A Small Fraction of Progenitors Differentiate Into Mature Adipocytes by Escaping the Constraints on the Cell Structure

Mahmoud Ahmed, Trang Huyen Lai and Deok Ryong Kim*

Department of Biochemistry and Convergence Medical Sciences, Institute of Health Sciences, Gyeongsang National University School of Medicine, Jinju, South Korea

Differentiating 3T3-L1 pre-adipocytes are a mixture of non-identical culture cells. It is vital to identify the cell types that respond to the induction stimulus to understand the pre-adipocyte potential and the mature adipocyte behavior. To test this hypothesis, we deconvoluted the gene expression profiles of the cell culture of MDI-induced 3T3-L1 cells. Then we estimated the fractions of the sub-populations and their changes in time. We characterized the sub-populations based on their specific expression profiles. Initial cell cultures comprised three distinct phenotypes. A small fraction of the starting cells responded to the induction and developed into mature adipocytes. Unresponsive cells were probably under structural constraints or were committed to differentiating into alternative phenotypes. Using the same population gene markers, similar proportions were found in induced human primary adipocyte cell cultures. The three sub-populations had diverse responses to treatment with various drugs and compounds. Only the response of the maturing sub-population resembled that estimated from the profiles of the mixture. We then showed that even at a low division rate, a small fraction of cells could increase its share in a dynamic two-populations model. Finally, we used a cell cycle expression index to validate that model. To sum, pre-adipocytes are a mixture of different cells of which a limited fraction become mature adipocytes.

Keywords: deconvolution, differentiation, adipocytes, 3T3-L1, subpopulation analysis

1. INTRODUCTION

Cell cultures contain populations of non-identical cells. Early on, pre-adipocytes cell cultures were recognized to be heterogeneous. Green *et al.* described the heterogeneous nature of differentiating 3T3-L1 cell (Green and Kehinde, 1975). Lee et al. (2019) showed that adipocyte progenitors within a single fat depot give rise to functionally distinct adipocytes. Cell cultures never attain a 100% differentiation efficiency; thus, only a fraction of the pre-adipocytes are responsive. A key question is how large or small this fraction is. Loo *et al.* showed that a minority of pre-adipocytes gives rise to mature adipocytes with characteristic gene expression and lipid droplets (Loo et al., 2009).

The varying cellular response to the induction media could be because of their intrinsic properties or is a stochastic process. Stochasticity and cell-cell interactions were suggested as key contributors to the differentiation of stem cells (Smith et al., 2015). If the pre-adipocytes are intrinsically different, we should be able, in principle, to find mechanisms that confer

responsiveness/resistance to the sub-populations. A study proposed that the adipocyte's ability to mature and accumulate fat is heritable. Modifying the key adipogenic transcription factor peroxisome proliferator-activated receptor gamma (*Pparg*) promoter region marks the cell capacity to differentiate, and re-differentiate (Katz et al., 2014).

Differentiation is a transcriptional and morphological development event. Cells proliferate only for 1 day after the 3T3-L1 induction in what is known as post-confluent mitosis followed by growth arrest (Bernlohr et al., 1985). Cell division and DNA replication are necessary to allow transcription factors into the open chromatin. This, in turn, drives the development of the new phenotype. Within days, the cell accumulates lipids in droplets and remodels its interior into a distinct fat cell.

Changes in the cytoskeleton and extracellular matrix rarely feature during the study of adipocyte differentiation. This is unfortunate given their prime importance to the morphological development of the adipocytes. It is clear from research done in stem cells with differentiation potentials; they have to respond to extracellular signals, not only by modulating the gene expression but their internal structure (Yourek et al., 2007; Guilak et al., 2009). Yang et al. (2013) showed that remodeling the cytoskeleton through the acetylation of alpha-tubulin is necessary for the differentiation of pre-adipocytes. Moreover, it was even suggested that actin cytoskeleton might interfere with the transcriptional regulation of the differentiation process by controlling the translocation of antagonistic nuclear factors (Nobusue et al., 2014).

Profiling by RNA-Seq is an efficient way to study gene expression changes between two or more conditions. However, the expression signal, in this case, is averaged over a mixed population of various cell types. Single-cell RNA-Seq (scRNA-Seq) is the most straightforward way to test the gene expression at the level of the individual cells. Because these datasets are not always available, deconvolution of bulk RNA-Seq data can be an alternative. In this study, the sub-populations in the mixture were estimated from their gene signatures and studied separately.

Convex analysis of mixtures (CAM) is an unsupervised method to deconvolute the profiles of samples from mixed phenotypes into their constituents (Wang et al., 2016). The creators of the method applied it to a dataset of gene expression in yeast at different cell cycle stages. CAM could detect a composite of sub-populations consistent with the true labels. In addition, the sub-populations gene markers identified by CAM were related to the corresponding phases. Herrington et al. (2018) used this method to derive distinct phenotypes in arterial tissues. Classification of the cell types based on their markers was consistent with tumor necrotic factor (TNF) alpha pathway activation.

Here we describe deconvoluting the gene expression of differentiating 3T3-L1 pre-adipocytes from bulk RNA-Seq data. The goal is to determine the distinct phenotypes in the cell cultures and estimate their fractions. Then we showed which of these fractions responded to the induction stimuli and became mature adipocytes. Finally, we attempted to explain the differences between the sub-populations as to which functions their specific markers involve in, which type of cells they give rise

to at the end of the differentiation course, and their response to treatment with various drugs and compounds.

2. MATERIALS AND METHODS

2.1. Cell Culture and Differentiation

3T3-L1 is a mouse pre-adipocyte cell line that can be induced to differentiate into mature adipocytes when treated with a chemical cocktail. A chemical cocktail of 1-Methyl-3-isobutylxanthine, Dexamethasone and Insulin (MDI) is known to induce the differentiation (Green and Kehinde, 1975). The treatment starts with a fully confluent 3T3-L1 cell culture (pre-adipocyte), accumulating lipid into lipid droplets within days. The time course comprises three stages; day 2–4, early; up to day 7 intermediate; and up to 14 days is the late-differentiation stage. A similar protocol produces a comparable differentiation course in human primary pre-adipocytes.

2.2. Data Curation and Processing

2.2.1. RNA-Seq Data

We surveyed GEO and SRA repositories for high-throughput sequencing data of MDI-induced 3T3-L1 and human primary pre-adipocyte samples at different time points (**Table 1**). Ninety-eight RNA-Seq samples were included in the study. This dataset was previously curated, and the processed data was packaged in a Bioconductor experimental data package curatedAdipoRNA (Ahmed and Kim, 2019). Briefly, raw reads were downloaded from the SRA ftp server using Wget. FASTQC was used to assess the quality of the raw reads (Andrews, 2020). For RNA-Seq, raw reads were aligned to UCSC mm10 mouse genome using HISAT2 (Kim et al., 2015). FeatureCounts was used to count the aligned reads (bam) in known genes (Liao et al., 2014).

2.2.2. Microarray Data

We conducted a similar survey of the literature for studies that subject the same cell line model for chemical or genetic perturbations (**Table 2**). Forty-two microarray of mature adipocytes (> 7 days post MDI-induction) treated with thiazolidinediones (TZD), four short-chain fatty acids (SFA), or three nonsteroidal anti-inflammatory drugs (NSAID) were previously collected and curated. The processed data were packaged in another experimental data package curatedAdipoArray (Ahmed et al., 2020). In addition, 24 microarray samples of MDI-induced human primary pre-adipocytes were downloaded from GEO in the form of processed probe intensities (GSE98680) (Verbanck et al., 2017).

Microarray data were downloaded from GEO in the form of probe intensities. Probe metadata (gene symbols) were used to collapse the intensities into gene level using the maximum average (Horvath and Dong, 2008). Data were normalized using normalize between arrays, and batch effects were removed using empirical Bayesian adjustments (Ritchie et al., 2015; Leek et al., 2020).

2.3. Identifying Adipocyte Populations

Convex Analysis of Mixtures (CAM) was used to deconvolute gene counts expression profiles (Wang et al., 2016). This analysis

TABLE 1 | RNA-Seq datasets of differentiating 3T3-L1 cells.

SRA ID	Time (h)	Ref.
SRP010905	192	Lo et al., 2013
SRP029602	−48–144	Al Adhami et al., 2015
SRP029985	0–168	Duteil et al., 2014
SRP033778	−48–240	Zhao et al., 2014
SRP041765	0–4	Siersbæk et al., 2014
SRP045772	0–48	Lim et al., 2015
SRP051775	168	Buchner et al., 2015
SRP066991	−96–168	Brunmeir et al., 2016
SRP078506	0–48	Brier et al., 2017
SRP090166	0–168	Park et al., 2017
SRP092793	240	Chaudhary et al., 2016
SRP100171	0–192	Chen et al., 2018
SRP100871	0–168	Siersbæk et al., 2017
SRP102079	0–4	Ryu et al., 2018
SRP109255	−48–24	Maroni et al., 2017
SRP119259	192	You et al., 2017

TABLE 2 | Microarray datasets of differentiating 3T3-L1 under pharmacological perturbations.

GEO ID	Type	Treatment	Ref.
GSE1458	TZD	pioglitazone/ rosiglitazone/ troglitazone	Hsiao et al., 2004
GSE42220	SFA	oleic acid/palmitic acid/palmitoleic acid/stearic acid	Shaw et al., 2013
GSE64075	NSAID	diclofenac/ ibuprofen/ indomethacin	Puhl et al., 2015

was applied using debCAM (Chen, 2020). Briefly, the scatter simplex of the cell mixture is a rotated and compressed version of the scatter simplex of the pure sub-populations (**Figure 1**). The sub-populations and their markers reside at the extremities of the shape. The one-vs.-everyone fold-change was calculated for all genes and used to select gene markers. Using the expression of the markers, the relative fractions of the sub-populations in the mixture were estimated by standardized averaging. The estimated fractions were used to deconvolute the mixed expressions into sub-populations specific profiles by least-square regressions.

2.4. Validating Markers and Populations

We validated the analysis using permutations of the identified markers and an independent dataset of gene expression profiles from primary adipocyte. To isolate the effect of the gene markers' set size and outliers, the fractions of the sub-populations were re-estimated using a smaller set chosen at random from the identified markers. The set size was chosen to be equal to that of the smallest sub-population. The same process was repeated several times for each perumutation of the markers. A supervised

version of CAM was applied using the expression profiles of human primary adipocytes using the full set of markers.

2.5. Gene Set Over-Representation

Lists of sub-populations specific markers were annotated using the gene ontology terms (org.Mm.eg.db) and tested for over-representation using ClusterProfiler (Yu et al., 2012; Carlson, 2020). For each gene set, the observed fraction of gene in the list was compared to the fraction expected by chance. *P*-values were calculated for each gene set and adjusted for multiple testing using false discovery rate (FDR).

2.6. Differential Expression Analysis

The effect of drugs and compounds treatment on gene expression of mature adipocytes was evaluated at the mixture and population levels. The differential gene expression between treatment and control groups was calculated using LIMMA (Ritchie et al., 2015). Population-specific gene expression analysis was applied to the same comparisons using PSEA (Kuhn et al., 2011). The previously identified markers of the three sub-populations were used to construct a reference population and perform one-to-one comparisons. *P*-values were calculated for each gene and adjusted for multiple testing using FDR.
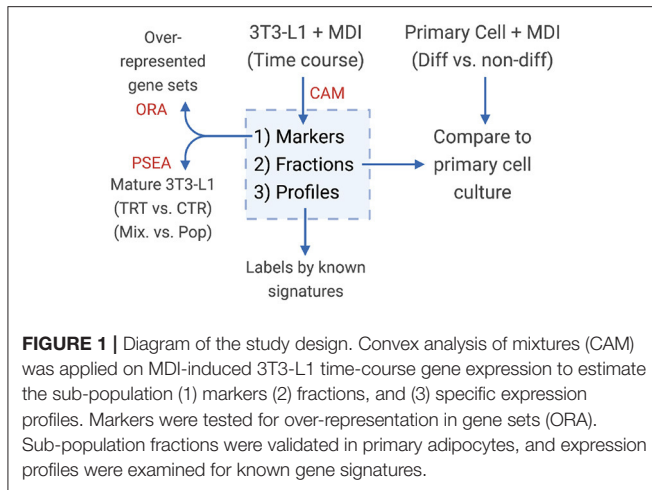
2.7. Software and Reproducibility

The analysis was conducted in an R and using Bioconductor packages (Huber et al., 2015; R Core Team, 2020). The software environment is available as a docker image (https://hub.docker.com/r/bcmslab/adipo_deconv). The code for reproducing the figures and tables in this document is open-source (https://github.com/BCMSLab/adipo_deconv).

3. RESULTS

3.1. Pre-adipocytes Cell Cultures Consist of Three Distinct Sub-populations

Averaging the gene expression in a given sample reflects the dominant phenotypes of cells during the course of differentiation. Pre-adipocytes were dissimilar to early and late-differentiated cells. We used as a measure of dissimilarity the distances among the samples within and in-between stage. The distances within the pre-adipocytes samples differed from early ($t = 24.07$; $P < 0.01$) and late-differentiated ($t = 30.05$; $P < 0.01$) samples. Moreover, the stage of differentiation explained about half the variance among the samples (**Figure 2A**). Un-differentiated and differentiating cells separated across the first two dimensions ($k = 2$; goodness-of-fit > 0.47) of the pairwise distances of the corresponding samples.

The 3T3-L1 cell cultures before and after the induction of differentiation consist of mixtures of cells. Average expression suffices to study the changes in gene expression over time, but the signal of each gene is potentially a composite of its expression in two or more cell types. These may differ in their phenotypes and/or genotypes. At each stage/time point of differentiation, the composition of the culture may differ, and the fractions of the sub-populations may change. Deconvoluting the gene expression of differentiating adipocytes would help address these points.

FIGURE 1 | Diagram of the study design. Convex analysis of mixtures (CAM) was applied on MDI-induced 3T3-L1 time-course gene expression to estimate the sub-population (1) markers (2) fractions, and (3) specific expression profiles. Markers were tested for over-representation in gene sets (ORA). Sub-population fractions were validated in primary adipocytes, and expression profiles were examined for known gene signatures.

FIGURE 2 | Dimensions reduction and deconvolution of differentiating adipocytes gene expression. Gene expression of MDI-induced 3T3-L1 pre-adipocytes ($n = 98$) and human primary adipocytes (HPP) ($n = 24$) was determined by RNA-Seq or microarrays at different times. **(A)** Euclidean distances among samples were calculated from the transformed gene counts. Samples were laid out in two dimensions to approximate the distances and colored by stage (S0, S1, or S2) for non-induced, early or, late differentiation. **(B)** The scatter simplex of the cell mixture is a rotated and compressed version of the scatter simplex of the pure sub-populations. In two dimensions, the sub-populations (P1, P2, and P3) and their markers reside at the extremities. **(C)** Using these markers, the relative fractions of the sub-populations in the mixture of differentiating 3T3-L1 were estimated by standardized averaging. **(D)** The relative fractions of the sub-populations in the mixture of differentiating HPP were estimated by standardized averaging of the same markers.

To determine the sub-populations in the cultures, we applied the convex analysis of mixtures (CAM) to the mixed gene expression profiles of differentiating adipocytes. The scatter simplex of the expression of the mixture is a transformed version of that of the pure sub-populations. When projected in low-dimensions, we observed three distinct archetypes characteristic of at least three groups of cells (**Figure 2B**). Gene markers of these sub-populations were identified as genes that are differentially expressed in one sub-populations vs. others (fold-change > 1; bootstrapped lower bound confidence interval > 0) (**Figure 3A** and **Supplementary File 1**). Markers were used to estimate the relative fractions and expression profiles of the subpopulations.

Various sets of sub-population markers gave patterns not very different from the above. We chose random sets of markers of the same size for the three sub-populations (bootstrapping). Lowering the number of markers and randomly re-sampling from the original set gave similar estimates of the fractions of the three populations at most time points (coefficient of variation < 10%; bias < 0.1) (**Figure 3B**).

3.2. Mature Adipocytes Originate From a Small Fraction of Pre-adipocytes

The starting cell culture is composed of disproportionate fractions of phenotypes that change with time (**Figure 2C**). One of the three groups (P1) makes up over 95%, and the rest is a mixture of the two other groups. Over the course of differentiation, the fraction of cells of the dominant group decreases to less than 40% of the late samples. In response to the induction stimulus, two factions of cells increase their share of the culture initially. Only one group (P2) continues to increase in size until the end and reaches more than 60% of the cells.
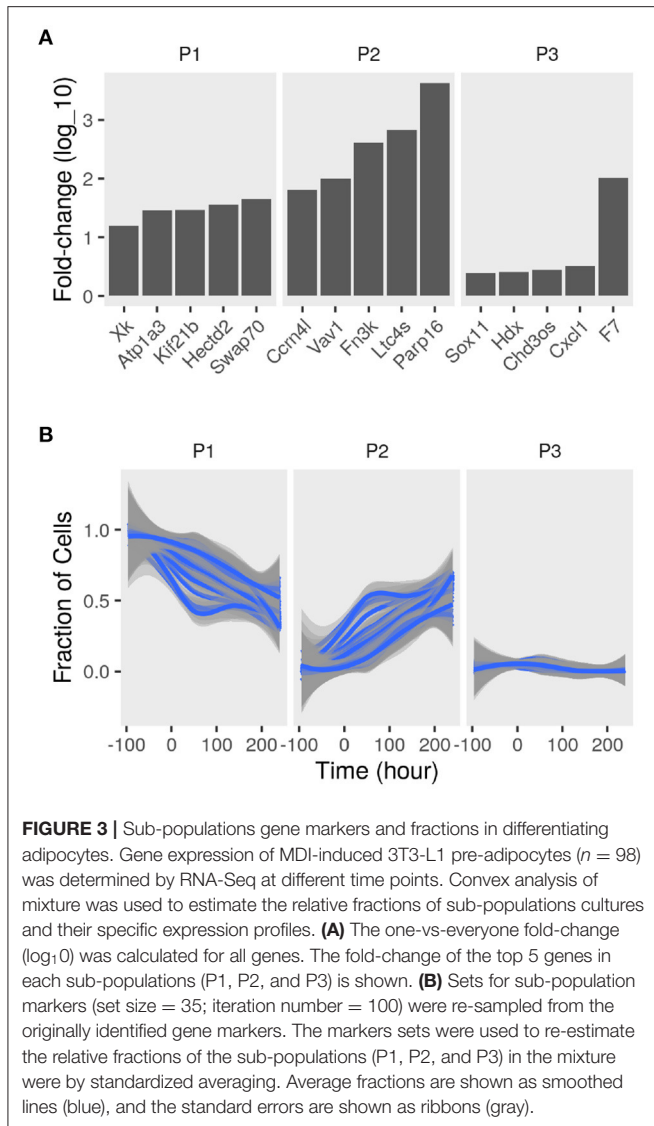
The three sub-populations were identified based on their specific gene expression profiles. Adipogenic and lipogenic genes marked the second sub-population (P2) as the ones responding to the induction stimulus and forming mature cells (**Figure 4A**). The first group of genes included the transcription factors *Pparg* and *Cebpb* involved in regulating the differentiation process. Those were highly expressed in P2 compared to the two other
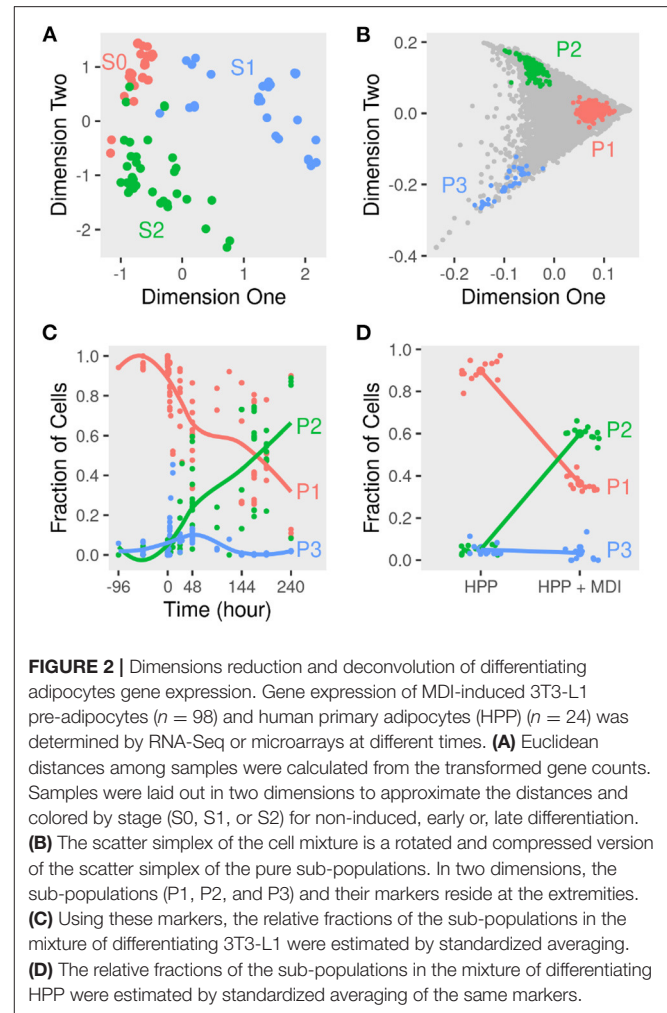
sub-populations (fold-change > 7; lower confidence interval > 0). The second group of genes code for key enzymes such as *Fasn* and *Lpl* in the lipid synthesis pathway and are essential for the accumulation of lipid. Those were also expressed at a higher level in P2 (fold-change > 1.5; lower confidence interval > 0).

The differences in gene expression between the two sub-populations extended to more than adipogenic and lipogenic related genes. Insulin metabolism-related genes were different expressed in the differentiating sub-population (P2) compared to the other two (fold-change > 3; lower confidence interval > 1.7) (**Figure 4B**). Although insulin receptor (*Insr*) and its substrate (*Irs1*) were expressed in the two major sub-populations, the glucose transporter solute carrier family 2 member 4 *(Slc2a4)* was expressed in P1 only. Platelet-derived growth factor subunit B (*Pdgfb*) exclusive expression in P2 indicates non-responsive cells to glucose.
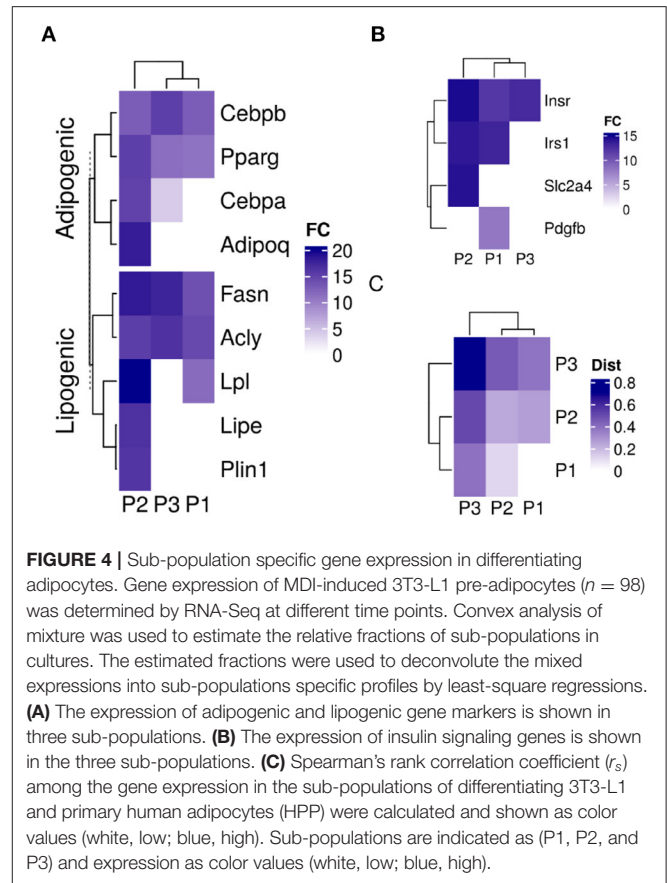
A similar mixture of sub-populations constitueed the cell culture of human primary pre-adipocytes. The fractions of primary cells corresponding to the previously identified markers resembled those estimated from 3T3-L1 cultures (**Figure 2D**). One sub-populations (P2) grew to dominate the final culture of MDI differentiated cells. The rest (P1 and P3) did not or

FIGURE 3 | Sub-populations gene markers and fractions in differentiating adipocytes. Gene expression of MDI-induced 3T3-L1 pre-adipocytes ($n = 98$) was determined by RNA-Seq at different time points. Convex analysis of mixture was used to estimate the relative fractions of sub-populations cultures and their specific expression profiles. **(A)** The one-vs-everyone fold-change ($\log_{10}$) was calculated for all genes. The fold-change of the top 5 genes in each sub-populations (P1, P2, and P3) is shown. **(B)** Sets for sub-population markers (set size = 35; iteration number = 100) were re-sampled from the originally identified gene markers. The markers sets were used to re-estimate the relative fractions of the sub-populations (P1, P2, and P3) in the mixture were by standardized averaging. Average fractions are shown as smoothed lines (blue), and the standard errors are shown as ribbons (gray).

FIGURE 4 | Sub-population specific gene expression in differentiating adipocytes. Gene expression of MDI-induced 3T3-L1 pre-adipocytes ($n = 98$) was determined by RNA-Seq at different time points. Convex analysis of mixture was used to estimate the relative fractions of sub-populations in cultures. The estimated fractions were used to deconvolute the mixed expressions into sub-populations specific profiles by least-square regressions. **(A)** The expression of adipogenic and lipogenic gene markers is shown in three sub-populations. **(B)** The expression of insulin signaling genes is shown in the three sub-populations. **(C)** Spearman's rank correlation coefficient (r_s) among the gene expression in the sub-populations of differentiating 3T3-L1 and primary human adipocytes (HPP) were calculated and shown as color values (white, low; blue, high). Sub-populations are indicated as (P1, P2, and P3) and expression as color values (white, low; blue, high).

only briefly responded. Gene expression profiles of the sub-populations in the primary cells were similar to that in the mouse cell model (**Figure 4C**). Because expression profiles come from two different types of cell lines, the expression of the sub-populations in each were strongly correlated (Spearman's rank correlation coefficient (r_s) > 0.6; P < 0.001). At the same time, the corresponding sub-populations were strongly correlated between the two cell lines. This was the case for P1 ($r_s = 0.47$; P < 0.001) and P2 ($r_s = 0.4$; P < 0.001). The only exception was the weaker correlations of P3 ($r_s = 0.2$) which was less than that of the mouse P3 with human P1 ($r_s = 0.4$; P < 0.001).

3.3. Sub-populations of Mature Adipocytes Respond Discrepantly to Various Compounds
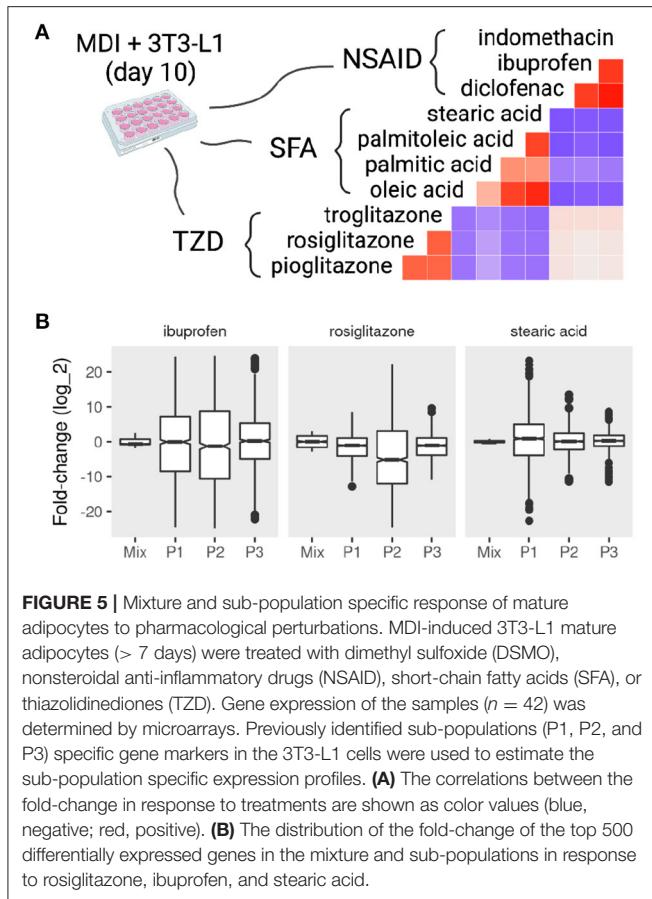
If, in fact, the mature adipocytes are a composite of different sub-populations, we expect to see a heterogeneous response to the treatment of different compounds. We collected data from MDI-induced adipocytes after day 7 and treated them with various compounds. These compounds included three thiazolidinediones (TZD), four short-chain fatty acids (SFA), and three nonsteroidal anti-inflammatory drugs (NSAID). The response of the cell mixture was evaluated by the regular differential expression, and that of the sub-populations was evaluated by population-specific expression analysis (PSEA). By comparing the response in the adipocyte mixture to that of the individual sub-populations, we could identify the effect of each compound on the mature adipocytes specifically.

We found that compounds from the same family elicited similar responses as expected (**Figure 5A**). The response to TZD: rosiglitazone, pioglitazone, troglitazone on the average gene expression was highly correlated (Pearson correlation coefficient (r) > 0.5; P < 0.001). Similarly, the fold-change in response to SFA and NSAID treatment was correlated (r > 0.4 and r > 0.6; P < 0.001). Only small correlations were found between the fold-change response between groups of compounds. TZD treatments were positively correlated with the NSAID and negatively correlated with that of SFA. NSAID and SFA were negatively correlated.

The response on the sub-population level to the treatment of different compounds often varied from the mixture and among each other (**Figure 5B** and **Supplementary File 2**). Treatment

FIGURE 5 | Mixture and sub-population specific response of mature adipocytes to pharmacological perturbations. MDI-induced 3T3-L1 mature adipocytes (> 7 days) were treated with dimethyl sulfoxide (DMSO), nonsteroidal anti-inflammatory drugs (NSAID), short-chain fatty acids (SFA), or thiazolidinediones (TZD). Gene expression of the samples (n = 42) was determined by microarrays. Previously identified sub-populations (P1, P2, and P3) specific gene markers in the 3T3-L1 cells were used to estimate the sub-population specific expression profiles. **(A)** The correlations between the fold-change in response to treatments are shown as color values (blue, negative; red, positive). **(B)** The distribution of the fold-change of the top 500 differentially expressed genes in the mixture and sub-populations in response to rosiglitazone, ibuprofen, and stearic acid.

with rosiglitazone produced more pronounced differential expression in P2 compared to the other populations (KS test $D^- < 0.15$; $P < 0.001$) as well as the mixture ($D^- < 0.24$; $P < 0.001$). A similar pattern was observed with the treatment of ibuprofen ($D^- < 0.13$; $P < 0.001$). By contrast, P1 was more different when adipocytes were treated with stearic acid ($D^- = 0.43$; $P < 0.001$).

3.4. A Dynamic Model of Two Sub-populations in a Fully Confluent Cell Culture

To show the validity of these observations, we constructed a model of sub-populations dynamics and compared it with further observations from the data. In this model, two sub-populations of different sizes grow at different rates in response to the environment in a fully confluent cell culture (**Figure 6A**). P2 is a small sub-population that grows at a rate r, $rP1$ while $P1$ is a large sub-population that decreases in size at a smaller or the same rate $-r$, $-rP1$. In this case, the total population size (N) can be estimated as $(P2 - P1)r$. This predicts that the total population (N) should decrease initially and bounces to the same size or higher after 2 days, given a small r. This system can be summarized as follows:

$$\frac{dP1}{dt} = -rP1 \quad (1)$$

$$\frac{dP2}{dt} = rP2 \quad (2)$$

$$\frac{dN}{dt} = (P2 - P1)r \quad (3)$$

A fully confluent cell culture imposes a constraint on the rate at which the sub-populations can grow or die (r). We chose a value that re-establishes the population size within a reasonable period (2 days). The rate at which one population can grow is bound by the rate at which the size of the other population shrinks. The values of the initial states ($P1$, $P2$, and N) were suggested by the previous analysis where the growing sub-population represents a small fraction of the total at the start of the cell culture.

3.5. Cell Cycle Index Fits the Dynamics of Two Sub-populations

We used a gene expression signature (*Mki67, Rb1, Hist1h2ae, Ccnb1, Cbx3, Gapdh, Ccnb2, and E2f1*) to estimate the cell cycle phase at population and sub-population levels (Dolatabadi et al., 2017). We found that, adipocytes at day 1 had higher cell cycle indecies than pre-adipocytes (KS test $D^- = 0.7$; $P < 0.01$) and adipocytes at day 2 ($D^- = 0.5$; $P > 0.05$) (**Figure 6B**). This pattern reflects the higher expression of cell cycle genes at day 1 compared to pre-adipocytes (enrichment score = 0.7; $P < 0.05$) and adipocytes at day 2 (enrichment score = 0.6; $P > 0.05$) which indicates a larger proportion of the cells in G0/S phase and, therefore, less active divisions (**Figure 6C**). This was consistent with the predicted population size of the toy model.

The sub-population specific differential expression confirmed that the two populations contribute differently to the expression of cell cycle genes and therefore to the active cell division (**Figure 6D**). This is clear from the correlation between the expression of the cell cycle genes and the reference (population signals). It is not clear from this analysis how to assign specific values to each sub-population at each of the two comparisons. This is because the expression is calculated for each of the eight genes individually. Moreover, the index reflects a continuous transition from G1 to S phase of the cell cycle rather than discrete categories. This would be affected by the size of the population at different time points and, therefore, their contribution to the overall expression of the genes constituting the index.

3.6. Resistance to Differentiation May Be Due to Structural Restrictions on the Cell and/or Commitment to Other Lineages

Several functional sets of genes were over-represented (count $\geq$ 3; false-discovery rate < 0.2) in the group of markers of the growing sub-population (P2). These were related to fat differentiation, nuclear receptors and, transcriptional activities (**Table 3**). More specifically, many of the characteristic markers were part of the process of *brown fat cell differentiation*. Molecular functions such as *nuclear receptor* and *steroid hormone activities* were over-represented in the same group. Several
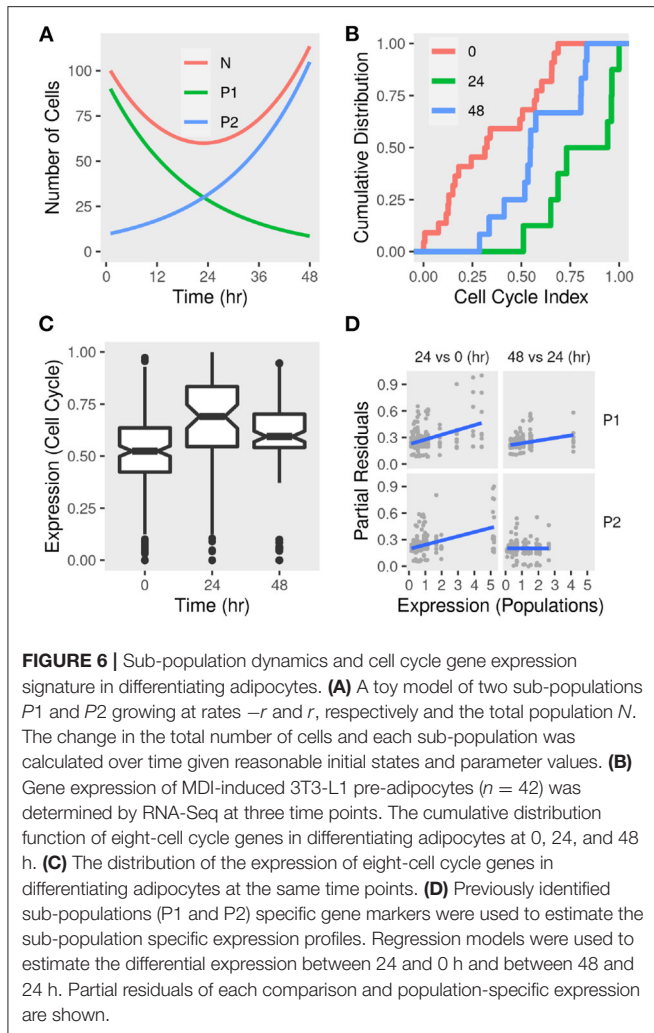
FIGURE 6 | Sub-population dynamics and cell cycle gene expression signature in differentiating adipocytes. **(A)** A toy model of two sub-populations $P1$ and $P2$ growing at rates $-r$ and r, respectively and the total population N. The change in the total number of cells and each sub-population was calculated over time given reasonable initial states and parameter values. **(B)** Gene expression of MDI-induced 3T3-L1 pre-adipocytes ($n = 42$) was determined by RNA-Seq at three time points. The cumulative distribution function of eight-cell cycle genes in differentiating adipocytes at 0, 24, and 48 h. **(C)** The distribution of the expression of eight-cell cycle genes in differentiating adipocytes at the same time points. **(D)** Previously identified sub-populations (P1 and P2) specific gene markers were used to estimate the sub-population specific expression profiles. Regression models were used to estimate the differential expression between 24 and 0 h and between 48 and 24 h. Partial residuals of each comparison and population-specific expression are shown.

TABLE 3 | Gene set over-representations in the sub-populations specific gene markers.

Pop.	Onto.	Term
P1	BP	establishment of mitotic spindle orientation
		actin filament bundle assembly
		actin cytoskeleton organization
		regulation of cell shape
		regulation of cell adhesion
	CC	spindle pole
		microtubule organizing center
		actin filament
		cell-cell junction
		bicellular tight junction
		actin cytoskeleton
	MF	actin binding
		actin filament binding
		transcription regulatory region DNA binding
		muscle alpha-actinin binding
P2	BP	actin cytoskeleton organization
		negative regulation of Rho protein signal transduction
		brown fat cell differentiation
	CC	stress fiber
		RNA polymerase II transcription factor complex
	MF	nuclear receptor activity
		steroid hormone receptor activity
		nuclear receptor transcription coactivator activity
P3	BP	positive regulation of gene expression
		neuron differentiation
		keratinocyte differentiation
		keratinization
	CC	cornified envelope
		extracellular matrix
	MF	signaling receptor binding
		structural molecule activity
		structural constituent of cytoskeleton

of these markers localized to cellular components such as *transcription factor complex*. Alternatively, the phenotype that did not show these characteristics could be thought of as resisting the induction stimulus. Indeed, completely different processes and functions were over-represented by the markers of the other two populations.

The markers of the other sub-populations were enriched in terms related to cellular structural processes and components (**Table 3**). The first group (P1) had multiple markers related to *mitotic spindles*, *microtubules*, *actin* and *cytoskeleton*. All are essential components of the cell structure and necessary for allowing morphological changes. In addition to these terms, several processes related to the assembly and organization of these components were enriched by P1 markers.

The third group (P3) had terms related to structural components, namely the *extracellular matrix* and molecular functions *structural constituent of cytoskeleton* enriched by its markers. But also biological processes terms related to the differentiation of cell types other than adipocytes, *neurons* and *keratinocytes differentiation* were over-represented in the P3 set of markers.
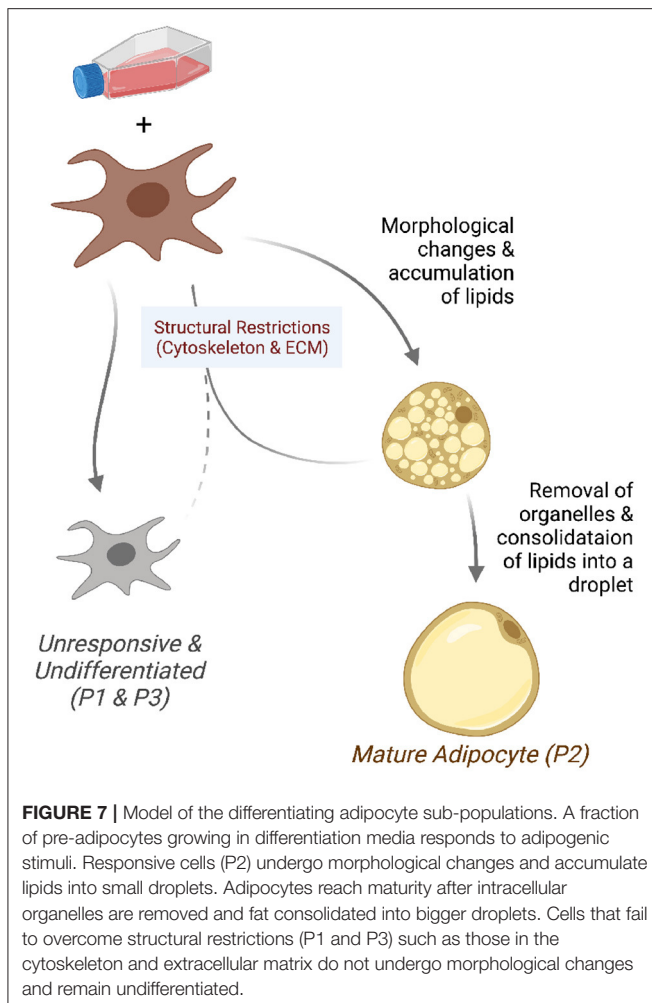
4. DISCUSSION

We found that the pre-adipocyte cell culture is composed of three distinct sub-populations. After the differentiation induction, a minority of cells grows and transforms into mature adipocytes. This fraction of cells have a gene expression profile conducive to activities related to the nuclear receptors, fat differentiation, and gene transcription. The sub-populations of cells that do not respond to induction are under structural constraints or are committed to differentiate into other cell types. **Figure 7** summarizes these observations in a graphical representation, which we further discuss below.

Identifying the factors that render fibroblasts responsive or resistant to the induction medium would help achieve

FIGURE 7 | Model of the differentiating adipocyte sub-populations. A fraction of pre-adipocytes growing in differentiation media responds to adipogenic stimuli. Responsive cells (P2) undergo morphological changes and accumulate lipids into small droplets. Adipocytes reach maturity after intracellular organelles are removed and fat consolidated into bigger droplets. Cells that fail to overcome structural restrictions (P1 and P3) such as those in the cytoskeleton and extracellular matrix do not undergo morphological changes and remain undifferentiated.

better differentiation efficiency. More importantly, it would improve our understanding of adipose tissue composition and how to encourage or discourage the terminal maturation of the adipocytes.

In the context of CAM, exclusively enriched genes in a subset of cells are the markers for the set (Wang et al., 2016). The expression of other genes (co-expressed) is the combination of their expression in the sub-populations. The expression of all genes in the mixture forms a k-dimensional simplex with a telling geometry. The simplex shape encloses the data, and the sub-populations with their markers are at the extremes.

To demonstrate the validity of these findings, one has to show that (1) pre-adipocyte cell cultures consist of distinct populations and that (2) some cells but not others can increase their population size. Differences intrinsic to the cells, in the ability to respond to MDI, in the kind of cell they become, or a combination of the above could give rise to the observed pattern. Although the size of the sub-populations is dissimilar to start with, it only changes after the differentiation induction. Because only the fractions of each sub-population rather than their absolute size were estimated, the inflection is possible to

achieve with either the duplication of the responsive cells or the shrinkage of the others. Indeed, pre-adipocytes can replicate after the induction (Bernlohr et al., 1985). Mature adipocytes can divide and redistribute the lipid droplets between the daughter cells (Nagayama et al., 2007).

Previous studies pointed out the heterogeneous nature of adipocyte cell cultures and adipose tissue progenitors. Often these studies focused on the phenotypic and molecular differences among the identified populations. Most attribute these differences to causes in two categories: asynchronous differentiation and lineage commitment. The first refers to cells that are essentially similar but only take longer to differentiate (Loo et al., 2009). Therefore, at any specific point, some cells would lack the characteristics of mature adipocytes. For example, poor-lipid accumulating cells would form larger droplets of lipid by merely prolonged simulation (Le and Cheng, 2009). While our study does not rule out this possibility, it makes counter observations. First, the differences between the sub-populations are stark from the very beginning (pre-adipocytes). Second, the lagging sub-populations do not show any sign of differentiation either in adipogenic regulation or lipid synthesis. Finally, the dataset extends up to 2 weeks, twice the period of typical maturation.

Studies based on the adipose tissue progenitor cells suggest that some cells are either incapable of differentiating into mature adipocytes, originate in certain depots, or can only differentiate into other cell types (brown/beige) adipocytes. One study suggested that mouse visceral, but not subcutaneous, white adipose tissue arises from cells expressing Wilms tumor (*Wt1*) gene (Chau and Hastie, 2014). Another found that the adipocyte's ability to accumulate fat is heritable (Katz et al., 2014). Histone modification of PPARG promoter marked cells which can de-differentiate and re-differentiate. It's even been shown that a minority sub-population lacking the typical adipocyte features might take on a regulatory role to influence its surroundings in a paracrine manner (Schwalie et al., 2018).

Here, we suggest that the mature adipocytes arise from a minority of the pre-adipocytes. Most pre-adipocytes cannot differentiate properly or accumulate lipids. This lack of responsiveness to the differentiation medium is either due to or associated with structural restraints. The major factors that differentiate the two types of cells are genes involved in the cellular cytoskeleton and actin filaments. Molecular differences also exist, most importantly in regulating adipogenesis, lipid synthesis, and insulin signaling.

This study was limited to publicly available gene expression data of the 3T3-L1 cell model and primary adipocytes. Experimental validation is required to identify the role of individual protein markers, their distributions in the sub-populations, and the cellular processes they are involved in. For example, would manipulating the proteins involved in organizing the cell structure or extracellular matrix affect the cell's ability to differentiate? These experiments should be done on single cells or separate the populations based on their potential to differentiate, which we plan to do in future investigations.

DATA AVAILABILITY STATEMENT

The data analyzed in the study are deposited in Figshare, https://doi.org/10.6084/m9.figshare.9906182 and https://doi.org/10.6084/m9.figshare.13088624.

AUTHOR CONTRIBUTIONS

MA conceived, performed, and reported the analysis. TL helped with the conceptualization and the writing of the manuscript. DK acquired the funding, supervised, and contributed to the concepts and the writing of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2021.753042/full#supplementary-material

REFERENCES

Ahmed, M., and Kim, D. R. (2019). Modelling the gene expression and the DNA-binding in the 3T3-L1 differentiating adipocytes. *Adipocyte* 8, 401–411. doi: 10.1080/21623945.2019.1697563

Ahmed, M., Min, D. S., and Kim, D. R. (2020). Curated gene expression dataset of differentiating 3T3-L1 adipocytes under pharmacological and genetic perturbations. *Adipocyte* 9, 600–608. doi: 10.1080/21623945.2020.1829852

Al Adhami, H., Evano, B., Le Digarcher, A., Gueydan, C., Dubois, E., Parrinello, H., et al. (2015). A systems-level approach to parental genomic imprinting: the imprinted gene network includes extracellular matrix genes and regulates cell cycle exit and differentiation. *Genome Res.* 25, 353–367. doi: 10.1101/gr.175919.114

Andrews, S. (2020). FastQC: a quality control tool for high throughput sequence data.

Bernlohr, D. A., Bolanowski, M. A., Kelly, T. J., and Lane, M. D. (1985). Evidence for an increase in transcription of specific mRNAs during differentiation of 3T3-L1 preadipocytes. *J. Biol. Chem.* 260, 5563–5567. doi: 10.1016/S0021-9258(18)89059-7

Brier, A. S. B., Loft, A., Madsen, J. G., Rosengren, T., Nielsen, R., Schmidt, S. F., et al. (2017). The KDM5 family is required for activation of pro-proliferative cell cycle genes during adipocyte differentiation. *Nucleic Acids Res.* 45, 1743–1759. doi: 10.1093/nar/gkw1156

Brunmeir, R., Wu, J., Peng, X., Kim, S. Y., Julien, S. G., Zhang, Q., et al. (2016). Comparative transcriptomic and epigenomic analyses reveal new regulators of murine brown adipogenesis. *PLoS Genet.* 12:e1006474. doi: 10.1371/journal.pgen.1006474

Buchner, D. A., Charrier, A., Srinivasan, E., Wang, L., Paulsen, M. T., Ljungman, M., et al. (2015). Zinc finger protein 407 (ZFP407) regulates insulin-stimulated glucose uptake and glucose transporter 4 (Glut4) mRNA. *J. Biol. Chem.* 290, 6376–6386. doi: 10.1074/jbc.M114.623736

Carlson, M. (2020). org.Mm.eg.db: Genome wide annotation for Human.

Chau, Y.-Y., and Hastie, N. (2014). Wt1, the mesothelium and the origins and heterogeneity of visceral fat progenitors. *Adipocyte* 4, 217–221. doi: 10.4161/21623945.2014.985009

Chaudhary, N., Gonzalez, E., Chang, S. H., Geng, F., Rafii, S., Altorki, N. K., et al. (2016). Adenovirus protein E4-ORF1 activation of PI3 kinase reveals differential regulation of downstream effector pathways in adipocytes. *Cell Rep.* 17, 3305–3318. doi: 10.1016/j.celrep.2016.11.082

Chen, L. (2020). debCAM: Deconvolution by Convex Analysis of Mixtures.

Chen, X., Ayala, I., Shannon, C., Fourcaudot, M., Acharya, N. K., Jenkinson, C. P., et al. (2018). The diabetes gene and wnt pathway effector TCF7L2 regulates adipocyte development and function. *Diabetes* 67, 554–568. doi: 10.2337/db17-0318

Dolatabadi, S., Candia, J., Akrap, N., Vannas, C., Tomic, T. T., Losert, W., et al. (2017). Cell cycle and cell size dependent gene expression reveals distinct subpopulations at single-cell level. *Front. Genet.* 8:1. doi: 10.3389/fgene.2017.00001

Duteil, D., Metzger, E., Willmann, D., Karagianni, P., Friedrichs, N., Greschik, H., et al. (2014). LSD1 promotes oxidative metabolism of white adipose tissue. *Nat. Commun.* 5:4093. doi: 10.1038/ncomms5093

Green, H., and Kehinde, O. (1975). An established preadipose cell line and its differentiation in culture II. Factors affecting the adipose conversion. *Cell* 5, 19–27. doi: 10.1016/0092-8674(75)90087-2

Guilak, F., Cohen, D. M., Estes, B. T., Gimble, J. M., Liedtke, W., and Chen, C. S. (2009). Control of stem cell fate by physical interactions with the extracellular matrix. *Cell Stem Cell.* 5, 17–26. doi: 10.1016/j.stem.2009.06.016

Herrington, D. M., Mao, C., Parker, S. J., Fu, Z., Yu, G., Chen, L., et al. (2018). Proteomic architecture of human coronary and aortic atherosclerosis. *Circulation* 137, 2741–2756. doi: 10.1161/CIRCULATIONAHA.118.034365

Horvath, S., and Dong, J. (2008). Geometric interpretation of gene coexpression network analysis. *PLoS Comput Biol.* 4:e1000117. doi: 10.1371/journal.pcbi.1000117

Hsiao, A., Worrall, D. S., Olefsky, J. M., and Subramaniam, S. (2004). Variance-modeled posterior inference of microarray data: Detecting gene-expression changes in 3T3-L1 adipocytes. *Bioinformatics* 20, 3108–3127. doi: 10.1093/bioinformatics/bth371

Huber, W., Carey, V. J., Gentleman, R., Anders, S., Carlson, M., Carvalho, B. S., et al. (2015). Orchestrating high-throughput genomic analysis with Bioconductor. *Nat. Methods* 12, 115–121. doi: 10.1038/nmeth.3252

Katz, L. S., Geras-Raaka, E., and Gershengorn, M. C. (2014). Heritability of fat accumulation in white adipocytes. *Am. J. Physiol. Endocrinol. Metabol.* 307, E335-E344. doi: 10.1152/ajpendo.00075.2014

Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12, 357–360. doi: 10.1038/nmeth.3317

Kuhn, A., Thu, D., Waldvogel, H. J., Faull, R. L., and Luthi-Carter, R. (2011). Population-specific expression analysis (PSEA) reveals molecular changes in diseased brain. *Nat. Methods* 8, 945–947. doi: 10.1038/nmeth.1710

Le, T. T., and Cheng, J. X. (2009). Single-cell profiling reveals the origin of phenotypic variability in adipogenesis. *PLoS ONE* 4:e5189. doi: 10.1371/journal.pone.0005189

Lee, K. Y., Luong, Q., Sharma, R., Dreyfuss, J. M., Ussar, S., and Kahn, C. R. (2019). Developmental and functional heterogeneity of white adipocytes within a single fat depot. *EMBO J.* 38, 1–19. doi: 10.15252/embj.201899291

Leek, J. T., Johnson, W. E., Parker, H. S., Fertig, E. J., Jaffe, A. E., Storey, J. D., et al. (2020). sva: surrogate variable analysis.

Liao, Y., Smyth, G. K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923–930. doi: 10.1093/bioinformatics/btt656

Lim, G. E., Albrecht, T., Piske, M., Sarai, K., Lee, J. T., Ramshaw, H. S., et al. (2015). 14-3-3ζ coordinates adipogenesis of visceral fat. *Nat. Commun.* 6:7671. doi: 10.1038/ncomms8671

Lo, K. A., Labadorf, A., Kennedy, N. J., Han, M. S., Yap, Y. S., Matthews, B., et al. (2013). Analysis of *in vitro* insulin-resistance models and their physiological relevance to *in vivo* diet-induced adipose insulin resistance. *Cell Rep.* 5, 259–270. doi: 10.1016/j.celrep.2013.08.039

Loo, L. H., Lin, H. J., Singh, D. K., Lyons, K. M., Altschuler, S. J., and Wu, L. F. (2009). Heterogeneity in the physiological states and pharmacological responses of differentiating 3T3-L1 preadipocytes. *J. Cell Biol.* 187, 375–384. doi: 10.1083/jcb.200904140

Maroni, G., Tkachuk, V. A., Egorov, A., Morelli, M. J., Luongo, R., Levantini, E., et al. (2017). Prep1 prevents premature adipogenesis of mesenchymal progenitors. *Sci. Rep.* 7, 15573. doi: 10.1038/s41598-017-15828-1

Nagayama, M., Uchida, T., and Gohara, K. (2007). Temporal and spatial variations of lipid droplets during adipocyte division and differentiation. *J. Lipid Res.* 48, 9–18. doi: 10.1194/jlr.M600155-JLR200

Nobusue, H., Onishi, N., Shimizu, T., Sugihara, E., Oki, Y., Sumikawa, Y., et al. (2014). Regulation of MKL1 via actin cytoskeleton dynamics drives adipocyte differentiation. *Na.t Commun.* 5:3368. doi: 10.1038/ncomms4368

Park, Y.-K., Wang, L., Giampietro, A., Lai, B., Lee, J.-E., and Ge, K. (2017). Distinct roles of transcription factors KLF4, Krox20, and peroxisome proliferator-activated receptor γ in adipogenesis. *Mol. Cell Biol.* 37, 00554–00516. doi: 10.1128/MCB.00554-16

Puhl, A. C., Milton, F. A., Cvoro, A., Sieglaff, D. H., Campos, J. C. L., Bernardes, A., et al. (2015). Mechanisms of peroxisome proliferator activated receptor γ regulation by non-steroidal anti-inflammatory drugs. *Nucl. Recept. Signal* 13:e004. doi: 10.1621/nrs.13004

R Core Team (2020). *R: A Language and Environment for Statistical Computing.*

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43, e47. doi: 10.1093/nar/gkv007

Ryu, K. W., Nandu, T., Kim, J., Challa, S., DeBerardinis, R. J., and Lee Kraus, W. (2018). Metabolic regulation of transcription through compartmentalized NAD+biosynthesis. *Science* 360:eaan5780. doi: 10.1126/science.aan5780

Schwalie, P. C., Dong, H., Zachara, M., Russeil, J., Alpern, D., Akchiche, N., et al. (2018). A stromal cell population that inhibits adipogenesis in mammalian fat depots. *Nature* 559, 103–108. doi: 10.1038/s41586-018-0226-8

Shaw, B., Lambert, S., Wong, M. H. T., Ralston, J. C., Stryjecki, C., and Mutch, D. M. (2013). Individual saturated and monounsaturated fatty acids trigger distinct transcriptional networks in differentiated 3T3-L1 preadipocytes. *J. Nutrigenet. Nutrigenomics* 6, 1–15. doi: 10.1159/000345913

Siersbæk, R., Baek, S., Rabiee, A., Nielsen, R., Traynor, S., Clark, N., et al. (2014). Molecular architecture of transcription factor hotspots in early adipogenesis. *Cell Rep.* 7, 1434–1442. doi: 10.1016/j.celrep.2014.04.043

Siersbæk, R., Madsen, J. G. S., Javierre, B. M., Nielsen, R., Bagge, E. K., Cairns, J., et al. (2017). Dynamic rewiring of promoter-anchored chromatin loops during adipocyte differentiation. *Mol. Cell.* 66, 420–435. doi: 10.1016/j.molcel.2017.04.010

Smith, Q., Stukalin, E., Kusuma, S., Gerecht, S., and Sun, S. X. (2015). Stochasticity and spatial interaction govern stem cell differentiation dynamics. *Sci. Rep.* 5, 1–10. doi: 10.1038/srep12617

Verbanck, M., Canouil, M., Leloire, A., Dhennin, V., Coumoul, X., Yengo, L., et al. (2017). Low-dose exposure to bisphenols A, F and S of human primary adipocyte impacts coding and non-coding RNA profiles. *PLoS ONE* 12:e0179583. doi: 10.1371/journal.pone.0179583

Wang, N., Hoffman, E. P., Chen, L., Chen, L., Zhang, Z., Liu, C., et al. (2016). Mathematical modelling of transcriptional heterogeneity identifies novel markers and subpopulations in complex tissues. *Sci. Rep.* 6:18909. doi: 10.1038/srep18909

Yang, W., Guo, X., Thein, S., Xu, F., Sugii, S., Baas, P. W., et al. (2013). Regulation of adipogenesis by cytoskeleton remodelling is facilitated by acetyltransferase MEC-17-dependent acetylation of α-tubulin. *Biochem. J.* 449, 606–612. doi: 10.1042/BJ20121121

You, D., Nilsson, E., Tenen, D. E., Lyubetskaya, A., Lo, J. C., Jiang, R., et al. (2017). Dnmt3a is an epigenetic mediator of adipose insulin resistance. *eLife* 6:e30766. doi: 10.7554/eLife.30766.042

Yourek, G., Hussain, M. A., and Mao, J. J. (2007). Cytoskeletal changes of mesenchymal stem cells during differentiation. *ASAIO J.* 53, 219–228. doi: 10.1097/MAT.0b013e31802deb2d

Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *OMICS* 16, 284–287. doi: 10.1089/omi.2011.0118

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