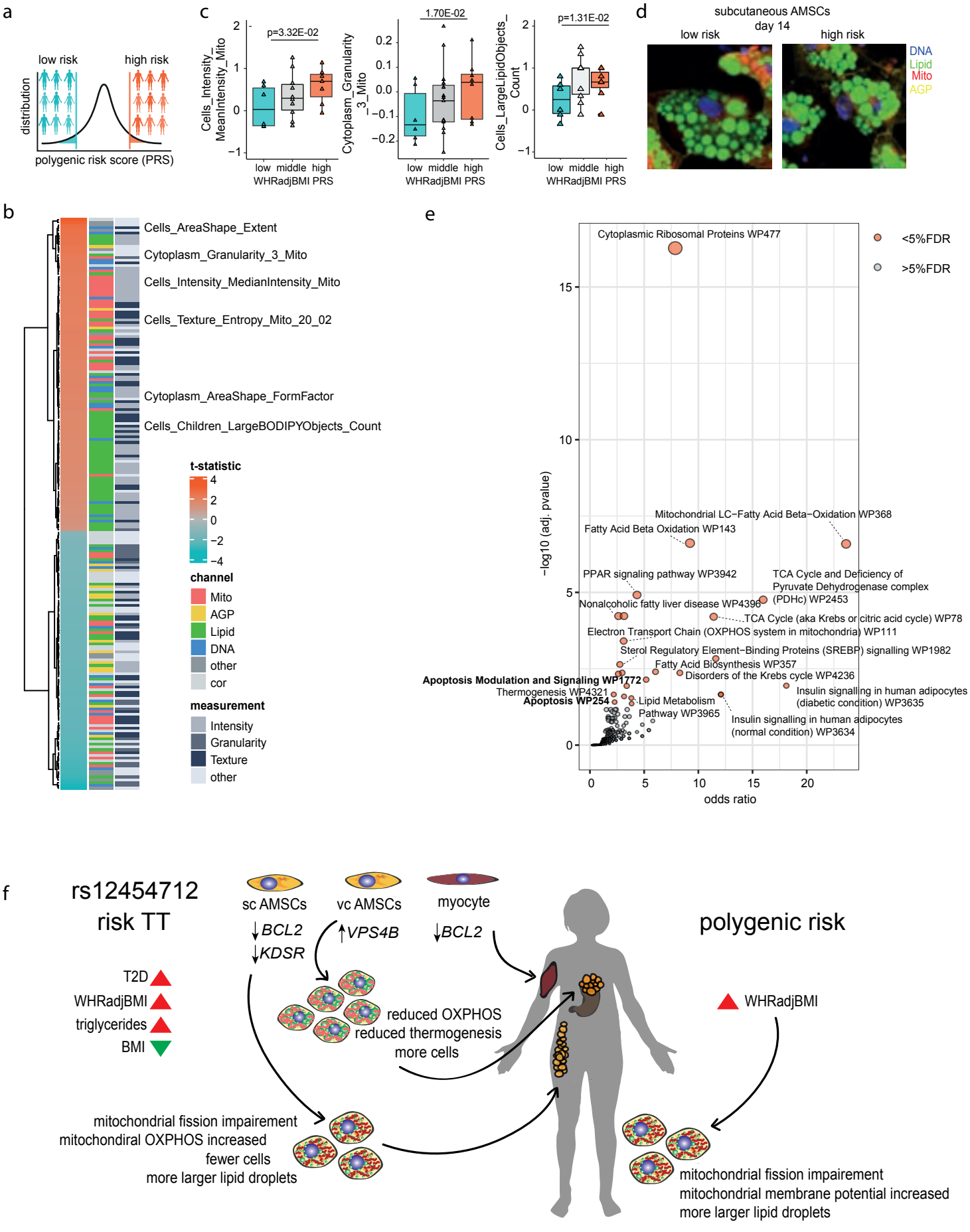
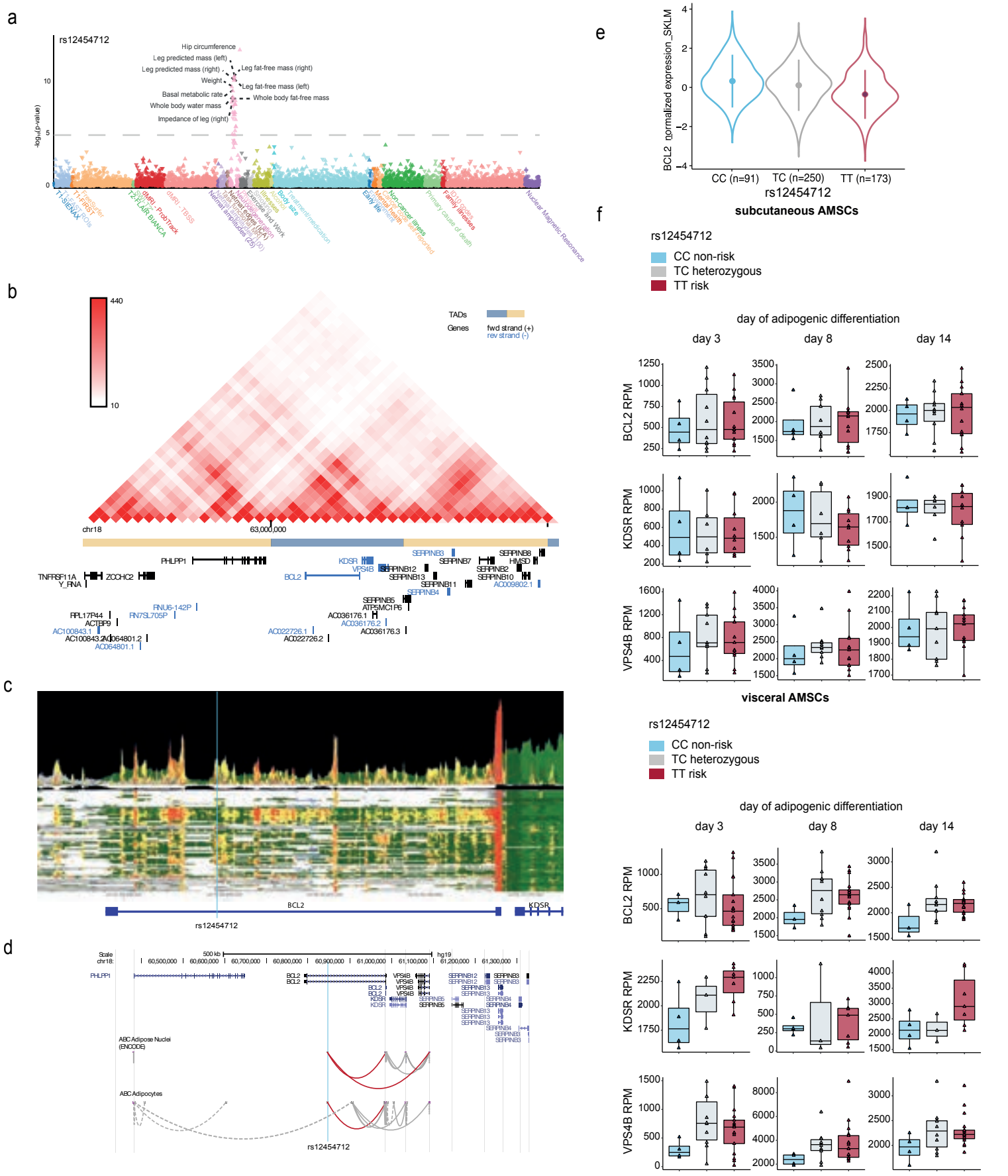
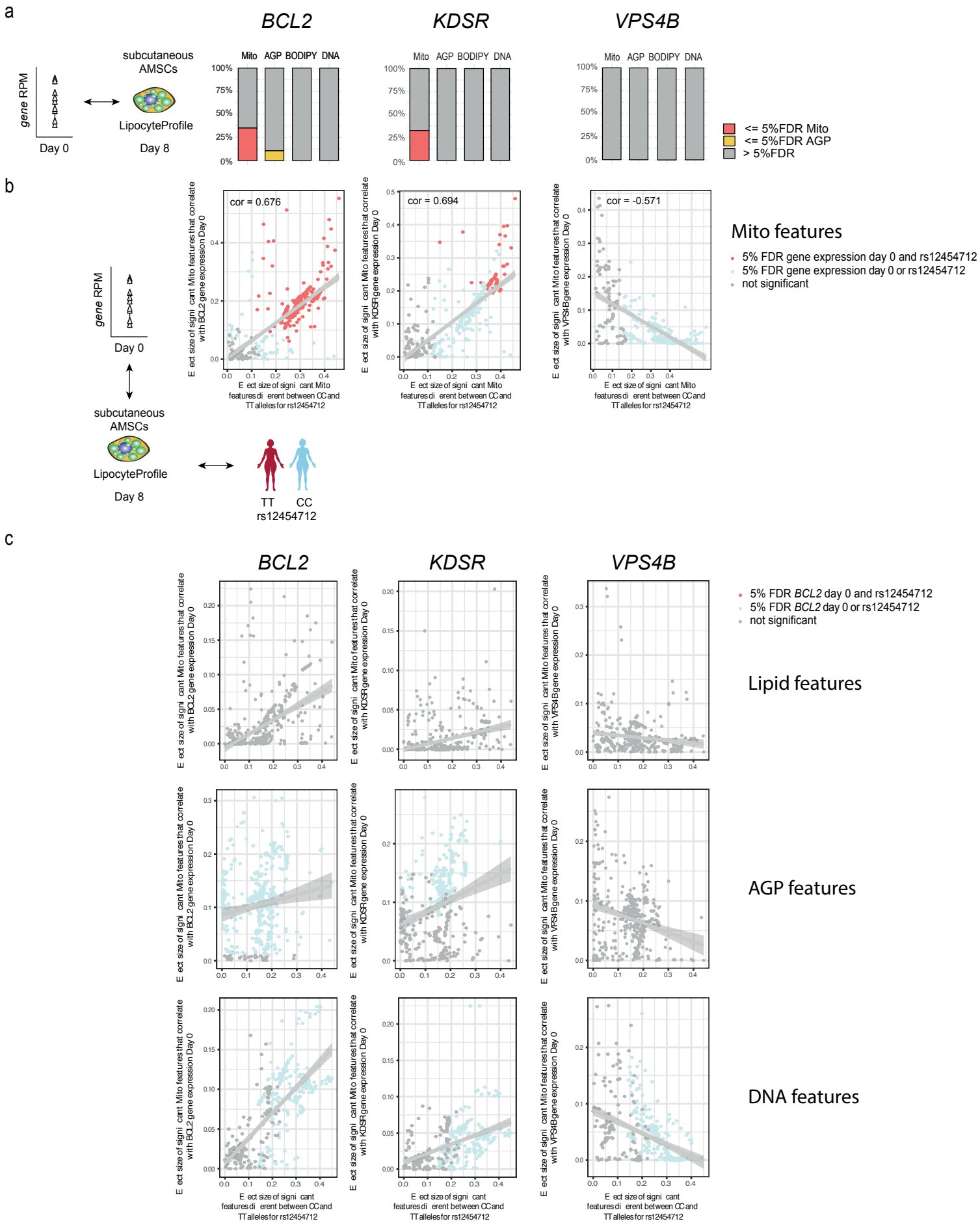

a
low risk
high risk
distribution
polygenic risk score (PRS)
c
p=3.32E-02
1.70E-02
p=1.31E-02
Cells_Intensity_MeanIntensity_Mito
Cytoplasm_Granularity_3_Mito
Cells_LargeLipidObjects_Count
low middle high
WHRadjBMI PRS
d
subcutaneous AMSCs
day 14
low risk
high risk
DNA
Lipid
Mito
AGP
b
Cells_AreaShape_Extent
Cytoplasm_Granularity_3_Mito
Cells_Intensity_MedianIntensity_Mito
Cells_Texture_Entropy_Mito_20_02
Cytoplasm_AreaShape_FormFactor
Cells_Children_LargeBODIPYObjects_Count
t-statistic
channel
Mito
AGP
Lipid
DNA
other
cor
measurement
Intensity
Granularity
Texture
other
e
<5%FDR
>5%FDR
Cytoplasmic Ribosomal Proteins WP477
Mitochondrial LC−Fatty Acid Beta−Oxidation WP368
Fatty Acid Beta Oxidation WP143
PPAR signaling pathway WP3942
TCA Cycle and Deficiency of Pyruvate Dehydrogenase complex (PDHc) WP2453
Nonalcoholic fatty liver disease WP4396
TCA Cycle (aka Krebs or citric acid cycle) WP78
Electron Transport Chain (OXPHOS system in mitochondria) WP111
Sterol Regulatory Element−Binding Proteins (SREBP) signalling WP1982
Fatty Acid Biosynthesis WP357
Apoptosis Modulation and Signaling WP1772
Disorders of the Krebs cycle WP4236
Thermogenesis WP4321
Insulin signalling in human adipocytes (diabetic condition) WP3635
Apoptosis WP254
Lipid Metabolism Pathway WP3965
Insulin signalling in human adipocytes (normal condition) WP3634
−log10 (adj. pvalue)
odds ratio
f
rs12454712
risk TT
T2D
WHRadjBMI
triglycerides
BMI
sc AMSCs
vc AMSCs
myocyte
↓BCL2
↓KDSR
↑VPS4B
↓BCL2
reduced OXPHOS
reduced thermogenesis
more cells
polygenic risk
WHRadjBMI
mitochondrial fission impairement
mitochondiral OXPHOS increased
fewer cells
more larger lipid droplets
mitochondrial fission impairement
mitochondrial membrane potential increased
more larger lipid droplets

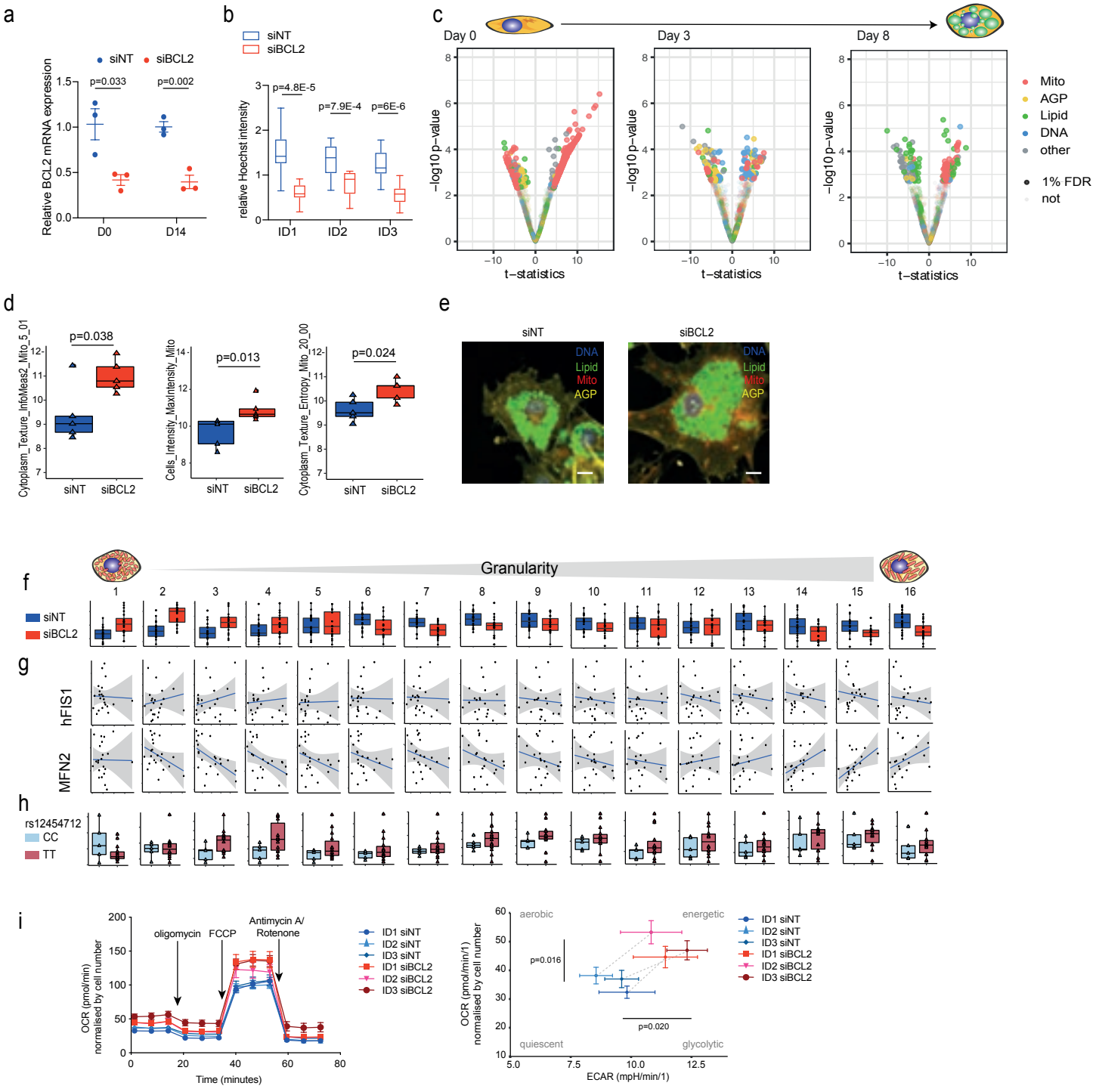
Figure 4



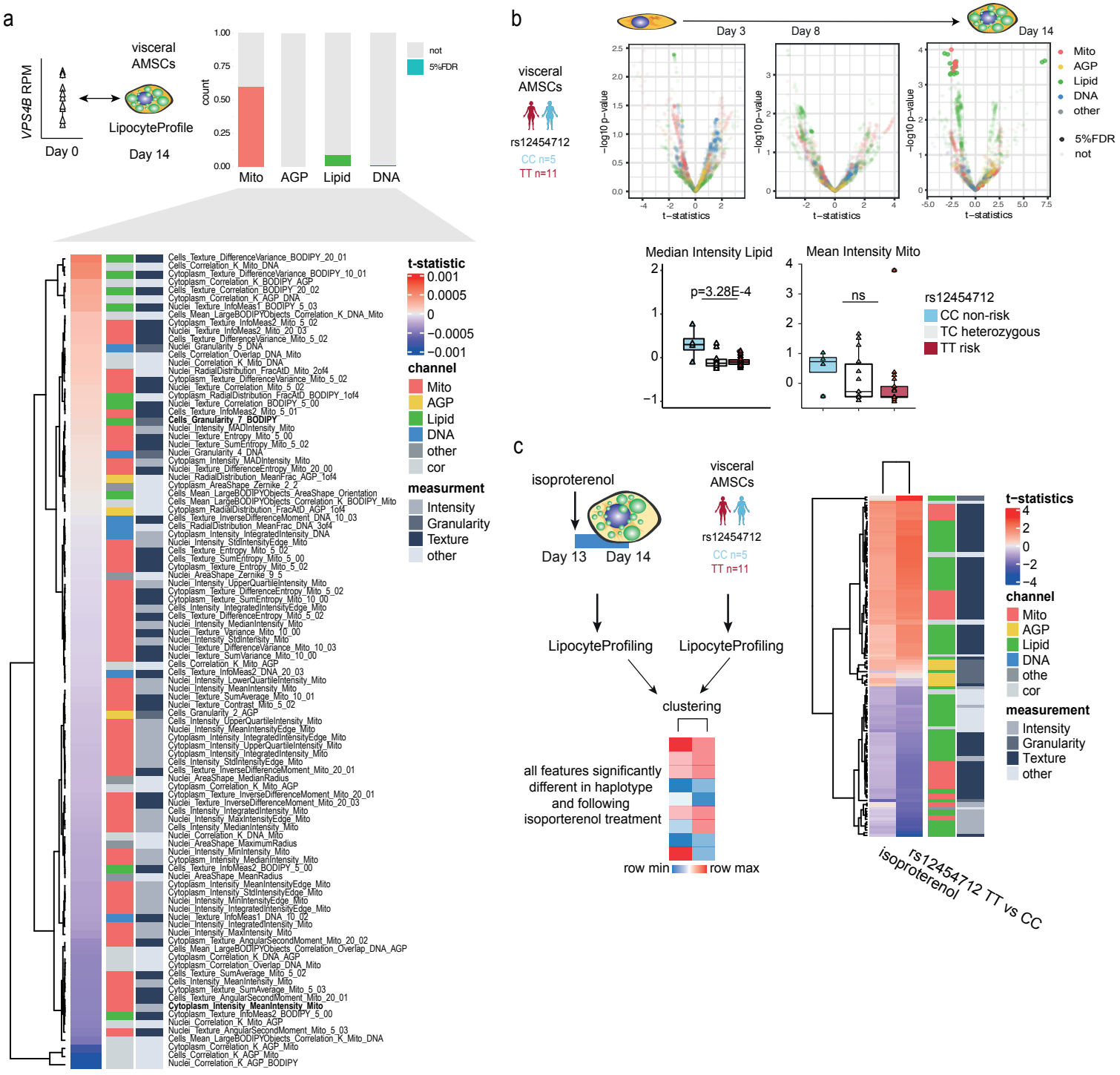
Supplementary Figure 1



Supplementary Figure 2



Supplementary Figure 3



Supplementary Figure 4